Universidade de Aveiro Departamento de Química
Ano 2017

Carolina Oliveira Pandeirada

Caracterização dos hidratos de carbono de *Nannochloropsis oculata* e sua utilização em *microarrays*

Characterization of carbohydrates from *Nannochloropsis oculata* and their use in microarrays

Universidade de Aveiro Departamento de Química
Ano 2017

Carolina Oliveira Pandeirada

Caracterização dos hidratos de carbono de *Nannochloropsis oculata* e sua utilização em *microarrays*

Characterization of carbohydrates from *Nannochloropsis oculata* and their use in microarrays

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Bioquímica, Ramo Alimentar, realizada sob a orientação científica da Doutora Cláudia Nunes, investigadora de pós-doutoramento do Departamento de Química da Universidade de Aveiro, e do Doutor Manuel A. Coimbra, Professor associado com agregação do Departamento de Química da Universidade de Aveiro.

Thanks are due to FCT through national founds and FEDER, within the PT2020 Partnership Agreement, for funding QOPNA (FCT UID/QUI/00062/2013)

To my family and friends

O júri/ The jury

Presidente/ President

Professora Doutora Rita Maria Pinho Ferreira
Professora auxiliar do Departamento de Química da Universidade de Aveiro

Arguente/ Arguent

Doutora Maria Angelina de Sá Palma
Ivestigadora da Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa

Orientador/ Supervisor

Professor Doutor Manuel António Coimbra Rodrigues da Silva
Professor associado com agregação do Departamento de Química da Universidade de Aveiro

Agradecimentos/ Acknowledgments

Os meus sinceros agradecimentos a/ My sinceriously thanks to:

À Drª. Cláudia Nunes, minha orientadora, por toda a confiança que depositou em mim ao longo deste ano, por toda a disponibilidade, por todas as interrogações e pelos inúmeros conhecimentos transmitidos. A si, Cláudia, agradeço todos os conselhos, o carinho e a força. Um grande Obrigada ☺

Ao Prof. Manuel A. Coimbra, meu co-orientador, pela oportunidade que me deu de poder ingressar num desafio que se tornou muito enriquecedor. Agradeço também toda a sabedoria partilhada e todas as interrogações feitas com olhar crítico ao meu trabalho, as quais me fizeram pensar, refletir e aprender. Obrigada!

À Drª. Angelina Palma da Universidade NOVA de Lisboa pela prestabilidade e acompanhamento ao longo dos ensaios de microarrays. Agradeço também a simpatia, os conhecimentos transmitidos e o olhar atento, crítico e preocupado sobre o meu trabalho. Obrigada ☺

À Élia, que em muito me ajudou a progredir e ganhar aptidões a nível laboratorial. Agradeço-te também toda a amizade. Obrigada ☺

À Viviana e à Drª. Benedita da Universidade NOVA de Lisboa por todo o acompanhamento, disponibilidade e conhecimentos que me transmitiram ao longo do tempo que passei na NOVA, foram incansáveis, Obrigada ☺ Em particular, quero agradecer aos membros dos grupos GlycoLab e X-tal da NOVA pelo acolhimento, por disponibilizarem os laboratboratórios para realizar os ensaios e pela partilha de conhecimentos. Obrigada!

Aos meus companheiros de laboratório, Joana, Pedro, Sónia, Angélica, Andreia, Guido e restantes membros do grupo ResPoStA, não só pelo companheirismo, diversão e "aturarem a minha criança", mas acima de tudo pelos conhecimentos e chamadas de atenção. Obrigada ☺

Aos meus amigos mais próximos, Cristiana, Rita, Sofia, Catarina, Catarina, Ana, Marco, Lisandra e Lino pelos momentos de diversão e apoio. Obrigada ☺

À família Calado, pelos conhecimentos transmitidos, pelo carinho e apoio. Obrigada ☺

À minha irmã, Mariana, por ser um exemplo e por me dar não só bons conselhos e visões, mas por toda a amizade e forças ao longo destes anos. Obrigada ☺

Á minha família, Pai, Mãe e Avôs, por até aos dias de hoje me acolherem e pela educação que me deram. A vós devo a pessoa que hoje sou! Obrigada ☺

E porque, "*Cada um é responsável por multiplicar os dons que Deus lhe deu*", eu agradeço ter tido a sorte de ter cada um de vós junto de mim para me ajudar numa parte desta difícil missão! Um grande bem-haja ☺

Palavras-chave

Nannochloropsis oculata; Polissacarídeos; Glicolípidos; Microarrays; Interações proteínas-hidratos de carbono.

Resumo

As microalgas marinhas são fáceis de cultivar e acumulam compostos, como polissacarídeos (PS) e glicolípidos (GL), aos quais têm sido atribuídas atividades biológicas. O presente trabalho teve como objetivo estudar as características estruturais dos PS e dos hidratos de carbono dos GL da microalga marinha *Nannochloropsis oculata*. A biomassa foi desengordurada, extraída com água quente (HW) e ultrafiltrada. Isto originou um resíduo rico em (1→4)-Glc, atribuído à presença de celulose na parede celular, um filtrado rico em manitol e um retentato (*Mw*>10 kDa) que continha os PS solúveis. Os PS solúveis foram fracionados por precipitação em etanol originando três frações, Et50, Et85 e EtSN. Os PS da fração Et85 foram sujeitos a fraccionamento por cromatografia de troca aniónica, obtendo-se as frações #1, #2 e #3. Todas as frações contiveram ácidos urónicos e grupos sulfato, sendo este conteúdo quase negligenciável na fração #1.

As análises de açúcares, juntamente com os *microarrays* usando proteínas de especificidades conhecidas, permitiram verificar que os PS da fração Et50 são constituídos por (β1→4)-Glc, exibindo também domínios de (β1→3, β1→4)-glucanas. Os PS das frações Et85 e EtSN incluíram maioritariamente Rha e Man, contendo também outros açúcares em menores quantidades, Xyl, Fuc e Gal. As frações Et85, #1, #2 e #3, e EtSN possuiram principalmente (1→2)-Rha, com uma ramificação na posição C-3. As frações Et85, #2 e #3 continham também o resíduo substituído 2,3,4-Rha. Na fração Et85 os grupos sulfato encontraram-se na posição C-3, e simultaneamente nas posições C-3 e C-4 de (1→2)-Rha. As frações #2 e #3, além de resíduos de Rha continham Xyl, Fuc e Gal, sugerindo que estes fazem parte de heteropolissacarídeos aniónicos sulfatados. A Man, nas frações Et85 e EtSN, encontrou-se maioritariamente como (1→3)- e (1→4)-Man, fazendo parte de PS não-aniónicos na fração Et85. A análise por *microarrays* evidenciou que a Man deve encontrar-se em configuração α devido à interação com a ConA. A proteína BT0996 da bactéria *Bacteroides thetaiotaomicron*, do sistema digestivo humano, usada nos *microarrays,* mostrou que os PS nas frações Et50, Et85, #1 e #3 podem conter um epítopo comum. Esta interação evidencia que os PS de *N. oculata* poderão ser metabolizados no sistema digestivo humano.

Os GL de *N. oculata* ocorrem principalmente como mono- e digalactolipídos, sendo a sua presença confirmada pelos *microarrays* por interação com a proteína RCA_{120}. A análise de açúcares evidenciou também que a fração lipídica deve conter sulfoquinovosil diacilglicerol.

Considerando os resultados obtidos, a combinação das análises de açúcares com os *microarrays* mostraram ser uma abordagem adequada para identificar as características estruturais dos PS e GL da microalga *N. oculata.*

Keywords

Nannochloropsis oculata; Polysaccharides; Glycolipids; Microarrays; Carbohydrate-protein interactions.

Abstract

Marine microalgae are of easy culturing and able to accumulate compounds, such as polysaccharides (PS) and glycolipids (GL), which have been attributed biological activities. The present work aimed to study the structural features of the PS and the carbohydrates of the GL of the marine microalga *Nannochloropsis oculata* using carbohydrate analyses and microarrays. Biomass was defatted, extracted with hot-water (HW), and ultrafiltered. This yielded a residue rich in (1→4)-Glc that was assigned to the presence of a cellulose cell-wall, one filtrate rich in mannitol, and one rentante (*Mw*>10 kDa) that comprised the soluble PS. The soluble PS were fractionated through ethanol precipitation giving three fractions, Et50, Et85 and EtSN. The PS in Et85 were further fractionated through an anion-exchange chromatography, yielding fractions #1, #2, and #3. All PS fractions comprised uronic acids and sulfate groups, being this content almost negligible in fraction #1.

Carbohydrate analyses together with microarrays experiments using proteins with known carbohydrate-binding specificities allowed to verify that the PS in Et50 are dominated by (β1→4)-Glc and exhibited also domains of (β1→3, β1→4)-glucans. PS in Et85 and EtSN fractions were composed by Rha and Man, comprising other sugars in lower content, Xyl, Fuc, and Gal. Fractions Et85, #1, #2, #3, and EtSN comprised mainly (1→2)-Rha with a branching point occurring at position C-3. Fractions Et85, #2, and #3 included also the substituted residue 2,3,4-Rha. In fraction Et85, the sulfate group occurred at position C-3, and simultaneously at positions C-3 and C-4 of (1→2)-Rha. Fractions #2 and #3, besides Rha residues, contained Xyl, Fuc, and Gal, suggesting that these are part of anionic sulfated heteropolysaccharides. The Man, in Et85 and EtSN, occurred mainly as (1→3)- and (1→4)-Man residues, being these part of non-anionic PS in fraction Et85. Carbohydrate microarrays highlighted that Man may be present in α-configuration due to the interaction with ConA. The protein BT0996 of the *Bacteroides thetaiotaomicron*, a human gut bacteria, used in microarrays, displayed that the PS in fractions Et50, Et85, #1, and #3 may share a common epitope. This interaction is remarkable as it shows that these PS are probably metabolized in the human gut.

The GL occur mainly as mono- and digalactolipids, being the presence of galactolipids confirmed through microarrays by interaction with RCA_{120}. Carbohydrate analyses highlighted that the lipid fraction may comprise sulfoquinovosyl diacylglycerol.

Considering the results obtained, the joint use of carbohydrate analyses with microarrays showed to be a suitable approach to identify the structural features of the PS and the GL of the marine microalga *N. oculata*.

Contents

List of Figures

List of Tables

List of Abbreviations

Ab (or Ig) – Antibody

ADHP - *N*-aminoacetyl-*N*-(9-anthracenylmethyl)-1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine

AF – Alexa-Fluor

AnGal – 3,6-anhydro-Gal

AOPE – Aminooxy-functionalized DHPE

AraOH – Arabinitol

BSA – Bovine serum albumin

C – Cholesterol

CBM – Carbohydrate-binding module

CBPs – Carbohydrate-binding proteins

CE – Carbohydrate esterase

ConA – Concanavalin A

CRD – Carbohydrate-recognition domain

C-type – Ca^{2+}-dependent

Cy – Cyanine

DAG – Diacylglycerol

DB – Degree of branching

DC-SIGN – Dendritic cell-specific ICAM-3 grabbing nonintegrin

DE – Degree of esterification

DGDG – Digalactosyl diacylglycerols

DHPE - 1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine

dH_2O – Destilled water

DMSO – Dimethyl sulfoxide

DP – Degree of polymerization

dw – Dry weight

EBV-EA – Epstein-Barr virus-associated early antigen

EPA – Eicosapentaenoic acid

EPS – Exopolysaccharides

ER – Endoplasmic reticulum

Et50 – Precipitate with 50 % EtOH (v/v)

Et85 – Precipitate with 85 % EtOH (v/v)

EtSN – Supernatant of the 85 % EtOH (v/v)

FA – Fatty acid

FI – Fluorescence intensity

FID – Flame ionization detector

FITC – Fluorescein isothiocyanate

Fraction #1 (Et85_#1) – Fraction eluted with 0 M NaCl from the anion-exchange chromatography of Et85

Fraction #2 (Et85_#2) – Fraction eluted with 0.125 M NaCl from the anion-exchange chromatography of Et85

Fraction #3 (Et85_#3) – Fraction eluted with 0.250 M NaCl from the anion-exchange chromatography of Et85

FucOH – Fucitol

GalA – Galacturonic acid

GalOH - Galactitol

GlcA – Glucuronic acid

GC – Gas chromatography

GH – Glycoside hydrolase

GIT – Gastrointestinal tract

GL – Glycolipids

GlcOH – Glucitol

GPC – Gel permeation chromatography

GT – Glycosyltransferase

HG – Homogalacturonnan
HMG – High methylated galacturonate
HW – Hot-water
IL – Interleukin
Ig – Immunoglobuline
IMAC – Immobilized-metal affinity
IPTG – β-D-1-thiogalactopyranoside
Kan – Kanamycin
LA – LB-agar
LacNAc – *N*-Acetyllactosamine
LB – Luria-Bertani
LC-PUFA – Long-chain polyunsaturated fatty acid
LMG – Low methylated galacturonate
LPS – Lipopolysaccharide
mAb – Monoclonal antibody
ManOH – Mannitol
MFF – *m*-phenylphenol
MGDG – Monogalactosyl diacylglycerols
Microbiota – Microorganisms in the large intestine
MS – Mass spectrometry
Mw – Molecular weight
N. oculata – *Nannochloropsis oculata*
Neu5Ac – *N*-acetylneuramic acid (sialic acid)
NGL – Neoglycolipid
NHS – *N*-hydroxysuccinimide
NMR – Nuclear magnetic resonance
NO – Nitric oxid
N-Prot factor – Nitrogen-to-nitrogen protein conversion
NS – Neutral sugars
OS – Oligosaccharides
PAGE – Polyacrylamide gel electrophoresis
PC – Phosphatidylcholine
PC1 - 1,2-Diacyl-*sn*-glycero-3-phosphocholine
PC2 - 1,2-Di-*O*-hexadecyl-*sn*-glycero-3-phosphocholine
PFPA – Perfluorophenylazide
PGA – Polygalacturonic acid
pI – Isoelectric point
PL – Polysaccharide lyase
PMII – Protein Marker II
PMAA – Permethylated alditol-acetate
PMT – Photomultiplier
PO – Polysaccharide oxidase
PS- Polysaccharides
PUL – Polysaccharide utilization loci
RCA-I – *Ricinus communis* agglutinin I
RCA_{120} – *R. communis* agglutinin 120
RGI – Rhamnogalacturonan I
RGII – Rhamnogalacturonan II
RhaOH – Rhamnitol
RibOH – Ribitol
SDS – Sodium-dodecyl sulfate
SEC – Size exclusion chromatography
Speedvac – Centrifuge concentrator at reduced pressure
SQDG – Sulfoquinovosyl diacylglycerols

Sus – Starch-utilization system

t – terminal

TCA – Trichloroacetic acid

TFA – Trifluoroacetic acid

TPA – 12-*O*-tetradecanoylphorbol-13-acetate

T_{room} – Room temperature

UA – Uronic acids

UEA-I – *Ulex europaeus* agglutinin I

WGA – Wheat germ agglutinin

XylOH - Xylitol

Chapter 1 – Introduction and Objectives

Introduction

1.1. Carbohydrates in marine microalgae

Marine microalgae are unicellular photosynthetic organisms of easy growing and culturing found in aquatic habitats. These organisms have also an interesting chemical composition that have been promoting their cultivation in large scale for biotechnological exploitation (1,2). In this field, microalgae may work as a source of biological/bioactive natural compounds with potential applications in various industries, such as food, biomedical and cosmetics (1–3). In food, these compounds are able to enhance its nutritional content and contribute positively to human health. One type of these compounds that have already proved to show various biological activities are carbohydrates (3,4).

Carbohydrates are a class of biological molecules composed of monosaccharide residues with highly diverse structures. Monosaccharides are linked to each other by glycosidic linkages and, depending on the number of residues that form carbohydrates, they are nominated as oligosaccharides (OS) or polysaccharides (PS) (large number of monosaccharide residues). In addition, carbohydrates can be incorporated into glycoconjugates, such as glycoproteins, proteoglycans, and glycolipids. The glycosyl portion of glycoconjugates, which are found largely on the cell surfaces, can mediate many biological processes, often participating as recognition elements (ligands) in cellular recognition systems through molecular interactions (binding events), like protein-carbohydrate interactions (5,6).

Regarding to carbohydrates in marine microalgae, they are found mainly as polysaccharides (PS) (7) and a minority amount can be found incorporated into glycolipids (GL) (8). Both PS and GL have been associated with various biological activities, such as immunomodulatory ability, antifungal, antiviral, anticoagulant, antitumor properties, among others (4,8). The biological activities of these compounds are mainly due to their chemical structures and physicochemical characteristics. However, establishing the relationship between the structures of these compounds and their bioactivities is difficult due to its high complexity and diversity among and inside microalgae species. As a result, the structure of these compounds and its relation in molecular interactions remains understudied in many species, like those that belong to the class Eustigmatophyceae. Thus, considering the potential biological roles of the carbohydrates in marine microalgae, it is important to

unravel the structural features of PS and GL in each species and understand how their structures can be involved in molecular interactions.

This chapter describes briefly the general structural characteristics of PS and GL present in marine microalgae, and an insight about the carbohydrates of *Nannochloropsis oculata*, a species of the class Eustigmatophyceae, it will be discussed in order to understand their properties (section 1.1.). An overview about a technology that allows to study interactions between carbohydrates and various molecules, the carbohydrate microarrays technology, will be also presented (setion 1.2.).

1.1.1. General characteristics of polysaccharides

The carbohydrate fraction in microalgae is present in a range of 5-30 %, and 80 to 95 % of the total carbohydrate fraction is composed of polysaccharides (PS) (7,9–12). Regarding to the characteristics of microalgae PS, including its sugars composition and chemical structure, some studies have been performed in order to unravel their features. The microalgae PS are variable in sugar composition but glucose tends to be the predominant sugar (20-90 % of total sugars). This is consistent with the fact that glucans are the major reserve PS in microalgae (7,13). Besides Gal and Man are also common sugars, and Rha, Fuc, Rib, Ara, Xyl, and uronic acids (UA) can be found in variable proportions. In addition, it should be referred that many of the PS in marine species can be found substituted with sulfate groups, which have been associated with several biological activities, such as immunomodulatory and antimicrobial properties (14–16). Therefore, this high diversity in sugars composition of microalgae PS suggests the presence of complex hetero-polymers.

The characterization of PS, in terms of composition and glycosidic linkages, has been described only for few PS. One example is the PS soluble in hot-water (HW) from microalgae that belong to the diatoms class. These PS were composed mainly of Glc, Gal, Man, Xyl and Rha as neutral sugars, with 3-Glc*p* residue as dominant glycosidic linkage, being found also other Glc residues, such as 3,6- and 2,3-Glc*p* residues. This indicates the presence of (β1→3)-linked glucan with low levels of branching at positions *O*-6 and *O*-2. The (β1→3)-glucans with a small degree of branching (DB) at C-6, (β1→3, β1→6)-branched glucans, is normally referred as chrysolaminarin (Fig. 1 A) (17–19).

In the diatom *Phaeodactylum tricornutum* other residues were also detected, such as t-Glc, t- and 6-Gal, being the terminal residues in high content. This suggests the presence of HW-polymers with low molecular weight (*Mw)*, since the ratio of terminal residues to 3-linked Glc*p* residues and 6-linked Gal was relatively high (17). Moreover, the PS in the HW soluble extracts from *P. tricornutum* contained sulfate groups and high levels of UA, in addition to the neutral sugar residues mentioned above for diatoms. The UA were found mostly as glucuronic acids (GlcA), which occurred as terminal (t-) and 2,4-glucuronosyl residues (19). This highlights that these PS are probably anionic sulfated PS, being the negative character conferred by the presence of UA and sulfate groups. These structures are possibly responsible for various biological activities. As example, Guzmán *et al.* (16) reported that the crude PS extracts with HW from *P. tricornutum*, which comprised UA and sulfate groups, showed anti-inflammatory effects in the paw edema test. This activity was related to the presence of sulfated PS, which probably were determined by its structural conformation.

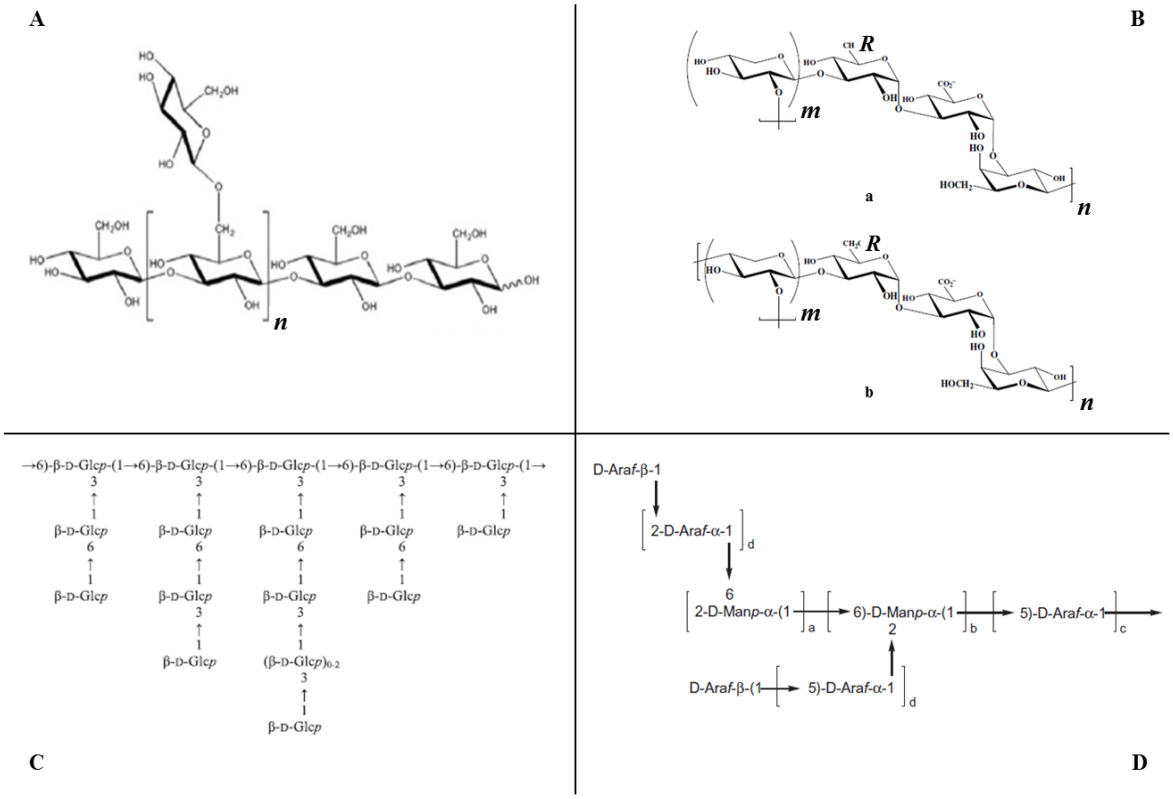


Figure 1. Examples of microalgae polysaccharide structures. **(A)** A model of (β1→3)-D-glucan branched at position *O*-6 of Glc residue from the diatom *P. tricornutum* (18). **(B)** Models **a** or **b** for the possible repeating unit in EPS from *Porphyridium sp*. R = H, SO_2O, terminal Gal or terminal Xyl, m = 2 or 3 (20). **(C)** The highly branched structure of the (β1→3,1→6)-D-glucan from *I. galbana* (21)**. (D)** The structure suggested for the arabinomannan extracted from the cell-wall of *C. vulgaris* (a+b=2, b>a, C=1, d=1) (22).

Besides t- and 2,4-glucuronosyl residues and sulfate groups, it was verified that the PS from *P. tricornutum* frustule are formed mainly by Man, which occurs predominantly as 2,3- and 3-Man linkages (19,23). This suggests the presence of a branched sulfated glucuronomannan, with a backbone composed of (β1→3)-mannan branched at C-2 with sulfate or OS chains with terminal glucuronosyl residues.

Another interesting characteristic of microalgae is that some species contain cells encapsulated within a complex cell-wall polysaccharide (20). The external part of this complex, designated as exopolysaccharides (EPS), may dissolves into the medium. This allows an easy recollection of PS, which represents an advantage for further application. Studies performed about the structure of *Porphyridium sp*. EPS, red microalgae of the Rhodophyta division, highlighted a high complexity of these PS. The EPS from *Porphyridium sp.* are normally sulfated polyanionic hetero-polymers. These tend to be formed by D- and L-Gal, D-Glc, D-Xyl, and D-GlcA, and some of these residues can be sulfated (20,24–26). They may also contain a few amount of methylated sugars, such as 2-*O*-Me-D-GlcA (24). In terms of glycosidic linkages, the methylation analysis revealed the presence mainly of t-, 2- or 4-Xyl*p*, t- and 3-Gal*p*, and 3- and 3,6-Glc*p* residues (20). This probably indicates that the linear backbone is composed of 2- or 4-Xyl*p*, 3-Gal*p* and 3-Glc*p* residues, with a branching point at the *O*-6 position of Glc residue, which may be ramified with t-Xyl*p* or t-Gal*p* (Fig. 1 B). Furthermore, Geresh *et al.* (20) verified the presence of 3-GlcA as a backbone component. In addition, after desulfation it was verified an increase of 3-Glc*p* and a decrease in 3,6-Glc*p*, proposing that sulfate groups are linked at the *O*-6 position of Glc residue. However, the NMR spectroscopy of EPS from *Porphyridium sp*. revealed that these PS were ramified at *O*-2 position of Gal with Xyl (25), instead of *O*-6 position of Glc as proposed before. These results display that the EPS structure is not invariable and may appear some differences even though inside the same microalgae species, which was explained as a consequence of the difference in the methodologies used during the analysis, or in the different origin of alga strains (20).

Remarkably, the EPS from *P. cruentum* showed antiviral activity especially against the *Vesicular stomatitis* virus (14). The antiviral activity was explained due to the anionic properties of the EPS conferred by the presence of sulfate groups along with carboxyl groups, making the EPS a good choice to protect against viruses infections. A possible mechanism proposed to explain this result is that the anionic polymers can interact with the

positive charges of the amino acids on the envelope surface of the viruses, inhibiting the virus penetration into the host cells (14,27). As a result, unravel the structural characteristics of the microalgae PS is important in order to understand/determine the biological processes that they can manage.

Other interesting PS have been recently described for various marine microalgae species (21,22,28,29). Some of these examples include the presence of highly branched (β1→3, 1→6)-D-glucans in *Isochrysis galbana* (21), which exhibited a potential anti-tumor activity (Fig. 1 C); the presence of one arabinomannan (Fig. 1 D) from the cell-wall of *Chlorella vulgaris* that was determined to possess a highly branched structure composed of 6- and 2-*O*-linked mannan as main chain, comprising also 5-Ara*f* residues. The last residues were organized as Ara*f*-[(β1→2)-Ara*f*-(α1→]$_n$, being side chains with an average chain length of two linked to the mannan main chain (22). A water-soluble α-glucan also present in *C. vulgaris* contains a backbone composed of (α1→6)-D-glucan branched at C-3 position with α-D-Glc*p* (29).

In addition, marine microalgae are emerging as a potential source of dietary fibres due to the presence of PS like chrysolaminarin and (β1→3, 1→6)-D-glucans (β-glucans), among others (30). This is particularly important due to the large number of beneficial effects that dietary fibres can display in the human health. This cames from the fact that these compounds are not able to be hydrolysed by the human enzymatic machinery, crossing along the gastrointestinal tract (GIT) without being metabolized and digested. These PS are not also fermented, at least completely, by the microorganisms (microbiota) in the large intestine. However, some of these PS can be selectively fermented by the colonic microbiota, working as prebiotics, to short-chain fatty acids which may promote various beneficial effects to the host's health. Although some chemical structures of the PS have come to light, the polysaccharide characterization of many species still remains unknown, especially due to high complexity of these polymers and the high diversity of microalgae.

1.1.2. General characteristics of glycolipids

Glycolipids (GL) are a class of lipids that comprise on its composition carbohydrates (6). In photosynthetic organisms, like microalgae, these compounds are found in photosynthetic organelle membranes such as in chloroplast and thylakoids (31). GL

comprise mostly galactolipids, such as monogalactosyl diacylglycerols (MGDGs) and digalactosyl diacylglyrcerols (DGDGs), and the sulfolipid sulfoquinovosyl diacylglycerol (SQDG), which play an important role in the membrane characteristics. Galactolipids are able to influence the ratio of bilayer to non-bilayer forming lipids, which determines the biophysical characteristics of membranes and it is essential for functional activity of membrane proteins (32). This is an important aspect in the photosynthetic process, since membrane lipids interact with the protein complexes of photosynthesis, like those present in (light harvesting complexes) photosystem I and photosystem II (33,34).

Relatively to the structure of these GL, they are composed of a glycerol backbone with two acyl chains, (R_1 and R_2) esterified at positions *sn*-1 and *sn*-2, and a sugar moiety, comprising the polar head group of the GL, located at position *sn*-3 (8,35) (Fig. 2). The acyl chains can accommodate fatty acids (FAs) with different chain lengths and degrees of unsaturation, being the FA composition highly influenced by the growth phase and the environmental conditions, like salinity, light and nutrient availability. In addition, this may not only influence the FA profile of the GL, but also alter the content in MGDG, DGDG and SQDG (36–38).

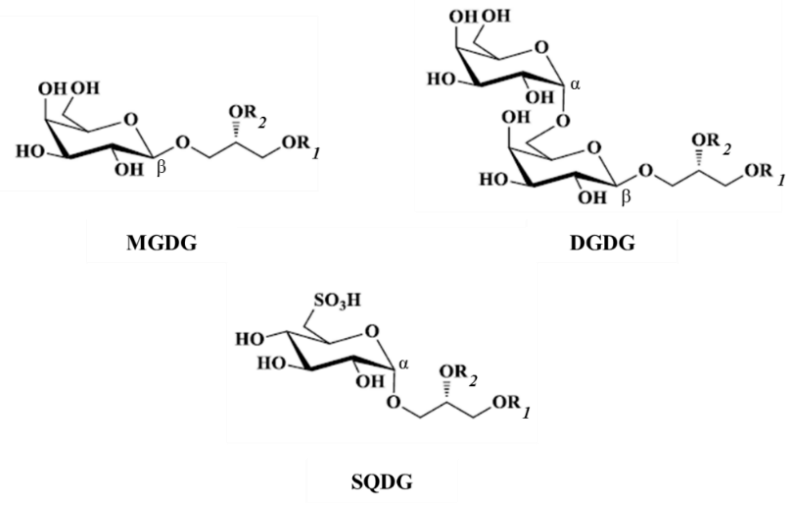


Figure 2. Structures of mono- and digalactosyl diacylglycerols (MGDG and DGDG) and sulfoquinovosyl diacylglycerol (SQDG) mainly found in photosynthetic membranes of microalgae. R_1 and R_2 represent the two fatty acid chains linked to the glycerol backbone. Image adapted from Zhang *et al.* (35).

The MGDG contains one Gal residue bound by a glycosidic linkage to the *sn*-3 position of the glycerol backbone in β-configuration, referred as 1,2-diacyl-3-*O*-(β-D-Gal*p*)-*sn*-glycerol (Fig. 2). Otherwise, DGDG comprises a disaccharide of Gal linked to one

diacylglycerol (DAG). The disaccharide is composed of a terminal α-Gal residue (1→6) linked to one β-Gal residue that binds the *sn*-3 of the glycerol backbone, forming the 1,2-diacyl-3-*O*-(Gal-(α1→6)-*O*-(β-D-Gal*p*)-*sn*-glycerol (DGDG) (Fig. 2). Unlike to MGDG and DGDG, SQDG has a negatively charged character. SQDG comprises one 6-deoxy-6-sulfo-α-D-Glc*p* (sulfoquinovose) as a polar head group linked to the DAG instead of Gal, referred as 1,2-diacyl-3-*O*-(6-sulfo-6-deoxy-α-D-Glc)-*sn*-glycerol (Fig. 2). The presence of the sulfonic acid linked at the position 6 of the Glc residue is responsible for the negative character of the SQDG (8,31,32,35).

Noteworthy, these type of GL extracted from various marine microalgae showed being able to mediate various biological activities, which are dependent not only on the glysosyl portion but also in their acyl chains (35). As example, the MGDG and DGDG extracted from *Chlorella vulgaris* exhibited antitumor activity through inhibition of the tumor promoting stage of Epstein-Barr virus-associated early antigen (EBV-EA) activation on Raji cells induced by 12-*O*-tetradecanoylphorbol-13-acetate (TPA) (39). Furthermore, MGDG and DGDG extracted from *Nannochloropsis granulata* (40) and *Pavlova lutheri* (41) exhibited anti-inflammatory properties. These were demonstrated through inhibition of nitric oxide (NO) production in RAW264.7 macrophage cells through the downregulation of *i*NOS expression, and through downregulation of the production of the pro-inflammatory interleukin (IL)-6 in response of inflammation stimulated with lipopolysaccharide (LPS) in human THP-1 macrophages, respectively.

Moreover, the SQDG extracted from the *Porphyridium sp*. showed anti-inflammatory, antioxidant and anti-proliferative effects (42). The anti-inflammatory and antioxidant activities were observed in *in vitro* assays through inhibition the production of the superoxide anion generated by phorbol myristate acetate in rat peritoneal leukocytes, which have a role in allergy and inflammation processes. The anti-proliferative effect was showed through the growth inhibition of human colon adenocarcinoma DLD-1 cells.

As a result, these biological activities assigned to GL various beneficial health effects with important nutritional significance. Therefore, detect and evaluate the structural characteristics of GL in each microalgae is important in order to understand the mechanisms in that they can be involved and try to establish its structure-activity relationship.

1.1.3. Carbohydrates of *Nannochloropsis oculata*

Nannochloropsis oculata is a unicellular marine microalga that belongs to the class Eustigmatophyceae. Their cells are spherical, green, and the size range from 1-9 µm (Fig. 3) (43,44). This microalga is considered one of the most promising feedstock for developing new food products, due to their high content in long-chain polyunsaturated fatty acids (LC-PUFAs) ω-3, particularly in eicosapentaenoic acid (C20:5, EPA) (12,43,45). Although their lipid fraction has been much studied, only few studies about the carbohydrate fraction of *N. oculata* have been done.

Figure 3. Illustrations of cells viewed by light microscopy of two strains of *N. oculata*, CCMP525 (**p**) and CCMP533 (**q**). Black droplet = red body (44).

The carbohydrates fraction represents 4 to 17 % of dry weight (dw) in *N. oculata* (9–12). However, this composition can change when the microalga are cultured under different conditions or harvested at different growth phases (46). *N. oculata* in stationary phase batch cultures seems to contain 2-4 times more carbohydrates than in logarithmic phase batch cultures. This result occurs due to nitrogen (N) limitation at the beginning of stationary phase, which should be responsible for the accumulation of carbohydrates and/or lipids at the expense of protein, being also recently verified for *N. oceanica* (47). So, this suggests that it is possible to influence the content in carbohydrates depending on the culture phase, which may represent an advantage to obtain carbohydrates for further exploration.

PS comprise the major part of carbohydrates in *N. oculata*, ranging from 74 to 88 % of total carbohydrates (10,11). In terms of sugar composition, the PS are composed mainly of Glc (45-68 weight % of PS), followed by Rha, Man, Gal, Rib, Fuc, and Xyl in about 4 to 10 % for each residue (9–11), suggesting a diversified composition of the PS and possible the presence of hetero-PS, as showed above for other marine microalgae. Moreover, recently Hafsa *et al.* (48) found that the total sugar content in the aqueous extract of *N. oculata* was

59 %, corresponding to 4 % of total dried matter. This extract was composed mainly of Glc (68 %) followed by inositol (20 %), containing also mannitol (6 % of extract). In addition, these authors verified the presence of Xyl, Man, and Gal in minor amounts and, notable, they verified the presence of sulfate groups in the *N. oculata* aqueous extract in about 6 %. As a result, this points out that the water soluble PS of *N. oculata* besides being hetero-PS, they are also probably sulfated and consequently negatively charged. In addition, these authors verified that the PS in this aqueous extract possessed antioxidant, antimicrobial, anticholinesterase and antiproliferation activities.

Regarding to chemical structure of PS from *N. oculata*, in terms of glycosidic linkages, to our knowledge, none, detailed study has been done. Therefore, despite being superficially known their sugar composition, the way that these residues are organised into the PS remains unclear. However, a recent study regarding to the carbohydrate constituents of *N. oculata* by ^{13}C NMR revealed the presence of resonances whose chemical shifts correspond to cellulose and both (β1→3)- and (β1→6)-glucans, which were assigned to the presence of chrysolaminarin (49). Additionally, in *N. oceanica*, a species that also belongs to the Eustigmatophyceae class, has been proposed the presence of (β1→3)-glucans with occasional (β1→6)-linked branches with some of the chains containing mannitol residues at the reducing end. These kind of polymers are storage PS normally designated as laminarin (47,50). Laminarin differs from chrysolaminarin since contains some of the chains with 1-linked mannitol residue at the reducing end, whereas chrysolaminarin does not contain mannitol (51). Furthermore, in both microalgae *N. oculata* and *N. oceanica* was highlighted the presence of one (β1→3)-glucan synthase gene (52). Therefore, the chemical shifts corresponding to (β1→3)- and (β1→6)-glucans in *N. oculata* might indicates the presence of chrysolaminarin or laminarin as storage polysaccharide.

Various studies propose that *N. oculata* contains GL comprising mostly galactolipids (about 25% of the total lipids (53)) and sulfolipid SQDG (45,49,53–55). The galactolipids were found mainly as MGDG and DGDG, in agreement with the reported previous for marine microalgae. Particularly, the presence of galactolipids are specially interesting since Banskota *et al.* (40) found that the MGDG and DGDG isolated from *N. granulata*, a species of the Eustigmatophyceae class, may work as anti-inflammatory agents. It should be mentioned that the content in GL depends on the growth and environmental conditions, similarly to PS. At the onset of stationary phase, MGDG, DGDG and SQDG account to 17,

12 and 6 % of the total of the lipid composition, respectively, but a highest content was obtained at the end of stationary phase (36). Notable, the content in SQDG, MGDG and DGDG decreased under N deprivation (38).

In terms of the fatty acids (FAs) profile of GL, they contain at least one site of unsaturation, with predominance of EPA, C16:0, C16:1, and C18:3 (53–55). Under N replete condition, it was observed that the most FAs present in MGDG were EPA, C14:0, C16:1, and C16:0, in DGDG were C16:0, EPA, and C16:1, and in SQDG C16:0, C16:1, C18:3, and C18:2 (38).

All in consideration, since *N. oculata* has been considered a promising source to develop new food products, due to the unclear knowledge of their PS and their interesting composition in GL, namely MGDG, DGDG and SQDG, this microalga seems challenging to study its carbohydrates, especially the PS and those present in GL. It will be interesting to study not only its structural characteristics but also possible molecular functions that these compounds can be involved.

1.2. Carbohydrate microarrays

Carbohydrate microarrays are emerging as a high-throughput screening technology that has boosted the detection of specific carbohydrates involved in interactions with various molecules. The biggest advantage of this technology is that a wide range of carbohydrate sequences, referred as probes, can be immobilized onto solid matrices using only few amounts of sample (in the range of pg or ng). Afterwards, these matrices can be probed with various molecules in order to elucidate the carbohydrate structures involved in such molecular interactions (binding events) (56).

The carbohydrate microarray technology is prepared in three main steps: 1º) *Immobilization (printing)* of probes, namely mono-, oligo- or polysaccharides, in spots onto an appropriate solid surface (slide); 2º) *Probing* with various molecules, specially with carbohydrate-binding proteins (CBPs), forming the binding assay; and 3º) *Detection and analysis of the interactions* between the carbohydrates immobilized and the probed molecule, mainly through fluorescence-measurements (an explanation about the fluorescence-based measurements used to detect and analyse carbohydrate-protein interactions in microarrays is summarized in supplementary information 1). Normally PS

probes are derived from natural sources, whereas monosaccharide and OS probes can be also obtained by chemical synthesis (56,57). In this section, an overview about the immobilization methods and the CBPs used during the binding assay it will be presented.

1.2.1. Immobilization methods

An efficient immobilization of the carbohydrate probes onto the solid surfaces is essential to study carbohydrate-protein interactions. In the last years, various immobilization methods have been developed (58,59). These techniques are widely classified into two categories, noncovalent and covalent immobilization.

1.2.1.1. Noncovalent immobilization

Noncovalent immobilization can be obtained through hydrophobic interactions established between the carbohydrates and the solid surface. An example of this kind of immobilization is the printing of unmodified PS, such as dextrans (20 to 2,000 kDa) and inulin (3.3 kDa), on nitrocellulose-coated glass slides by adsorption to the surface (60). The fluorescent intensities detected after extensive washing of the printed slides showed an efficient immobilization of the PS without chemical modification on the nitrocellulose-coated glass slide. Nonetheless, it should be noted that the PS are attached to the surface randomly, since they are non-specifically adsorbed on the solid surface. Furthermore, Wang *et al.* (60) found that after the washing steps the dextrans with higher molecular weight (*Mw*) were better retained to the solid surface than the smaller dextrans. The result can be explained because polymers with large contact areas contain a higher hydrophobic character that allows an efficient adsorption on the surface, whereas the smaller ones are more hydrophilic. This suggests that the immobilization through hydrophobic interactions may be influenced by the polymer *Mw*, being suitable for PS immobilization.

In order to overcome the poor retention of carbohydrates with low *Mw* (*e.g.* OS) by hydrophobic interactions on the solid surface, they can be conjugated with hydrophobic tags, resulting in a site-specific immobilization. One example of this approach is the noncovalent immobilization of lipid-conjugated OS, named as neoglycolipids (NGLs) (61–64) onto nitrocellulose surfaces. The lipid tag allows to obtain NGLs probes with amphipathic properties appropriate for printing on solid matrices by adsorption.

The NGLs can be generated by reductive amination (Fig. 4), a reaction that occurs between the C-1 aldehyde of the OS reducing sugar and an amino phospholipid, such as *N*-aminoacetyl-*N*-(9-anthracenylmethyl)-1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine (ADHP) or 1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine (DHPE) (64). This reaction results in a ring opening of the core of monosaccharide residues at the reducing ends. Although the majority of recognition motifs are maintained, this may influence the functions of the OS, especially if they are short-chain OS, where recognition systems tend to require an unmodified core monosaccharide (65).

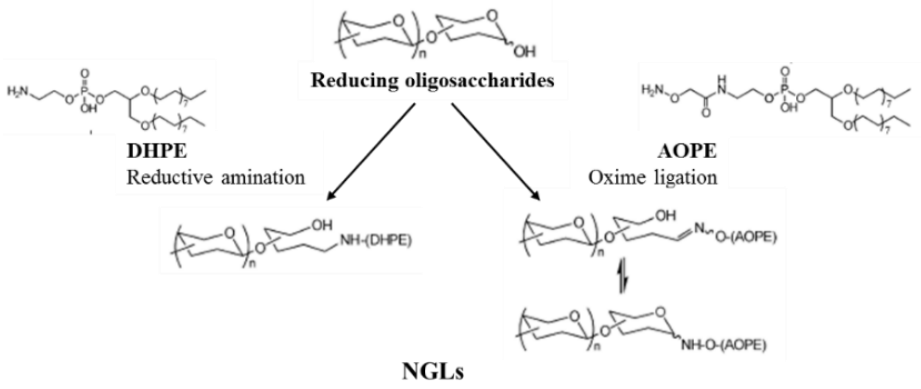


Figure 4. Preparation of neoglycolipids (NGLs) from reducing oligosaccharides by reductive amination or oxime ligation reactions (63).

An alternative to reductive amination for OS derivatization as NGLs is the oxime ligation. The oxime-linked NGL is obtained from the reaction that occurs between the C-1 aldehyde of the OS reducing sugar and an aminooxy group of a functionalized lipid, such as aminooxy-functionalized DHPE (AOPE) (Fig. 4). The oxime-linked core monosaccharide is found predominantly in the ring-closed form, opposite to reductive amination. This allows not only an efficient noncovalent immobilization of the OS by adsorption onto nitrocellulose slides, but also keep its structure and functions in carbohydrate-protein interactions (63).

Other strategies can be used to noncovalent immobilization of OS, such as through electrostatic interactions between polyanionic glycopolymers and Au-substrates coated with a layer-by-layer ending with a cationic polymer (66), or through fluorous-fluorous interactions between sugars conjugated with a single C_8F_{17}-tail and fluorous-derivatized glass slides (67–69). It should be noted that these approaches require chemical modification of the carbohydrates and the solid surfaces.

1.2.1.2. Covalent immobilization

Construction of carbohydrate microarrays by covalent immobilization on the solid surface uses both chemically modified or unmodified carbohydrates and derivatized surfaces. This technique relies on the reaction that occurs between a functional group present on carbohydrates and those present on the solid surface.

- *Chemically modified carbohydrates*

A wide range of covalent immobilization methods of functionalized carbohydrates on the derivatized surface have been developed. Examples of this approach include covalent attachment through reaction that occurs between maleimide and thiol functional groups by printing maleimide-functionalized mono- and disaccharide probes onto thiol-derivatized glass slides or thiol-functionalized OS probes onto maleimide-derivatized glass slides (58,59). Nonetheless, Park *et al.* (70) found that maleimide-linked carbohydrate probes prepared by one-pot amination give only β-anomeric carbohydrate probes, whereas maleimide-linked carbohydrate probes prepared by allylation generates both α- and β-anomeric configurations. This suggests that depending on the strategy employed to prepare the maleimide-functionalized carbohydrate probes may affect the nature of the anomeric linkage. Moreover, the authors showed that this result may influence their protein binding affinities. Thus, the configuration acquired by the carbohydrates during chemical modification should be taken in account during the microarray construction.

Many other functional groups can be used in order to obtain chemically modified carbohydrates that react selectively with those groups present on the solid surface. These include the printing of fluorescently labelled OS with a secondary aromatic amine group on the epoxide-activated glass slides, through the reaction that occurs between the secondary amine group and the epoxide group (71). Amino-derivatized OS can be also printed on an amine-reactive *N*-hydroxysuccinimide (NHS)-activated glass slide by formation of an amide bond (72,73). In addition, coupling monosaccharides with a propargyl group allows its immobilization onto a glass slide covered with chains ending in a terminal azide group. This covalent immobilization is achieved through the formation of a covalent bond between the propargyl and azide groups (74).

The covalent immobilization methods present here indicate the site-specific attachment of the mono- and oligosaccharide probes on the solid surface, since the covalent

bond is formed between a reactive group of the modified carbohydrate and those present on the slide. Further, this approach seems appropriate to print mono- and oligosaccharides. Nonetheless, the preparation of modified carbohydrates is time-consuming and may influence its functions (70). Thus, it is desirable to use methods that allow the printing without chemical modification.

- *Chemically unmodified carbohydrates*

Covalent immobilization of unmodified carbohydrates requires a surface derivatization with functional groups that selective and efficiently reacts with carbohydrates in order to obtain a covalent bond. One possible approach is the use of hydrazide- or aminooxy-coated glass slides. This method seems suitable to print mono-, oligo- and polysaccharides due to a chemoselectively reaction that occurs between hydrazide or aminooxy groups and the reducing end of the carbohydrates (75). However, Lee and Shin (75) verified that the carbohydrates immobilized on hydrazide-coated slides allowed to obtain higher fluorescent signals after the binding assay with lectins than those obtained from aminooxy-coated slides. This was explained as a consequence of the predominance of acyclic form obtained from reaction of carbohydrates with the aminooxy group, whereas in the reaction with hydrazide group prevailed the cyclic form with β-configuration. This result suggests that the configuration acquired by the residue involved in covalent linkage may affect the binding properties of the carbohydrates, as verified by Park *et al.* (70). For that reason, apart from ensure an efficient immobilization, it is important to keep the structure of the carbohydrates.

Another covalent immobilization method is based on irradiation, which allows the attachment of unmodified carbohydrates to derivatized surfaces. Examples of this approach include the covalent binding of the PS on dextran-coated glass slides derivatized with a photolabile group (aryl-trifluoromethyl-diazirine). The irradiation of the aryl-trifluoromethyl-diazirine group forms a reactive carbene which is able to react with vicinal molecules, like PS, to form covalent bonds (76). A similar strategy is the slides derivatization with a photoactive polymer containing a photoactive group (perfluorophenylazide (PFPA)). The light activation promote a reaction between the surface azide groups present on PFPA and the carbohydrates, with consequent covalent immobilization (77–79). However, the attachment of the carbohydrates to the surface by irradiation is site-nonspecific, since the

reactive carbene or the azide group may react with any residue (76–79). This is an important issue because may have impact in their interaction with the binding partners (79).

In sum, the noncovalent and covalent immobilization methods present in this section, which are summarized in figure 5, display several options to get an efficient print of the carbohydrates on the solid surface. This allows the possibility of choosing the most suitable immobilization method to print the carbohydrates depending on its features.

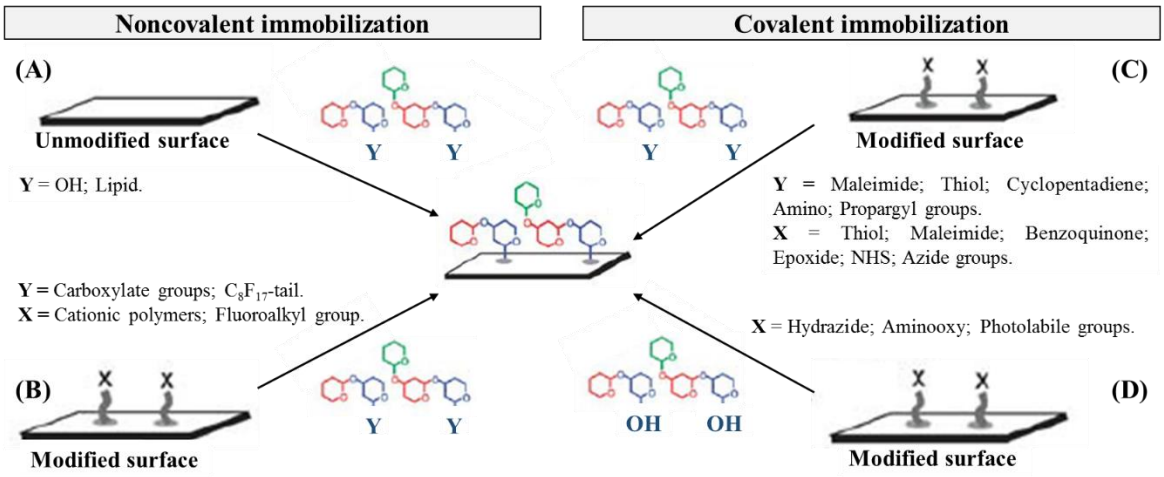


Figure 5. Immobilization methods for the construction of carbohydrate microarrays. Noncovalent immobilization: attachment of carbohydrates **(A)** through hydrophobic interactions on the underivatized surface, and **(B)** through electrostatic or fluorous-fluorous interactions on the derivatized surface. Covalent immobilization: attachment of **(C)** chemically modified carbohydrates onto derivatized surface, and **(D)** unmodified carbohydrates onto derivatized surface. Image adapted from Park *et al.* (58).

1.2.2. Carbohydrate-binding proteins

There is an increase awareness that specific carbohydrate sequences are able to mediate various biological processes through working as ligands in carbohydrate-protein interactions (57,80). As a result, after carbohydrates printing on the solid surface, the binding assay has been performed by probing the array with various carbohydrate-binding proteins (CBPs) that interact noncovalently with carbohydrates, such as lectins, antibodies and carbohydrate-binding modules (81,82).

1.2.2.1. Lectins

Lectins are proteins that bind carbohydrates reversibly and with high affinity and specificity, being suitable to detect specific carbohydrates and distinguish identical sugars (6). These proteins may contain two or more protein modules within the lectin polypeptide

that are responsible for its sugar-binding activity. These modules normally are referred as carbohydrate-recognition domains (CRDs). Lectins have a major role in a broad range of biological processes due to its sugar-binding activity, such as its participation in the movement of glycoconjugates to the cell surface, in cell-cell recognition events and in cell adhesion. In addition, lectins are found almost in all organisms (6,81,83).

Regarding to plant lectins, due to their popularity as laboratory tools they are suitable to detect and characterize carbohydrates, since many of them have well-known specificities (Table 1) (84). Among plant lectins, some of the most known are: concanavalin A (ConA) of the Jack Bean (*Canavalia ensiformis*), which is specific for α-linked mannose but it is able to recognise α-D-Glc*p* and α-D-GlcNAc*p* with lower specificity; *Ricinus communis* agglutinin (RCA) that is specific to interact with β-D-Gal*p* residues and with low affinity to α-linked D-Gal*p* residues; Wheat germ agglutinin (WGA) binds to Neu5Ac(α2→3)Gal(β1→4)Glc and GlcNAc(β1→4)GlcNAc; and *Ulex europaeus* agglutinin-I (UEA-I) that binds Fuc(α1→2)Gal(β1→4)GlcNAc and Fuc(α1→2)Gal(β1→4)[Fuc(α1→3)]GlcNAc (61,81,84–87).

The CRDs in animal lectins are divided into different families due to its distinct structures. Eight families are already well-established, four are intracellular lectins, namely calnexin, M-type, L-type, and P-type, whereas the other four are extracellular lectins, C-type, R-type, singlecs and galectins. The intracellular lectins are found in luminal compartments of the secretory pathway, such as endoplasmic reticulum (ER) and Golgi complex, and work in the glycoprotein trafficking, protein sorting, degradation of glycoproteins, among other functions. Otherwise, the extracellular lectins are located in the extracellular matrix or body fluids, or in the cell membrane, working as example in cell adhesion, cell signalling, innate immune, enzyme targeting, pathogen recognition, and glycan crosslinking in the extracellular matrix (83,88).

Table 1. Typical carbohydrate ligands reported for various plant and animal lectins.

		Typical carbohydrate ligands	Reference
Plant Lectins			
Concanavalin A (Con A)		α-linked mannose (α1→2)- and (α1→3)-mannose oligosaccharides Glc residues	(81,84,85,89, 90)
***Ricinus Communis* Agglutinin** (RCA)		β-Gal > α-Gal	(85)
	RCA_{120}	Terminal Gal(β1→4) > terminal Gal(β1→3) and Gal(β1→6) Terminal Gal(β1→4) in *N*-glycans Gal(6S)(β1→4)AnGal(α1→3)Gal(6S)(β1→4)AnGal(α1→3)Gal(6S) Gal(2S)(β1→[4)-Gal(2S6S)(α1→3)Gal(2S)(β1→]24)-Gal(2S6S)	(61,62,85,91)
	RCA-I	Gal(β1→4)GlcNAc Neu5Ac(α2→6)Gal(β1→4)GlcNAc	(87,91)
Wheat germ agglutinin (WGA)		Neu5Ac(α2→3)Gal(β1→4)Glc GlcNAc(β1→4)GlcNAc D-GlcNAc*p*(β1→4)-D-GlcNAc*p*(β1→4)-D-GlcNAc > D-GlcNAc*p*(β1→4)-D-GlcNAc	(81,85)
***Ulex europaeus* agglutinin-I** (UEA-I)		α-L-Fuc Fuc(α1→2)Gal(β1→4)GlcNAc Fuc(α1→2)Gal(β1→4)[Fuc(α1→3)]GlcNAc	(84–87)
Animal Lectins			
C-type lectin family	**Dectin-1**	(β1→3)-glucooligosaccharides with DP-11 and DP-13	(92)
	Dectin-2	branched $Man_{8/9}GlcNAc$ in α-configuration	(93)
	DC-SIGN	Branched $Man_{6\text{-}9}GlcNAc$ in α-configuration; (β1→2)-gluco-OS with DP-5 and DP-6; (α1→4)-gluco-OS with DP-2 and DP-3; Glc(β1→3)Glc(α1→3)Glc(β1→3)Glc- Glc(β1→6) Glc(β1→6)	(92,94)
	Langerin	Linear Man_3 and Man_4 in α-configuration; [6-SO_4]Gal(β1→4)GlcNAc.	(94,95)
	L- and P-selectins	→3)GalNAc(4S6S)(β1→4)[Fuc*X*(α1→3)]GlcA(β1→ *X* = different sulfation patterns of fucose	(96)
Calnexin		Glc_1Man_9	(83,88)
M-type		Man_8	
P-type		Mannose 6-phosphate	
Singlecs		Sialic acid (Neu5Ac)	
Galectins		β-galactosides Gal(β1→4)Glc (lactose) Gal(β1→4)GlcNAcβ (*N*-acetyllactosamine)	(81,97)

Remarkably, L-type, R-type and C-type lectins may display various carbohydrate ligands, whereas the other families have typical carbohydrate ligands (83,88). As an example, the C-type lectins Dectin-1 and Dectin-2 showed different carbohydrate ligands. These are defined as "C-type" since the binding of carbohydrate ligands to its CRD is Ca^{2+}-dependent (C-type). It was found that Dectin-1 is able to recognise (β1→3)-glucooligosaccharides with a degree of polymerization (DP) of 11 to 13 (92), whereas Dectin-2 is able to recognise branched $Man_{8/9}GlcNAc$ in α-configuration (93). Consequently, due to the diversity of ligands that the lectins in these families may possess, various studies have been done through carbohydrate microarrays in order to unravel its carbohydrate ligands, for example for the C-type lectins dendritic cell-specific ICAM-3 grabbing nonintegrin (DC-SIGN), Langerin, and Selectins (Table 1) (92,94–96).

The typical carbohydrate ligands of the remaining five animal lectin families are described in table 1 (83,88). Due to this feature, these lectin families are suitable to identify and determine carbohydrate sequences printed on the array. As example, galectins, a family of proteins implicated in immune regulation, which are expressed by gut epithelial cells, are known to binding β-galactosides, combining preferentially with lactose and *N*-acetyllactosamine (LacNAc) (81,97). Thus, galectins are useful lectins to identify carbohydrate sequences that are able to mediate the immune system through recognition of β-Gal residues. Noteworthy, these type of lectins will be suitable to understand if the β-Gal residues of the MGDG and DGDG extracted from microalgae are able to mediate the immune system.

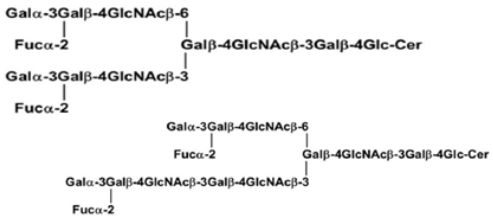
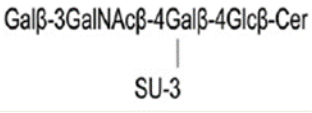
1.2.2.2. Antibodies

Antibodies (Abs) or immunoglobulins (Igs) are soluble proteins produced by the immune system with specificity to recognise a specific epitope on an antigen. Due to the high binding specificity and affinity of the Abs, they have been prepared as analytical reagents to polyclonal and monoclonal antibodies. Polyclonal antibodies are a mixture of Abs that bind different specific epitopes inside the antigen. Unlike, monoclonal antibodies (mAbs) are homogeneous Abs that all bind the same epitope (6). Thus, mAbs are valuable proteins for carbohydrate research. However, it should be noted that its use is high dependent on the knowledge of the epitopes they recognise (82,98).

Among the most used mAbs for carbohydrate research are the mAbs with specificities for plant cell-wall components (98,99). Some examples include: mAb LM5 that is specific

for (β1→4)-galactan (72,100,101); mAb LM6 specifics to (α1→5)-arabinan (99–101); mAb anti-(β1→4)-mannan (LM21) that recognises (β1→4)-mannan (100,101); and the mAb anti-(β1→3, 1→4)-glucan (BS-400-3) that binds mixed (β1→3)-(β1→4)-glucans without cross-reactivity to (β1→3)-glucans (100,102) (Table 2).

Table 2. Carbohydrate ligands recognised by various monoclonal antibodies (mAbs) through carbohydrate microarrays.

mAb	Carbohydrate ligands	Reference
LM5	(β1→4)-galactan	(72,100,101)
LM6	(α1→5)-arabinan	(99–101)
anti-(β1→4)-mannan (LM21)	(β1→4)-mannan	(100,101)
anti-(β1→3, β1→4)-glucan	mixed (β1→3, β1→4)-glucans	(100,102)
F77		(103)
AE3		(62)

Noteworthy, carbohydrate microarrays probed with mAbs known to raise against cancer cells have allowed to identify cancer-specific cell surface antigens as carbohydrates and unravel its sequences. One example is the mAb F77, previously found to label primary prostate cancers, that revealed a selective binding to dodecaglycosyl- and tetradecaglycosylceramide. These probes comprised the same glycosyl backbone (Gal(β1→4)GlcNAc(β1→3)Gal(β1→4)Glc-Ceramide) and two branches containing LacNAc at positions C-6 and C-3 of backbone terminal β-Gal, differing only in one of the branches (Table 2). The common branch located at position C-6 contained Gal(β1→4)GlcNAc (type 2) substituted with (α1→2)-Fuc in β-Gal. Lack of 6-linked branch in the backbone removed the recognition by mAb F77. Otherwise, the absence of 3-linked branch did not affect this interaction. This may indicate that mAb F77 requires a 6-linked branch of a poly-*N*-LacNAc backbone substituted with (α1→2)-Fuc to interact with OS

(103). In addition, the carcinoma specific mAb AE3 was found to bind more intensively to monosulfated tetra-glycosyl ceramide than to the unsulfated form Gal(β1→3)GalNAc(β1→4)Gal(β1→4)Glcβ1-Ceramide, when the sulfate group was present at position C-3 of internal β-Gal (Table 2) (62). Thus, these results obtained from carbohydrate microarrays analyses using sequence-defined carbohydrate probes allowed to highlight carbohydrate sequences that may work as cancer biomarkers.

1.2.2.3. Carbohydrate-binding modules

Carbohydrate-binding modules (CBMs) are non-catalytic protein domains with carbohydrate-binding function that are normally appended to Carbohydrate-Active EnZymes (CAZymes). CAZymes catalyse the breakdown, biosynthesis or modification of complex carbohydrates, as well as glycoconjugates. Some of these enzymes are: glycoside hydrolases (GHs), polysaccharide lyases (PLs), carbohydrate esterases (CEs), polysaccharide oxidases (POs) and glycosyltransferases (GTs). CBMs are connected to the catalytic site through a flexible linker that promote the association between the catalytic domain of the enzyme and the substrate, potentiating the enzyme activity (104–106).

CBMs have been identified and classified into families based on its amino acid sequence similarity in the CAZyme database (www.CAZy.org, (107)). Currently, they are grouped into 81 CBM families and can exhibit diverse ligand specificity. CBM nomenclature may include, besides the family number, the organism name or even the enzyme from it is derived. As example, the family 17 CBM from *Clostridium cellulovorans* Cel5A is called as *Cc*Cel5ACBM17. In addition, when the enzyme contains various CBMs that belong to the same family, the correspondent number of the position of the CBM in the enzyme relatively to the *N*-terminus may be included. Exemplifying, *C. stercorarium* has an enzyme with a triplet of family 6 CBM so the three modules are referred as *Cs*CBM6-1 (first CBM), *Cs*CBM6-2 (second CBM) and *Cs*CBM6-3 (third CBM) (104).

Additionally, depending on the way that CBMs interact with carbohydrates, these modules are classified into three distinct types: A, B and C. CBMs from Type A, also referred as ‘surface-binding’, interact to the surface of crystalline PS and they show little or no affinity for soluble carbohydrates. This interaction is achieved due to the presence of planar hydrophobic platforms. CBMs from Type B or *endo*-type are those that interact with internal OS chains. These CBMs contain a groove or cleft that accommodates the OS chain, with preferentially binding to OS with a DP between 3 and 6. Finally, CBMs from Type C

or *exo*-type interact with the terminal of an OS, binding mono-, di- or trisaccharides, and so they do not comprise grooves on binding-sites (104,106). Examples of the structures of the CBMs from type A, B and C are shown in figure 6.

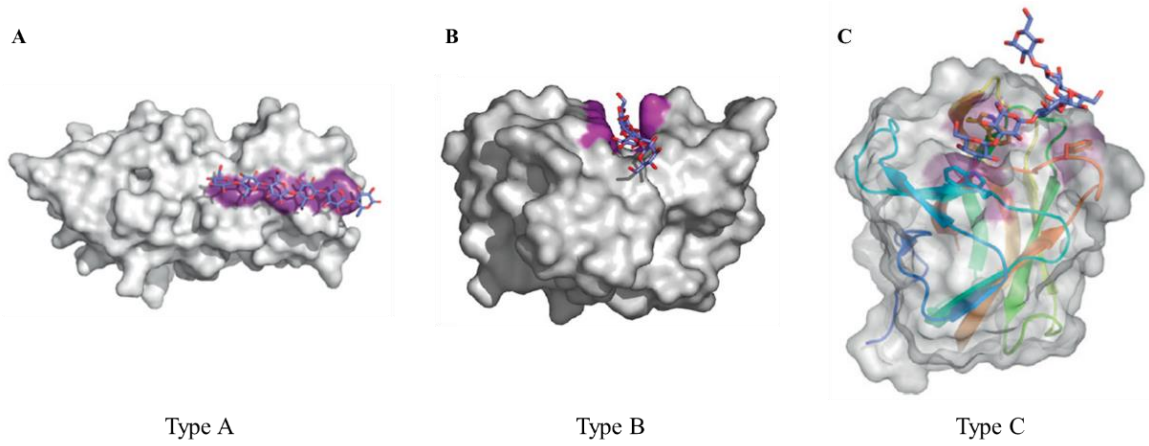


Figure 6. Structural examples of CBMs from Type A, B and C. (**A**) The crystallographic dimer of CBM63 from *Bacillus subtilis* (PDB ID 4FER) in complex with cellohexaose (blue sticks). Surface composed by the aromatic amino acids that form the planar hydrophobic surface are shown in purple (106). (**B**) The crystallographic structure of the CBM4-2 from *Celllulomonas fimi* (PDB ID 1GU3] in complex with a cellopentaose molecule. The amino acid side chain involved in binding are shown in purple (104). (**C**) The X-ray crystal structure of the CBM6 from *Bacilus halodurans* in complex with laminarihexaose (PDB ID 1W9W). The CBM specifically recognizes the non-reducing end of the sugar. The secondary structures are shown as colored ramped cartoons with relevant amino acid side chains involved in carbohydrate binding shown as sticks. Solvent accessible surfaces are shown in gray with the surfaces contributed by the aromatic residues colored purple (106).

It should be noted that although there has been an increase in the number of entries of CBMs in CAZy database, the determination of their specificities to carbohydrate ligands by which they potentiate the enzyme activity still remains understudied (106,107). Many CBMs studied are from microbial origin, and the biggest advantages in working with these CBMs is that the gene/protein sequences are easily obtained and the protein structures are often known (82). With this in mind, microarrays with carbohydrate-sequence defined probes are a suitable tool to obtain more detailed information about the carbohydrate specificities of these CBMs. Various CBM families with some recognition features known were used in microarrays comprising carbohydrate-sequence defined probes (92,100). This allowed not only to confirm the CBM specificity, but also to provide new insights in its carbohydrate-binding specificities. These experiments included the following microbial CBMs: *Rm*CBM4-2, *Tm*CBM41, *Ct*CBM11, *Cb*CBM32-2, and *Cm*CBM6-2. The carbohydrate ligands identified by carbohydrate microarrays are summarized in table 3.

Table 3. Carbohydrate ligands of various microbial carbohydrate-binding modules (CBMs) ascertained through carbohydrate microarrays.

Microbial CBM	Carbohydrate ligands	Reference
***Rm*CBM4-2**	Linear (β1→4)-linked D-xylose with DP 2 to 5	(100)
***Tm*CBM41**	(α1→4)-gluco-oligosaccharides Linear (α1→4)-(α1→6)-linked oligosaccharides derived from pullulan	(92)
***Ct*CBM11**	Mixed (β1→3, β1→4)-glucans with DP 7 or longer	(92)
***Cb*CBM32-2**	Linear (β1→3)-, (β1→2)-gluco-oligosaccharides (β1→3, β1→6)-branched gluco-oligosaccharides	(92)
***Cm*CBM6-2**	Linear (β1→2)-, (β1→3)- and (β1→4)- gluco-oligosaccharides Mixed (β1→3, β1→4)-gluco-oligosaccharides	(92)

Noteworthy, one type of microbial CBMs that are of particular interest to study are those that belong to glycan-degrading enzymes of the human gut microbiome, since these may have roles in human health (108).

- *Human gut microbiome*

The microbes that reside in the human lower gastrointestinal tract (GIT) (gut microbiota) may display important biological functions absent in human cells. One of these functions is to improve the energy capture from dietary components (108,109). This is due to its function as metabolic organ with enzymatic properties, including the action of CAZymes and CBMs, that allows to degrade dietary complex carbohydrates into mono- or oligosaccharides that cannot be hydrolysed by host enzymes, providing nutrients (110). One of the bacteria in the human gut involved in these metabolic processes is *Bacteroides thetaiotaomicron* (109).

B. thetaiotaomicron is a gram-negative bacteria that can utilize PS as source of carbon and energy (111). Their ability to target these PS is due to the presence of gene clusters termed as polysaccharide utilization locis (PULs) that are selectively activated by the PS and encode several outermembrane protein complexes. These proteins are responsible for capturing, degrading, and importing the PS (110,111).

B. thetaiotaomicron comprises an enormous diversity of PULs and each PUL is highly specific for a defined polymer (111,112). Together, the protein and its respective PUL are part of the human gut microbiome, which is defined as all the genes and genomes of the gut

microbiota and the products that they encode (113). The first described PUL-encoded carbohydrate-uptake system in *B. thetaiotaomicron* was the starch utilization system (Sus) (Fig. 7). The Sus system is composed of eight Sus proteins (SusRBACDGEF) that are required to bind and degrade starch. SusR is responsible for *sus* transcriptional activation (transcriptional regulator) in response to the accumulation of maltose in the periplasm. SusB and SusA are located in the periplasm and show activity of neopullulanase and α-glucosidade, respectively, to process maltooligosaccharides that arrived in the periplasm into fermentable glucose. The maltooligosaccharides are derived from the action of the remaining Sus proteins, SusCDGEF, which are located in the outer membrane. SusDEF are responsible for bind the starch at the surface of the cell and this is subsequently hydrolysed to maltooligosaccharides by SusG, one α-amylase of the GH13 family. Thereafter, these maltooligosaccharides transit to the periplasm through the SusC transporter (111).

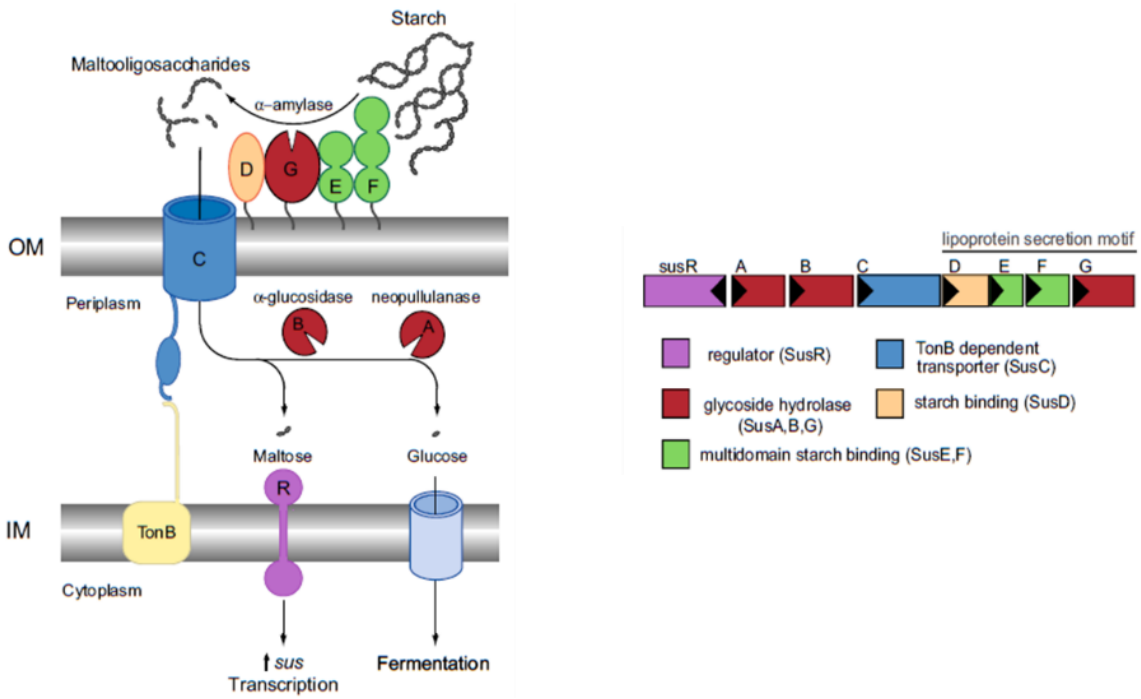


Figure 7. Overview of the starch utilization system (Sus) in *B. thetaiotaomicron*. The inner membrane-spanning protein SusR to sense the disaccharide inducer, maltose, in the periplasm and subsequently drives the transcription of the locus susABCDEFG. Starch is bound to the surface of the cell by the starch-binding outer membrane lipoproteins SusDEF. Subsequent hydrolysis by a similarly membrane-tethered α-amylase, SusG, generates oligosaccharides small enough to transit through the TonB-dependent transporter (SusC). Once in the periplasm, SusA and SusB, a neopullulanase and α-glucosidase, respectively, process oligosaccharides into glucose. The monosaccharide is then shuttled into the cytoplasm by an unknown transporter. Image adapted from Foley *et al.* (111).

Interestingly, it was found that the structure of SusG is composed of A, B, and C domains, containing the catalytic site in the A-domain and, within the B-domain, was verified the insertion of a CBM58. Furthermore, it was observed the ability of the CBM to bind maltooligosaccharides. This is an important finding since the CBM58 can enhance the ability of the SusG protein to bind starch and improve the catalytic efficacy, showing the importance of the presence of CBM to digest PS (110).

Besides the activation of the PUL that encodes the SusG protein, named also as BT3698, in response to pullulan, other PULs encoding carbohydrate-degrading enzymes have been activated in response to various PS. Notable, the expression of the PUL that encodes the BT0996 protein in response to rhamnogalacturonan I (RG-I) and rhamnogalacturonan II (RG-II) (112,114). BT0996 comprises three domains, two families of GH (GH2 and GH137) and the 57 family CBM (www.cazy.org and (114)).

In particular, Ndeh *et al.* (114) found that BT0996 participated in degradation of RG-II through cleavage of GlcA(β1→4)Fuc, displaying β-D-glucuronidase (GH2) activity, and cleavage of the terminal Ara*f*(β1→2)Rha by β-L-arabinofuranosidase (GH137). This shows that *B. thetaiotaomicron* is able to degrade the most complex PS in nature, RG-II, emphasizing even more the importance of the human gut microbiome. Nevertheless, the knowledge of the fine specificity of the CBM domain of BT0996 to its carbohydrate epitope remains unclear, since the epitope recognised by the CBM can be different from the epitope recognised by the catalytic domain, being so important doing research on this area to fully understand its mechanism of action. Moreover, considering the high complexity of the PS in marine microalgae, it will be interesting understand if these PS can be metabolized by the human gut microbiome. This is particularly important, since one of the future possible applications to microalgae PS is its employment in food, being therefore important to understand their influence in the human gut microbiome.

Taking the above considerations for lectins, Abs, CBMs, and in particular the microbial CBMs, it was possible to highlight that these proteins can be used in carbohydrate microarrays with two objectives. Proteins with well-known carbohydrate binding specificities are suitable to do a general screening of the array in order to identify and obtain structural detail of the carbohydrates printed on the array. Otherwise, constructing a microarray with sequence-defined carbohydrates allows to study proteins with unknown

specificities towards carbohydrates and determine its recognition motifs through the binding patterns highlighted during the experiment, in other words, unravel carbohydrate-protein interactions.

1.3. Objectives of the thesis

Marine microalgae seem an interesting source of carbohydrates, in particular of polysaccharides (PS) and glycolipids (GL), to employ in various areas, such as in food or biomedical applications. This is due to the various biological activities that these compounds may exhibit and, regarding to PS, these may work in particular as dietary fibres. These carbohydrate properties are dependent on their structural features. However, the structure of the most PS in marine species and the structure-function relationships still remains unclear. Thus, the present thesis proposes to study the carbohydrates of the marine microalga *Nannhochloropsis oculata*, which has been considered one promising source to develop new food products.

In order to study the carbohydrates of *N. oculata*, in particular its PS and the glycosyl portion of its GL, this thesis is divided into two main sections:

1. Characterization of Polysaccharides from *N. oculata* (Section 3.1.)

The aims of this section are:

- To perform a sequential extraction of the PS present in the biomass of *N. oculata* and characterize their structural characteristics using carbohydrate analyses.
- To employ the PS extracted from the biomass of *N. oculata* as PS probes to develop carbohydrate microarrays. Carbohydrate microarrays will be used with two objectives:
    1. As a supplementary method to obtain a detail knowledge of the structural characteristics of the PS through analysis with two lectins, two mAbs, and one microbial CBM with known carbohydrate-binding specificities;
    2. As a tool to evaluate interactions of the PS from *N. oculata* with the two human microbiota proteins produced through recombinant expression (BT3698 and BT0996). For this the CBM domains of two proteins belonging to the human gut bacteria *Bacteroides thetaiotaomicron*, BT3698 (CBM58) and BT0996 (CBM57), were produced.

2. Characterization of Glycolipids from *N. oculata* (Section 3.2.)

The aims of this section are:

- To perform the extraction of the lipid fraction from the biomass of *N. oculata* and characterize the carbohydrate portion that belong to the GL using carbohydrate analyses.
- To employ the lipid fraction of *N. oculata* comprising the GL as liposome probes to develop carbohydrate microarrays, with two main objectives:
    1. To evaluate the ability/suitability of using the GL as natural probes to develop microarrays by analysis with one lectin with known specificity (validation);
    2. To study carbohydrate-protein interactions of the carbohydrates present in GL with one animal lectin with functions in the immune system.

The laboratorial work related to the extraction and carbohydrate analyses of PS and GL of *N. oculata* were performed at the University of Aveiro, whereas the work related to recombinant expression of proteins and microarray experiments were performed at the FCT Universidade Nova de Lisboa.

Chapter 2 – Material and Methods

2.1. Characterization of polysaccharides from *N. oculata*

2.1.1. Sample

Lyophilized biomass of the green marine microalga *Nannochloropsis oculata* (*N. oculata)* cells were used to perform the sequential extraction of polysaccharides. This microalga was purchased from Necton S.A. The chemical composition indicated for *N. oculata* biomass by Necton S.A. is 30-35 % of dry weight (dw) in protein and 13-20 % dw of lipids. It is also indicated that their biomass is rich in omega-3 fatty acids (eicosapentaenoic acid (EPA), arachidonic acid and α-linolenic acid), vitamins, amino acids, minerals and carotenoids.

2.1.2. Sequential extraction of polysaccharides from biomass of *N. oculata*

- Biomass defatting

Biomass of *N. oculata* was defatted with 2:1 (v/v) chloroform/methanol. The biomass (in about 40 g) was suspended in 160 mL of $CHCl_3$:MeOH (2:1, v/v), dispersed with ultra-turrax (IKA T25 basic with a dispersion tool S25N-18G) for 15 min at 13 500 rpm and then stirred for 1 h and centrifuged at 15 000 rpm for 30 min at 0 ºC. The residue obtained was added again to 160 mL of $CHCl_3$:MeOH (2:1, v/v) stirred for 1 h and centrifuged at 15 000 rpm for 30 min at 0 ºC. This step was repeated until the solvent was colorless. This yielded a residue enriched in carbohydrates, named as defatted biomass.

- Extraction of hot-water soluble polysaccharides

The defatted biomass from *N. oculata* was used for extraction of hot-water (HW) soluble polysaccharides (PS). 20 g of defatted biomass were suspended in 200 mL of distilled water (dH_2O), using a ratio of 1:10 (w/v). This solution was heated at 65 ºC for 3,5 h, then centrifuged at 15 000 rpm for 30 min at 0 ºC. The supernatant, containing the HW soluble PS, was recovered and the residue was submitted to a new extraction, as mentioned above. The residue of HW extraction was kept for analysis and the supernatants containing the HW soluble PS were ultra-filtered on a 10 kDa membrane filter, in order to isolate the PS. The

filtrate, with a molecular weight (*Mw*) under 10 kDa, and the retentate (*Mw* > 10 kDa) were concentrated and lyophilized.

- Sequential ethanol precipitation of the HW soluble PS

A portion of lyophilized retentate from the ultrafiltration of the HW soluble PS with a *Mw*>10 kDa was dissolved in dH_2O and fractionated by sequential ethanol (EtOH) precipitation, in order to separate and purify different PS soluble in HW. Firstly, absolute EtOH was added to the retentate solution to reach 50 % (v/v) of EtOH and the material precipitated (Et50) was recovered by centrifugation (at 1500 rpm for 30 min at 0 ºC). Then, absolute EtOH was added to the remained supernatant to reach 85 % (v/v) of EtOH and the material precipitated (Et85) was separated from the supernatant (EtSN) by centrifugation. Each precipitate and the remaining supernatant were dissolved in dH_2O, concentrated by rotary evaporation at 30 ºC and lyophilized. The lyophilized precipitates and supernatant were further dissolved in dH_2O, dialyzed (cut-off 12-14 kDa), and lyophilized again, yielding three fractions rich in PS (Et50, Et85 and EtSN). The Et50 and Et85 fractions corresponding to the precipitates obtained with 50 % and 85 % of EtOH, respectively, and the fraction EtSN to the remaining material containing in the supernatant.

- Anion-exchange chromatography of the fraction Et85

Polysaccharides present in the fraction Et85 were further separated by an anion-exchange chromatography as described by Wang *et al.* (115). Crude Et85 fraction (100 mg) was dissolved in Tris-HCl buffer (0.01 M, pH 7.6, 2.7 mL) and applied to a column (2.5 x 1.0 cm) of DEAE-Trisacryl M (Sigma-Aldrich) pre-equilibrated with the Tris-HCl buffer. The column was eluted with the initial buffer giving a positively or neutral charged fraction, named as Fraction #1 (or Et85_#1). The retained material on the column was eluted with a stepwise elution with Tris-HCl buffers containing an increase concentration of NaCl (0.125, 0.250, 0.500, and 1.0 M), giving four fractions named as Fraction #2 (or Et85_#2), #3 (or Et85_#3), #4, and #5. The column was washed with three column volumes for each elution buffer at a flow rate of 3 mL/h. Each fraction was dialyzed (cut-off 12-14 kDa) against dH_2O and lyophilized.

A diagram of the procedure employed to extract sequentially the PS from the biomass of *N. oculata* is schematized in figure 8.

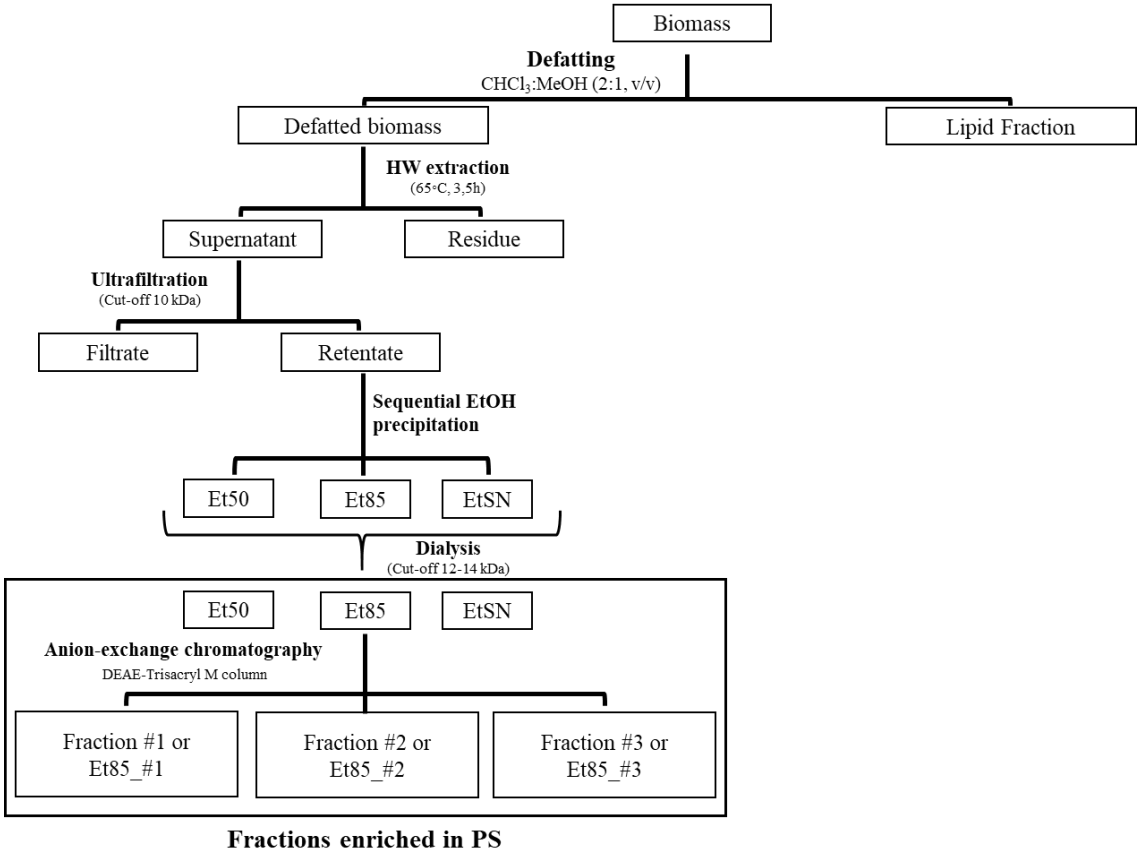


Figure 8. Diagram of the procedure used to extract sequentially the fractions enriched in polysaccharides from *N. oculata*.

Biomass of *N. oculata*, defatted biomass and the fractions from the extraction with HW and from the sequential EtOH precipitation were analysed to their content in protein. In addition, all fractions (biomass, defatted biomass, residue, filtrate, retentate, Et50, Et85, EtSN, and fractions #1, #2 and #3 from the anion-exchange chromatography of Et85) obtaining during the sequential extraction of PS were analysed to carbohydrates, in terms of neutral sugars, and content in uronic acids and sulfate groups. Filtrate was also analysed in terms of free alditols. Fractions from the sequential EtOH precipitation (Et50, Et85 and EtSN) and the fractions from the anion-exchange chromatography of Et85 (Fraction #1, #2 and #3) were more deeply studied in terms of PS characterization, being analysed also relatively to their constitution in non-free and free alditols, glycosidic linkages and determination of the *Mw* of the polymers. The glycosidic linkages present in the PS of the residue were also analysed.

2.1.3. Protein content

The content in protein was determined through elemental analysis converting the total content in nitrogen (% N) into protein through multiplication of % N by the nitrogen-to-protein conversion factor (*N*-Prot factor) 4.95 determined to *N. oculata* (116). The elemental analysis was performed in a Truspec 630-200-200 using 2 mg of each sample (duplicate). The operating temperatures of the combustion furnace was 1075 ^{o}C and the afterburner temperature was 850 ^{o}C. Thermal conductivity was used to detect the nitrogen.

2.1.4. RNase digestion of the fraction Et50

Part of the fraction Et50 obtained from the sequential EtOH precipitation was submitted to a treatment with RNase. The fraction was dissolved in dH_2O at a final concentration of 1.0 mg/mL (5 mg of Et50 in 5 mL of dH_2O). Then, it was added to the solution 500 µL of the Buffer 500 mM Tris-HCl at pH 7.5, 0.1 mM $CaCl_2$, 2.5 mM $MgCl_2$ with 0.06 % NaN_3, and the solution was agitated in a vortex for 30 s. Posteriorly, 10 mg/mL of RNase with a specific activity of 5000 U/mg protein (ThermoFischer Scientific – Molecular biology) was added to the solution to reach at final concentration of 10 µg/mL (5 µL). The solution was incubated at 37 ^{o}C in a water bath for about 15 h. Afterwards, the solution was dialysed (cut-off 12-14 000 Da) against dH_2O and the treatment with RNase was repeated to fully digest the RNA present in Et50. After second dialysis, the digested Et50 fraction was lyophilised and analysed in terms of neutral sugars (procedure described in section 2.1.5.).

2.1.5. Carbohydrate analyses

The procedures used to analyse the carbohydrates in terms of composition in **neutral sugars** (neutral monosaccharides), which are polyols containing one carbonyl functional group at the anomeric carbon, **non-free alditols**, which represent the polyols form of the respective monosaccharide linked to the PS, and **free alditols**, which represent the polyol form of the respective monosaccharide that is not linked to any compound, are presented in this section. In addition, the procedures used to quantify the **uronic acids (UA)**, which are

sugars that comprise both carbonyl and carboxyl acid functional groups, and **sulfate groups** are present, as well as the procedure used to determine the composition in **glycosidic linkages** present in the PS fractions. Furthermore, the procedure used to determine the **sulfate group position** in the PS and the method used to determine their ***Mw*** are presented.

2.1.5.1 Neutral sugars, non-free and free alditols analyses

Neutral sugars (NS) were analysed and quantified as their alditol acetates by gas chromatography (GC) coupled to a flame ionization detector (FID), using 2-deoxyglucose 1.0 mg/mL as internal standard (117,118). The neutral sugars present in 1-2 mg of the PS fractions (Et50, Et85, EtSN, fractions #1, #2 and #3) were obtained by a pre-hydrolysis in 0.2 mL of 72 % H_2SO_4 (w/w) for 3 h at room temperature (T_{room}) followed by 2.5 h hydrolysis with 1 M H_2SO_4 at 100 ºC. After hydrolysis, 200 µL of internal standard was added to the hydrolysed material. Neutral sugars were converted into their alditol acetates after reduction with $NaBH_4$ (15 % (w/v) in 3 M NH_3) (100 µL) at 30 ºC for 1 h, and acetylation with anhydride acetic (3 mL) in the presence of 1-methylimidazol (450 µL) at 30 ºC for 30 min.

Non-free and free alditols were also analysed as their alditol acetates by GC-FID but using 2-deoxyglucitol 1.0 mg/mL as internal standard. Non-free alditols present in 1-2 mg of samples were obtained by pre-hydrolysis followed by hydrolysis, as mentioned above for neutral sugars. On the other hand, free-alditols present in 1-2 mg of samples were directly dissolved in dH_2O (2.4 mL), without the pre-hydrolysis and hydrolysis step. Then, the internal standard (0.2 mg) was added to the mixtures and non-free and free alditols were converted into their alditol acetates after acetylation with anhydride acetic (3 mL) in presence of 1-methylimidazol (450 µL) at 30 ºC for 30 min.

Alditol acetates were extracted with dichloromethane (CH_2Cl_2) through a liquid-liquid extraction. dH_2O (3.0 mL) and CH_2Cl_2 (2.5 mL) were added to the mixtures and these were agitated by hand, centrifuged (3000 rpm for 30 s), and the aqueous phase was removed by suction. The organic phase was washed three times with 3.0 mL dH_2O, agitated, centrifuged and the aqueous phase was completely removed by suction. The organic phase was evaporated in a centrifugal concentrator at reduced pressure (Speedvac), anhydrous acetone (1 mL) was added and evaporated in Speedvac (twice), allowing to obtain the alditol acetates. Then alditol acetates were dissolved in anhydrous acetone and 2 µL from each sample were injected and analysed by GC-FID. A Perkin-Elmer Clarus 400 gas chromatograph with a

split injector and a FID detector was equipped with a 30 m x 0.25 mm DB-225 capilar column (Agilent J&W GC columns, USA) with a film thickness of 0.15 μm. The operating temperatures of the injector and the detector were 220 ºC and 230 ºC, respectively. The GC oven temperature program was set for an initial temperature of 200 ºC for 1 min, raised to 220 ºC at 40 ºC/min, holding for 7 min at this temperature, then raised to 230 ºC at 20 ºC/min, and held for 1 min. The flow rate of the carrier gas (H_2) was 1.7 mL/min.

2.1.5.2. Uronic acids content

The uronic acids (UA) present in the PS samples were obtained in simultaneous during neutral sugars analysis. After 1 h of hydrolysis at 100 ºC, an aliquot of 0.5 mL of the hydrolysed was removed and diluted in dH_2O for further UA analysis (117).

UA were quantified through the *m*-phenylphenol (MFF) colourimetric method as described by Coimbra *et al.* (118). Standards of D-galacturonic acid (D-GalA) in a concentration range of 0 to 80 μg/mL were prepared in order to build a calibration curve. To 500 μL of sample (hydrolysed) or standard (triplicates, one blank and two replicas) were added in an ice bath, 3 mL of 50 mM sodium borate prepared in concentrated sulphuric acid. Then, the mixtures were heated in a boiling water bath (100 ºC) for 10 min, cold down in ice, and in the dark 100 μL of MFF 0.15 % (w/v) in 0.5 % (w/v) NaOH was added to the replicas of each sample or standard. After reaction for 30 min in the dark with MFF, the absorbance was read at 520 nm. UA concentrations were estimated from the calibration curve built with D-GalA as standard.

2.1.5.3. Determination of sulfate content

The sulfate was quantified by turbidimetry for determination of the sulfate ester of sulfated PS, as described by Dodgson and Price (119). The turbidimetry of the solution has origin in a reaction that occurs between the sulfate groups present on the PS and barium. The intensity of the turbidimetry is directly proportional to the amount of sulfates in solution. The cloudy produced can be quantified by spectrophotometry, reading the absorbance at 360 nm, in comparison with a calibration curve built with sulfate potassium (K_2SO_4) solutions as standard, in a range of 10 to 300 μg/mL of SO_4^{2-}.

Firstly, 2 g of gelatine were dissolved in 400 mL of hot water (60-70 ºC) and let *overnight* at 5 ºC. The gelatine solution was divided into two parts (200 mL) 2-3 h before

using and to one of the parts was added 1 g of barium chloride ($BaCl_2$). Barium chloride-gelatin reagent was used as a cloud-stabilizer of the barium sulfate formed.

Depending on the total content in sugars (NS and UA), a mass of each sample corresponding to 10 % of sugars was dissolved in 1 M HCl (1 mL) and the solution was hydrolysed at 105–110 ºC for 5 h. After hydrolysis, one aliquot of sample or standard of K_2SO_4 solution (0.2 mL) was transferred to a tube containing 3.8 mL of 3 % (w/v) trichloroacetic acid (TCA) and 1 mL of a $BaCl_2$-gelatin reagent, reacting for 20 min under stirring. Then, the released barium sulfate suspension absorbance was measured at 360 nm against a reagent blank consisting of 0.2 mL of water, 3.8 mL of 3% (w/v) TCA, and 1 mL of the $BaCl_2$-gelatin reagent. A second aliquot (0.2 mL) of the hydrolysed or standard was transferred to another tube containing 3.8 mL of 3% (w/v) TCA and 1 mL of gelatin solution without $BaCl_2$. The absorbance of this 'control' solution was then measured at 360 nm against the blank. This control represents the ultraviolet-absorbing compounds produced during hydrolysis. The resulting value was subtracted from those analysed in presence of $BaCl_2$-gelatin reagent. The sulfate content as SO_3^- was determined from the calibration curve built with the K_2SO_4 standard solutions.

2.1.5.4. Glycosidic linkages analysis

Methylation

Glycosidic linkages composition was determined by GC coupled to mass spectrometry (MS) of the partially methylated alditol acetates (PMAAs). These were obtained from methylated PS that were further hydrolysed, reduced and acetylated (120,121). Before methylation of PS, samples (1-2 mg) were dried *overnight* at T_{room} in a vacuum oven in presence of P_2O_5. Then, the samples were dissolved in 1.0 mL of dimethyl sulfoxide (DMSO) under low agitation *overnight*. Powdered NaOH (40 mg) was added under argon (Ar) atmosphere to the samples and let 30 min under magnetic agitation. The methylation reaction was performed by adding to the samples CH_3I (80 µL). Samples were methylated three times, reacting 20 min under vigorous agitation between each addition of CH_3I in order to methylate all the free hydroxyl groups of the PS. Afterwards, the methylated PS were dissolved in 3 mL of $CHCl_3$:MeOH (1:1, v/v) and dialysed against 50 % EtOH:H_2O (three times). After dialysis, samples were concentrated by Speedvac and submitted to another methylation procedure and subsequent dialysis, allowing to obtain the methylated PS.

Reduction of carboxyl groups of methylated PS

Part of the dried methylated fractions from the PS fractions were carboxyl-reduced to determine the UA present. The methylated fraction was heated in a sealed tube with a mixture of $LiAlD_4$ (20 mg) in tetrahydrofuran (THF) (1 mL) at 65 ºC for 4 h. The excess of reagent was then destroyed with ethanol (2-3 drops) and the pH of the mixture was adjusted to neutrality with 2 M H_3PO_4. The reduced polymer (supernatant) was isolated by addition of 2 mL of $CHCl_3$:MeOH (2:1, v/v), and the precipitate was removed by centrifugation (3000 rpm for 30 s). The remaining precipitate was washed three times with $CHCl_3$:MeOH (2:1, v/v) and centrifuged. The supernatants were combined and evaporated to dryness in Speedvac.

Hydrolysis, Reduction and Acetylation of methylated PS

The methylated PS and the carboxyl-reduced methylated PS were hydrolysed with 500 µL of 2 M trifluoroacetic acid (TFA) at 121 ºC for 1 h, and dried in Speedvac. The reduction of partially methylated monosaccharides was performed with 20 mg of $NaBD_4$ in 3 M NH_3 (200 µL) and dH_2O (100 µL) for 1 h at 30 ºC. Then acetylated with anhydride acetic (3.0 mL) in presence of 1-methylimidazol (450 µL) at 30 ºC for 30 min. PMAAs were extracted with CH_2Cl_2, as for the alditol-acetates in the section of neutral sugars analysis, and PMAAs (1 µL) were analysed by GC-MS. A gas chromatograph mass spectrometer (GCMS-QP2010, SHIMADZU) with a split injector was used, equipped with a 30 m DB-1 column (Agilent J&W GC columns, USA) with diameter and film thickness of 0.25 mm and 0.10 µm, respectively, and connected to an Agilent 5973 quadrupole mass selective detector. Samples were injected in a split mode (2.0 min) with the injector operating at 250 ºC. The GC oven temperature program was set at an initial temperature of 80 ºC, raised to 140 ºC at 10 ºC/min, holding for 5 min, raised to 150 ºC at 0.20 ºC/min, then raised to 250 ºC at 60 ºC/min, holding for 2 min. The flow rate of the carrier gas (He) was set at 1.84 mL/min. The mass spectrometer was operated in the electron impact mode (EI) at 70 eV scanning the range 50–700 m/z, in a full scan acquisition mode. Chromatogram peaks were identified comparing all mass spectra with a laboratory made database of PMAAs. A representative scheme of the GC-MS analysis of the PMAAs derived from fraction Et85_#1 to determine the glycosidic linkage is shown in supplementary material, supplementary information 2.

2.1.5.5. Desulfation of Et85 fraction

Desulfation of the Et85 fraction was performed in order to determine the linkage position of sulfate esters in PS. The procedure was followed using a modified method described by Miller *et al*. (122) and Geresh *et al.* (20). 3 mg of sample were dissolved in dried DMSO (1.8 mL). Then, 100 µL of pyridine was added followed by addition of 6.5 mg of pyromellitic acid (Sigma-Aldrich), 12 mg of flourine sodium (NaF) (Fluka), and additional 200 µL of pyridine (Sigma-Aldrich). The mixture was stirred and incubated at 120 ^{o}C for 3 h. Posteriorly 3 % (w/v) sodium bicarbonate ($NaHCO_3$) solution (1 mL) was added and the desulfated PS in the reaction mixture solution were purified by dialysis against dH_2O and then recovered by lyophilisation. The desulfated PS were then analysed by methylation, using the exchange of solvent for extraction of methylated PS and without reduction of carboxyl groups, as described above in the glycosidic linkages analysis.

2.1.5.6. Determination of molecular weight of the PS

The average *Mw* of the polymers present on EtOH fractions and on the fractions #1, #2 and #3 from the anion-exchange chromatography were determined by a size exclusion chromatography (SEC) analysis. The SEC analysis was carried out using two PL aquagel-OH MIXED 8 µm, 300 x 7.5 mm, columns protected by a PL aquagel-OH Guard 8 µm pre-column on a PL-Gel Permeation Chromatograph (GPC) 110 system (Polymer Laboratories, UK). The pre-column, the SEC columns, the injection system, and the refractive index detector were maintained at 36 ^{o}C during the analysis. Standards and samples were dissolved in 0.1 M $NaNO_3$ aqueous solutions to a concentration of about 0.5 % (5 mg/mL). The eluent (aqueous 0.1 M $NaNO_3$ solution) was pumped at a flow rate of 0.9 mL/min. The analytical columns were calibrated with pullulan standards (Polymer Laboratories, UK) in the range of 1–100 kDa.

2.1.6. Statistical analysis

Statistical analysis was performed through two statistic tests using the Excel software (Microsoft office 2013). The F-test was used to analyse the variance and the Student's T-test was used to detect significant differences between samples. Differences were deemed significant at p-value<0.05.

2.1.7. Recombinant expression of proteins from *Bacteroides thetaiotaomicron* VPI-5482: BT3698 and BT0996

Carbohydrate Active EnZymes (CAZymes) may contain non-catalytic domains that are able to bind carbohydrates and increase the enzyme activity (www.CAZy.org, (107)). These domains are named as carbohydrate-binding modules (CBMs). Although it has been increasing the discovery of CBMs, its study in terms of carbohydrate-binding specificities still remains understudied. The bacteria *Bacteroides thetaiotaomicron* resident in the human gut expresses enzymes involved in degradation of complex PS that comprise CBMs. Here, it will be presented the procedure employed to produce two CBM domains belonging to two different proteins in order to use them in microarrays.

2.1.7.1. DNA, bacterial strains and plasmids

The plasmid encoding the different proteins used in this study were previously cloned by Correia *et al.* in collaboration with NZYTech® (Lisboa, Portugal). The cloned domains of the proteins BT3698 and BT0996 correspond to the appended non-catalytic CBM modules of each protein. In BT3698 the CBM domain cloned was from the family 58 CBM (CBM58), natively appended to the glycoside hydrolase (GH) family 13 (GH13), whereas in BT0996 the CBM domain cloned was the CBM57 appended to the GH2 (Molecular architecture of each native protein expressed by *B. thetaiotaomicron*, and the protein parameters is described with more detail in supplementary material, supplementary Fig. 3 and Table 1). The selected domains (CBM58 and CBM57) were cloned into a pHTP1 vector with kanamycin (Kan) resistance. The produced plasmids encode for recombinant proteins containing an N-terminal tag with 6 histidines with the following sequence: MGSSHHHHHHSSGPQQGLR. A representative scheme of the recombinant proteins overexpressed in *Escherichia coli* competent cells (described in section 2.1.7.2) is shown in supplementary Fig. 4.

Expression of BT3698 and BT0996 is under control of the T7 phage promotor, induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) or auto-induction culture medium (described in more detail in section 2.1.7.4.).

2.1.7.2 Competent cells

Competence is the ability of a cell to take up extracellular DNA in the process called transformation. To obtain the competence state, cells have to be chemically treated to make them transiently permeable to DNA. Competent cells are used to produce recombinant proteins because they are able to incorporate strange DNA without degrading it and to express the information contained in this DNA. The competent cells used in this study to incorporate the DNA plasmids that encodes for the recombinant proteins BT3698 and BT0996 were the expression strains *Escherichia coli* BL21 (DE3) and *E. coli* TUNER (DE3), respectively. The used batch of competent cells were purchased from NZYTech® (Lisboa, Portugal).

2.1.7.3. Transformation

The transformation of *E. coli* BL21 (DE3) and TUNER (DE3) cells used for expression of the proteins was performed in a sterile environment. For the transformation process, 5 µL of DNA plasmid of BT3698 was added to 50 µL of competent cells *E. coli* BL21 (DE3) and 5 µl of DNA plasmid of BT0996 was added to 100 µL of competent cells *E. coli* TUNER (DE3). The cells were kept on ice for 30 min, then heated at 42 ºC on a heat block for 1 min and then incubated on ice for 5 and 2 min to BL21 and TUNER cells, respectively. Then, the cells were re-suspended in 1 mL of Luria-Bertani (LB) medium (Medium culture composition described in supplementary information 3) and incubated at 37 ºC and 200 rpm for 1h30min (Orbital ShakerIncubator ES-20, from Grant.bio). The cell suspension was centrifuged for 5 min at 6.7 rpm and 900 µL of the supernatant was discarded. The remaining cells were re-suspended in the remaining supernatant and spread in LB-agar (LA) plate with Kan antibiotic at 50 µg/mL. The DNA plasmids contain a gene with Kan resistance and consequently only cells that contain the plasmid will be able to grow and form colonies.

2.1.7.4. Expression of proteins BT3698 and BT0996

In order to obtain the most favourable expression condition for each CBM cloned domain of the BT3698 and BT0996 proteins, two different expression protocols were used (Table 4). One protocol referred to as Auto-induction for BT3968 and the other as IPTG-induction for BT0996. The main differences between the two procedures are: 1) different

cell culture medium, 2) the use of IPTG for inducing the expression of protein and 3) the temperature of expression.

Table 4. Main differences in the auto-induction and IPTG-induction protocols for protein expression.

Protocol	Cell culture medium	IPTG induction	Temperature of protein expression (ºC)	Period (Hours)
Auto-induction	Auto-induction[a]	No	25	*Overnight* (in about 15 H)
IPTG-induction	LB[b]	Yes	37	3

[a]- Auto-induction medium from NZYTech® (composition in Supplementary material, supplementary information 3)
[b]- LB – Luria-Bertani medium culture (composition in Supplementary material, supplementary information 3)

- Protocol for expression of BT3698 with auto-induction

After transformation of *E. coli* BL21 (DE3) cells (section 2.1.7.3.), one colony was picked up from LA plates with Kan antibiotic 50 μg/mL and pre-inoculated in 10 mL of LB medium 50 μg/mL Kan. The pre-inoculum was incubated at 37 ºC and 200 rpm *overnight* (Bench-top shaker-incubator ES-20, Grant Bio). Then, the pre-inoculum was transferred to 250 mL of auto-induction medium containing 250 μL of 50 mg/mL Kan and incubated at 37 ºC until OD_{600nm} reached 1.5 and further incubated at 25 ºC and 180 rpm (Floor refrigerated shaker-incubator IS-971R, Lab Companion) for protein expression.

- Protocol for expression of BT0996 with IPTG-induction

After transformation of *E. coli* TUNER (DE3) cells (section 2.1.7.3.), one colony was pre-inoculated in 10 mL of LB medium with 10 μg/mL Kan. The pre-inoculum was incubated at 37 ºC and 150 rpm *overnight* (Bench-top shaker-incubator ES-20, Grant Bio). In the following day, 5 mL from the pre-inoculum was added to 500 mL of LB medium with 10 μg/mL Kan. The cell culture was incubated at 37 ºC and 180 rpm (Floor refrigerated shaker-incubator IS-971R, Lab Companion) until OD_{600nm} reached 0.6-0.8. Once cell culture reached this density, 0.2 mM IPTG was added and the incubation was carried out at 180 rpm using 37 ºC and 3 h as desired temperature and period of induction.

2.1.7.5. Cell harvesting and lysis

The cells were harvested by centrifugation at 5000 rpm for 10 min at 4 ºC (Refrigerated centrifuge Avanti J-26 XPI with JA-10 rotor, Beckman Coulter). After discarding the supernatant, the cells *E. coli* BL21 (DE3) were re-suspended in lysis buffer containing 50

mM Hepes pH 7.5, 1 M NaCl, 5 mM $CaCl_2$ and 10 mM imidazole, and the *E. coli* TUNER (DE3) cells were re-suspended in lysis buffer that contained 50 mM Hepes pH 7.5, 1 M NaCl, 2.5 mM $CaCl_2$, and 10 mM imidazole (10 mL per gram of cell pellet).

Cells were disrupted using a sonicator (UP100H, Hielsher) with 3-5 cycles of 1 min each cycle. This method was used to disrupt cell membranes and release cellular content, including the recombinant proteins. The soluble fraction, where remained the soluble proteins, was clarified by centrifugation (Refrigerated centrifuge Avanti J-26 XPI with JA25.20 rotor, Beckman Coulter) at 15000 rpm and 4 ºC for 30 min and continued to purification protocol.

2.1.7.6. Purification of BT3698 and BT0996 by affinity chromatography

An efficient method to purify the recombinant proteins (BT3698 and BT0996) from other bacterial proteins is the immobilized-metal affinity chromatography (IMAC) and that is the reason for including a N-terminal His-tag (Section 2.1.7.1.) in the cloned domain of the recombinant protein.

IMAC is based on the strong affinity and selectivity that transition bivalent metal ions such as Zn^{2+}, Cu^{2+}, Ni^{2+}, and Co^{2+} have to bind histidine and cysteine residues in aqueous solutions. These metals are immobilized in the chromatography column and bind to His-tagged and Cys-tagged proteins (123).

Purification of BT3698 and BT0996 proteins were carried out through different methods. BT3698 was purified through the affinity chromatography, using His Gravitrap™ columns, and BT0996 using ÄKTA START.

- Affinity purification with His Gravitrap™ columns

His GraviTrap™ (GE Healthcare) is a prepared, single-use column for purification of histidine-tagged proteins by IMAC that allows fast and simple gravity-flow purifications.

This procedure can be divided into four steps: equilibration, sample application, washing and elution. For this procedure 3 different buffers were used: Buffer A (Lysis buffer) – 50 mM HEPES, 1 M NaCl, 5 mM $CaCl_2$ and 10 mM imidazole at a final pH of 7.5; Buffer B - 50 mM HEPES at pH 7.5,1 M NaCl, 5 mM $CaCl_2$, and 60 mM imidazole at a final pH of 7.5; Buffer C - 50 mM HEPES at pH 7.5, 1 M NaCl, 5 mM $CaCl_2$ and 300 mM imidazole at a final pH of 7.5.

Firstly the columns loaded with Ni^{2+} were washed with 5 mL of 20 % (v/v) of EtOH and then with 10 mL of Milli-Q water to remove the EtOH. After this, the columns were equilibrated in buffer A. The extracted sample (obtain in section 2.1.7.5.) was then loaded into the column and the His-tag from the protein binds to the Ni^{2+} of the column. In order to wash other *E. coli* proteins, and thus prevent contamination of the sample, 10 mL of buffer A and then another 10 mL of buffer B were loaded. The recombinant protein BT3698 was eluted with 5 mL of buffer C. The fractions were eluted in separate and the fraction eluted with buffer C, containing the protein BT3698, was saved for quantification, analysis in sodium-dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) for confirmation of the purity of the protein, and posterior use in carbohydrate microarrays.

- Purification with ÄKTA START

For purification of BT0996 protein, this was submitted to an IMAC using a 5 mL His Trap^{TM} sepharose column loaded with Ni^{2+}, which is coupled to the chromatograph ÄKTA START, both from GE-Healthcare.

In this protocol two different buffer solutions were used for the concentration gradient: Buffer A (lysis buffer) – 50 mM HEPES, 1 M NaCl, 2.5 mM $CaCl_2$ and 10 mM imidazole at a final pH of 7.5; Buffer B (elution buffer) – 50 mM HEPES at pH 7.5, 1 M NaCl, 5 mM $CaCl_2$ and 500 mM imidazole at a final pH of 7.5. The column was firstly washed with approximately 50 mL of Milli-Q water and subsequently equilibrated with 20 mL of buffer A. The cell extract (soluble fraction obtained in section 2.1.7.5.) was then loaded into the column. A first wash with 15 % of buffer B was applied in order to separate the constituents with low affinity for Ni^{2+}. Then, a concentration gradient of imidazole, from 10 % to 100 % of 100 mL of buffer B, was setup in order to collect BT0996. The collected fractions were subsequently quantified and analysed by SDS-PAGE, and saved to posterior use in carbohydrate microarrays.

2.1.7.7. Protein quantification

The proteins were delivered in a micro-well plate at a final volume of 2 μL each in 50 mM HEPES at a pH of 7.5, 1 M NaCl, 5 mM $CaCl_2$ and 300 mM imidazole. The protein concentration was determined with a SpectraMax®190 Absorbance Plate Reader from Molecular Devices. Absorbance was measured at 280 nm, which corresponds to the wavelength in which proteins absorb due to their aromatic amino acid residues. As a control,

the protein buffer (50 mM HEPES at a pH of 7.5, 1 M NaCl, 5 mM $CaCl_2$ and 300 mM imidazole) was used. Concentration was calculated by applying the Lambert-Beer equation: $A = \varepsilon cl$, considering the reading absorbance (A), the molar extinction coefficient (ε) and the length (l = 0.05 cm) of the microplate. The ε_{280nm} of BT3698 is 26930 $M^{-1}.cm^{-1}$ and of BT0996 is 16960 $M^{-1}.cm^{-1}$.

2.1.7.8. SDS-PAGE analysis

In this work, a polyacrylamide gel electrophoresis (PAGE) was performed in denaturing conditions in the presence of the anionic detergent sodium-dodecyl sulfate, SDS-PAGE. SDS disrupts the protein structure to produce a linear polypeptide chain coated with a uniform layer of negatively charged SDS molecules. As a consequence, the proteins will be separated according to their *Mw*, without the influence of the intrinsic negative charge. The protocol followed using resolving gel of 10 % acrylamide.

The preparation of protein samples was carried out by mixing the respective protein (BT3698 and BT0996) solution (20 µL) with 5 µL of loading buffer (1 M Tris-HCl pH 6.8, 4 % (w/v) SDS, 55 % (v/v) glycerol, 5 % (w/v) DTT, and 0.02 % (w/v) Comassie Blue R250) and boiling in a heat block at 100 ºC for 5 min. After a fast spinning, 20 µL of the denatured protein sample and 4 µL of protein marker was delivered to each well. The marker used in this study was the NZYColour Protein Marker II (PMII). The run parameters for the electrophoresis were setup to a fixed amperage of 60 mA. Electrophoresis gels were stained with NZYTech BlueSafe.

2.1.8. Development of microarrays with the PS extracted from *N. oculata*

To further characterize the chemical structures of the HW soluble PS extracted from *N. oculata* and search for carbohydrate-protein interactions, fractions Et50, Et85, EtSN, Et85_#1, Et85_#2 and Et85_#3 (Section 2.1.2.) were used as polysaccharide probes in microarrays. The printed array was probed with proteins with well-known carbohydrate-binding specificities and with the two recombinant proteins produced in section 2.1.7.

2.1.8.1. Construction of the carbohydrate microarrays

Carbohydrate microarrays developed with the PS fractions from *N. oculata* was designated as Microalgae PS set 1 and comprised a total of 16 carbohydrate probes, which

are described in table 5. Six of these probes were the PS fractions from *N. oculata* and the other ten were commercial PS used as control for the proteins tested in binding assays (described in section 2.1.8.2.).

Table 5. List of all carbohydrate probes that composed the Microalgae polysaccharides (PS) set 1, which were analysed in the binding charts and in the matrix (heat-map). The microarray comprises PS from *N. oculata* and from commercial sources. The main sugar composition in the commercial polysaccharide probes is described.

Carbohydrate Probes Set 1	**Source**	**Main sugar composition**
Et50	*N. oculata*	-
Et85	*N. oculata*	-
Et85_#1	*N. oculata*	-
Et85_#2	*N. oculata*	-
Et85_#3	*N. oculata*	-
EtSN	*N. oculata*	-
Pullulan	- (Megazyme)	Linear mixed-linked (α1→4, α1→6)-D-Glucan
α-Mannan	*S. cereviseae* (Sigma-Aldrich)	Linear (α1→6)-D-Mannan
Glucomannan	- (Megazyme)	Linear mixed-linked (β1→4)-D-Mannan with (β1→4)-D-Glc
Galactomannan	- (Megazyme)	Branched (β1→4)-D-Mannan with (β1→4)-D-Gal
Rhamnogalacturonan I (RGI)	Soybean (Megazyme)	(α1→4)-GalA with (α1→2)-Rha
Low Methylated Galacturonate (LMG)**	Apple (OligoTech)	Low methylated (α1→4)-GalA, DE* < 5%
High methylated Galacturonate (HMG)**	Apple (OligoTech)	High methylated (α1→4)-GalA, DE > 70%
Polygalacturonic acid (PGA)**	Apple (Sigma-Aldrich)	Gal (4.5 % dw) as t- and (1→4)-Gal, and (1→4)-GalA Glc (2 % dw) Ara (0.9 % dw) as (1→4)-Ara*p* or (1→5)-Ara*f* Rha (0.7 % dw) as (1→2)-Rha
PGA**	Citrus (Megazyme)	Gal (4.5 % dw) as (1→4)- and t-Gal, and (1→4)-GalA Glc (3 % dw) Ara (0.8 % dw) as (1→4)-Ara*p* or (1→5)-Ara*f* and t-Ara*f* Rha (0.5 % dw) as (1→2)-Rha
Mixed β-Glucan**	Oat (Megazyme)	Mixed-linked (β1→3, β1→4)-D-glucan

*DE – Degree of esterification.
**- Carbohydrate composition described in supplementary material, supplementary information 4.

For the construction of the microarray, probes were non-covalently immobilized onto 2-pad nitrocellulose-coated FAST® (Whatman) glass slide (example of a single pad slide in figure 9), using the MicroCaster™ System. This is an arrayer system alternative to robotic arrayers that can be used to manually-arraying molecules. The non-covalent immobilization

of the PS on the slide occurs by contact, unlike the robotic arrayers. The MicroCaster allows to print spots with 500 µm diameter each and with a volume per spot between 20 to 70 nL.

Each carbohydrate probe was printed in triplicate at two levels: 0.1 and 0.5 mg (dw) /mL (30 and 150 pg/spot). Sample solutions for printing were prepared in the following way: each carbohydrate probe was dissolved in H_2O Milli-Q with 0.02 % (w/v) NaN_3 at a final concentration of 0.1 or 0.5 mg/mL containing 0.2 µg/mL cyanine 3 (Cy3) at a final volume of 100 µL. Cy3 is a fluorophore that was added as a marker for spot location of the printed probes and for quantitation analysis (56). Then, sample solutions (20 µL) were transferred into a 384-well plate named as source plate (AB-1384/G 384-well PCR Plates, Thermo Scientific) for manual printing according to the planned microarray layout. A representative scheme of the printing/arraying of probes onto a single pad slide (identical process used in 2-pad slides) is shown in figure 9.

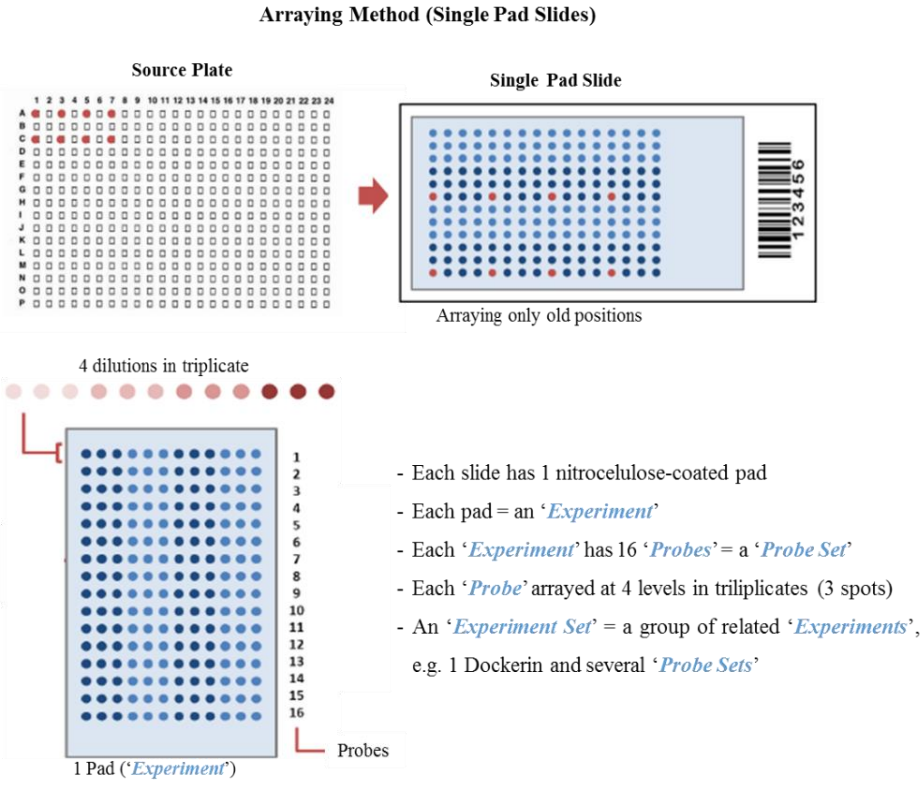


Figure 9. Scheme of the arraying method using single pad slides. An example of a layout of one single pad slide with 16 probes printed at 4 different dilutions in triplicate each is represented, as well as designations used during a carbohydrate microarray experiment. Image adapted from ISBio training course, Microarray screening analysis supporting material, 2015, Caparica, Portugal (124).

2.1.8.2. Carbohydrate microarrays binding assay

In this work, a total of seven proteins were analysed using the Microalgae PS set 1, two lectins, two antibodies (Abs), and three microbial CBMs (a summary of the conditions used during the binding assay for each protein analysed in this work is described in supplementary table 2). In addition, the two CBM domains of the proteins BT3698 and BT0996 produced in section 2.1.7. will be employed. Description of the reported specificities for each protein used in microarrays is described in table 6.

Table 6. List of the lectins, antibodies and CBMs used in the carbohydrate microarrays developed with Microalgae PS set 1 and their reported specificities.

Proteins	Specificity	Reference
Lectins		
Concanavalin A (ConA)	α-Man α-mannose-linked oligosaccharides (α1→2)- and (α1→3)-mannose oligosaccharides	(85,90)
Ricinus Communis Agglutinin 120 (RCA_{120})	β-Gal > α-Gal Terminal Gal(β1→4) > terminal Gal(β1→3) and Gal(β1→6) Terminal Gal(β1→4) in *N*-glycans Gal(6S)(β1→4)AnGal(α1→3)Gal(6S)(β1→4)AnGal(α1→3)Gal(6S) Gal(2S)(β1→[4)-Gal(2S6S)(α1→3)Gal(2S)(β1→]24)-Gal(2S6S)	(61,62,85,91)
Antibodies		
Anti-Heteromannan	(β1→4)-linked mannan; Glucomannan and galactomannan polysaccharides; (β1→4)-manno-oligosaccharides (DP2 to DP5).	(100,101,125)
Anti-(β1→3, β1→4)--D-glucan	Linear mixed-linked (β1→3, β1→4)-Glc	(99,102)
Microbial CBMs		
*Cm*CBM6-2	(β1→4)-, (β1→3)-, (β1→3, 1→4)- and (β1→2)-Glc or (β1→4)-Xyl; Linear (β1→3)- and (β1→4)-glucans Mixed-linked (β1→3, β1→4)-Glc Linear (β1→2)-gluco-oligosaccharides with DP-2 and longer	(92)
BT3698 (CBM58)	Starch	(112)
BT0996 (CBM57)	RG-I and RG-II (putative)	(112,114)

Firstly, the printed slides with the Microalgae PS probes of set 1 were scanned for Cy3 (532 nm; 10 µm resolution, 5-10 % laser power, 350 photomultiplier (PMT) gain) using GenePix® 4300A microarrays scanner (Molecular Devices), for visualization of printed probe spots (Fig. 10). After scanning, each pad of a nitrocellulose-coated 2-pad glass slide was wetted with 500 µL of water. The slide was gently tilting to remove the excess of water and the respective blocking solution (500 µL) for each protein was added to each pad and incubated for 60 min on dark at T_{room} (Blocking solutions of each protein are described in

supplementary table 2). This step protects the printed array from smearing and the unspecific linkage of the proteins onto surface without spots. The blocking solution was comprised of a protein solution such as Casein 0.02 % with 1 % BSA made in HEPES-buffered saline (HBS) with 5 mM $CaCl_2$, 5 mM HEPES pH 7.4, 150 mM NaCl, 5 mM $CaCl_2$, or 3% BSA (w/v) also made in HBS with 5 mM $CaCl_2$. The stock solutions used to prepare the different blocking solutions were 30 % (w/v) BSA (Sigma-Aldrich, A8577) and 1 % Casein (Thermo 37583).

After 60 min of incubation, the blocking solution was removed and the pads were washed once with HBS (500 µL). After the HBS was removed, the protocol followed for the binding step of lectins, antibodies and CBMs was different.

The biotinylated **lectins** ConA and RCA_{120} were analysed by one step, adding 500 µL onto the pad at the desired concentration (prepared in the diluent solution) and incubated for 90 min.

For **antibodies**, 500 µL of the primary antibody (Rat IgM and Mouse IgG) (prepared at the desired concentration in the diluent solution) was added to each pad and incubated for 90 min at T_{room}. After this period, the biological sample was removed through tilting the slide gently and the pad was washed 4 times with 500 µL of HBS buffer. The excess of solution was removed and 500 µL of biotinylated detection Abs (secondary Abs), prepared at the desired concentration in the diluent solution, were added to the pad and incubated for 60 min.

The **His-tagged CBMs** were analysed using a pre-complex (1:3:3, protein:Ab1:Ab2) formed with detection Abs as follows: protein-antibody complexes were prepared by pre-incubating the primary detection Ab (mouse anti-His), Ab1, with the secondary biotinylated detection Ab (biotinylated anti-mouse IgG), Ab2, for 15 min at T_{room}, followed by addition of the protein (His-tagged *Cm*CBM6-2, BT3698 and BT0996**)** for further 15 min. After this period, the protein-antibody complexes were diluted to a final volume of 500 µL with diluent solution and the pre-complexes solution was added to the pads.

For the detection of binding signals, 500 µL of Alexa-Fluor (AF)-647-labeled streptavidin (1 µg/mL in the respective blocker solution for each protein) was added to each pad and incubated for 45 min. Between incubations, each pad was washed four times with HBS buffer (500 µL). After the last incubation with streptavidin, pads were washed four times with HBS buffer and two times with Milli-Q water and dried in the dark for 5 min.

Each microarray slide was then scanned for AF-647 (635 nm; 10 µm resolution; 5-90 % laser power, 350 PMT gain) for detection of the binding event using the GenePix® 4300A microarray. The Cy3 and AF-647 fluorescence intensity for each spot was quantified using the GenePix® Pro Software (Molecular Devices). A representative scheme of the carbohydrate microarray binding experiment is shown in figure 10.

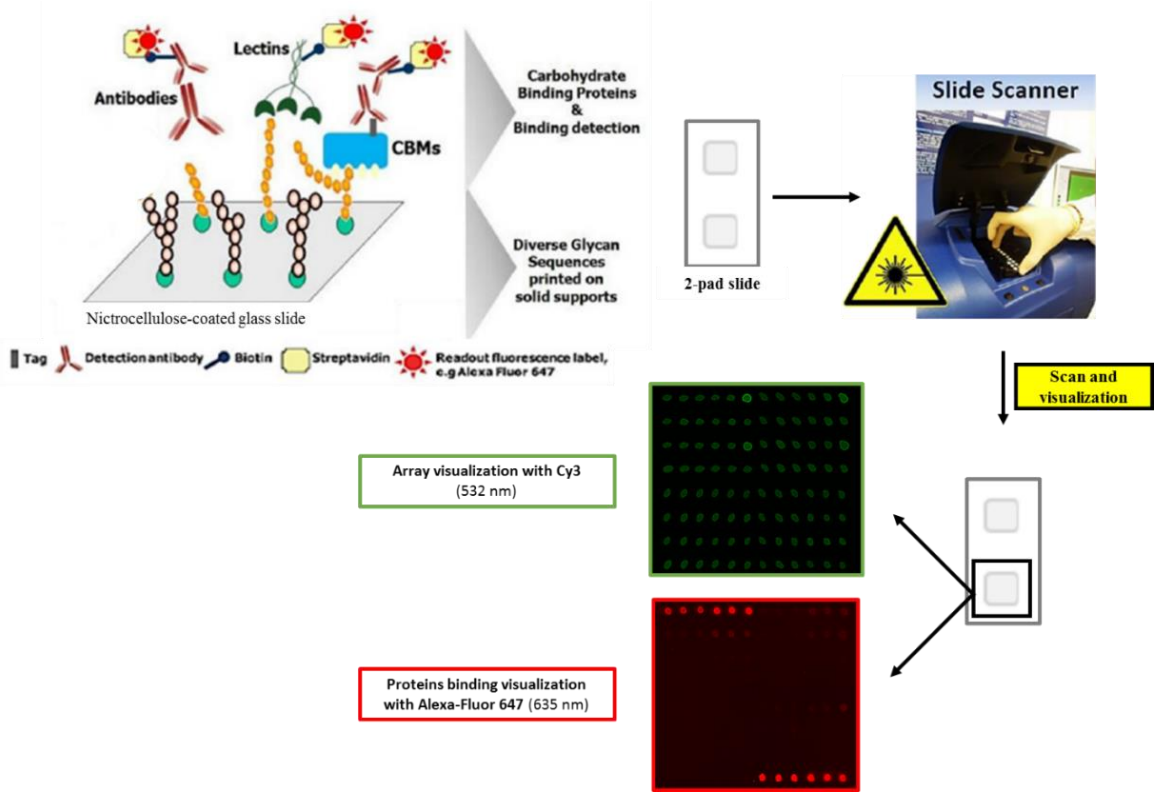


Figure 10. Scheme of the carbohydrate microarray binding experiment using slides of 2 pads. Adapted from ISBio training course, Microarray screening analysis supporting material, 2015, Caparica, Portugal (124).

2.1.8.3. Carbohydrate microarrays data analysis

The scan of the printed slides for the Cy3 (532 nm) before performing the binding assay and the scan for AF-647 (635 nm) after the binding assay are essential for the quantification of the results. The first scan allows to visualize and locate the probe spots, whereas the scan for AF-647 allows to visualize the spots where occurs the binding between the protein and the printed probe. Analyses and quantitation of the data was performed using the GenePix® Pro Software (Molecular devices), a program for imaging analysis and quantitation. For scanning the slides, it is necessary to choose the following parameters: laser wavelength (532 nm for Cy3 and 635 nm for AF-647) and corresponding emission filters;

photomultiplier (PMT) gain and laser power (%) according to the intensity of the fluorescence spot signals versus the background.

After the image acquisition, the process for data analysis can be divided into three steps: 1) construction of the grid for each slide, which locates all the spots in each microarray pad referred as '*block'*; 2) quantitation of the fluorescence intensity; and 3) processing and presentation of the quantified data in form of charts, tables or matrices (heatmaps). These three steps are described/explained below:

1) and 2) *Construction of the grid* and *Quantitation of the fluorescence intensity*. The grid is constructed using the Cy3 image (.TIFF) scan as reference and then is adjusted to the AF-647 slide scan (Fig. 10) for posterior quantitation of the results. The grid is saved as a GenePix settings files (.gps) and allows the location and quantitation of the fluorescence intensity of the spots in each microarray block.

The program calculates the fluorescence intensity associated to each spot and outside the spot, which is referred as local background. The fluorescence intensity considered for each spot is subtracted to local background. After this analysis the program saves all the scanning parameters, raw images and the numerical values in a GenePix results file (.gpr).

3) *Processing and presentation of the quantified results*. Charts and one heatmap were constructed for processing the results. The charts were constructed using the GraphPad Prism® 6 program by input the quantified data, whereas the heatmap was manually constructed using the Excel (Microsoft office 2013).

2.2. Characterization of glycolipids from *N. oculata*

2.2.1. Extraction of lipids from biomass of *N. oculata*

Biomass of *N. oculata* purchased from Necton S.A. (description in section 2.1.1.) was defatted with 2:1 (v/v) $CHCl_3$:MeOH using a modified method described by Wu *et al.* (126). The biomass of *N. oculata* (in about 40 mg) was suspended in 3.0 mL of $CHCl_3$:MeOH (2:1, v/v), stirred for 10 min and centrifuged at 4000 rpm for 10 min. The solvent phase, comprising the lipid compounds, was separated from the defatted biomass. This procedure was repeated three times until the total lipids were extracted. The solvent phases from each extraction were combined and evaporated at T_{room} in Speedvac. Then, total lipids were quantified by gravimetry using an analytical balance (Precisa, 40SM-200A). The lipids were dissolved in $CHCl_3$:MeOH (2:1, v/v), dialyzed (cut-off 12-14 kDa) against 50 % of EtOH:dH_2O (three times) and concentrated under nitrogen atmosphere, yielding the dialysed lipid fraction containing the glycolipids.

2.2.1.1. Carbohydrate analyses

Lipid fraction was analysed to their constitution in sugars before and after dialysis, in terms of: neutral sugars, free alditols, uronic acids, sulfate groups, and glycosidic linkages, as described in section 2.1.5.

The glycosidic linkages were analysed without reduction of the carboxyl groups and using a different method to extract the methylated PS. These were extracted through a liquid-liquid extraction using the following procedure:

The lipid fraction containing methylated PS were dissolved in 2.0 mL of dH_2O and neutralized with 1 M HCl. CH_2Cl_2 (3.0 mL) was added and the solution was handling agitated and centrifuged (3000 rpm for 30 s) in order to separate the two phases. The aqueous phase was removed by suction. The organic phase was washed with 2 mL of dH_2O, stirred, centrifuged and the aqueous phase was completely removed as mentioned previously. The wash was repeated with 3 mL of dH_2O. The CH_2Cl_2 phase was concentrated by Speedvac and then submitted to a new methylation and subsequent liquid-liquid extraction, allowing to extract the methylated PS.

Statistical analysis was performed as described in section 2.1.6.

2.2.2. Development of microarrays with the dialysed lipid fraction extracted from *N. oculata*

To further characterize the carbohydrate portion of the glycolipids (GL) present in *N. oculta*, the dialysed lipid fraction extracted from *N. oculata* in section 2.2.1. was employed in microarrays as liposome probes. The microarray was probed using proteins with well-known carbohydrate-binding specificities.

2.2.2.1. Construction of the carbohydrate microarrays

Carbohydrate microarray assays developed with the lipid fraction extracted from *N. oculata* with $CHCl_3$:MeOH, which comprised the GL, were named as Microalgae GL set 2. This set comprised a total of 14 probes printed as liposomes. The liposomes were composed by the lipid fraction from *N. oculata*, referred as GL, and two carrier lipids, one phosphatidylcholine (PC) and cholesterol (C) (≥ 99 % pure), which were purchased from Sigma-Aldrich as lyophilized powder. Two PC were used to form the liposomes, 1,2-Diacyl-*sn*-glycero-3-phosphocholine (PC1) (Lecithin from egg yolk, ≥ 99 % pure) or 1,2-Di-*O*-hexadecyl-*sn*-glycero-3-phosphocholine (PC2) (≥ 99 % pure) (Chemical structures of the carrier lipids are depicted in supplementary figure 5). The 14 liposome probes used to construct the carbohydrate microarray with the GL from *N. oculata* differed in the following molar ratio: PC (1 or 2):C:GL. The GL were printed with and without carrier lipids at two levels: 50 pmol (low level) and 150 pmol (high level). In addition, the carrier lipids were also analysed at two levels, 500 and 1000 pmol to PC1 or PC2, and 300 and 1000 pmol to cholesterol. The characteristics/composition of the liposome probes used in microarray are described in table 7.

Table 7. List of all liposome probes that composed the Microalgae glycolipids (GL) set 2, which were analysed in the binding charts. The microarray comprised liposomes that were composed by the GL present in the dialysed lipid fraction from *N. oculata*, one phosphatidylcoline (PC), 1,2-Diacyl-*sn*-glycero-3-phosphocholine (PC1) or 1,2-Di-*O*-hexadecyl-*sn*-glycero-3-phosphocholine (PC2), and the cholesterol (C). Composition of the probes differed in the mol ratio PC:C:GL.

Liposome Probes (Set 2)	**Ratio PC:C:GL (pmol)**	**Liposome composition**
1	1000:1000:150	PC1:C:GL
2	1000:1000:50	
3	1000:1000:0	
4	1000:1000:150	PC2:C:GL
5	1000:1000:50	
6	1000:1000:0	
7	500:300:150	PC1:C:GL
8	500:300:50	
9	500:300:0	
10	500:300:150	PC2:C:GL
11	500:300:50	
12	500:300:0	
13	0:0:150	PC:C:GL
14	0:0:50	

For the construction of the microarray, liposome probes were immobilized non-covalently onto 2-pad nitrocellulose-coated FAST® (Whatman) glass slide, using the MicroCaster™ System, following the procedure used in section 2.1.8.1.

Each liposome probe was printed in triplicate. Liposome solutions for printing were prepared in the following way (64): 1) *Preparation of working solutions.* Working solutions of carrier lipids (PC1, PC2 and C) were prepared in screw-cap glass vials at a final concentration of 400 µM in MeOH, and the working solution of GL (dialysed lipid fraction) from *N. oculata* was also prepared in glass vials at a final concentration of 10 µM in $CHCl_3$:MeOH (2:1, v/v). 2) *Liposomes formation*. To prepare the liposome probes with a mol ratio of 1000:1000:150 or 1000:1000:50 or 1000:1000:0 (PC:C:G), 5 µL of MeOH and 5 µL of the working solution of PC (1 or 2) and C were added to one microtube, followed by 30 µL (for the 150 pmol level) or 10 µL (for the 50 pmol level) or 0 µL (for absence of GL) of the work solution of the GL at 10 µM. Otherwise, to prepare the probes with a ratio of 500:300:150 or 500:300:50 or 500:300:0 (PC:C:GL), 5 µL of MeOH, 2.5 µL of the working solution of PC (1 or 2) and C, and 30 µL (for the 150 pmol level) or 10 µL (for the 50 pmol level) or 0 µL (for absence of GL) of the work solution of the GL at 10 µM were added to the microtube. The probes without carrier lipids (0:0:150 or 0:0:50 PC:C:GL) were

prepared by adding to one microtube 5 µL of MeOH and 30 or 10 µL of the work solution of the GL. After preparation of the solutions containing the carrier lipids and the GL in MeOH, the microtubes were left uncapped in a drying oven *overnight* at 37 ºC to evaporate completely the solvent. Thereafter, each dried mixture was re-suspended in Milli-Q water containing 0.2 µg/mL Cy3 (20 µL), agitated in vortex (approximately 5 s), and centrifuged briefly (spin down). Then, the microtubes were sonicated at 37 ºC for 15 min in an ultrasonic water bath (Ultrasonic Cleaning System Starsonic 35, Liarre) in order to form the micelles soluble in water (formation of liposomes). The liposome solutions (20 µL) were transferred into a plastic 384-well plates (AB-1384/G 384-well PCR Plates, Thermo Scientific), according to the planned microarray layout, and the manual printing was performed in nitrocellulose-coated 2-pad slides as described in the printing of PS (section 2.1.8.1.).

2.2.2.2. Carbohydrate microarrays binding assay

Two proteins were probed using the Microalgae GL set 2, the plant lectin RCA_{120} and the animal lectin hGalectin-3. These lectins have well-known carbohydrate-binding specificities towards Gal residues. This will allow to understand if the dialysed lipid fraction from *N. oculata* contains galactolipids and if they can be efficiently retained onto the slide surface without any chemical modification. In other words, if they are suitable natural probes to develop microarrays with GL and further study carbohydrate-protein interactions in which these GL may be involved. The lectins used in the binding assays with the microarray developed with the Microalgae GL set 2 and its reported specificities is shown in table 8.

Table 8. Lectins used in the carbohydrate microarrays developed with Microalgae GL set 2 and their reported specificities.

Lectins	Specificity	Reference
Ricinus Communis Agglutinin 120 (RCA_{120})	β-Gal > α-Gal Terminal Gal(β1→4) > terminal Gal(β1→3) and Gal(β1→6) Terminal Gal(β1→4) in *N*-glycans Gal(6S)(β1→4)AnGal(α1→3)Gal(6S)(β1→4)AnGal(α1→3)Gal(6S) Gal(2S)(β1→[4)-Gal(2S6S)(α1→3)Gal(2S)(β1→]24)-Gal(2S6S)	(61,62,85,91)
hGalectin-3	β-galactosides Gal(β1→4)GlcNAcβ	(88,97)

Firstly, the printed slides with the Microalgae GL set 2 probes were scanned for Cy3 (532 nm; 10 µm resolution, 5-10 % laser power, 350 PMT gain) using GenePix® 4300A

microarrays scanner (Molecular Devices) (Fig. 10). After scanning, each pad of a nitrocellulose-coated 2-pad glass slide was wetted with 500 µL of water. The slide was gently tilting to remove the excess of water and the respective blocking solution (500 µL) for each protein (Blocking solutions and the conditions used during the binding assay for each protein are described in supplementary table 3) was added to each pad and incubated for 60 min on dark at T_{room}. The blocker/diluent solutions used were 3 % BSA in HBS with 5 mM $CaCl_2$, and 0.5 % Casein in HBS with 5 mM $CaCl_2$ for RCA_{120} and hGalectin-3, respectively. The stock solutions used to prepare the different blocking solutions were the same used in section 2.1.8.2.

After 60 min of incubation, the blocking solution was removed and the pads were washed once with HBS (500 µL). After the HBS was removed, the protocol followed for the binding step of the lectins was different.

The biotinylated **lectin RCA_{120}** were analysed by one step, adding 500 µL onto the pad at the desired concentration (prepared in the diluent solution) and incubated for 90 min.

For **hGalectin-3**, 500 µL of the protein was added to the pad and incubated for 90 min at T_{room}. After this period, the excess of protein was removed and the pad was washed four times with HBS buffer (500 µL). Then, in order to detect the binding of the protein to the liposome probes 500 µL of a pre-complex formed by Abs (1:1, Ab1:Ab2) was added to the pad and incubated for 60 min at T_{room}. The pre-complex was prepared as follows: Ab complexes were prepared by pre-incubation of the primary Ab Manti-His, Ab1, with the secondary biotinylated detection Ab BI anti-mouse IgG, Ab2, for 15 min at T_{room}. After this period, the Ab complexes were diluted to a final volume of 500 µL (prepared at the desired concentration) with diluent solution and the pre-complex solution was added to the pad. After the 60 min of incubation with the pre-complex, the pad was washed 4 times with 500 µL of HBS buffer.

For the detection of binding signals, 500 µL of AF-647-labeled streptavidin (1 µg/mL in the respective blocker solution for each protein) was added to each pad and incubated for 45 min.

Between incubations, each pad was washed four times with HBS buffer (500 µL). After the last incubation with streptavidin, pads were washed four times with HBS buffer and two times with Milli-Q water and dried in the dark for 5 min.

Each microarray slide was then scanned for AF-647 (635 nm; 10 μm resolution; 5-90 % laser power, 350 PMT gain) for detection of the binding event using the GenePix® 4300A microarray. The Cy3 and AF-647 fluorescence intensity for each spot was quantified using the GenePix® Pro Software (Molecular Devices).

2.2.2.3. Carbohydrate microarrays data analysis

The data obtained from the carbohydrate microarrays developed with the dialysed lipid fraction from *N. oculata* were analysed as described in section 2.1.8.3. For microarrays developed with the GL only charts/tables were constructed for processing and presenting the quantified results, using the GraphPad Prism® 6 program.

Chapter 3 – Results and Discussion

3.1. Characterization of polysaccharides from *N. oculata*

Polysaccharides (PS) in marine microalgae have been found to exhibit various biological activities (4), which are assigned to their structural features. However, the detailed knowledge of their structures still remains understudied due to the large number of marine species. With this in mind, the present section will focus on the study of the structural characteristics of the PS sequentially extracted from the marine microalga *Nannochloropsis oculata*, which was purchased from Necton S.A.

3.1.1. Characterization of the fractions derived from sequential extraction of PS from *N. oculata*

3.1.1.1. Biomass of *Nannochloropsis oculata*

Total content in protein, neutral sugars (NS), uronic acids (UA) and sulfate (SO_3^-) groups present in the biomass of *Nannochloropsis oculata* and in its biomass after being defatted with $CHCl_3$:MeOH (2:1, v/v) (defatted biomass) is summarized in table 9. The total content in protein present in biomass of *N. oculata* was 34 % dw. The result is in agreement with the reported by other authors, which indicate a content in a range of 20 to 42 % (10–12,46). In addition, Necton S.A. indicates in the biomass package label where *N. oculata* is commercialised a content between 30-35 % dw, being thus in accordance. Regarding to defatted biomass, there was a decrease in 4 % in the protein content (34 to 30 % dw) compared to the initial biomass (table 9).

Table 9. Content in protein, neutral sugars (NS), uronic acids (UA), and sulfate groups (SO_3^-) present in the biomass of *N. oculata* and in their defatted biomass. Results are expressed in percentage of dry weight (dw).

	(mean ± SD) % dw*				
Nannochloropsis oculata	**Protein****	**NS**	**UA**	**SO_3^-**	**∑ NS+UA+SO_3^-**
Biomass	33.9 ± 0.7[a]	17.9 ± 0.2[a]	0.5 ± 0.2[a]	0.6 ± 0.1[a]	19.0 ± 0.1[a]
Defatted biomass	29.6 ± 0.0[b]	16.6 ± 0.0[b]	0.9 ± 0.1[b]	2.6 ± 0.8[b]	20.3 ± 0.5[a]

*- Data represent the mean values ± SD in percentage of dry weight (dw) from duplicate (n=2), n=2, n=4 and n=3 analyses for protein, NS, UA and SO_3^-, respectively.
Values in the same column with different superscript letters are significantly different ($p<0.05$).
**- Content in protein was determined through elemental analysis converting the total content in nitrogen (% N) into protein through multiplication of % N by the nitrogen-to-protein conversion factor (*N*-Prot factor) 4.95 determined for *N. oculata* (116).

NS comprise 17.9 % dw in biomass of *N. oculata* and 16.6 % in defatted biomass (Table 9). The result shows a significant ($p<0.05$) decrease in the NS content after defatting. This can be explained due to the use of MeOH in the extraction of lipid compounds, which has a polar character able to extract/retain some mono- and/or oligosaccharides together with lipids (9–11). In addition, the decrease in NS after defatting can also be a consequence of simultaneous extraction of glycoconjugates with lipid compounds, like glycolipids found in photosynthetic organelle membranes in microalgae (31).

Relatively to UA, it was verified an increase in its amount in the biomass of *N. oculata* after treatment with $CHCl_3$:MeOH (2:1, v/v) in about 0.4 % dw ($p<0.05$) (Table 9). The result shows that UA are retained in the defatted biomass, which consists in a residue enriched in carbohydrates.

Additionally, it was verified that the total content in sugars, represented as the sum of NS and UA, in biomass and in their defatted form represents about 18.4 and 17.6 % dw, respectively. This shows a decrease in the total content in sugars ($p<0.05$), mainly due to the loss of NS. The total content in sugars obtained in this work is in agreement with which is reported for microalgae in general, whose content range from 5 to 30 % (7), and in particular for *N. oculata* between 5 and 17 % dw (9–12,48). However, it should be emphasised that the content in carbohydrates may depend on the environmental growing conditions (46,47).

Regarding to the sulfate (SO_3^-) groups content, these represented 0.6 and 2.6 % dw in biomass and in defatted biomass, respectively (Table 9). The result shows an increase in SO_3^- content in biomass with defatting ($p<0.05$), displaying that these are not extracted with $CHCl_3$:MeOH (2:1, v/v). This should be due to the low hydrophilic character of this solution that it is unable to extract hydrophilic compounds, like carbohydrates, especially PS with SO_3^- groups. Furthermore, the result suggests that these groups may be part of the carbohydrate composition, since they are retained in the defatted biomass that was enriched in PS. As a result, considering that the carbohydrates in *N. oculata* are composed not only by NS and UA but may also contain SO_3^-, likely as components of the PS, the total content in carbohydrates comprises an identical amount (19 and 20 % dw) in biomass of *N. oculata* and in their defatted form ($p<0.05$) (Table 9). This means that NS represent in about 94 and 82 % of the total carbohydrates content in biomass and in defatted biomass of *N. oculata*, respectively, and the UA comprise in about 3 and 4 % of the carbohydrates. Relatively to SO_3^- groups, they represent 3 and 13 % of the carbohydrates in biomass and in defatted

biomass. As recently found by Hafsa *et al.* (48), the microalga aqueous extract of *N. oculata* comprises sulfate groups in about 6 % of the extract. This result supports the sulfate content determined in this work. Additionally, in other marine microalgae have also been found the presence of sulfated PS (20,23).

Carbohydrates composition of *N. oculata*, in terms of neutral monosaccharides and content in UA and SO_3^- groups, is represented in figure 11. Glucose (Glc) was the predominant neutral sugar together with Man (in about 35 mol %) in the biomass of *N. oculata*, followed by Gal (10 %). In addition, Rha (4.3 %), Rib (3.7 %), and Xyl (1.8 %) were found in lower amounts, being the content in Ara almost negligible, in agreement with which is reported (9–11). In addition, comparing the NS composition of the carbohydrates obtained in this study with the reported by literature for the constitution of PS, they are in agreement for all sugars, except for Man. Literature indicates that the PS of *N. oculata* and of their class (Eustigmatophyceae) are composed mainly by Glc (45-68 %), followed by Rha (8-12 %), Fuc, Xyl, Rib, and Man (2-6 %) in varying proportions (9–11). The higher content in Man obtained in this work is likely due to the fact that defatted biomass comprise not only PS, but may also include other types of carbohydrates, like mono- and oligosaccharides. In microalgae it is possible that the Man detected may derive from mannitol (127), including in *N. oculata* (48). In addition, the sugar composition of carbohydrates depends on the growth conditions of microalgae (46,47). Furthermore, Brown *et al.* (11) reported that *N. oculata* has a PS composition similar to *P. tricornutum*, which is rich in Man. Relatively to the other sugars, and in particular to Glc, the high content obtained may be due to the presence of a cell wall rich in cellulose that protects *N. oculata*, or even due to the presence of chrysolaminarin or laminarin as storage PS (49).

Analysing the carbohydrates composition of defatted biomass, a residue enriched in PS, it was observed that the sugar profile suffered a slight decrease in Ara, and a high loss of Gal (10 to 1 %) ($p<0.05$), compared to the initial biomass (Fig. 11). The decrease in Gal may be a consequence of the presence of galactolipids in *N. oculata* (45,49,53–55), which probably were extracted with $CHCl_3$:MeOH (2:1, v/v) during the defatting step.

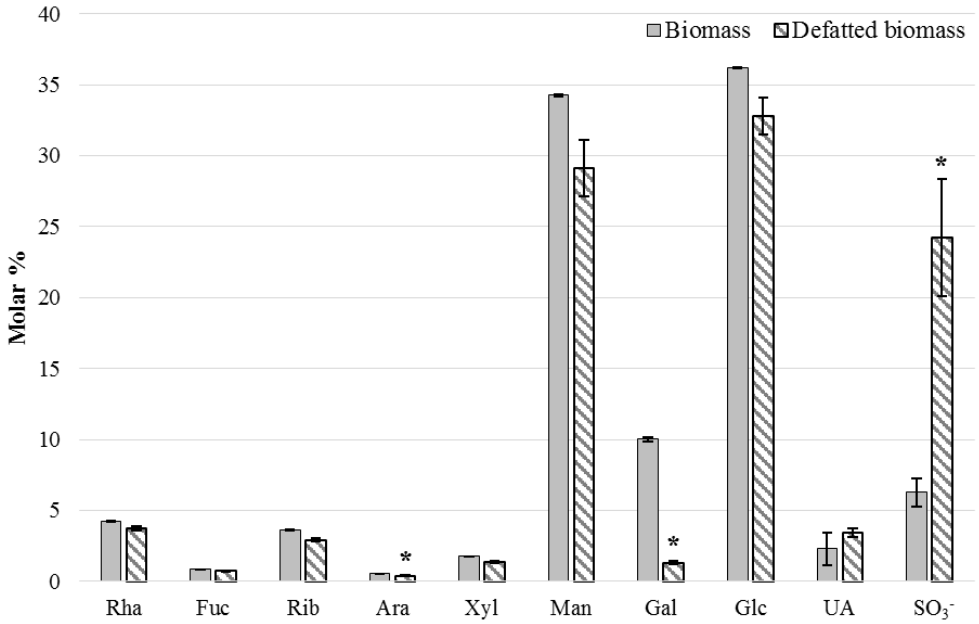


Figure 11. Composition in neutral and acidic sugars (UA), and in sulfate (SO_3^-) groups in the biomass and in defatted biomass of *N. oculata*. Results are expressed in molar percentage (mol %). Asterisks represent values significantly different ($p<0.05$) between the biomass and defatted biomass.

Regarding to UA, their content in carbohydrates was bellow 5 mol % (Fig. 11) in biomass and in defatted biomass, as well as Rha, Rib, and Xyl, being these sugars in similar amounts (2 to 4 %). Opposite to NS, the content in UA increased in biomass defatted. To our knowledge, nothing is mentioned in literature about the presence of UA in *N. oculata*. Nonetheless, in marine microalgae have been reported the presence of acidic sugars, like glucuronic acids (GlcA) in diatoms and *Porphyridium sp.* (19,20). Comparatively to UA, the content in SO_3^- groups also increased in defatted biomass relatively to the initial biomass, in about 18 % (from 6 % to 24 %) (Fig. 11).

3.1.1.2. Hot-water (HW) extracts

Defatted biomass of *N. oculata* was used to extract their hot-water (HW) soluble polysaccharides (PS). The extraction with HW yielded a supernatant containing mono-, oligo- and polysaccharides soluble in HW. The supernatant was further separated through ultrafiltration (cut-off 10 kDa), yielding a fraction with a *Mw* lower than 10 kDa (Filtrate) and a fraction with a *Mw* higher than 10 kDa (Retentate), which contained the HW soluble PS. The total material recovered was 90.9 % relatively to defatted biomass, remaining 67.7, 17.3 and 5.9 % of the material in the Residue (material insoluble in HW), Filtrate and Retentate, respectively (Table 10). Relatively to the protein content present in each fraction,

this is in higher amount in Residue (55 %), followed by Retentate and Filtrate (18 %) (Table 10).

Table 10. Yields of extraction with hot-water (HW) after ultrafiltration (cut-off 10 kDa) from the defatted biomass, results are expressed as weight percentage (%) of defatted biomass. Content in protein, neutral sugars (NS), uronic acids (UA) and sulfate groups (SO_3^-) present in the Residue (material insoluble in HW), Filtrate (F *Mw*<10 kDa) and Retentate (Ret *Mw*>10 kDa) after ultrafiltration obtained from the HW extraction. Results are expressed as weight % of fraction.

		(mean ± SD) %*				
HW extraction	**Yield (w/w) %**	**Protein****	**NS**	**UA**	**SO_3^-**	**∑ NS+UA+SO_3^-**
Residue	67.7	54.6 ± 0.1	14.4 ± 3.9	1.4 ± 0.2	1.8 ± 0.5	17.7 ± 3.6
F *Mw*<10 kDa	17.3	17.5 ± 0.2	65.2 ± 10.2	0.7 ± 0.1	3.5 ± 1.9	68.7 ± 12.1
Ret *Mw*>10 kDa	5.9	40.6 ± 0.2	33.7 ± 0.4	3.2 ± 0.5	4.9 ± 1.7	41.8 ± 1.7

*- Data represent the mean values ± SD in weight % of each fraction from duplicate (n=2), n=4, n=3, and n=2 analyses for protein, NS, UA, and SO_3^-, respectively.
**- Content in protein was determined through elemental analysis converting the total content in nitrogen (% N) into protein through multiplication of % N by the nitrogen-to-protein conversion factor (*N*-Prot factor) 4.95 determined for *N. oculata* (116).

In terms of content in NS, UA and SO_3^- groups, it was verified that carbohydrates are composed mainly by NS, representing 81.4, 94.9 and 80.6 % of the total content in carbohydrates in the Residue, Filtrate and Retentate, respectively (Table 10). Relatively to residue, this fraction was the poorest in carbohydrates (17.7 %), due to the lowest content in NS (14 %) and SO_3^- groups (1.8 %). On the other side, filtrate was the richest in carbohydrates (68.7 %), being these composed mainly by NS (65.2 %).

Retentate was composed mainly by NS (33.7 %), being the richest fraction in UA (3.2 %) and in SO_3^- groups (4.9 %) (Table 10). As a result, this suggests that the HW soluble PS with a *Mw*>10 kDa from *N. oculata* should be sulfated and contain an anionic character, as it has been described for other marine microalgae (20,23,25,26) corroborating the mentioned before for defatted biomass. Moreover, the result for sulfate is in accordance with the reported by Hafsa *et al.* (48) that verified the presence of SO_3^-groups in aqueous extracts, which represented 6 % of *N. oculata* extract.

The total NS content in the supernatant of HW (retentate and filtrate), comprising the HW soluble PS, mono- and oligosaccharides, represents 8.9 ± 1.2 % of dw. Comparatively, Hafsa *et al.* (48) verified that the total sugar content in aqueous extract of *N. oculata* represents 4.1 % of total dried matter. Thus, in the present work the total sugar content was higher than obtained by these authors. This may be a consequence of some differences during

extraction of soluble carbohydrates. As example, the different solvent used to degrease the biomass, since in this work was used a solution of $CHCl_3$:MeOH (2:1, v/v) and these authors used only MeOH. In addition, in this work the soluble carbohydrates were extracted with HW at 65 ^oC, whereas Hafsa *et al.* (48) used water probably at T_{room}, since they only referred that the biomass was extracted with water. This suggests that likely MeOH alone is able to remove more sugars from the biomass, and using HW allows to extract more sugars.

With respect to carbohydrates composition of each fraction from the HW extraction, it was verified that the high content in Glc present in defatted biomass remained predominantly in the residue, being the richest fraction in Glc (58 mol % of the total carbohydrates) (Fig. 12). This suggests the presence of an insoluble polymer composed of Glc, such as cellulose, since the residue contains HW insoluble PS. As a result, in order to understand if the high content in Glc is from cellulose, the residue was studied to their constitution in glycosidic linkages. The glycosidic linkages analysis showed the predominance of 4-Glc (64 mol % of the total of NS), followed by 2-Rha*p* (6.8 %), and terminal (t)-Glc (4.5 %), containing other linkages of Glc, Man, Rha and Xyl in lower amounts. Therefore, the high amount of 4-Glc in the residue suggests the presence of cellulose, since this polymer is an insoluble fibre composed of a linear non-branched (β1→4)-linked Glc, consistent with the presence of a cellulose wall in *N. oculata* (49). In addition, it was verified that UA and SO_3^- groups represent 7 and 20 mol % in carbohydrates of the residue.

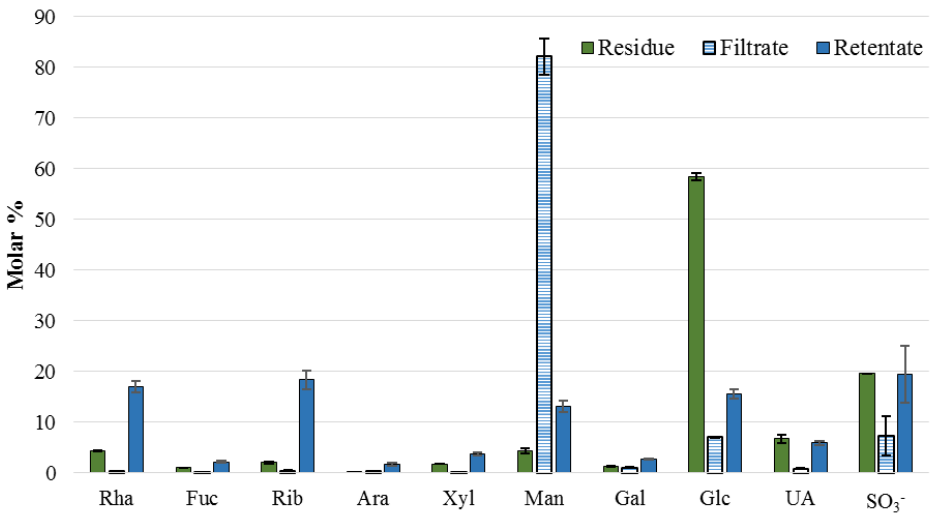


Figure 12. Composition in neutral and acidic sugars (UA), and in sulfate (SO_3^-) groups present in the Residue, Filtrate (*Mw*<10 kDa) and Retentate (*Mw*>10 kDa) fractions obtained from the hot-water (HW) extraction after ultrafiltration (cut-off 10 kDa). Results are expressed in molar percentage (mol %).

Relatively to filtrate, the fraction containing material with *Mw* under 10 kDa, it was verified that Man was the predominant NS (in about 82 mol %) with presence of a residual amount of Glc (7 %), accounting these two residues almost the total content in NS (Fig. 12). Consequently, this suggests that the high content in Man obtained previously in the biomass of *N. oculata* (determined in section 3.1.1.1.) was not derived from PS, in agreement with reported for their PS composition (9–11). Additionally, it was verified that the filtrate did not contain UA and was the fraction with lower content in SO_3^- (7 %), suggesting that the UA and SO_3^- make part of larger polymers, like PS.

Due to the high content in Man obtained in the filtrate, this fraction was analysed in terms of free alditols. It was verified the presence of mannitol in about 31.24 ± 3.52 weight % of the filtrate, corresponding to 3.8 % of dw in *N. oculata*. Comparatively to the total Man content in filtrate obtained from the NS analysis (58.48 ± 8.95 % w/w), which comprises in this case the sum of Man and mannitol, only 27.2 % of the total Man determined by NS analysis was effectively Man. As a result, the high content in Man in filtrate fraction may be explained as a consequence of the presence of mannitol in *N. oculata*. It has been reported that mannitol is an osmoregulatory compound of the genus *Nannochloropsis* that allow a rapid acclimation to variable osmolarity (128). This compound can change its intracellular content in response to variations in external salinity, exhibiting a role as an osmotic effector (127). In addition, it has been reported its function as storage compound in *Emiliania huxleyi,* and in other algae and green plants (129).

Carbohydrate composition of the retentate, a fraction that contained the HW soluble PS higher than 10 kDa is represented in figure 12. This fraction contained Rib (18 mol %), Rha, Glc and Man (13 %) as predominant NS, comprising also Xyl (4 %), Fuc and Gal (3 %) in minor amounts. This profile suggests a diversified monosaccharides composition of the PS, being this composition in agreement with the reported by literature, which indicates the presence of Rha (8-12 %), Fuc, Xyl, Rib, and Man in various proportions (9–11). In particular, it was recently reported that the aqueous extract of *N. oculata* besides comprise mannitol, contained Glc as predominant sugar, and Xyl, Man and Gal in minor amounts (48), being not found Rib and Rha, unlike to the results obtained in this study.

Additionally, in similar amounts ($p<0.05$) of Rib, Rha, Man and Glc were found SO_3^- groups (19 %) as component of the retentate (Fig. 12), in accordance with the reported by Hafsa *et al.* (48) for the presence of SO_3^- groups in the aqueous extract of *N. oculata*. This

suggests the presence of sulfated PS in *N. oculata*, as described for various marine microalgae, such as *P. tricornutum* and *P. cruetum* (20,23,25). Furthermore, retentate also comprised UA in about 6 %. Thus, the retentate fraction seems contain a diversified composition in neutral monosaccharides and may be negatively charged due to the presence of SO_3^- groups and/or UA.

These results showed that the extraction with HW allowed to obtain a high content in Glc present in *N. oculata* that should be due to the presence of cellulose, since Glc remained in the residue and this fraction contained 4-Glc as predominant sugar. In addition, the HW extraction with posterior ultrafiltration allowed to verify the presence of mannitol in *N. oculata* (filtrate), likely working as an osmotic effector (130). Finally, other sugars, including Rha, Fuc, Xyl, Rib, and Man reported (9–11) as components of the PS in *N. oculata* are part of the HW soluble PS composition, since they remained in retentate.

3.1.1.3. Fractions of the sequential ethanol (EtOH) precipitation

Sequential ethanol (EtOH) precipitation of the retentate fraction from the HW extraction after ultrafiltration (cut-off 10 kDa) was performed in order to separate different types of polysaccharides (PS). These were separated due to their different solubility in EtOH, which depends on various features, such as molecular weight (*Mw*), degree of branching (DB), charge and arrangement of the polymers. Exemplifying, the solubility in EtOH tends to increase to polymers with low *Mw*, high DB and charged.

As a result of the sequential EtOH precipitation were obtained three fractions enriched in PS, one precipitate from precipitation with 50 % EtOH (v/v) referred as Et50, and another one with 85 % EtOH (v/v), referred as Et85, and one supernatant (named as EtSN). These fractions were obtained with a recovery of 60.2, 24.4, and 9.2 % of the retentate to Et50, Et85 and EtSN, respectively (Table 11). The yield of sequential EtOH precipitation showed that the major part of the polymeric material in retentate was insoluble in 50 % EtOH.

Et50 was the richest fraction in protein (52 % w/w), followed by EtSN (28 %) and Et85 (22 %) fractions (Table 11). Although Et50 was the fraction where the polymeric material from the retentate preferentially remained (yield of 60 %), their high yield of extraction was mostly due to the presence of protein and not to PS, since this fraction showed the lowest content in PS (29 %). Consequently, Et50 fraction contained also the lowest content in NS (25 %), UA (4 %) and SO_3^- groups (1 %).

Table 11. Yields of sequential EtOH precipitation of the retentate to each fraction (Et50, Et85 and EtSN), expressed as weight percentage (%) of retentate. Content in protein, and in neutral sugars (NS), uronic acids (UA) and sulfate (SO_3^-) groups that compose the polysaccharides (PS) present in Et50, Et85 and EtSN. Results expressed as weight % of fraction.

		(mean ± SD) %*				
Fraction	**Yield (w/w) %**	**Protein****	**NS**	**UA**	**SO_3^-**	**∑ NS+UA+SO_3^-**
Et50	60.2	51.8 ± 0.9	24.9 ± 1.0	3.9 ± 0.2	0.6 ± 0.0	29.4 ± 0.8
Et85	24.4	21.6 ± 0.1	42.1 ± 0.1	9.5 ± 0.8	3.2 ± 0.7	54.9 ± 0.2
EtSN	9.2	27.6 ± 0.0	33.6 ± 0.9	7.3 ± 0.3	1.5 ± 0.1	42.4 ± 0.6

*- Data represent the mean values ± SD in weight % of each fraction from duplicate (n=2), n=4, n=3, and n=6 analyses for protein, NS, UA, and SO_3^-, respectively.
**- Content in protein was determined through elemental analysis converting the total content in nitrogen (% N) into protein through multiplication of % N by the nitrogen-to-protein conversion factor (N-Prot factor) 4.95 determined for *N. oculata* by (116).

In relation to Et85, its polymeric material was dominated by PS (55 %), being the richest fraction in PS compared to Et50 and EtSN (Table 11). Et85 displayed the highest amount in NS (42 %), UA (10 %) and SO_3^- (3 %). Relatively to EtSN, it was also composed preferentially by PS (42 %) and comprised 34 % of NS, 7 % of UA and 2 % of SO_3^- groups (Table 11). The presence of UA and SO_3^- groups in Et85 and EtSN suggests that some of the HW soluble PS from *N. oculata* may be negatively charged and sulfated.

Comparing the PS composition of the three EtOH fractions, it was observed that the PS in Et50 were composed mainly by NS (85 % of total PS), and contained the lowest amount of UA (14 %) and SO_3^- groups (2 %). This low content in UA and SO_3^- groups proposes the presence of polymers slightly charged, which is consistent with the expected, since was the fraction more insoluble in EtOH (precipitate with 50 % EtOH v/v), thereby with a lower hydrophilic character. Regarding to Et85 and EtSN, their PS were also composed mainly by NS, representing 77 and 79 % of the PS, respectively. UA were detected in same proportion in PS of Et85 and EtSN (17 %). Otherwise, PS in Et85 fraction had the highest proportion in SO_3^- groups, accounting 6 % of their PS, whereas EtSN accounted of 3 %, suggesting that PS in Et85 were more sulfated. Moreover, the higher negative character of Et85 compared to EtSN, mainly due to the higher amount in SO_3^- groups, was unexpected. The negative charge should confer a high hydrophilic character. Nonetheless, the PS in Et85 were insoluble in 85 % of EtOH (v/v). As a result, this probably may be a consequence of the arrangement of the polymers in Et85 compared to those in EtSN. The arrangement/organization of the polymers may restrict/prevent the exhibition of SO_3^- groups

at the surface of the polymer, making the PS in Et85 insoluble in EtOH. In addition, the *Mw* of the polymers and their DB may also have influence.

A gel permeation chromatography (GPC) analysis was performed to determine the *Mw* of the polymers present in each EtOH fraction. Chromatogram from GPC analysis is represented in figure 13 and the respective parameters related to the weight are summarized in table 12. The GPC analysis showed that Et50 and EtSN comprised only one population of polymers with an average of *Mw* of 630 and 26 kDa, respectively, whereas Et85 comprised two populations, one with an average *Mw* of 5 162 kDa and another one with 65 kDa (Fig. 13 and Table 12).

Et50 comprised polymers with high *Mw* (table 12), explaining in part their insolubility in 50 % EtOH (v/v) together with its low charged character. In addition, Et50 contained a high polydispersity index (*Mw/Mn*), indicating that their polymers are non-uniform (polydisperse), in other words, they have an inconsistent size, shape and/or mass distribution. Unlike, EtSN comprised the polymers with the lowest *Mw*, being also one of the reasons for their solubility in 85 % EtOH (v/v), in addition to its negatively charged character as described before.

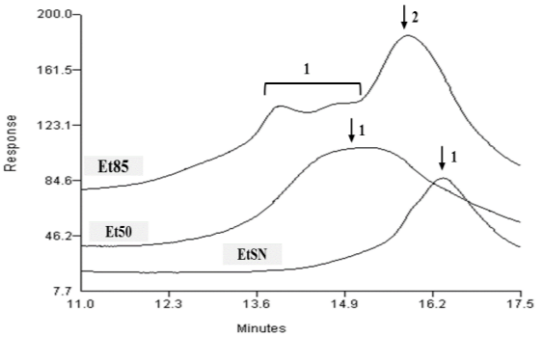


Figure 13. Chromatogram from the GPC analysis of the fractions obtained during the sequential EtOH precipitation (Et50, Et85 and EtSN). Numbers 1 and 2 represent the populations of polymers present in each fraction with different molecular weights (*Mw*).

Table 12. Estimation of the weight average molecular weight (*Mw*) and number average molecular weight (*Mn)* present in the EtOH fractions by GPC. Data are expressed in Dalton (*Da*) unit.

	Population	***Mw***	***Mn***	***Mw/Mn***
Et50	1	630 020	13 370	47.1
Et85	1	5 162 224	1 297 380	4.0
	2	65 117	7 030	9.3
EtSN	1	26 480	3 861	6.9

GPC analysis of Et85 revealed that this fraction comprised one population of polymers with a very high *Mw* (5 162 kDa) (Table 12). The chromatogram belonging to Et85 (Fig. 13) showed a fraction, represented with number 1, composed of two populations. However, they were not separate efficiently. As a result, the high *Mw* determined in the first fraction of Et85

might be a consequence of the presence of an aggregate of polymers that prevent their retention into the column and consequently their separation, explaining the high *Mw* obtained in the first population of Et85 fraction (Table 12). On the other hand, the second population in Et85 comprised an amount of polymers with identical *Mw* in average (Table 12). The *Mw* determined for the second population of polymers in Et85 (65 kDa) was lower than those in Et50 (630 kDa) and higher than those in EtSN (26 kDa). Thus, the *Mw* of the polymers in Et85 may be also one of the reasons for their precipitation only with 85 % of EtOH (v/v), being in accordance with the expected for the solubility in EtOH depending on the *Mw* of the polymers.

The PS composition of the EtOH fractions, in terms of neutral monosaccharides and its respective alditol form, and UA and SO_3^- groups is represented in figure 14. PS in Et50 comprised as major sugars Rib (35 mol %) and Glc (33 %) in identical amounts ($p<0.05$), being this the fraction where the Rib and Glc from retentate essentially remained. In addition, other sugars were found in lower amounts, like Rha and Man (5 %), and the PS in this fraction did not contain alditols in its composition (Fig. 14 B). Regarding to UA and SO_3^- groups, accounted of 11 % and 4 % of the PS in Et50, respectively, being the fraction with lower content in UA and SO_3^-.

In relation to PS composition of Et85 fraction, Rha and Man were the main NS, being in identical amounts (25 mol %) ($p<0.05$) (Fig. 14 A). Xyl, Glc, Fuc and Gal were also found in lower amount (in about 5 %). Besides, PS in Et85 contained alditols, which included mainly mannitol (ManOH) and rhamnitol (RhaOH) (3 mol %) (Fig. 14 B), and other alditols in residual amounts. Alditols in Et85 accounted 12 mol % of the total sugars, alditols and sulfate content. The total content in non-free (polymeric) alditols was relatively high for making part of the PS because they must be linked to PS through their reducing ends (51). This would imply the presence of many reducing ends and consequently the presence of short PS, leading to think that the alditols may be linked to PS through another binding point, since their PS were relatively large as ascertained by GPC (65 and 5 162 kDa).

In addition, PS in Et85 also contained UA (15 %) and SO_3^- groups (11 %), being the richest fraction in sulfated PS (Fig. 14). As already mentioned, the PS in this fraction should exhibit a higher hydrophilic character than those in EtSN, due to their negative character. However, the PS in Et85 were insoluble in 85 % of EtOH (v/v), contrarily to PS in EtSN, possibly due to their higher *Mw* (determined before by GPC). Interestingly, it was verified

that the total content in non-free alditols was similar to the content in SO_3^- groups (11 mol %). Consequently, this observation led to propose that if the alditol was linked to SO_3^- group present in the PS, the sulfate was in its neutral form, becoming hydrophobic. Thus, the neutral sulfated PS would be hydrophobic and insoluble in EtOH, being one possible reason for its precipitation in 85 % EtOH (v/v), together with its higher *Mw* compared to PS in EtSN. At the same time, the sulfate would work as an additional binding point to link alditols, justifying the high content obtained. To our knowledge, to date nothing was reported to the presence of neutral sulfate PS in marine microalgae.

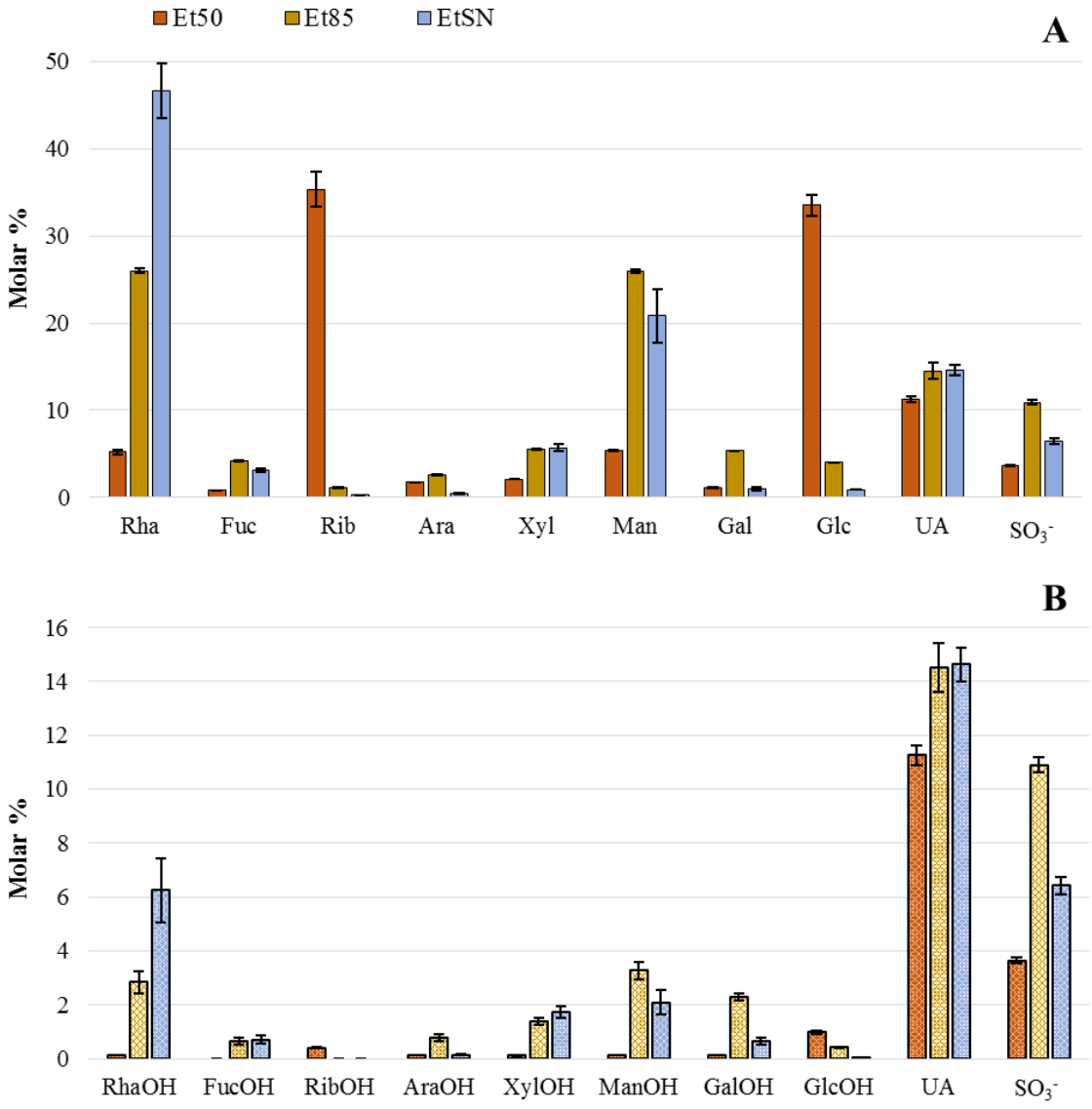


Figure 14. Polysaccharides (PS) composition of the EtOH fractions (Et50, Et85 and EtSN). (**A**) Composition in neutral monosaccharides, uronic acids (UA), and in sulfate (SO_3^-) groups. (**B**) Composition in alditols, UA and SO_3^- of the PS. Results are expressed in molar percentage (mol %) of the total sugars, alditols, and sulfate content.

Similarly to Et85, PS in EtSN were composed mainly by Rha (47 %), and Man (20 %), containing also lower amounts of Xyl (5 %) and Fuc (3 %) (Fig. 14 A). Their PS also comprised alditols mainly as RhaOH (6 %) and ManOH (2 %), accounting the non-free alditols a total of 12 mol % (Fig. 14 B). Furthermore, it was verified the presence of UA (15 %) and SO_3^- groups (6 %). In this fraction, contrarily to Et85, the content in alditols was higher than the content in SO_3^- groups. This likely suggests that alditols were preferentially linked through the reducing end, since the PS in EtSN were found to be smaller than those in Et85, having consequently more binding points (reducing ends).

The glycosidic linkages of the sugars that compose the PS in the EtOH fractions (Table 13) were determined from methylation analysis. The predominant sugar in Et50 was the 4-linked Glc (accounting 60 mol % of the total sugars content), suggesting the presence of a backbone composed of (1→4)-Glc in their PS. In addition, it was observed other kind of Glc residues in lower levels, such as 3- (6 %) and 2-linked (1 %) Glc, proposing that these residues may be part of the main or a linear chain. It was found the presence of 4-glucuronic acid (4-GlcA) in about 2 %.

Table 13. Glycosyl linkage composition (mol % of the total sugars) of the fractions obtained from sequential EtOH precipitation (Et50, Et85 and EtSN) and after desulfation of the fraction Et85 (Et85 desulfated). Glycosidic linkages of the PS in Et50 and EtSN were determined from the carboxyl-reduced partially methylated alditol acetates (PMAAs), and the Et85 and Et85 desulfated through PMAAs without reduction of carboxyl groups.

	% molar			
Glycosidic Linkage*	**Et50**	**Et85**	**Et85 desulfated**	**EtSN**
t-Rha	2.5	5.3	8.9	13.4
2-Rha	13.3	18.4	27.9	41.8
3-Rha	-	1.4	0.7	1.5
3,4-Rha	-	2.0	1.5	-
2,3-Rha	1.9	3.8	2.0	5.6
2,4-Rha	-	0.5	-	0.5
2,3,4-Rha	-	1.9	-	-
Total	17.7	33.3	41.0	62.8
t-Fuc	-	1.0	2.4	-
3-Fuc	-	-	0.7	-
Total	0.0	1.0	3.1	0.0
t-Rib*f*	-	0.6	0.9	-
Total	0.0	0.6	0.9	0.0
2-Ara*f*	-	-	1.2	-
4-Ara*p*=5-Ara*f*	-	3.1	7.3	-
3,4-Ara*p*=3,5-Ara*f*	-	1.8	-	-

Table 13. Continuation

	% molar			
Glycosidic Linkage	**Et50**	**Et85**	**Et85 desulfated**	**EtSN**
2,3,4-Ara*p*=2,3,5-Ara*f*	1.8	1.8	-	1.8
Total	1.8	6.7	8.4	1.8
t-Xyl	2.4	2.9	7.9	5.3
4-Xyl	2.0	2.5	-	-
2-Xyl	-	2.5	5.2	-
Total	4.4	7.9	16.3	5.3
t-Man	1.3	1.4	2.0	2.3
2-Man	-	1.0	-	-
4-Man	3.5	6.7	5.6	5.2
3-Man	-	16.7	12.2	6.8
3,4-Man	-	0.6	-	-
2,3-Man	-	0.7	-	1.1
2,4-Man	-	0.5	-	-
4,6-Man	-	0.5	0.6	
2,3,4-Man	-	-	-	1.3
3,4,6-Man	-	0.7	-	-
2,3,4,6-Man	-	2.3	-	-
Total	4.8	31.1	20.4	16.7
t-Gal	-	1.9	2.0	0.5
3-Gal	-	1.9	1.0	-
6-Gal	-	1.2	0.8	-
3,4-Gal	-	0.7	0.4	-
2,3,4,6-Gal	-	3.5	0.5	0.6
Total	0.0	9.2	4.7	1.1
t-Glc	1.4	0.6	1.0	-
3-Glc	6.0	-	-	-
2-Glc	1.1	-	2.0	4.6
4-Glc	59.7	1.3	1.7	-
6-Glc	0.5	-	-	2.4
2,3,4,6-Glc	-	4.4	-	-
4-GlcA	1.9	-	-	3.6
Total	68.7	6.3	4.7	7.0
Total (NS+UA)	70.6	6.3	4.7	10.6

* - Other linkages were found in minor amounts.

Besides Glc residues, methylation analysis of Et50 also allowed to verify the presence of 2-Rha (13 %) and other residues of Rha appearing as terminal t- (4 %) and 2,3-Rha (2 %). The presence of 2,3-linked Rha may indicate that 2-Rha was substituted at position *O*-3 of Rha residue. Furthermore, it was detected the residues 4- and t-Man, and t- and 4-Xyl (2 %).

Relatively to Rib, which together with Glc was one of the predominant sugars in Et50, none Rib residue was found. As a result, the methylation analysis did not allow to identify the way that Rib was linked in the PS. Similarly, Chiovitti *et al.* (17) also verified that the PS in *P. tricornutum* had as main NS Rib and they also verified the loss of Rib during linkage analysis. The authors assumed that the Rib was derived from contamination with RNA, which was lost during methylation analysis due to the high labile character of the ribofuranosyl residues. This happens because RNA is unstable/degraded in alkaline medium and the methylation analysis is performed in basic pH.

A treatment of Et50 with RNase was performed in order to ascertain if Rib was derived from RNA. The composition in NS before and after digestion with RNase of Et50 is summarized in table 14. Rib in Et50 accounted for 42 mol % of the total NS and after digestion with RNase accounted only to 34 mol %. The result indicates a decrease in 8 mol % of Rib after treatment with RNase. In addition, Ara, Xyl, and Man slightly increased after treatment with RNase, as well as Glc. Thus, this suggests that the Rib in Et50 was not derived from RNA, which does not bear the proposed before for the loss of Rib during methylation analysis.

Table 14. Neutral sugars (NS) composition of the fraction Et50 before and after digestion with RNase. Results are expressed as mean ± SD in molar percentage (%) of the total NS from duplicate analyses.

	mol %	
	Et50	**Et50 digested with RNase**
Rha	6.1 ± 0.4	7.7 ± 0.4
Fuc	0.9 ± 0.0	1.1 ± 0.1
Rib	41.5 ± 2.1	33.5 ± 1.3
Ara	2.1 ± 0.0	2.6 ± 0.1
Xyl	2.4 ± 0.0	3.3 ± 0.0
Man	6.3 ± 0.1	8.0 ± 0.1
Gal	1.3 ± 0.1	1.0 ± 0.9
Glc	39.4 ± 1.7	42.7 ± 0.9

Microalgae generally contain high levels of water-soluble vitamins, which include riboflavin (Vitamin B_2), but their content is high variable among algal species (131). In *N. oculata* has been reported the presence of riboflavin as one of its water-soluble vitamins (9,132,133). The structure of riboflavin includes ribitol as lateral chain (Fig. 15). Interestingly, degradation of riboflavin results in a cleavage of the ribitol, releasing it into

the medium (134). As a result, besides RNA, also riboflavin could be an explanation as a source of the alditol acetate correspondent to Rib. However, as seen previous in the PS composition of Et50 (Fig. 14 B), this fraction did not comprise ribitol as non-free alditol, suggesting that the high content in Rib of Et50 did not derive from riboflavin.

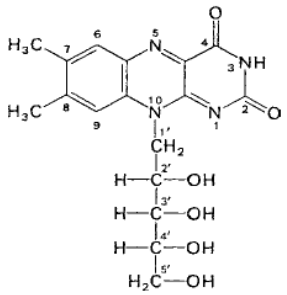


Figure 15. Chemical structure of riboflavin (134).

Regarding to the glycosidic linkages that compose the PS of the fraction Et85, the methylation analysis allowed to detect 2-Rha (18 mol % of the total sugars), 3- and 4-Man in about 17 and 7 %, respectively, as dominant sugar residues (Table 13). The result proposed the presence of PS with linear chains composed of (1→2)-Rha, as well as (1→3)- and (1→4)-Man. In addition, other residues of Rha were identified, which included t- (5 %), 2,3- (4 %), 3,4-, 2,3,4- (2 %), and 3-Rha (1 %). The 2,3- and 2,3,4-Rha residues suggest that (1→2)-linked Rha is probably substituted at positions *O*-3, and simultaneously at *O*-3 and *O*-4, respectively. Moreover, methylation analysis of Et85 showed other residues belonging to Man corresponding to substituted sugars in about 0.5-0.7 %, such as 2,3-, 3,4-, 2,4-, and 4,6-Man, and sugars residues belonging to Gal (t-, 3-, and 6-Gal) and Xyl (t-, 2-, and 4-Xyl). The high presence of substituted residues of Rha and Man may indicate that the PS in Et85 display a high DB. The total content in terminals (non-reducing ends) was 13.7 %, which is identical to the content in alditols (12 mol %). However, considering that these polymers are quite big and may display a high DB, the terminals are probably derived preferentially from ramifications and not from the backbone. Thus, there are probably much less reducing ends than terminal residues. As a result, this suggests that there are not enough reducing ends to bind the alditol, supporting the proposed before, that the alditol could be linked through another binding point, maybe the sulfate group.

Since Et85 was the fraction with the highest content in SO_3^- groups (Table 11 and Fig. 14), the desulfation of their PS before methylation analysis was performed in order to ascertain about its linkage position on the PS. The glycosidic linkages analysis after desulfation of the PS in Et85 showed a decrease in 2,3-Rha (from 4 to 2 %) and the loss of 2,3,4-Rha (2 to 0 %), with an increase in 2-Rha (from 19 to 28 %) (Table 13). This suggests that SO_3^- groups were probably linked at the position *O*-3 of (1→2,3)-linked Rha residues, and at positions *O*-3 and *O*-4 of (1→2,3,4)-linked Rha. Consequently, only 2 mol % of the total 2,3-Rha (4 mol %) present in Et85 is effectively one branched residue, since the other 2 % derived from the substitution at position *O*-3 with sulfate group. Comparatively, 2,3,4-Rha does not represent a branched residue, instead it represents a disubstituted residue with sulfate. To our knowledge nothing was reported about substitution of Rha with sulfate groups in marine microalgae, as observed in this study. Nonetheless, in marine green algae have already been reported the presence of sulfated rhamnan. In *Monostroma latissimum* was found the presence of SO_3^- groups at position C-2 or C-3 of (α1→3)-L-Rha*p* or (α1→2)-L-Rha*p*, respectively (135). Similarly, in *Gayralia oxysperma* was determined that (α1→3)-L-Rha*p* may be substituted with SO_3^- groups at positions C-2 or C-4, and (α1→2)-L-Rha*p* at positions C-4 or C-3,4 (136).

EtSN comprised 2-Rha (42 mol % of total sugars) as predominant sugar residue (Table 13), followed by t-Rha (13 %), 3-Man (7 %), 2,3-Rha (6 %), t-Xyl, 4-Man and 2-Glc in identical amounts (5 %), including 4-GlcA (4 %) as acidic sugar. This finding suggests that PS in EtSN probably contained a backbone composed of (1→2)-linked Rha, with some points of branching at *O*-3 position of Rha residue, due to the presence of 2,3-Rha, similarly to Et85. Besides, the Man from PS in this fraction occured predominantly as linear (1→3)- and (1→4)-linked Man, and it was highlighted the predominance of terminal residues (t-Rha, t-Xyl and t-Man), accounting a total of 21.5 mol %. The high content in terminal residues suggests that these polymers are short, as found by GPC analysis, and may display a lower DB than Et85, since this fraction did not comprise much substituted/branched residues. Noteworthy, the content in terminal residues (22 mol %) was much higher than the total content in alditols (12 mol %) and sulfate groups (6 mol %) (Fig. 14). This result, along with the lower *Mw* in EtSN, highlights that the PS in this fraction may display more reducing ends and so more binding points to attach the alditols, supporting the proposed before. Furthermore, this may indicate that the SO_3^- groups were negatively charged attached to the

PS, conferring a higher hydrophilic character and explaining their solubility in 85 % EtOH (v/v) together with the low *Mw*.

Additionally, in about 1 to 2 % it was verified the presence of 3-Rha, and t-, 2,3- and 2,3,4-Man, suggesting some branching points become from Man. Relatively to acidic sugars, it was verified the presence of 4-GlcA (4 %), proposing that this residue is part of the linear or the main chain. Although the glycosidic linkages in EtSN seem identical to those in Et85, it was notable that EtSN contained a higher amount of 2-, t- and 2,3-Rha, and a lower content in 3-Man. The higher content in t-Rha suggests that these PS are shorter than those in Et85, in agreement with the determined by GPC, and have a lower DB due to the reduced presence of substituted Rha residues.

Comparing the PS of the three EtOH fractions, it was verified that Et50 comprised the PS with higher *Mw* (630 kDa) (Table 12), and Glc and Rib in the HW soluble PS higher than 10 kDa (retentate) remained predominantly in this fraction (Fig. 14). Additionally, the glycosidic linkages analysis showed that Glc occurred predominantly as 4-Glc (60 mol % of total sugars) (Table 13), suggesting that the PS in Et50 contained a backbone composed of (1→4)-linked Glc. Green microalgae can employ (α1→4)-glucan-based starch as a carbon storage (47), being already reported the accumulation of starch in the microalga *Chlorella fusca* (137). Linear (α1→4)-D-glucan similar to amylose was also reported in EPS from *Dunaliella tertiolecta* (138).

Additionally, other residues of Glc were found in the PS of Et50, including 3- and 2-Glc that may be part of the backbone or linear chains, and 4-GlcA as acidic sugar (Table 13). GlcA residues were found in other marine microalgae, such as in *P. tricornutum* and *Porphirydium sp.* (19,20). 3- and 2-Glc may also indicate possible points of interruption inside the (1→4)-linked Glc backbone, suggesting the presence of (1→3, 1→4)-glucans. This would promote a more flexible structure of the PS, explaining their solubility in HW even though containing a high *Mw*. Some glucans have been found in microalgae, which included the highly branched (β1→3, β1→6)-D-glucans in *I. galbana* (21) and it was reported the presence of laminarin as storage PS in *Nannochloropsis* (47,49). However, the structure of these Glc polymers are different from the structure highlighted by methylation analysis in this work for the PS in Et50.

As recently reviewed, microalgae may be a source of dietary fibres and prebiotics (30), and a recent study determined the content in β-glucans in forty microalgae, including in *N. salina*, but the study was not centred on their structures (139). Regarding to the presence of mixed linked (β1→3, β1→4)-glucans in microalgae, little is known. Popper and Tuohy (140) reported that brown algae and diatoms may probably display (β1→3, β1→4)-D-glucans as components of cell wall. In particular, Salmeán *et al.* (141) verified the presence of mixed linked glucans in brown algae cell walls, and these polymers were also found as component of the secondary cell wall in the green alga *Micrasterias* (142).

Brown algae is included in the Stramenopile lineage, whose lineage comprises also diatoms and the class Eustigmatophyceae that includes the microalga *N. oculata* (143,144). Noteworthy, genes that encode to cellulose synthase-like (Csl) proteins from the glycosyltransferase GT2 family have been identified in land plants may participate in mixed linked synthesis (145). Although mixed linked glucans have not been reported in the class Eustigamatophyceae, some GT2 genes are present in *Nannochloropsis sp.* and they are related to those present in brown algae (50,146). This pumps out that the class Eustigmatophyceae and the brown algae may share common cell wall components, in particular mixed linked glucans. As a result, this proposes that the 3-linked Glc determined in Et50 probably may derive from mixed linked (1→3, 1→4)-glucans present in its cell wall.

Besides Glc, other residues were found as components of Et50, such as 2-, t- and 2,3-Rha, 4- and t-Man, and t- and 4-Xyl. Regarding to Rib, which was one of the main NS, it was not possible to identify the way that it was linked in the PS through methylation analysis. In addition, it was verified that this Rib did not derive neither from RNA nor riboflavin, suggesting that Rib belongs to PS.

Et85 comprised two populations of polymers with *Mw* of 5 162 and 65 kDa in average, whereas EtSN comprised only one population of polymers with lower *Mw* (26 kDa) (Table 12). Although these fractions exhibited different *Mw*, the sugar profile of their PS were identical and different from the PS in Et50. It was verified that the other two predominant sugars in retentate, Rha and Man, were essentially distributed by Et85 and EtSN. In Et85, Rha and Man were found in identical content (25 mol %), whereas EtSN was richer in Rha (47 %) and contained a lower amount of Man (20 %) (Fig. 14). Regarding to the glycosidic linkages of these sugars (Table 13), it was found that Rha occurred mainly as (1→2)-linked Rha in both fractions, being found in higher content in EtSN (42 %) than in Et85 (18 %).

Furthermore, it was verified that (1→2)-Rha may be branched mainly at position *O*-3 of Rha. Notable, in EtSN was observed a higher content in t-Rha and a lower content in substituted sugars than in Et85, proposing that its PS are shorter, explaining their solubility in 85 % EtOH (v/v). Moreover, this result is in accordance with the result obtained from GPC analysis, which showed that EtSN comprised the smallest PS. In addition, desulfation of PS of Et85 showed that (1→2)-linked Rha may be monosubstituted with SO_3^- groups at position C-3, or disubstitutted at positions C-3 and C-4 of Rha. Similarly, in green algae were reported the presence of sulfated rhamnan with a backbone of (1→2)- or (1→3)-linked Rha substituted with SO_3^- groups in various positions, such as C-2, C-3 or C-4 of Rha (135,136,147). This shows that green algae and the green marine microalga *N. oculata* may share some structural characteristics of their PS.

Relatively to Man, this sugar was detected mainly as 3- and 4-Man in Et85 and EtSN, indicating the presence of linear (1→3)- and (1→4)-Man (Table 13). As substituted residues of Man were detected 3,4-, 2,3-, 2,4- and 4,6-Man (0.5-0.7 %) in Et85, and 2,3- and 2,3,4-Man (1 %) in EtSN. Comparatively, in *P. tricornutum* was reported the presence of 3-linked Man with some branching points at position C-2 of Man (19,23), similarly to the determined in this work, since it was found (1→3)-Man as linear chain, comprising one of the branching points at position C-2 (2,3-Man residue). However, the glycosidic linkages of Man reported for *P. tricornutum* derived from its frustule (19,23), and nothing is mentioned about the presence of a frustule in *N. oculata*.

Besides Rha and Man, the Xyl and Fuc from retentate remained also in the PS of Et85 and EtSN, in a range of 3 to 5 mol % (Fig. 14), occurring preferentially as terminals (Table 8). This likely indicates that these residues were found at the ramifications of the PS. Similarly, it was detected in *Porphyridium sp.* that their EPS were ramified with t-Xyl (20), showing some identical patterns between the EPS and the HW soluble PS from *N. oculata*. Unlike to EtSN, the PS in Et85 comprised also Gal and Glc (Fig. 14). As a result, this highlighted that Et85 and EtSN are probably composed by heteropolysaccharides.

In addition to NS, the PS in Et85 and EtSN contained alditols on its composition, mainly as RhaOH and ManOH (Fig. 14 B). Polymeric ManOH has been found in *N. oceanica* as component of the storage PS laminarin, which consists in (β1→3)-glucans with occasional (β1→6)-linked branches and some of the chains terminated at the reducing end with mannitol (47). However, in this study was not detected the presence neither (1→3)- nor

(1→6)-Glc in Et85 and EtSN fractions (Table 13). This finding proposes that ManOH did not derive from laminarin. Furthermore, Hafsa *et al.* (48) recently found the presence of ManOH as component of the PS in the *N. oculata* aqueous extract, supporting the results obtained in this work.

Regarding to the composition in acidic sugars (UA), which were in approximate content in all fractions, from 11 to 15 mol % (Fig. 14), was found for Et50 and EtSN the presence of 4-GlcA residue (Table 13), conferring to the PS a slight negative character. UA have been reported to occur in various marine microalgae, mostly as GlcA (20,25). In addition, SO_3^- groups were found mainly as part of the PS of Et85 (11 mol %) (Fig. 14). The presence of SO_3^- was also verified in *N. oculata* (48). Sulfated PS are found particularly in marine organisms and not in land organisms and they have been related to be involved in various biological activities (4,14). Although sulfated PS have been detected in various marine microalgae (3,4), their structures have been less studied than those from seaweeds. Some structural characteristics of sulfated PS have been highlighted for *P. tricornutum* and *Porphyridium sp*. In *P. tricornutum* was found the presence of a branched sulfated glucuronomannan with a backbone composed of (β1→3)-mannan branched at position C-2 with sulfate or with oligosaccharides chains with terminal GlcA (19,23). On the other hand, in *Porphyridium sp*. was verified by methylation analysis that SO_3^- groups were linked at the *O*-6 position of (1→3,6)-linked glucopyranosyl residues after desulfation (20).

All in consideration, the results suggest that the PS in Et50 are different from those in Et85 and EtSN. PS in Et50 were the largest and they were composed preferentially by (1→4)-linked Glc and Rib. On the other hand, PS in Et85 and EtSN contained an approximate sugar composition and more diversified (Rha, Man, Xyl and Fuc) than those in Et50. These were composed mainly by Rha as (1→2)-linked Rha and (1→2,3)-Rha, and Man as (1→3)- or (1→4)-linked Man. In addition, Et85 and EtSN comprised UA and SO_3^- groups, proposing the presence of anionic heteropolysaccharides. Et85 showed the highest sulfated character and was found that the SO_3^- groups may be present at C-3, and C-3 and C-4 positions of (1→2)-linked Rha. At the same time, Et85 showed one population of polymers with a very high *Mw* that may indicate an aggregate of polymers. Unlike, EtSN comprised the polymers with the lowest *Mw* and a lower DB than Et85. Therefore, these

results seem indicate that sequential EtOH precipitation allowed to separate different kinds of HW soluble PS present in the constitution of *N. oculata*.

3.1.1.4. Fractions derived from the anion-exchange chromatography of Et85

The fraction Et85 obtained from the sequential EtOH precipitation was the fraction richest in SO_3^- groups, being so the fraction where remained the HW sulfated soluble PS from *N. oculata*. The total content in SO_3^- groups (11 mol %) was identical to the total content in polymeric alditols, and the total content in terminal sugar residues was low, suggesting the absence of enough reducing ends to bind the alditols. This led to hypothesize that the alditol may be linked to the PS through another binding point instead of the reducing end, probably the sulfate group, becoming a neutral sulfated PS. In addition, the GPC analysis of this fraction revealed one population of polymers with a very high *Mw* (5 162 kDa), maybe due to the presence of an aggregate of polymers that were not efficiently separated. In order to verify if the sulfated PS exists in its neutral form and to further separate the polymers present in Et85 by charge, an anion-exchange chromatography was performed.

Five fractions were obtained from the anion-exchange chromatography with a total recovery yield of 73.4 % (w/w): Fraction #1, #2, #3, #4, and #5 from 0, 0.125, 0.250, 0.5 and 1.0 M NaCl elution, respectively. Almost no PS were detected in fractions #4 and #5, being consequently only the fractions #1, #2, and #3 analysed, named as Et85_#1, Et85_#2 and Et85_#3. The yield of these last three fractions were 28.8, 26.5 and 12.8 %, respectively (Table 15). The results showed that the most part of the polymeric material remained in fractions #1 and #2, being the fraction #3 the poorest in polymeric material.

The chemical composition of the PS in the fractions obtained from the anion-exchange chromatography of Et85, in terms of content in NS, UA and SO_3^- groups, are shown in table 15. Fraction #1 was the richest in NS (65 % w/w), followed by fraction #2 and #3, which comprised 33 and 20 % of NS, respectively. Fraction #3 had the highest content in UA (22 %) and SO_3^- groups (6 %), whereas fraction #1 had the lowest content in UA (6 %) and SO_3^- groups (0.5 %). This indicates that fraction #1 contained the lowest negative character and the fraction #3 the highest. Fraction #1 was eluted without NaCl, so did not interact with the stationary phase positively charged, being composed of neutral or positively charged material, in agreement with low negative charged obtained. Otherwise, fraction #3 was

eluted with 0.250 M NaCl, and so interacted more with the stationary phase, explaining the higher negative character obtained. The UA and sulfate content in fraction #2 was 10 and 3 %, respectively, having these PS a lower negative character than those in Et85_#3, as expected, since was eluted with a lower concentration of NaCl and thus less retained in the positively charged column.

Table 15. Yields of the Fractions #1, #2 and #3 obtained from the anion-exchange chromatography of the fraction Et85 with 0, 0.125 and 0.250 M NaCl, respectively, results are expressed as weight percentage (%) of Et85. Content in neutral sugars (NS), uronic acids (UA) and sulfate (SO_3^-) groups that compose the polysaccharides (PS) present in fractions #1, #2, and #3. Results expressed as weight % of fraction.

		(mean ± SD) %*			
Fraction	**Yield %**	**NS**	**UA**	**SO_3^-**	**∑ NS+UA+SO_3^-**
#1	28.8	64.8 ± 3.3	6.1 ± 0.5	0.5 ± 0.01	71.5 ± 2.8[a]
#2	26.5	32.5 ± 3.2	10.4 ± 0.7	2.5 ± 0.2	45.2 ± 2.4[b]
#3	12.8	20.1 ± 3.9	22.3 ± 0.4	5.9 ± 0.3	48.1 ± 4.0[b]

*- Data represent the mean values ± SD in weight % of each fraction from duplicate (n=2), n=4, and n=3 analyses for NS, UA, and SO_3^-, respectively.
Different superscripts in same column mean results significantly different ($p<0.05$).

The total content in PS, as a result of the sum of the content in NS, UA and sulfate, was 72, 45 and 48 % for fraction #1, #2, and #3, respectively (Table 15). Interestingly, the result showed that although fractions #1 and #2 had an identical yield of extraction, it was in fraction #1 where the PS preferentially remained, since their content in PS was higher in about 26 % relatively to fraction #2 ($p<0.05$). In addition, it should be highlighted that in spite of the lower yield of fraction #3, the PS content in fraction #2 and #3 was similar ($p<0.05$). PS in fraction #1 comprised mainly NS (90 % of total PS), 9 % of UA and a negligible content of sulfate (0.7 %), showing the preferential non-negatively charged character of these PS. This suggests the absence of neutral sulfated PS, opposing the hypothesis proposed for the presence of neutral sulfated PS due to the binding of alditols to the sulfate group. On the other hand, PS in fraction #3 comprised an identical content in NS (42 %) and UA (46 %), and had the highest content in sulfate groups (13 %), retaining thus the main content in anionic PS. The PS in fraction #2 included 6 and 23 % of sulfate and UA, being composed mainly by NS (72 %), as fraction #1.

NS and non-free alditols analyses allowed to detect the distribution of monosaccharides of Et85 to the fractions #1, #2, and #3. Composition in monosaccharides, alitols, UA and SO_3^- groups that compose the PS in fractions #1, #2 and #3 is shown in figure

16. The results showed that PS in fraction #1 were composed mainly by Rha (39 mol %), Man (36 %), and mannitol (11 %), being other sugars found in residual amounts. This suggests that the Man and mannitol present in Et85 may be part of neutral or positively charged PS. Although being in residual content, PS in fraction #1 displayed also UA (19 %) and sulfate groups (1 %), as already referred.

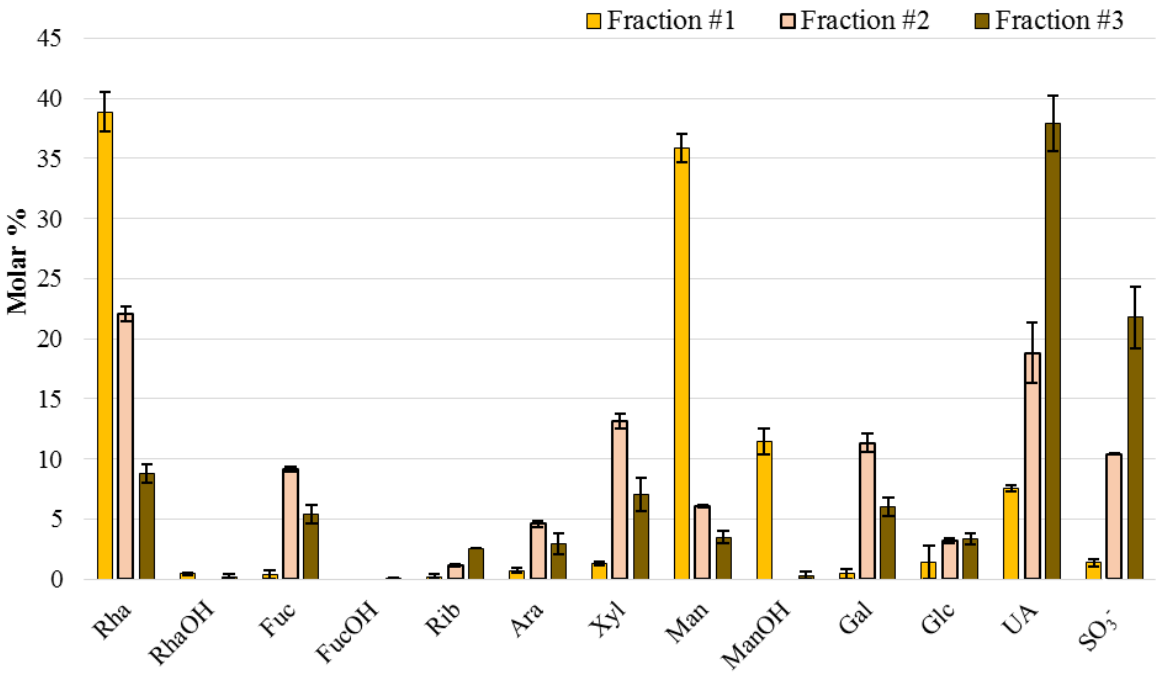


Figure 16. Composition in neutral monosaccharides, alditols, uronic acids (UA), and in sulfate (SO_3^-) groups of the polysaccharides (PS) in Fraction #1, #2, and #3 obtained from the anion-exchange chromatography of Et85 with 0, 0.125 and 0.250 M NaCl, respectively. Results are expressed in molar percentage (mol %) of the total content in sugars, alditols and sulfate.

Regarding to the monosaccharide composition of the PS in fractions #2 and #3, these were composed mainly by Rha, Xyl, Gal and Fuc, being more representative in PS of fraction #2 than in fraction #3 (Fig. 16). Unlike, PS in fraction #3 had a higher proportion in UA (38 mol %) and sulfate (22 %) than fraction #2. In addition, it was not detected any alditol in the constitution of fraction #2 and #3, suggesting that they are not present in anionic PS.

The glycosidic linkages that compose the PS retained in each fraction from the anion-exchange chromatography of Et85 are described in table 16. The results showed that the high content in Rha and Man found in fraction #1 occurred in PS mainly as 2-Rha (27 mol %), 3- and 4-Man (30 and 20 %, respectively), t-Rha (8 %), and displayed 2,3-Rha (4 %) as substituted residue. This suggests the presence of a backbone composed of (1→2)-linked

Rha that may be substituted preferentially at position C-3 of Rha, as found previously for the PS present in the fraction Et85 (section 3.1.1.3.). The high content in t-Rha may indicate that the polymers in this fraction were small, since this content was higher than the content in the substituted residue 2,3-Rha (4 %), and this fraction did not contain other substituted sugar residues. In addition, the 3- and 4-Man residues detected suggest the presence of linear polymers of (1→3)- and (1→4)-linked Man or even mixed (1→3,1→4)-linked Man, being these glycosidic linkages in agreement with the previously determined from analysis of Et85.

Table 16. Glycosyl linkage composition (mol % of the total sugars) of the fractions Et85_#1, Et85_#2 and Et85_#3 obtained from the anion-exchange chromatography of the fraction Et85 with 0, 0.125 and 0.250 M NaCl, respectively, and of the initial Et85 fraction. Glycosidic linkages were determined from the carboxyl-reduced partially methylated alditol acetates (PMAAs), except for Et85.

		% molar		
Glycosidic Linkage*	**Et85**	**Et85_#1**	**Et85_#2**	**Et85_#3**
t-Rha	5.3	7.9	6.9	4.6
2-Rha	18.4	26.8	22.5	13.0
3-Rha	1.4	1.1	1.1	0.5
3,4-Rha	2.0	-	1.2	1.3
2,3-Rha	3.8	4.1	3.8	1.3
2,4-Rha	0.5	-	0.6	-
2,3,4-Rha	1.9	-	3.5	1.6
Total	33.3	39.9	39.6	22.3
t-Fuc	1.0	-	3.5	-
3-Fuc	-	-	0.5	0.9
2-Fuc	-	-	0.5	-
Total	1.0	0.0	4.5	0.9
t-Rib*f*	0.6	-	0.7	0.7
Total	0.6	0.0	0.7	0.7
4-Ara*p*=5-Ara*f*	3.1	-	2.2	2.9
3,4-Ara*p*=3,5-Ara*f*	1.8	-	-	-
2,3,4-Ara*p*=2,3,5-Ara*f*	1.8	-	-	6.6
Total	6.7	0.0	2.2	9.5
t-Xyl	2.9	1.3	3.8	9.5
3-Xyl	-	-	-	1.4
4-Xyl	2.5	-	1.0	5.0
2-Xyl	2.5	-	8.7	3.3
Total	7.9	1.3	13.5	19.2
t-Man	1.4	2.0	1.1	0.7
2-Man	1.0	-	-	-
4-Man	6.7	20.3	1.3	-
3-Man	16.7	29.6	-	-

Table 16. Continuation

	% molar			
Glycosidic Linkage	**Et85**	**Et85_#1**	**Et85_#2**	**Et85_#3**
3,4-Man	0.6	-	-	-
2,3-Man	0.7	0.5	-	-
2,4-Man	0.5	-	0.5	-
4,6-Man	0.5	-	0.5	-
3,4,6-Man	0.7	-	-	-
2,3,4,6-Man	2.3	-	-	2.2
Total	31.1	52.4	3.4	2.9
t-Gal	1.9	-	5.3	1.7
3-Gal	1.9	-	8.0	1.2
6-Gal	1.2	-	6.7	-
3,4-Gal	0.7	-	-	0.5
2,3,4-Gal	3.5	-	-	-
2,3,4,6-Gal	-	-	-	12.4
Total	9.2	0.0	20.0	15.8
t-Glc	0.6	-	0.6	3.6
2-Glc	-	1.0	1.3	-
4-Glc	1.3	2.8	4.8	3.9
2,3,4,6-Glc	4.4	-	-	18.7
t-GlcA	-	-	4.7	1.7
4-GlcA	-	-	1.6	0.7
Total	6.3	3.8	6.7	26.2
Total (NS+UA)	6.3	3.8	13.0	28.6

* - Other glycosidic linkages were found in minor amounts.

Fraction #2 named also as Et85_#2, which comprised Rha, Xyl, Fuc and Gal as main NS, displayed a high diversity in glycosidic linkages, especially regarding to Rha residues (Table 16). Their PS comprised as main residue 2-Rha (23 mol %), suggesting the presence of a backbone composed of (1→2)-linked Rha. Other residues were found in a range between 3 to 8 %, such as 2-Xyl (9 %), 3-Gal, t-Rha, 6-Gal, t-Gal, 4-Glc, t-Xyl, 2,3- and 2,3,4-Rha, and t-Fuc (3.5 %). Besides 2,3- and 2,3,4-Rha residues, also 3- and 3,4-Rha residues were identified in minor amounts (1 %) and as acidic sugars were observed t- and 4-GlcA residues. The high predominance of terminal sugars (Rha, Gal, Xyl, Fuc and Man), which accounted a total of 21 mol %, may indicate that the PS in fraction #2 are small and/or contain a high DB due to the presence of substituted residues, displaying a very complex structure. Interestingly, it should be noted that this fraction retained the 2,3- and 2,3,4-Rha residues, which were found from the methylation analysis of the desulfated Et85 to be substituted at

positions C-3, and C-3 and C-4 with sulfate groups. This result evidences that probably the sulfated rhamnan PS were present in this fraction. Thus, 2,3- and 2,3,4-Rha in Et85_#2 are probably not only points of branching, but also points of substitution with sulfate groups.

Although fraction #2 and #3 were similar in terms of composition in monosaccharides, fraction #3 did not comprise substituted Rha sugars, which accounted a total of 4 mol % of the total sugars lower in comparation with fraction #2 (9 %) (Table 16). The predominant residue found in fraction #3 was 2-Rha (13 %), suggesting the presence of a backbone composed of (1→2)-linked Rha, similarly to fraction #2, followed by t-Xyl (9 %), which suggests the presence of short or highly branched PS with Xyl. Other residues were found in a range between 2 and 5 %, such as 4-Xyl, t-Rha, 4- and t-Glc, 2-Xyl, 4-Ara*p*=5-Ara*f*, t-Gal, and 2,3,4-Rha, being found as acidic sugars t-GlcA and 4-GlcA. Due to the high amount of terminal residues (a total of 21 %) and the lower content in substituted residues comparing to fraction #2, especially regarding to Rha residues, this probably indicates that the polymers in fraction #3 are shorter and display a lower DB. Further, it should be noted that the sulfate group was proposed to be linked on the PS of Et85 at position C-3 of (1→2,3)-Rha, or at positions C-3 and C-4 of (1→2,3,4)-Rha. These Rha residues were found in Et85_#3, which was the richest in SO_3^- groups, but their content was lower than those found in Et85_#2, and in particular, 2,3-Rha was higher in Et85_#1. The result may indicate that the SO_3^- groups linked at these positions of Rha might occur preferentially in fraction #2, or even in residual amount in fraction #1 at position C-3 of (1→2,3)-Rha.

In order to understand if the anion-exchange chromatography besides separate the polymers present in Et85 by charge allowed to obtain polymers with different *Mw*, a GPC analysis was performed. This experiment showed that fraction #1 comprised only one population of polymers with *Mw* of 19 kDa in average (Fig. 17 and Table 17). Unlike, fraction #2 and #3 contained two populations of polymers with different *Mw*. The *Mw* of the two populations in fraction #2 was 448 and 47 kDa, and in fraction #3 339 and 14 kDa. The results showed that the polymers with lower *Mw* present in the initial Et85 fraction remained in fraction #1 (19 kDa) and #3 (14 kDa) (Table 17). This supports the proposed before for the presence of short polymers in the composition of fraction #1 and #3 due to the high presence of terminals, especially Rha and Xyl residues, respectively.

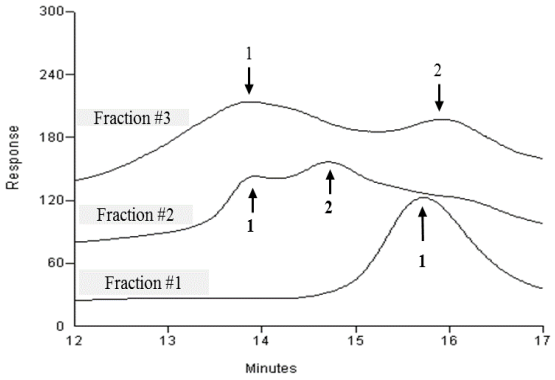


Figure 17. Chromatogram from the GPC analysis of the fractions #1, #2, and #3 obtained from the anion-exchange chromatography of the fraction Et85. Numbers 1 and 2 represent the populations of polymers present in each fraction with different molecular weights (*Mw*).

Table 17. Estimation of the weight average molecular weights (*Mw*) and number average molecular weight (*Mn)* present in the fractions obtained from the anion-exchange chromatography (#1, #2 and #3) by GPC.

Fraction	Population	*Mw*	*Mn*	*Mw/Mn*
#1	1	19 330	12 835	1.5
#2	1	448 420	274 115	1.6
	2	47 340	14 630	3.2
#3	1	339 440	195 675	1.7
	2	13 920	9 413	1.5

Fractions #2 and #3 comprised the PS more negatively charged and two populations of polymers with high *Mw* (Table 17). Noteworthy, the high *Mw* obtained for the population 1 in fractions #2 (448 kDa) and #3 (339 kDa) may be explained due to the presence of polymers with higher DB compared to fraction #1, due to the presence of terminal and substituted sugar residues. Additionally, fraction #2 comprised a population with higher *Mw* than fraction #3, evidencing the proposed above that fraction #2 may display a higher DB than fraction #3. In particular, this displayed that the polymers in fraction #2 and #3 have a higher DB compared to fraction #1. Moreover, this could suggest that carboxyl and sulfate groups may be present preferentially in branched PS, since they remained predominantly in fractions #2 and #3.

As a result, the GPC analysis allowed to separate the aggregate of polymers present in the first population of polymers in Et85 with a *Mw* of 5 162 kDa (Fig. 17, table 17) by populations 1 and 2 of fraction #2, and population 1 of fraction #3. On the other side, the second population of polymers previous found in Et85 with *Mw* of 65 kDa were distributed into population 1 of fraction #1 and population 2 of fraction #3. Thus, the GPC analysis showed that the three fractions obtained from the anion-exchange chromatography, besides being differently charged, contained polymers with different *Mw*.

Comparing the three fractions (#1, #2 and #3) obtained from the anion-exchange chromatography of fraction Et85, it was verified that fraction #3 was the richest in UA and

SO_3^- groups, followed by fraction #2, being their content almost negligible in fraction #1. Thus, the hypothesis of the presence of neutral sulfated PS due to the linkage of one alditol to the sulfate group on the PS still remains unclear.

Man and ManOH were isolated from the fraction Et85 into the fraction #1. Man occurred preferentially as (1→3)- and (1→4)-linked Man in agreement with the previous found for Et85. This suggested that Man polymers were component of neutral or positively charged HW soluble PS in *N. oculata* and should not be a point of substitution with sulfate. Additionally, Rha, which together with Man accounting the predominant sugars in Et85, was distributed by the three fractions, being found in higher content in PS of the fraction #1 (39 mol %) together with Man. In fraction #1, Rha should be part of a backbone composed of (1→2)-linked Rha with a branching point occurring at position C-3 of Rha, as found for Et85. In the other fractions, Rha had a high diversity of glycosidic linkages, especially in fraction #2, which comprised the 2,3- and 2,3,4-Rha residues that were found to be substituted with sulfate groups in Et85.

The other sugars found in lower amounts in Et85, Xyl, Gal, and Fuc, were distributed mainly by the PS of the fractions #2 and #3, occurring with various types of linkages and with presence of terminal residues. The predominance of terminals in fraction #2 and #3 suggests that these polymers were small or high branched. In addition, fraction #2 comprised a higher amount of substituted residues than fraction #1 and #3, suggesting that its PS contained a higher DB. Moreover, this highlighted that fraction #2 and #3 were formed by anionic sulfated heteropolysaccharides with a very complex structures, unlike fraction #1.

The GPC analysis revealed the presence of polymers with different *Mw* in the three fractions. Polymers in fraction #1 comprised a *Mw* of 19 kDa in average, whereas fractions #2 and #3 showed two different populations of polymers with 47 and 448 kDa in fraction #2, and 14 and 339 kDa in fraction #3. As a result, considering the content in UA and SO_3^- groups in fractions #1, #2 and #3, their composition in monosaccharides and glycosidic linkages, and the *Mw* of their polymers, the anion-exchange chromatography allowed to separate different kinds of PS present in Et85, especially between fraction #1 and the other two fractions.

Considering all the results obtained from the sequential procedure employed to study the HW soluble PS present in *N. oculata*, it was possible to highlight some structural features

of their PS, in particular those that remained in the following fractions: Et50, Et85, EtSN, Et85_#1, Et85_#2 and Et85_#3. This initial characterization of the HW soluble PS constitutes a functional starting point to develop carbohydrate microarrays, since allows to select proteins (CBMs, antibodies and lectins) with well-defined specificities that may recognize these structures and others, allowing to get more information about the structural characteristics of the PS. In addition, the knowledge of the chemical structures of these PS is also important in order to search for new possible carbohydrate-protein interactions that they can manage, and structurally characterize novel protein-carbohydrate complexes.

3.1.2. Recombinant expression of proteins from *Bacteroides thetaiotaomicron* VPI-5482: BT3698 and BT0996

The bacteria *Bacteroides thetaiotaomicron* resident in the human gut has polysaccharide utilization locis (PULs) that encode enzymes involved in degradation of complex PS (112). These enzymes comprise CBMs but the detailed knowledge of its interaction with carbohydrates remains unclear. With this in mind, to study some of these modules, two CBM domains of two proteins, BT3698 and BT0996, were produced through recombinant expression.

3.1.2.1. Proteins quantification and SDS-PAGE analysis

The proteins BT3698 and BT0996 belonging to *Bacteroides thetaiotaomicron* were produced by recombinant expression in order to be posteriorly use in carbohydrate microarrays analysis. After production, these proteins were quantified (Table 18) and analysed by SDS-PAGE (Fig. 18) to infer about their *Mw* and purity, respectively.

The cloned domain of BT3698, the appended non-catalytic CBM module (CBM58) of a SusG protein, was produced with a concentration of 2.0 mg/mL, whereas the cloned domain of BT0996, a CBM57 appended to the GH2, was produced at a concentration of 1.0 mg/mL (Table 18). The results showed that both proteins were obtained in good amounts to further employment in microarrays.

Table 18. Characteristics of the recombinant proteins BT3698 and BT0996. Organism that they belong, cloned domain and concentration of each protein (µg/mL).

Protein	Organism	Cloned domain	$C_{protein}$ (µg/mL)
BT3698	*Bacteroides thetaiotaomicron VPI-5482*	CBM58	1955.7
BT0996	*Bacteroides thetaiotaomicron VPI-5482*	CBM57	1000.0

The results of the SDS-PAGE analysis (Fig. 18) confirmed the predicted *Mw* for each protein, 15.34 and 15.89 kDa for BT3698 and BT0996, respectively (Protein parameters of each protein is described in supplementary table 1). Further, this analysis also showed the purity of the proteins in the samples and the absence of significant degradation, suggesting the absence of contamination with other proteins. Thus, the affinity chromatography allowed to obtain these proteins purified, being prepared to use in carbohydrate microarrays.

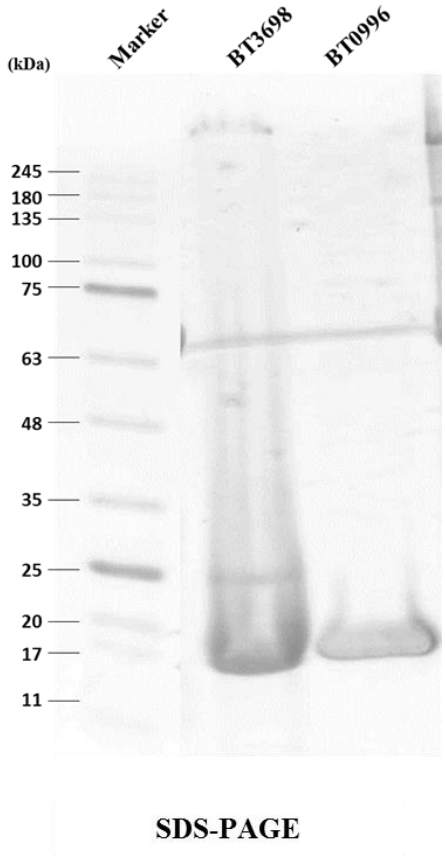


Figure 18. SDS-PAGE (10 % acrylamide) results showing purification of the proteins BT3698 and BT0996. Results of purification with His Gravitrap columns for BT3698 and with ÄKTA START for BT0996. BT3698 and BT0996 proteins were eluted with 5 ml of 50 mM HEPES at pH 7.5, 1 M NaCl, 5 mM $CaCl_2$ and 300 mM imidazole at a final pH of 7.5, and with 50 mM HEPES at pH 7.5, 1 M NaCl, 5 mM $CaCl_2$ and 500 mM imidazole at a final pH of 7.5, respectively. **SDS-PAGE**: Gel was performed with a constant amperage of 60 mA. Visualization of proteins molecular weight (*Mw*) was assed by coloration with NZYTech BlueSafe. For SDS gel was used the respective concentrations: BT3698 33 µg and BT0996 16 µg. A Marker II from NZYTech® was used (Marker) as reference.

The recombinant expression procedure was an important step that allowed to obtain the proteins BT3698 and BT0996 from *B. thetaiotaomicron* in good amounts (2.0 and 1.0 mg/mL, respectively) and without contamination with other proteins for posterior probing

in microarrays developed with the PS fractions derived from *N. oculata*. These proteins will be used to further characterize the PS of *N. oculata* and understand if they are recognized by *B. thetaiotaomicron*.

3.1.3. Carbohydrate microarrays developed with PS from *N. oculata*

In order to further characterize and understand the chemical structures of the HW soluble PS extracted from *N. oculata*, the fractions Et50, Et85, EtSN, Et85_#1, Et85_#2 and Et85_#3 were printed non-covalently in microarrays. The carbohydrate microarray developed with PS fractions extracted from *N. oculata* and PS control from commercial sources, which together constituted the Microalgae PS set 1 (Table 19), were analysed using a total of seven proteins. Five of these proteins included two lectins (ConA, RCA_{120}), two antibodies (Anti-heteromannan and Anti-(β1→3, β1→3)-D-glucan), and one microbial CBM (*Cm*CBM6-2) with well-known carbohydrate-binding specificities, allowing to obtain more information about the structural characteristics of the microalga PS.

The remaining two proteins investigated were the CBM domains of the proteins BT3698 and BT0996 that belong to *B. thetaiotaomicron* produced by recombinant expression (features of the proteins produced in section 3.1.2.). Although the carbohydrate-binding specificity of these CBMs are not fully studied, they were included in this experiment since *B. thetaiotaomicron* is a bacteria resident in the human gut, being important understand if the PS from *N. oculata* may be recognized by this bacteria through their structural proteins and consequently being metabolized if included in the human diet. The probes used in microarrays are described in table 19. The carbohydrate-binding features of each protein analysed is described in table 6 (section 2.1.8.2.).

Table 19. List of all carbohydrate probes that composed the Microalgae polysaccharides (PS) set 1 that were printed in microarrays. The main sugar composition is described for each probe.

Carbohydrate Probes Set 1	Main sugar composition
Et50	4-Glc (60 mol%)>2-Rha (13%)>3-Glc (6%)> 4-Man (4%)> t-Rha (3%)> t and 4-Xyl = 4-GlcA (2%)
Et85	2-Rha (18 mol%) >3-Man (17%)> 4-Man (7%)> t-Rha (5%)>2,3-Rha (4%)> 4-Ara*p*=5-Ara*f* = t-, 2- and 4-Xyl (3%)> 3,4- and 2,3,4-Rha (2%)
Et85_#1	3-Man (30 mol%)> 2-Rha (27%)> 4-Man (20%)> t-Rha (8%) > t-GlcA (5%)> 2,3-Rha (4%)> 4-Glc (3%)> t-Man = 4-GlcA (2%)
Et85_#2	2-Rha (23 mol%)> 2-Xyl (9%)> 3-Gal (8%)> t-Rha (7%)> 6-Gal (7%)> t-Gal = 4-Glc (5%)> t-Xyl = 2,3- and 2,3,4-Rha = t-Fuc (4%)
Et85_#3	2-Rha (13 mol%) > t-Xyl (10%)> 4-Xyl > t-Rha (5%)> 4- and t-Glc (4%)> 2-Xyl (3%) > 4-Ara*p*=5-Ara*f* (3%)
EtSN	2-Rha (42 mol%)> t-Rha (13%)> 3-Man (7%)> 2,3-Rha (6%)> t-Xyl = 4-Man =2-Glc (5%)> 4-GlcA (4%)> t-Man (2%)
Pullulan	Linear mixed-linked (α1→4, α1→6)-D-Glucan
α-Mannan	Linear (α1→6)-D-Mannan
Glucomannan	Linear mixed-linked (β1→4)-D-Mannan with (β1→4)-D-Glc
Galactomannan	Branched (β1→4)-D-Mannan with (β1→4)-D-Gal
Rhamnogalacturonan I (RGI)	(α1→4)-GalA with (α1→2)-Rha
Low Methylated Galacturonate (LMG)**	Low methylated (α1→4)-GalA DE* < 5%
High methylated Galacturonate (HMG)**	High methylated (α1→4)-GalA DE > 70%
Polygalacturonic acid (PGA) from apple**	Gal (4.5 % dw) as t- and (1→4)-Gal, and (1→4)-GalA Glc (2 % dw) Ara (0.9 % dw) as (1→4)-Ara*p* or (1→5)-Ara*f* Rha (0.7 % dw) as (1→2)-Rha
PGA from citrus**	Gal (4.5 % dw) as (1→4)- and t-Gal, and (1→4)-GalA Glc (3 % dw) Ara (0.8 % dw) as (1→4)-Ara*p* or (1→5)-Ara*f* and t-Ara*f* Rha (0.5 % dw) as (1→2)-Rha
Mixed β-Glucan**	Mixed-linked (β1→3, β1→4)-D-glucan

*DE – Degree of esterification.
**- Carbohydrate composition described in supplementary material, supplementary information 4.

3.1.3.1. Plant lectins

Microarray screening analysis with ConA showed a preferential binding for the respective polysaccharide control used in this experiment (α-mannan) (Fig. 19 and 23). The result is in accord with the reported for ConA specificity, since this lectin is highly specific to recognise α-linked mannose (85,90). However, the fluorescence signal obtained from the interaction between ConA and α-mannan at 0.1 mg/mL of concentration was very low, being better observed the specificity of this lectin for α-mannan when printed at 0.5 mg/mL but still with a low signal. It should be mentioned that good/strong carbohydrate-protein interactions detected through microarray experiments are considered when the fluorescence signals are above 10 000. As a result, in order to obtain a better fluorescence signals from the analysis with ConA, possible solutions would be increasing the protein concentration or changing the blocking solution used during binding assay.

Additionally, slight interactions (fluorescence intensities under 1000) were observed between ConA and the PS fractions from *N. oculata* printed at 0.5 mg/mL, Et50 > Et85 > Et85_#1 > Et85_#3 > Et85_#2 > EtSN (Fig. 19 and 23). Considering that the interaction between ConA and its respective positive polysaccharide control (α-mannan) was also low, and no interactions were observed between ConA and galactomannan or glucomannan that comprising Man in β-configuration, this likely indicates that the Man residues found previously in the PS fractions from *N. oculata* should occur predominantly in α-configuration, especially the Man present as 3- and 4-linked Man. In order to confirm this result this microarray experiment should be repeated to optimize the experiment conditions. Some approaches that could be tested are: increasing the protein concentration, changing the blocking solution, or increasing the concentration of the probes.

Relatively to the carbohydrate microarrays analysed with RCA_{120}, this lectin bound preferentially to polygalacturonic acid (PGA) from citrus at 0.5 mg/mL, followed by binding to low methylated galacturonate (LMG), also at 0.5 mg/mL, but with very low signals. In addition, negligible interactions with the PS probes belong to *N. oculata* (Fig. 19 and 23) were found.

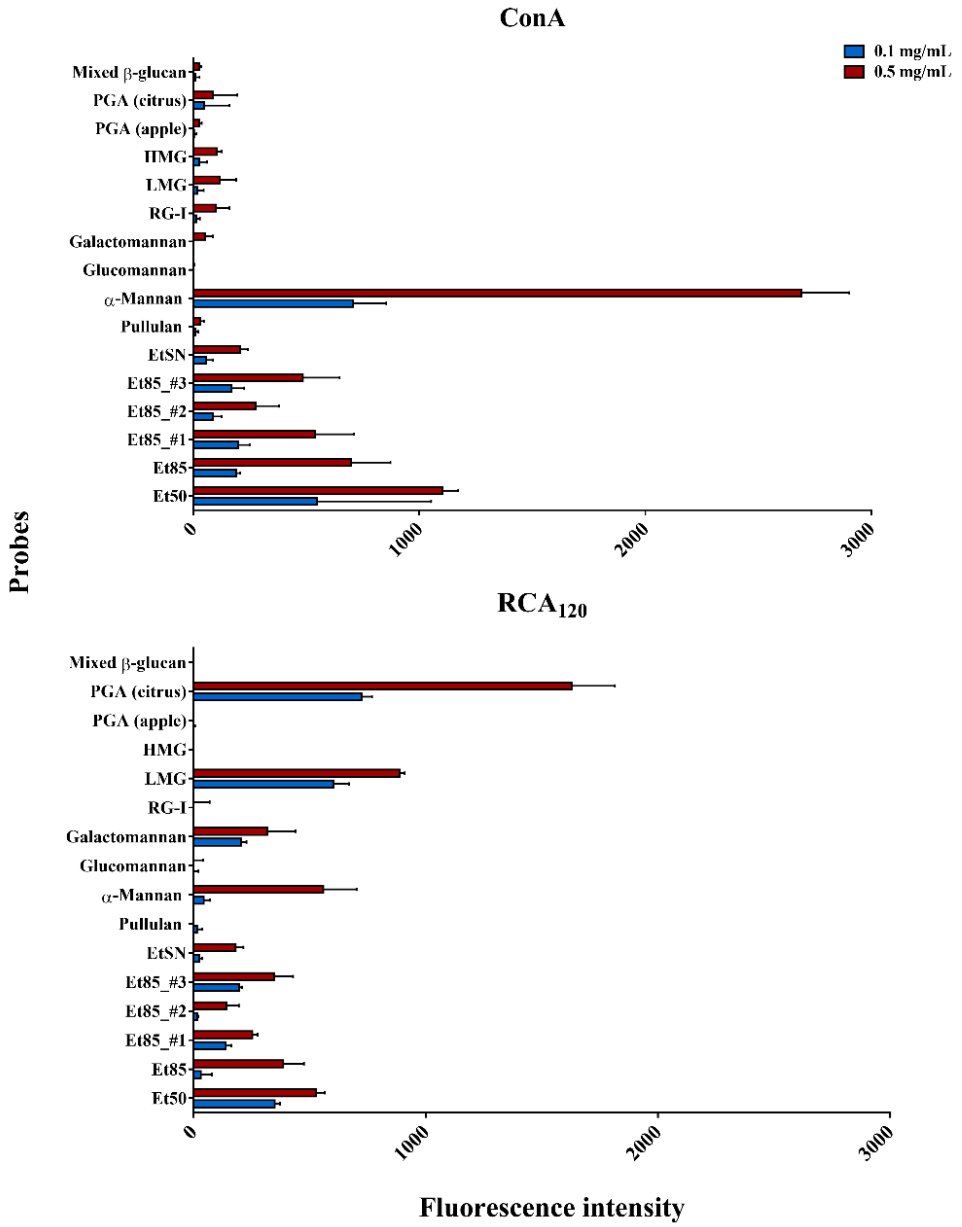


Figure 19. Screening of the microarrays developed with Microalgae PS set 1 probes by analysis with two plant lectins: ConA and RCA_{120}. The binding signals are depicted as means of fluorescence intensities of triplicate spots of probe arrayed (with error bars). Each probe was printed in triplicate at two levels: 0.1 (blue bars) and 0.5 (red bars) mg/mL (30 and 150 pg/spot). Quantified fluorescence intensity is plotted on the x-axis. Carbohydrate probes are plotted on the y-axis.

Although two PGAs were included in this study, one from citrus and another one from apple, no interaction was observed between RCA_{120} and the PGA from apple (Fig. 19 and 23). Both PGAs contain an identical sugar composition (composition described in supplemental information 4) but their degree of esterification (DE) is unknown. The different interaction with RCA_{120} should be due to the extraction of PGAs in different phases of the fruit ripening that may influence its constitution and probably due to different DEs.

Exemplifying, pectins from a ripe fruit exhibit a lower DE, a lower *Mw*, and a decrease in neutral sugars content compared to pectins in unripe fruits (148). In addition, the absence of interaction with a polysaccharide composed of high methylated galacturonate residues (HMG) and interaction only with LMG also suggests that the DE influences the interaction with RCA_{120}. Furthermore, this result displays that RCA_{120} recognises preferentially pectins with a backbone composed of (α1→4)-linked GalA without esterification. This proposes that the esterification may restrict the interaction of the lectin to PS.

RCA_{120} is known to bind Gal residues, especially terminal Gal moieties in β-configuration rather than in α-configuration (85). In β-configuration RCA_{120} recognises preferentially terminal Gal(β1→4) to Gal(β1→3) and Gal(β1→6), without any requirement for the sub-terminal residues and is able to recognise sulfated galactose sequences with 2-*O*- and/or 6-*O*-sulfation (61). In addition, RCA_{120} may also bind to *N*-glycans, and preferentially when they contain Gal(β1→4) terminal (62). Thus, the low interactions found with the PGA from citrus and LMG should be due to the presence of α-Gal residues, since normally in pectins they occur in α-configuration and it was found Gal residues as components of these PS (supplementary information 2). In addition, this highlighted that RCA_{120} only recognises these residues when the backbone composed of (α1→4)-linked GalA displays a low DE. Furthermore, the negligible interactions observed from the spots printed with the PS fractions from *N. oculata* suggests the lack of Gal in their HW soluble PS, particularly in β-configuration as Gal(β1→4) terminal, in accordance with the previous reported in section 3.1.1. The previous characterization showed the presence of Gal only in low amounts and it was not found the 4-Gal residue on its composition.

Nevertheless, the results from the analysis with RCA_{120} are not considered conclusive because the fluorescence signals obtained were too low, being the background intensity so high that may overlap possible interactions. Possible solutions to obtain higher fluorescence signals and to get results more conclusive are: increase the protein concentration, change the blocker solution used during the binding assay (information of concentration and blocker solution used to RCA_{120} is shown in supplementary table 2), or increase the concentration of the probes.

3.1.3.2. Monoclonal antibodies

Carbohydrate microarrays probed with the monoclonal antibody (mAb) anti-heteromannan LM21 (Rat IgM) showed predominant interactions to galactomannan and glucomannan printed at both concentrations (0.1 and 0.5 mg/mL) (Fig. 20 and 23), confirming the specificity reported for this antibody towards (β1→4)-linked mannans (100,101,125). In addition, it was observed unexpected fluorescence signals from the spots printed with Et50 and mixed β-glucan probes. These fluorescence signals were not considered because they might derived from interaction with the mAb anti-(β1→3, β1→4)-D-glucan (highly specific to (β1→3, β1→4)-D-Glc), due to contamination with this mAb during binding assay. As a result, this highlighted that no interactions were found between the mAb anti-heteromannan and the PS fractions of *N. oculata*, suggesting that Man, in particular the 4-Man, found previously mainly in Et85, Et85_#1 and EtSN were not in β-configuration. This corroborates the results obtained with ConA, which indicated the presence of Man in α-configuration despite the low fluorescence signals. However, to further confirm these results, the microarray experiment probed with the mAb anti-heteromannan must be repeated.

The mAb anti-(β1→3, β1→4)-D-glucan interacted preferentially with the mixed β-glucan from Oat (Fig. 20 and 23) that is composed of mixed-linked (β1→3, β1→4)-D-glucan, in accord with the specificity reported for this mAb (102). Remarkably, this mAb showed a higher interaction with the mixed β-glucan printed at low concentration (0.1 mg/mL) than at high concentration (0.5 mg/mL), suggesting that at low concentration may be in saturation.

Additionally, it was observed a strong interaction between the anti-(β1→3, β1→4)-D-glucan and Et50, being the interaction stronger at high concentration (Fig. 20 and 23). This indicates that the high amount of Glc found in the Et50 fraction from *N. oculata*, especially as 4-Glc (60 mol % of the total sugars) and containing also 3-Glc (6 %), should occur in β-configuration. Noteworthy, this mAb has a strict specificity to (β1→3, β1→4)-D-glucan and does not present cross-reactivity to (β1→3)-D-glucan (102). Consequently, this indicates the presence of mixed-linked (β1→3, β1→4)-D-glucan as component of Et50. This result confirms the proposed in section 3.1.1.3., where it was referred that the 3-Glc in Et50 might derive from mixed-linked (β1→3, β1→4)-D-glucans, which were already found as components of the cell wall in brown and green algae (141,142), emphasising that *N. oculata*

may share common features with brown and green algae. In particular, Salmeán *et al.* (141) verified through a glycan array analysed with the mAb BS-400-3 (anti-(β1→3, β1→4)-D-glucan) the presence of mixed linked glucans in brown algae cell walls. These authors also detected that the mixed linked glucans had a regular block structure of cellotriose units linked by single (β1→3)-D-linkages (Glc(β1→4)Glc(β1→3)Glc). In addition, the presence of mixed linked glucans in Et50 explains the solubility of these PS in HW in spite of their high *Mw*, due to the interruption inside the backbone that promotes the flexibility of the polymer, becoming more soluble.

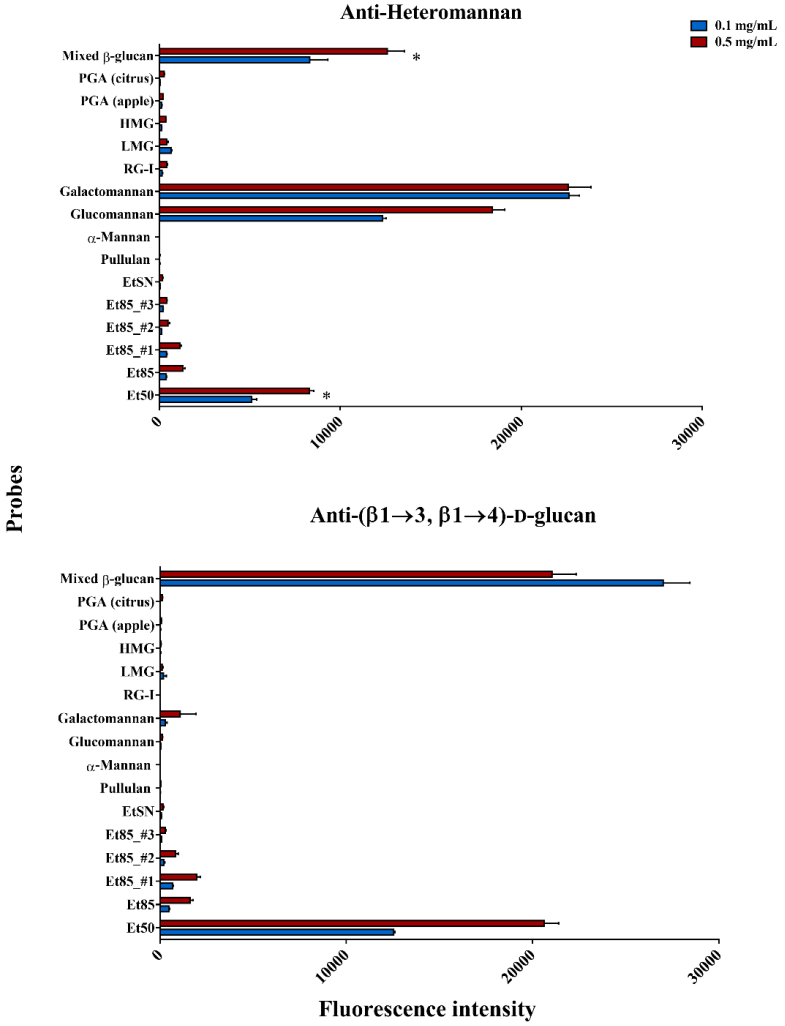


Figure 20. Screening of the microarrays developed with Microalgae PS set 1 probes by analysis with two antibodies (Abs). Ab names are depicted on top of each graph and both Abs have different specificities (Table 6). Carbohydrate sequence information of the probes is in Table 19. The binding signals are depicted as means of fluorescence intensities of triplicate spots of probe arrayed (with error bars). Each probe was printed in triplicate at two levels: 0.1 (blue bars) and 0.5 (red bars) mg/mL (30 and 150 pg/spot). Quantified fluorescence intensity is plotted on the x-axis. Carbohydrate probes are plotted on the y-axis. Fluorescene intensity signals of the bars marked with an asterisk in the Anti-heteromann graph are not considered due to contamination of this microarray experiment with the mAb anti-(β1→3, β1→4)-D-glucan during the binding assay.

The presence of mixed-linked (β1→3, β1→4)- D-glucans in *N. oculata* constitutes an interesting finding. This is particularly important because these polymers are considered as dietary fibres or may work as prebiotics (30), making *N. oculata* a potential source of these compounds to employ in various areas, such as in food or medical applications.

3.1.3.3. Microbial carbohydrate-binding modules (CBMs)

- *CmCBM6-2*

*Cm*CBM6-2 bound strongly to mixed-linked (β1→3, β1→4)- D-glucan from Oat (Fig. 21 and 23), in agreement with the binding reported for this CBM (92,149) and for the family 6 CBM (150). *Cm*CBM6-2 contains two binding clefts (A and B) with different carbohydrate specificities. Cleft A is able to recognise (β1→3)- and (β1→4)-Glc oligosaccharides, whereas cleft B recognises (β1→4)- and mixed (β1→3, β1→4)-linked Glc oligosaccharides (92). In addition, one of these clefts, the cleft A, is also responsible for recognise (β1→4)-xylans (149,150).

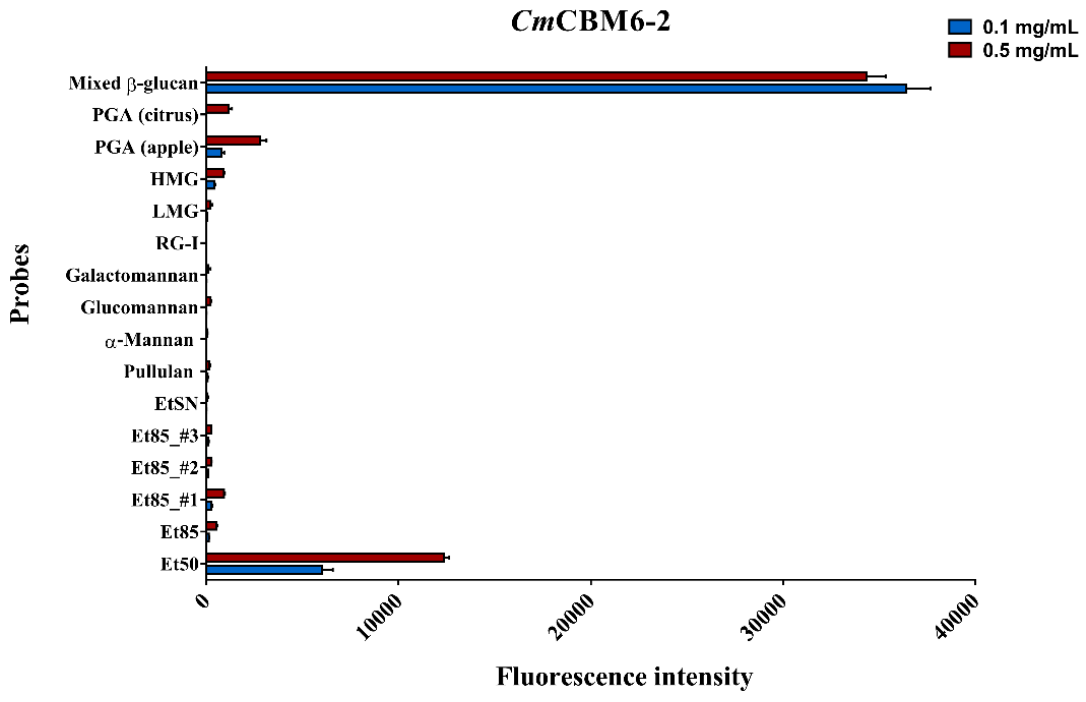


Figure 21. Screening of the microarray developed with Microalgae PS set 1 probes by analysis with the family 6 CBM CBM6-2 from *Cellvibrio mixtus* (*Cm*CBM6-2). Cabohydrate-binding specificity of *Cm*CBM6-2 is in Table 6, and the carbohydrate sequence information of the probes is in Table 19. The binding signals are depicted as means of fluorescence intensities of triplicate spots of probe arrayed (with error bars). Each probe was printed in triplicate at two levels: 0.1 (blue bars) and 0.5 (red bars) mg (dw)/mL (30 and 150 pg/spot). Quantified fluorescence intensity is plotted on the x-axis. Carbohydrate probes are plotted on the y-axis.

Relatively to the PS fractions from *N. oculata*, it was observed that *Cm*CBM6-2 bound Et50, being the binding stronger at 0.5 mg/mL (Fig. 21 and 23). As a result, the interaction with the *Cm*CBM6-2 suggests that the high amount of 4-Glc previously found in Et50 (section 3.1.1.3.) was present in β-configuration, as well as the other Glc residues, like 3-Glc, confirming the results obtained with the mAb anti-(β1→3, β1→4)-D-glucan. In addition, this may also indicate the presence of mixed-linked (β1→3, β1→4)-D-glucan in *N. oculata*, which is supported by the results obtained before with the mAb anti-(β1→3, β1→4)-D-glucan, since this mAb has strict specificity to (β1→3, β1→4)-D-glucan, unlike to *Cm*CBM6-2 that has a broad specificity to β-linked Glc residues.

- *CBMs belonging to B. thetaiotaomicron*

The protein BT3698, which comprised as cloned domain the CBM58 of a SusG protein, bound only to the linear (α1→4, α1→6)-D-glucan polymer (pullulan) (Fig. 22 and 23), in accordance with the specificity reported for this protein (110,112). Structural studies showed that the CBM58 of the SusG protein is able to bind maltooligosaccharides (composed of Glc(α1→4)Glc) (110), explaining the interaction of BT3698 to pullulan. Furthermore, in *B. thetaiotaomicron* was observed the expression of the PUL that encodes BT3698 in response to their growing in a medium with the starch-like molecule pullulan, suggesting its ability to use starch (112).

Additionally, no interactions were found between BT3698 and the PS fractions from *N. oculata* (Fig. 22 and 23). This allows to ascertain the absence of starch-like molecules in the HW soluble PS of *N. oculata*, in other words, sequences of (α1→4)-D-linked Glc residues with units linked of (α1→6)-D-Glc. Moreover, this result supports the previous observed with the mAb anti-(β1→3, β1→4)-D-glucan and the *Cm*CBM6-2 for the fraction Et50, where it was found that Glc was in β-configuration.

In respect to BT0996, which comprised as cloned domain the CBM57 appended to the GH2 whose carbohydrate-binding specificity is unknown, it was verified its preferential binding towards LMG and PGA from citrus probes, both printed at high concentration (Fig. 22 and 23). Moreover, no interaction was observed between BT0996 and HMG that differs from the LMG in the DE, suggesting that BT0996 needs the carboxylate groups of the (α1→4)-linked GalA residues without esterification to interact with these probes. Interestingly, BT0996 bound PGA from citrus but did not bind PGA from apple, despite both PGAs contain an identical sugar composition (sugar composition in supplementary

information 2). This difference probably derives from different DEs, as verified for HMG and LMG probes. As mentioned previously for RCA_{120}, this may be due to the extraction of PGAs in different stages of ripening, since pectins from a ripe fruit exhibit a lower DE, a lower *Mw*, and a decrease in NS content compared to pectins in unripe fruits (148).

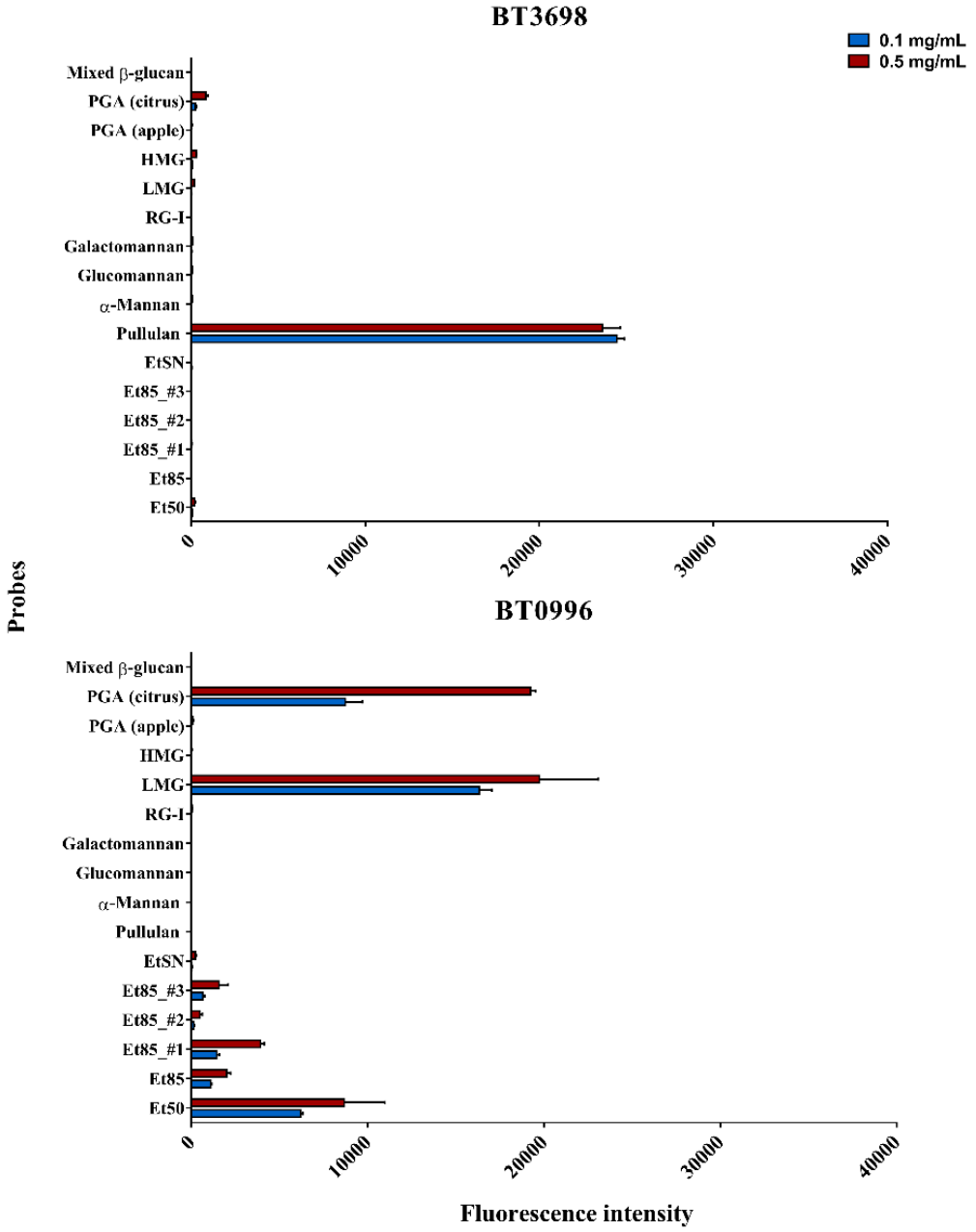


Figure 22. Analysis of the carbohydrate microarrays developed with Microalgae PS set 1 probes with the CBM domains of BT3698 and BT0996, from the bacterium *Bacteroides thetaiotaomicron*. Carbohydrate-binding specificity of each protein is in Table 6, and carbohydrate sequence information of the probes included in the microarray is in Table 19. The binding signals are depicted as means of fluorescence intensities of triplicate spots of probe arrayed (with error bars). Each probe was printed in triplicate at two levels: 0.1 (blue bars) and 0.5 (red bars) mg (dw)/mL (30 and 150 pg/spot). Quantified fluorescence intensity is plotted on the y-axis. Carbohydrate probes are plotted on the x-axis.

Although the specificity of the BT0996 towards carbohydrates still remains understudied, Martens *et al.* (112) reported the expression of the PUL that encodes BT0996 in response to the pectins rhamnogalacturonan I and II (RG-I and RG-II), without its expression in presence of pectic galactans and homogalacturonnans (HGs). This proposes the ability of this protein to recognise RG-I and RG-II, but their specific epitope is unknown.

RG-I comprises a backbone composed of the repeating disaccharide units →4)-D-GalA-(α1→2)-L-Rha-(α1→. The GalA residues normally are not substituted, being the Rha residues the branching points found in RG-I. The Rha residues tend to attach neutral sugars in its side chains, like galactans, arabinans, arabinogalactan-I or arabinogalacta-II. Otherwise, RG-II has an extremely complex structure not related to the RG-I (148). RG-II contains a high proportion of Rha residues occurring as terminal, (1→3), and branched (1→2,3,4) units. In addition, RG-II contains a HG backbone composed of (α1→4)-linked D-GalA with side chains that include rare monosaccharides, like apiose, 2-*O*-methyl-α-D-Xyl, aceric acid, and others (148). As a result, this suggests that the PUL encoding BT0996 is expressed in response to a common epitope present in RG-I and RG-II that does not occur in pectin galactan and HG. Nevertheless, it is difficult to know the exactly epitope due to the high complexity of RG-II.

Considering the absence of Rha residues in galactan and HG, and that these are present in RG-I and RG-II, probably the epitope recognised by BT0996 are the Rha residues. Additionaly, taking in consideration that the LMG and PGA from citrus probes are pectins extracted from natural sources, these may contain some residual amounts of other sugars found in pectins (RG-I and RG-II), such as Rha and Ara. Effectively, these sugars were found as components of its carbohydrate composition (supplementary information 4), which may explain the interaction found with BT0996 (Fig. 22 and 23) in addition to its low DE. Moreover, this further suggests that BT0996 might recognise Rha residues, since these are common among RG-I, RG-II, LMG and PGA from citrus.

Recently, it was showed that BT0996 contains two distinct catalytic domains that contribute to degradation of RG-II through targeting its side chains, which is important since highligths that BT0996 is able to degrade the most complex PS in nature. This is particularly interesting since allows to consider that BT0996 is probably able to recognise the high complex PS present in microalgae. One of the catalytic domains in BT996 is a β-D-glucuronidase (GH2) that hydrolysis between GlcA(β1→4)Fuc, and the other one is a β-L-

arabinofuranosidase (GH137) that hydrolysis the terminal Ara*f*(β1→2)Rha (114). This reinforces that BT0996 should bind some carbohydrate present in RG-II, directing the catalytic domain to perform its function, but the specific epitope remains unknown.

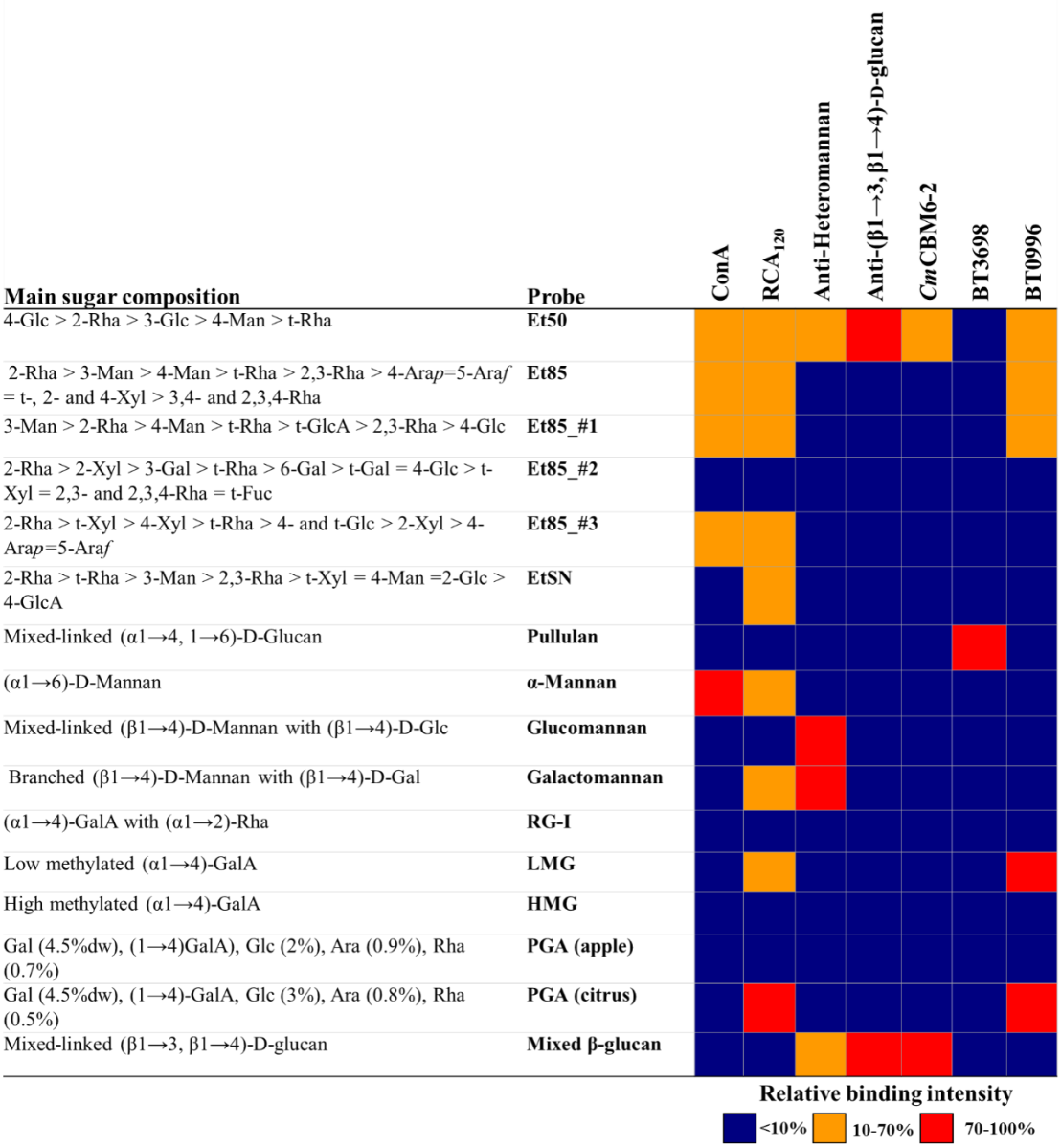


Figure 23. Comparison of the relative binding intensities of the proteins investigated in carbohydrate microarrays: ConA, RCA_{120}, Anti-heteromannan (Rat IgM), Anti-(β1→3, β1→4)-D-glucan, *Cm*CBM6-2, BT3698 and BT0996. The heat map representation shows the different polysaccharides (PS) binding patterns revealed by the microarray analysis containing the Microalgae PS set 1 (Table 19). The relative binding intensities were calculated as the percentage of the mean fluorescence signal intensity at 0.5 mg/mL PS probe given by the probe most strongly bound by each protein (normalized as 100%).

Besides interaction with LMG and PGA (citrus), BT0996 exhibited strong interaction with the fraction Et50 from *N. oculata* (Fig. 22 and 23). Lower interactions were also observed with the fractions Et85_#1, Et85 and Et85_#3, especially when printed at high

concentration (Fig. 22 and 23). This showed a binding pattern of BT0996 among LMG, PGA from citrus, Et50, Et85_#1, Et85 and Et85_#3 probes. Considering that the fraction Et50 was predominantly composed of (1→4)-linked Glc, probably as backbone, and the fractions Et85, Et85_#1 and Et85_#3 were apparently different from Et50, as well as from the LMG and PGA (citrus), the interaction with BT0996 should be due to the recognition of one specific epitope or common branching point present in these PS.

Looking at the previous structural analysis of the PS fractions from *N. oculata* (section 3.1.1.), the major common residue among Et50, Et85_#1, Et85, and Et85_#3 that is also found as component of RG-I and RG-II, and was also present in LMG and PGA from citrus probes (supplementary table 5), was the (1→2)-Rha. This probably indicates that (1→2)-Rha may be one sugar residue recognised by BT0996.

Remarkably, 2-Rha was found in higher amount in the Et85 fractions than in Et50 but the interaction observed was higher for Et50 probe. A possible explanation for this event may be due to the higher *Mw* of the PS in Et50 (630 kDa) compared to the *Mw* of the polymers in the Et85 fractions. The high *Mw* promotes a better adsorption of the PS on the surface of the slide, as verified by Wang *et al.* (60). This allows to exhibit more epitopes to the protein analysed, and consequently a higher binding between the protein and PS, since its epitope should be more available. Furthermore, although Et85_#2 and EtSN seem identical to the other Et85 fractions in terms of neutral sugars composition and also contain on its constitution 2-Rha, no interaction with these probes were observed. The result may be due to the apparently higher content in substituted Rha residues of Et85_#2, and in particular relatively to (1→2)-Rha, which highlighted a higher DB of the PS in this fraction that may restrict the access of the CBM domain to its recognition epitope. Regarding to EtSN, the absence of interaction may be due to the low *Mw* of their PS. One possible approach that could be used to obtain a better adsorption of the PS with low *Mw* at the surface of the slide is their conjugation with hydrophobic tags, like lipids, forming neoglycolipids (63,64). Another possibility would be to increase the probes concentration.

The results highlighted that the PS fractions of *N. oculata* did not comprise starch on its composition, as found from the analysis with BT3698. The fractions Et50, Et85_#1, Et85_#3 and Et85 might share an identical epitope that is recognised by BT0996. Thus, the PS fractions from *N. oculata* may be probably metabolized in the human gut by *B. thetaiotaomicron* through the involvement of BT0996.

In summary, these results allowed to verify that the high amount of Glc previously found in the Et50 fraction might occur in β-configuration, due to its interaction with the *Cm*CBM6-2 and the mAb anti-(β1→3, β1→4)-D-glucan. Moreover, the interaction with the last protein indicated the presence of domains composed of mixed-linked (β1→3, β1→4)-glucans in the Et50 fraction of *N. oculata*, since this protein has strict specificity. The presence of these sort of PS in Et50 is particularly important since they may work as dietary fibres or prebiotics. In addition, it was proposed that the Man present especially in the Et85 fractions might occur in α-configuration due to the interaction with ConA and absence of interaction with the mAb anti-heteromonnan. However, in order to better analyse the results with ConA the experimental conditions should be optimized.

The analysis with two proteins from *B. thetaiotaomicron*, BT3698 (CBM58) and BT0996 (CBM57), showed that the PS fractions from *N. oculata* did not comprise starch-like molecules, due to the absence of interaction with BT3698. Moreover, the PS present in Et50, Et85_#1, Et85 and Et85_#3 should share a common epitope that it is recognised by BT0996. Thus, the carbohydrate microarrays developed with the PS fractions from *N. oculata* allowed to get more knowledge about its structural characteristics and to understand that these PS are able to be recognised by *B. thetaiotaomicron*, in particular by BT0996 that is able to recognise one of the most complexes PS in nature, the RG-II. From this work, it was possible emphasising even more the ability of BT0996 to recognise very complex PS, due to its interaction with the PS of *N. oculata*, which displayed a structure very complex. Finally, this result proposes that some of the HW soluble PS present in *N. oculata* probably can be metabolized in the human gut.

3.2. Characterization of glycolipids from *N. oculata*

During sequential extraction of polysaccharides from *N. oculata*, it was observed the loss of Gal from their biomass with defatting using $CHCl_3$:MeOH (2:1, v/v). It has been reported the presence of glycolipids (GL) in photosynthetic organelle membranes, especially as galactolipids and sulfolipids (31). This may justify the loss of Gal, as a consequence of the presence of galactolipids in *N. oculata*, which were extracted together with other lipid compounds. Interestingly, GL from microalgae have been found to show various biological activities, such as antiviral, antitumor, anti-proliferative, among others (8). Their mechanism of action is not only due to their acyl chains but also due to the glycosyl portion (35), being so important to study possible carbohydrate-protein interactions in which they can be involved. As a result, in order to verify if the loss of Gal was due to the presence of galactolipids, the lipid fraction of *N. oculata* was extracted and characterized in terms of carbohydrates. This section describes the carbohydrate features of the GL present in the lipid fraction of *N. oculata*.

3.2.1. Carbohydrate characterization of the lipid fraction

N. oculata comprised 29 % dw of lipid compounds extracted with $CHCl_3$:MeOH (2:1, v/v). The result is in agreement with the reported in literature, since the lipid content in *N. oculata* may vary between 16 and 32 % dw but was found in higher content in other studies and its content may depends on the growth conditions (12,43,151,152). As a result, this work showed that biomass of *N. oculata* contained 19 % dw of carbohydrates, 34 % dw of protein, and 29 % dw of lipid compounds.

The carbohydrate composition of the lipid fraction of *N. oculata* is shown in table 20. The results showed that the lipid fraction comprised as main neutral sugar (NS) Man (7.7 weight % of lipid fraction), followed by Gal (2.5 %) and contained a residual amount of Glc (0.3 %). The NS accounted a total of 11 % of the composition of the lipid fraction, which represents 3 % dw in the biomass of *N. oculata*. In addition, the lipid fraction also contained sulfate groups in about 4 % of the lipid fraction. The low content in Glc together with the presence of sulfate groups proposes the presence of the sulfolipid SQDG in *N. oculata*, as found in other studies (45,49,53–55). SQDG is a negatively charged glycolipid composed of

a monoglycosyl diacylglycerol with a sulfonic acid linked at position 6 of the monosaccharide moiety (1,2-diacyl-3-*O*-(6-sulfo-6-deoxy-α-D-glucosyl)-*sn*-glycerol) (8).

As initially found in section 3.1.1.1., after defatting treatment of *N. oculata* the content in Gal decreased from 2 % dw in biomass to 0.22 % dw, suggesting that 1.88 % dw of Gal remained in the lipid fraction. From the carbohydrate analysis of lipid fraction, it was found that the Gal represents in about 0.73 % dw in the biomass of *N. oculata*, suggesting that the Gal was extracted together with the lipids, justifying the decrease in Gal during the defatting treatment. Thus, the result suggests the presence of galactolipids in *N. oculata*, as reported in various studies (45,49,53–55).

Table 20. Composition in neutral sugars (NS), alditols and sulfate (SO_3^-) groups of the lipid fraction extracted with $CHCl_3$:MeOH (2:1, v/v) from *N. oculata*, and the lipid fraction after being dialysed. Results are expressed in weight percentage (%) of the lipid fraction.

	(mean ± SD) %*					
	Man	Gal	Glc	**NS**	**SO_3^-**	**Free Mannitol**
Lipid fraction	7.7 ± 0.2^a	2.5 ± 0.1^a	0.3 ± 0.2^a	10.6 ± 0.3^a	3.8 ± 0.2	11.5 ± 0.8
Dialysed lipid fraction	$-^b$	2.8 ± 1.0^a	0.3 ± 0.2^a	3.2 ± 1.1^b	N.D.	N.D.

*- Data represent the mean values ± SD in weight % of the lipid fraction from duplicate (n=2), n=6, and n=2 analyses for NS, SO_3^-, and free mannitol, respectively.
N.D. – Not determined.
Different superscripts in same column indicate values significantly different ($p<0.05$).

The lipid fraction comprised a higher content in Man than in Gal (Table 20). This result was unexpected because it had not been observed the loss of Man during defatting step of *N. oculata* biomass in the sequential extraction of PS (section 3.1.1.1.). However, it was verified that *N. oculata* contained free mannitol in its composition, which remained as component of the filtrate fraction (supernatant with *Mw* under 10 kDa) obtained during the HW extraction, representing *ca*. 3.9 % dw of *N. oculata* biomass. Considering this result, the lipid fraction was analysed to free alditols in order to understand if this high content in Man was derived from mannitol.

The free alditols analysis showed the presence of mannitol in about 12 % as component of the lipid fraction (Table 20), which represents 3.3 % dw of *N. oculata* biomass. The result suggests that the high content in Man present in the lipid fraction was due to the extraction of mannitol along with the lipids, which previous remained in the filtrate fraction (3.9 % dw). This likely occurred due to the presence of MeOH in the mixture of solvents used in this study to extract the lipid compounds, which is able to increase the polar character of the

solution, allowing to extract not only GL but also mono- or oligosaccharides. Thus, in addition to galactolipids, $CHCl_3$:MeOH (2:1, v/v) was able to extract mannitol that is found in the cytosol of microalgae and may work as an osmotic regulator (130). However, the reason for the presence of mannitol in the lipid fraction during the extraction of lipids and its presence in filtrate during the sequential extraction of PS is unclear. A possible explanation may be due to the different amount of biomass used to isolate the PS and the GL. The biomass used to study the PS was much higher (40 g) than the biomass used to extract the GL (about 40 mg each replica), which might prevent the extraction of mannitol with the lipid compounds during defatting.

In order to remove the mannitol from the lipid fraction and obtain a lipid fraction rich in galactolipids, the lipid fraction was re-suspended in $CHCl_3$:MeOH (2:1, v/v) and dialysed against 50 % EtOH:H_20 (v/v) (cut-off 12-14 kDa). The yield of dialysis was 43 % (w/w), giving a fraction named as dialysed lipid fraction. The carbohydrate composition of the dialysed lipid fraction is shown in table 20 and their proportion in carbohydrates in figure 24.

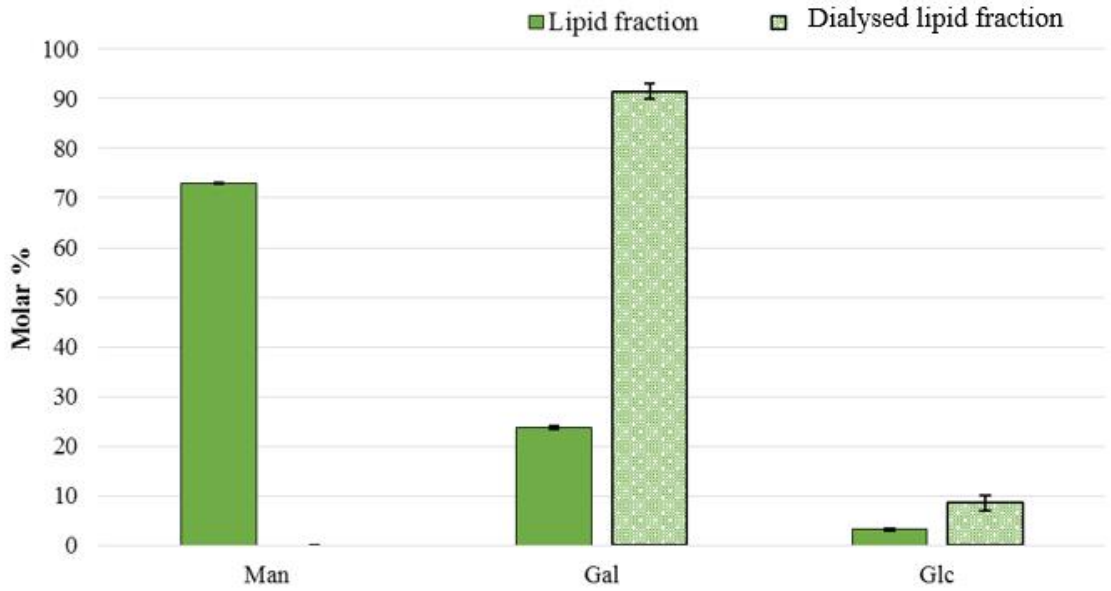


Figure 24. Monosaccharides composition of the lipid fraction extracted with $CHCl_3$:MeOH (2:1, v/v) from *N. oculata* before (Lipid fraction) and after dialysis (Dialysed lipid fraction). Results are expressed in molar percentage (mol %) of the total content in sugars.

The dialysed lipid fraction comprised only Gal (3 % of the lipid fraction) and Glc (0.3 %) as monosaccharides, being found in identical content than in the initial lipid fraction,

since the results were not significantly different ($p<0.05$) (Table 20). Noteworthy, the lipid fraction dialysed did not contain Man, and consequently did not contain mannitol, being 3.2 % the total content in sugars in the dialysed lipid fraction (Table 20).

Comparing the lipid fraction before and after dialysis, it was verified that carbohydrates were dominated by Gal (91 mol %) in dialysed lipid fraction, containing also Glc (9 %) (Fig. 24). Unlike, the carbohydrates in initial lipid fraction were governed by Man (73 %), followed by Gal (23 %) and Glc (3 %). Therefore, the dialysis allowed to remove the mannitol from the lipid fraction and retain the Gal and Glc, suggesting that these monosaccharides are part of GL.

Glycosidic linkages analysis allowed to verify the way that the sugars were linked in the GL present in the lipidic fraction, these results are shown in table 21. Gal occurred mostly as terminal t-Gal (63 mol %) and 6-Gal (25 %), being also found residual amounts of 4-Gal (1 %). In addition, the lipid fraction comprised in about 2 % of 6-Glc, which indicates that Glc occurred in GL linked/substituted at position 6. This was probably due to the linked of sulfonic acid at the position 6 of Glc moiety, confirming the proposed above for the presence of SQDG in *N. oculata,* as reported in other studies (45,49,53–55).

Table 21. Glycosyl linkage composition (mol % of the total neutral sugars) of the lipid fraction extracted with $CHCl_3$:MeOH (2:1, v/v) from the biomass of *N. oculata*.

Glycosidic linkage*	**Molar %**
t-Gal	63.1
4-Gal	1.1
6-Gal	25.0
Total	89.2
t-Glc	0.9
6-Glc	2.4
Total	3.3

*- Other linkages were found in minor amount.

Considering that MGDG contains one Gal in β-configuration linked to the *sn*-3 position of the glycerol and DGDG contains a disaccharide of Gal composed of a terminal α-Gal residue (1→6) linked to the inner β-Gal (8), this indicates that the 6-Gal determined through glycosidic linkages analysis is derived from DGDG, meaning that *ca* 28 % of the total Gal was found in DGDG. Furthermore, each disaccharide of Gal contains one 6-Gal moiety and one t-Gal moiety (in same amount). Thus, the remaining t-Gal is derived from

MGDG. Consequently, this suggests that Gal was found preferentially as MGDG in galactolipids, representing *ca* 43 %.

The glycosidic linkages analysis of the lipid fraction confirmed that the GL present in the lipid fraction of *N. oculata* occurred mainly as galactolipids, preferentially as MGDG (43 %) and DGDG (28 %), and contained also the sulfolipid SQDG, corroborating the mentioned by other authors (45,49,53–55). The presence of SQDG and galactolipids in *N. oculata* is particularly relevant because these type of GL are able to mediate various biological activities. As example, SQDG from *Porphyridium sp*. displayed anti-proliferative, antioxidant, and anti-inflammatory activities (42), and MGDG and DGDG from *Nannochloropsis granulata* showed a potential anti-inflammatory activity (40). Therefore, the lipid fraction of *N. oculata* seems a source of GL with potential to employ in several areas, including in food, due to the properties that they can exhibit. However, the direct carbohydrate-protein interactions that these compounds may be involved still remains unclear, being important perform research on this field.

These results highlighted some structural characteristics of the carbohydrates found as component of the GL in the lipid fraction of *N. oculata*, suggesting the presence of MGDG, DGDG and SQDG in their composition. This constitutes a functional starting point to develop carbohydrate microarrays with GL from natural sources. Moreover, this allows to select proteins with well-defined specificities that may recognise these structures. These proteins may not only confirm the presence of GL as components of the lipid fraction of *N. oculata* but also search for carbohydrate-protein interaction that they can manage.

3.2.2. Carbohydrate microarray developed with the dialysed lipid fraction extracted from *N. oculata*

The dialysed lipid fraction extracted from *N. oculata* was used as liposome probes to develop carbohydrate microarrays, being this experiment referred as Microalgae GL set 2. Liposomes were formed by two carrier lipids (1,2-Diacyl-*sn*-glycero-3-phosphocholine (lecithin from egg yolk) *(*PC1) or 1,2-Di-*O*-hexadecyl-*sn*-glycero-3-phosphocholine (PC2) and cholesterol (C)), whose chemical structures are shown in supplementary figure 5, and the dialysed lipid fraction extracted from *N. oculata*, named as GL.

Microarrays were probed with two lectins, the plant lectin RCA_{120} and the animal lectin hGalectin-3. The two lectins used have already well-known carbohydrate specificities towards Gal residues, which were found as components of the GL in *N. oculata*, probably as MGDG and DGDG. With this in mind, RCA_{120} will allow to understand simultaneously if the GL from *N. oculata* comprised galactolipids in its composition and consequently if they are suitable natural probes to be efficiently printed on the slide surface of the microarrays. In addition, hGalectin-3 was used in this study to understand if the glysosyl portion of the galactolipids are able to mediate the immunologic system through this lectin in particular.

The liposome probes used in microarrays are described in table 22 and the carbohydrate-binding features of each protein analysed in these experiments are described in table 8.

Table 22. List of all liposome probes that composed the Microalgae glycolipids (GL) set 2. The microarray comprised liposomes composed of the GL present in the dialysed lipid fraction from *N. oculata*, one phosphatidylcholine (PC), 1,2-Diacyl-*sn*-glycero-3-phosphocholine (PC1) or 1,2-Di-*O*-hexadecyl-*sn*-glycero-3-phosphocholine (PC2), and the cholesterol (C). Composition of the probes differed in the mol ratio PC:C:GL. The main sugar composition is depicted for each probe.

Liposome Probes (Set 2)	**Ratio PC:C:GL (pmol)**	**Liposome composition**	**Main sugar composition**
1	1000:1000:150	PC1:C:GL	t-Gal (43 %) Gal(1→6)Gal (28 mol %)
2	1000:1000:50		
3	1000:1000:0		-
4	1000:1000:150	PC2:C:GL	t-Gal (43 %) Gal(1→6)Gal (28 mol %)
5	1000:1000:50		
6	1000:1000:0		-
7	500:300:150	PC1:C:GL	t-Gal (43 %) Gal(1→6)Gal (28 mol %)
8	500:300:50		
9	500:300:0		-
10	500:300:150	PC2:C:GL	t-Gal (43 %) Gal(1→6)Gal (28 mol %)
11	500:300:50		
12	500:300:0		-
13	0:0:150	PC:C:GL	t-Gal (43 %) Gal(1→6)Gal (28 mol %)
14	0:0:50		

3.2.2.1. RCA_{120}

The carbohydrate microarrays probed with RCA_{120} showed the preferential binding of this protein to liposome probes comprising the GL at the highest level (150 pmol) in presence of the PC1 carrier lipid (PC1:C:GL, 500:300:150 and 1000:1000:150) (Fig. 25). The result suggests that the lipid fraction dialysed from *N. oculata* referred to as GL contained in its composition β-Gal residues, due to the specificity of RCA_{120} to bind terminal β-Gal moieties (61,62,85). This indicates that the GL comprises galactolipids, confirming the previously proposed (section 3.2.1). Interestingly, the binding was stronger to liposome probes comprising a lower amount in PC1 and C (500:300:150, PC1:C:GL) than those with higher amount in PC1 and C (1000:1000:150, PC1:C:GL). The result suggests that the high content in PC1 and C may restrict the presentation of the glycosyl portion of the galactolipids in the liposomes to the protein.

Although RCA_{120} bound mostly liposome probes containing PC1, it was also verified fluorescence signals derived from the binding between RCA_{120} and the liposome probes formed with the carrier lipid PC2 in a ratio of 500:300:150, and between RCA_{120} and the liposomes using only the GL at the highest level (0:0:150) (Fig. 25). The last result suggests that the GL in the lipid fraction dialysed from *N. oculata* can be efficiently non-covalently immobilized onto nitrocellulose-coated glass slides without using carrier lipids. In addition, the difference in the fluorescence intensities obtained between the probes comprising the carrier lipid PC1 and PC2 may be due to the difference on its acyl chains composition, since PC1 has unsaturated acyl chains whereas the acyl chains in PC2 are saturated. This promotes a different mobility/fluidity of the liposome membrane, which may influence the presentation of the GL to the protein. Consequently, the result suggests that PC1 improves the interaction between the protein and the glysosyl portion of the galactolipids likely because of the presence of unsaturated acyl chains, which may increase the membrane mobility.

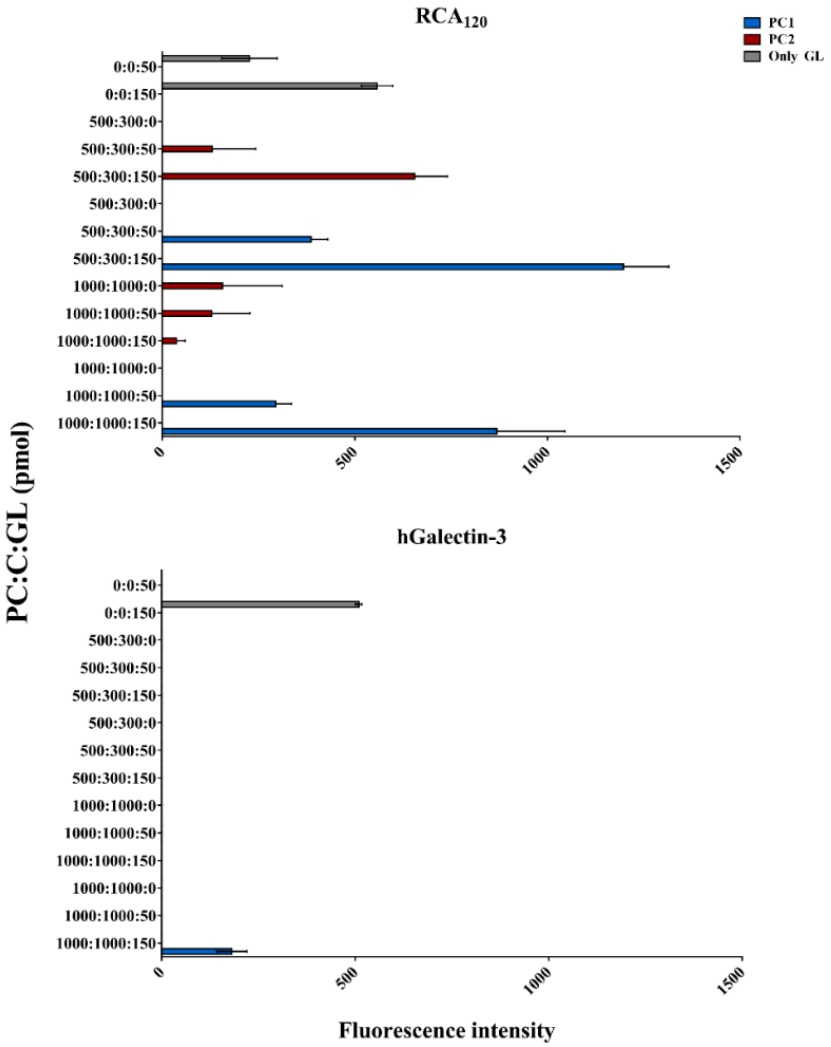


Figure 25. Screening of the carbohydrate microarrays developed with Microalgae GL set 2 probes, which comprise the dialysed lipid fraction (GL) from *N. oculata,* by analysis with two lectins, the plant lectin RCA_{120} and the animal lectin hGalectin-3. Protein names are depicted on top of each graph and the specificity of each protein towards Gal residues is described in table 8. Carbohydrate composition of the liposome probes included in the microarray is in table 22. Binding signals are depicted as means of fluorescence intensities of triplicate spots of probe arrayed (with error bars). Quantified fluorescence intensity is plotted on the x-axis. Liposome probes comprising the GL are plotted on the y-axis.

3.2.2.2. hGalectin-3

From analysis of the carbohydrate microarray experiment with hGalectin-3 it was not identified interaction with any liposome probe (Fig. 25). Although the analysis with RCA_{120} indicated the presence of β-Gal residues as component of the GL in the lipid fraction of *N. oculata*, these residues were not bound by hGalectin-3, which is reported to bind Gal-rich glycans (88,97). The result may be due to the fact that hGalectin-3 tends to recognise β-Gal residues linked to GlcNAc, like Gal(β1→4/3)GlcNAc, explaining the absence of interaction

with the galactolipids from *N. oculata*. Additionally, the result suggests that the galactolipids of *N. oculata* will not promote any immunologic response activated by hGalectin-3, since galectins are a family of glycan-binding proteins implicated in immune regulation (97). This is particularly relevant if these compounds are thinking to employ in human nutrition or biomedical applications.

In summary, the results obtained from the carbohydrate microarrays developed with the dialysed lipid fraction extracted from *N. oculata* allowed to confirm the presence of galactolipids in its composition, due to the interaction with RCA_{120}. The presence of galactolipids in *N. oculata* is particularly important since it has been assigned to the galactolipids from marine microalgae various biological activities. In addition, the result obtained with RCA_{120} allowed to verify that the dialysed lipid fraction from *N. oculata*, which comprises GL, are a suitable natural source of liposome probes to develop carbohydrate microarrays. This will allow to further study carbohydrate-protein interactions in which the GL can be involved. Finally, it was highlighted that the Gal residues of the galactolipids were not recognised by hGalectin-3, suggesting that galactolipids are not able to mediate the immune system by action of this galectin.

Chapter 4. – Conclusions and future perspectives

Conclusion and Future work/perspectives

The work performed in this thesis was centred in the study of the carbohydrates of the marine microalga *Nannochloropsis oculata*, and in particular in their polysaccharides (PS) and glycolipids (GL). The PS and GL were studied through carbohydrate analyses and carbohydrate microarrays.

The sequential procedure employed to extract and study the PS of *N. oculata* allowed to verify that the carbohydrates present in the crude biomass of *N. oculata* accounted to 19 % dw and were composed mainly of Glc, Man and Gal, containing also lower amounts of Rha, Rib and Xyl, and acidic sugars as uronic acids (UA). It was also found on its composition sulfate groups. The defatted biomass comprised an identical content and composition in carbohydrates but it was verified the loss of galactose, due to the presence of galactolipids that remained in the lipid fraction.

The hot-water (HW) extraction employed to the defatted biomass with subsequent ultrafiltration (cut-off 10 kDa) allowed to obtain one residue, one filtrate (*Mw* < 10 kDa) and one retentate (*Mw* > 10 kDa) that encompassed the HW soluble PS. Carbohydrate analyses of these fractions highlighted that the residue was rich in (1→4)-linked Glc moiety, which was assigned to the presence of a cell-wall rich in cellulose in *N. oculata*. Moreover, it was verified that the filtrate comprised a high content in free mannitol, suggesting a possible role of mannitol as an osmotic regulator in *N. oculata*, as reported for other green marine microalgae. Regarding to the retentate, a fraction that retained the HW soluble PS, it was composed mainly of Rib, Rha, Glc, and Man, comprising also minor amounts of Xyl, Fuc and Gal. Besides this sugars, it was also verified the presence of UA and sulfate groups, suggesting that the HW soluble PS in *N. oculata* are anionic sulfated PS.

The sequential ethanol (EtOH) precipitation with an increase in EtOH concentration (50 % and 85 % EtOH v/v) employed to the retentate yielded three fractions enriched in PS: Et50, Et85 and EtSN, which were deeply studied in terms of structural characteristics of their PS. PS in fraction Et50 had a *Mw* of 630 kDa in average and showed as main monosaccharides Rib and Glc. Glc was found mainly as 4-Glc (60 mol % of total PS), suggesting the presence of a backbone composed of (1→4)-glucan, followed by minor amounts of 3-Glc (6 %) and the acidic sugar 4-GclA (2 %), which may be also part of linear chains and/or may occur inside the backbone. The presence of 3-linked Glc suggested the presence of mixed glucans, like (β1→3, β1→4)-glucans, which were confirmed from the

microarray probed with the mAb anti-(β1→3, β1→4)-D-glucans, since this has strict specificity to (β1→3, β1→4)-glucans. Furthermore, carbohydrate microarrays showed that the Glc found in Et50 might be present in β-configuration not only due to the interaction with the mAb anti-(β1→3, β1→4)-D-glucans, but also with the CBM *Cm*CBM6-2. In particular, the finding of the presence of mixed glucans in *N. oculata* is important because these PS may work as dietary fibres or prebiotics. Additionally, some UA (10 %) and sulfate groups (4 %) were found as components of the PS in Et50, as well as other sugar residues like (1→2)-Rha.

Et85 comprised two populations of polymers with 65 and 5 162 kDa, and EtSN one with 26 kDa. Both fractions showed as main sugars Rha and Man, being Rha in higher amount in EtSN (40 mol %) than in Et85 (23 %), followed by Xyl (5 %) and Fuc. Et85 and EtSN both showed the presence of PS composed of (1→2)-Rha with a branching point occurring at position C-3 of Rha. Additionally, EtSN showed a higher content in terminals, especially in t-Rha (13 %), and a lower content in substituted residues, demonstrating that PS were smaller than those in Et85, confirming the verified through GPC analysis, and contained a lower DB.

Man was preferentially found as (1→3)- and (1→4)-Man in Et85 and EtSN fractions but in higher amount in Et85. In particular, the carbohydrate microarrays analysis proposed that Man might be present in α-configuration due to interaction with ConA (highly specific to α-mannan) and absence of interaction with the mAb anti-heteromannan (specific to (β1→4)-mannan).

In Et85 and EtSN, Xyl and Fuc were found preferentially as terminals, proposing that they were present at the non-reducing end of short polymers/chains or as substituent residues, being also found the 2- and 4-linked Xyl in Et85. Finally, both fractions contained UA in identical content (15 mol %) and sulfate groups. Notable, Et85 had the highest sulfate content (11 %) and it was found that the sulfate group may be present at position C-3, and simultaneously at positions C-3 and C-4 of (1→2)-linked Rha. As a result, Et85 and EtSN fractions seemed to be formed by anionic sulfated heteropolysaccharides due to the presence of UA and sulfate groups and the high diversity in sugars, suggesting a very complex structure of these PS.

The PS in Et85 were further separate through an anion-exchange chromatography in three fractions: Fraction #1, #2 and #3, eluted with 0, 0.125 and 0.250 M NaCl, respectively.

Fraction #1 comprised PS with 19 kDa in average and had the lowest content in UA and sulfate groups. PS in fraction #1 were composed mainly by (1→2)-linked Rha with a branching point at position C-3 of Rha, containing also a high content in terminal Rha residues, showing the presence of short polymers. This fraction retained the Man present in the initial Et85 fraction as (1→3)- and (1→4)-linked Man, indicating that these Man polymers are part of non-anionic and unsulfated PS.

Fraction #2 and #3 comprised two populations of polymers with 448 and 47 kDa, and 339 and 14 kDa, respectively. In terms of sugar composition both fractions were composed mainly by Rha, Xyl, Gal and Fuc, occurring preferentially as terminals, and with high diversity in substituted residues, especially fraction #2. This proposed that the PS in fraction #2 contain a high DB and may show a very complex structure. In addition, fraction #3 had the highest negative character, which was conferred by the presence of UA, which were found as t- and 4-GlcA, and sulfate groups. Thus, the conventional carbohydrate analyses done to the fractions from sequential EtOH precipitation and anion-exchange chromatography showed that *N. oculata* comprises different types of HW soluble PS, particularly between the PS in Et50 from those in Et85 and EtSN, and between fraction #1 from those in fraction #2 and #3.

Regarding to the microarrays experiments developed with the PS fractions extracted from *N. oculata* and analysed with the CBM domains of the two proteins produced by recombinant expression, BT3698 (CBM58) and BT0996 (CBM57), belonging to the human gut bacteria *Bacteroides thetaiotaomicron*, it was highlighted the absence of starch-like molecules in the HW soluble PS, since it was not verified any interaction with BT3698 (specific to starch molecules). BT0996 displayed a binding pattern among the following probes: low methylated galacturonate (LMG), polygalacturonate acid (PGA) from citrus, Et50, Et85_#1, Et85 and Et85_#3. This result suggested that these probes should share a common epitope recognised by BT0996, which was proposed to be the (1→2)-Rha residue. In this study was also proposed that BT0996 requires a low DE of (α1→4)-GlcA to interact with pectins, since no interactions were observed with high methylated galacturonate.

Relatively to the study of GL, it was verified the predominant presence of terminal (63 mol % of the total neutral sugars) and 6-linked (25 %) Gal residues in the lipid fraction of *N. oculata*, which highlighted the presence of galactolipids. These were also confirmed as components of the lipid fraction of *N. oculata* through microarrays, due to the interaction

observed with RCA_{120} that is specific to β-Gal residues. Galactolipids were found preferentially as MGDG followed by DGDG, since 43 mol % of the total Gal (t-Gal) came from MGDG and 28 % (6-Gal) from DGDG.

Additionally, the lipid fraction also comprised the 6-linked Glc moiety and sulfate groups, which were assigned to the presence of the sulfolipid sulfoquinovosyl diacyglycerol (SQDG). The presence of these galactolipids and the SQDG in *N. oculata* is of special interest since these compounds have been found to show various biological activities.

Remarkably, the microarrays developed with the liposome probes comprising the dialysed lipid fraction extracted from *N. oculata* and analysed with RCA_{120} highlighted that the lipid fraction is a suitable natural source of GL probes to develop microarrays, due to the interaction observed with $RCA_{120.}$ As a result, this shows that employing the lipid fraction of *N. oculata* in microarrays is probably a suitable approach that can be used to study carbohydrate-protein interactions determined by the glycosyl portion of GL. Finally, the carbohydrate microarrays experiments printed with the liposome probes showed that the Gal residues of the galactolipids present in *N. oculata* are not recognised by hGalectin-3, suggesting that galactolipids are not able to mediate the immune system through this galectin.

Considering all the results obtained in this thesis, the joint use of carbohydrate analyses and microarrays seems to be a suitable approach not only to unravel the structural characteristics of the PS and the carbohydrates of the GL present in the marine microalga *N. oculata*, but also to demystify molecular interactions determined by these compounds. In particular, this thesis ascertained the presence of mixed-linked (β1→3, β1→4)-glucans in *N. oculata*, which is particularly important since allows to considerer *N. oculata* as a potential source of dietary fibres or prebiotics. Additionally, it was verified that some of the HW soluble PS in *N. oculata* may be anionic sulfated heteropolysaccharides with a very complex structure. It should be emphasized that some of the HW soluble PS may share a specific epitope that is recognised by BT0996, a protein produced by the human gut bacteria *B. thetaiotaomicron*, suggesting that these PS are probably metabolized in the human gut. Finally, this study confirmed the presence of galactolipids as component of *N. oculata*, being their lipid fraction an appropriated natural source of GL probes to develop microarrays.

As future perspectives, it would be interesting to further understand and determine how the sulfate groups are linked to the PS and if they are able to mediate any biological activity. In addition, it would be also interesting to explore the use of *N. oculata* as a source of dietary fibres to develop new food products, since it has already been used in this field but due to their high content in EPA. Finally, an interesting future work it would be to further isolate and characterize the galactolipids in order to do further studies through carbohydrate microarrays and investigate structure-function relationships determined by its carbohydrate portion.

References

References

1. Spolaore P, Joannis-Cassan C, Duran E, and Isambert A. Commercial applications of microalgae. J Biosci Bioeng. 2006; 101:87–96.
2. Odjadjare EC, Mutanda T, and Olaniran AO. Potential biotechnological application of microalgae: a critical review. Crit Rev Biotechnol. 2017; 37:37–52.
3. de Jesus Raposo MF, de Morais AMB, and de Morais RMSC. Marine Polysaccharides from Algae with Potential Biomedical Applications. Mar Drugs. 2015; 13:2967–3028.
4. Raposo MF de J, de Morais RMSC, and Bernardo de Morais AMM. Bioactivity and Applications of Sulphated Polysaccharides from Marine Microalgae. Mar Drugs. 2013; 11:233–52.
5. Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, and Etzler ME. Essentials of Glycobiology. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009.
6. Nelson DL and Cox MM. Lehninger Principles of Biochemistry. 3rd ed. New York, New York and Basingstoke: W. H. Freeman and Company; 2002.
7. Volkaman JK and Brown MR. Nutritional Value of Microalgae and Applications. Algal Cultures, Analogues of blooms and applications. 2005; 407-457.
8. da Costa E, Silva J, Mendonça SH, Abreu MH, and Domingues MR. Lipidomic Approaches towards Deciphering Glycolipids from Microalgae as a Reservoir of Bioactive Lipids. Mar Drugs. 2016; 14:101.
9. Brown MR, Jeffrey SW, Volkman JK, and Dunstan GA. Nutritional properties of microalgae for mariculture. Aquaculture. 1997; 151:315–31.
10. Volkman JK, Brown MR, Dunstan GA, and Jeffrey SW. The Biochemical Composition of Marine Microalgae from the Class Eustigmatophyceae. J Phycol. 1993; 29:69–78.
11. Brown MR. The amino-acid and sugar composition of 16 species of microalgae used in mariculture. J Exp Mar Biol Ecol. 1991; 145:79–99.
12. Matos ÂP, Feller R, Moecke EHS, Oliveira JV de, Junior AF, Derner RB, and Sant'Anna ES. Chemical Characterization of Six Microalgae with Potential Utility for Food Application. J Am Oil Chem Soc. 2016; 93:963–72.
13. Handa N, and Yanagi K. Studies on water-extractable carbohydrates of the particulate matter from the northwest Pacific Ocean. Mar Biol. 1969; 4:197–207.
14. Raposo MF de J, de Morais AMMB, and de Morais RMSC. Influence of sulphate on the composition and antibacterial and antiviral properties of the exopolysaccharide from *Porphyridium cruentum*. Life Sci. 2014; 101:56–63.
15. Yim JH, Son E, Pyo S, and Lee HK. Novel sulfated polysaccharide derived from red-tide microalga *Gyrodinium impudicum* strain KG03 with immunostimulating activity in vivo. Mar Biotechnol. 2005; 7:331–8.
16. Guzmán S, Gato A, Lamela M, Freire-Garabal M, and Calleja JM. Anti-inflammatory and immunomodulatory activities of polysaccharide from *Chlorella stigmatophora* and *Phaeodactylum tricornutum*. Phytother Res. 2003; 17:665–70.
17. Chiovitti A, Molino P, Crawford SA, Teng R, Spurck T, and Wetherbee R. The glucans extracted with warm water from diatoms are mainly derived from intracellular chrysolaminaran and not extracellular polysaccharides. Eur J Phycol. 2004; 39:117–28.
18. Caballero MA, Jallet D, Shi L, Rithner C, Zhang Y, and Peers G. Quantification of chrysolaminarin from the model diatom *Phaeodactylum tricornutum*. Algal Res. 2016; 20:180–8.
19. Abdullahi AS, Underwood GJC, and Gretz MR. Extracellular Matrix Assembly in Diatoms (bacillariophyceae). V. Environmental Effects on Polysaccharide Synthesis in the Model Diatom, *Phaeodactylum Tricornutum*. J Phycol. 2006; 42:363–78.

20. Geresh S, Arad S (Malis), Levy-Ontman O, Zhang W, Tekoah Y, and Glaser R. Isolation and characterization of poly- and oligosaccharides from the red microalga *Porphyridium sp.* Carbohydr Res. 2009; 344:343–9.

21. Sadovskaya I, Souissi A, Souissi S, Grard T, Lencel P, Greene CM, Duin S, Dmitrenok PS, Chizhov AO, Shashkov AS, and Usov AI. Chemical structure and biological activity of a highly branched (1→ 3,1→6)-β-D-glucan from *Isochrysis galban*a. Carbohydr Polym. 2014; 111:139–48.

22. Pieper S, Unterieser I, Mann F, and Mischnick P. A new arabinomannan from the cell wall of the chlorococcal algae *Chlorella vulgaris*. Carbohydr Res. 2012; 352:166–76.

23. Ford CW, and Percival E. 1299. Carbohydrates of *Phaeodactylum tricornutum*. Part II. A sulphated glucuronomannan. J Chem Soc. 1965; 7042–6.

24. Heaney-Kieras J, and J. Chapman D. Structural studies on the extracellular polysaccharide of the red alga, *Porphyridium cruentum*. Carbohydr Res. 1976; 52:169–77.

25. Gloaguen V, Ruiz G, Morvan H, Mouradi-Givernaud A, Maes E, Krausz P, and Strecker G. The extracellular polysaccharide of *Porphyridium sp.*: an NMR study of lithium-resistant oligosaccharidic fragments. Carbohydr Res. 2004; 339(1):97–103.

26. Patel AK, Laroche C, Marcati A, Ursu AV, Jubeau S, Marchal L, Petit E, Djelveh G, and Michaud P. Separation and fractionation of exopolysaccharides from *Porphyridium cruentum*. Bioresour Technol. 2013; 145:345–50.

27. Witvrouw M, Este JA, Mateu MQ, Reymen D, Andrei G, Snoeck R, Ikeda S, Pauwels R, Bianchini NV, Desmyter J, and De Clercq E. Activity of a Sulfated Polysaccharide Extracted from the Red Seaweed *Aghardhiella tenera* against Human Immunodeficiency Virus and Other Enveloped Viruses. Antivir Chem Chemother. 1994; 5:297–303.

28. Qi J, and Kim SM. Characterization and immunomodulatory activities of polysaccharides extracted from green alga *Chlorella ellipsoidea*. Int J Biol Macromol. 2017; 95:106–14.

29. Tabarsa M, Shin I-S, Lee JH, Surayot U, Park W, and You S. An immune-enhancing water-soluble α-glucan from *Chlorella vulgaris* and structural characteristics. Food Sci Biotechnol. 2015; 24:1933–41.

30. de Jesus Raposo MF, de Morais AMMB, and de Morais RMSC. Emergent Sources of Prebiotics: Seaweeds and Microalgae. Mar Drugs. 2016; 14:1-27.

31. Thompson GA. Lipids and membrane function in green algae. Biochim Biophys Acta BBA - Lipids Lipid Metab. 1996; 1302:17–45.

32. Hölzl G, and Dörmann P. Structure and function of glycoglycerolipids in plants and bacteria. Prog Lipid Res. 2007; 46:225–43.

33. Páli T, Garab G, Horváth LI, and Kóta Z. Functional significance of the lipid-protein interface in photosynthetic membranes. Cell Mol Life Sci CMLS. 2003; 60:1591–606.

34. Jones MR. Lipids in photosynthetic reaction centres: Structural roles and functional holes. Prog Lipid Res. 2007; 46:56–87.

35. Zhang J, Li C, Yu G, and Guan H. Total Synthesis and Structure-Activity Relationship of Glycoglycerolipids from Marine Organisms. Mar Drugs. 2014; 12:3634–59.

36. Li S, Xu J, Chen J, Chen J, Zhou C, and Yan X. The major lipid changes of some important diet microalgae during the entire growth phase. Aquaculture. 2014; 428–429:104–10.

37. Kim S-H, Liu K-H, Lee S-Y, Hong S-J, Cho B-K, Lee H, Lee C-G, and Choi H-K. Effects of Light Intensity and Nitrogen Starvation on Glycerolipid, Glycerophospholipid, and Carotenoid Composition in *Dunaliella tertiolecta* Culture. PLOS ONE. 2013; 8:e72415.

38. Martin GJO, Hill DRA, Olmstead ILD, Bergamin A, Shears MJ, Dias DA, Kentish SE, Scales PJ, Botté CY, and Callahan D. Lipid Profile Remodeling in Response to Nitrogen Deprivation in the Microalgae *Chlorella sp.* (Trebouxiophyceae) and *Nannochloropsis sp.* (Eustigmatophyceae). PLOS ONE. 2014; 9:e103389.

39. Morimoto T, Nagatsu A, Murakami N, Sakakibara J, Tokuda H, Nishino H, and Iwashima A. Anti-tumour-promoting glyceroglycolipids from the green alga, *Chlorella vulgaris*. Phytochemistry. 1995; 40:1433–7.

40. Banskota AH, Stefanova R, Gallant P, and McGinn PJ. Mono- and digalactosyldiacylglycerols: potent nitric oxide inhibitors from the marine microalga *Nannochloropsis granulata*. J Appl Phycol. 2012; 25:349–57.

41. Robertson RC, Guihéneuf F, Bahar B, Schmid M, Stengel DB, Fitzgerald GF, Ross RP, and Stanton C. The Anti-Inflammatory Effect of Algae-Derived Lipid Extracts on Lipopolysaccharide (LPS)-Stimulated Human THP-1 Macrophages. Mar Drugs. 2015; 13:5402–24.

42. Bergé JP, Debiton E, Dumay J, Durand P, and Barthomeuf C. In Vitro Anti-inflammatory and Anti-proliferative Activity of Sulfolipids from the Red Alga *Porphyridium cruentum*. J Agric Food Chem. 2002; 50:6227–32.

43. Zhu Y, and Dunford NT. Growth and Biomass Characteristics of *Picochlorum oklahomensis* and *Nannochloropsis oculata*. J Am Oil Chem Soc. 2013; 90:841–9.

44. Andersen RA, Brett RW, Potter D, and Sexton JP. Phylogeny of the Eustigmatophyceae based upon 18s rDNA, with emphasis on *Nannochloropsis*. Protist. 1998; 149:61–74.

45. Kagan ML, Levy A, and Leikin-Frenkel A. Comparative study of tissue deposition of omega-3 fatty acids from polar-lipid rich oil of the microalgae *Nannochloropsis oculata* with krill oil in rats. Food Funct. 2015; 6:185–91.

46. Brown MR, Garland CD, Jeffrey SW, Jameson ID, and Leroi JM. The gross and amino acid compositions of batch and semi-continuous cultures of *Isochrysis sp*. (clone T.ISO), *Pavlova lutheri* and *Nannochloropsis oculata*. J Appl Phycol. 1993; 5:285–96.

47. Jia J, Han D, Gerken HG, Li Y, Sommerfeld M, Hu Q, and Xu Jian. Molecular mechanisms for photosynthetic carbon partitioning into storage neutral lipids in *Nannochloropsis oceanica* under nitrogen-depletion conditions. Algal Res. 2015; 7:66–77.

48. Hafsa MB, Ismail MB, Garrab M, Aly R, Gagnon J, and Naghmouchi K. Antimicrobial, antioxidant1, cytotoxic and anticholinesterase activities of water-soluble polysaccharides extracted from microalgae *Isochrysis galbana* and *Nannochloropsis oculata*. J Serb Chem Soc [Serial on the Internet]. 2017 [cited 2017 Jun 9]; 0. Available from: http://shd-pub.org.rs/index.php/JSCS/article/view/4136

49. Arnold AA, Genard B, Zito F, Tremblay R, Warschawski DE, and Marcotte I. Identification of lipid and saccharide constituents of whole microalgal cells by ^{13}C solid-state NMR. Biochim Biophys Acta BBA - Biomembr. 2015; 1848:369–77.

50. Vieler A, Wu G, Tsai C-H, Bullard B, Cornish AJ, Harvey C, Reca I-B, Thornburg C, Achawanantakun R, Buehl CJ, Campbell MS, Cavalier D, Childs KL, Clark TJ, Deshpande R, Erickson E, Fergunson AA, Handee W, Kong Q, Li X, Liu B, Lundback S, Peng C, Roston RL, Sanjaya, Simpson JP, TerBush A, Warakanont J, Zäuner S, Farre EM, Hegg EL, Jiang N, Kuo M-H, Lu Y, Niyogi KK, Ohlrogge J, Osteryoung KW, Shachar-Hill Y, Sears BB, Sun Y, Takahashi H, Yandell M, Shiu S-H, and Benning C. Genome, Functional Gene Annotation, and Nuclear Transformation of the Heterokont Oleaginous Alga *Nannochloropsis oceanica* CCMP1779. PLOS Genet. 2012; 8:e1003064.

51. Beattie A, Hirst EL, and Percival E. Studies on the metabolism of the Chrysophyceae. Comparative structural investigations on leucosin (chrysolaminarin) separated from diatoms and laminarin from the brown algae. Biochem J. 1961; 79:531–7.

52. Wang D, Ning K, Li J, Hu J, Han D, Wang H, Zeng X, Jing X, Zhou Q, Su X, Chang X, Wang A, Wang W, Jia J, Wei L, Xin Y, Qiao Y, Huang R, Chen J, Han B, Yoon K, Hill, RT, Zohar Y, Chen F, Hu Q, and Xu J. *Nannochloropsis* Genomes Reveal Evolution of Microalgal Oleaginous Traits. PLOS Genet. 2014; 10(1):e1004094.

53. Hodgson PA, Henderson RJ, Sargent JR, and Leftley JW. Patterns of variation in the lipid class and fatty acid composition of *Nannochloropsis oculata* (Eustigmatophyceae) during batch culture. J Appl Phycol. 1991; 3:169–81.

54. Servaes K, Maesen M, Prandi B, Sforza S, and Elst K. Polar Lipid Profile of *Nannochloropsis oculata* Determined Using a Variety of Lipid Extraction Procedures. J Agric Food Chem. 2015; 63:3931–41.

55. Liu P, Corilo YE, and Marshall AG. Polar Lipid Composition of Biodiesel Algae Candidates *Nannochloropsis oculata* and *Haematococcus pluvialis* from Nano Liquid Chromatography Coupled with Negative Electrospray Ionization 14.5 T Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Energy Fuels. 2016; 30:8270–6.

56. Liu Y, Palma AS, and Feizi T. Carbohydrate microarrays: key developments in glycobiology. Biol Chem. 2009; 390:647–656.

57. Liang C-H, and Wu C-Y. Glycan array: a powerful tool for glycomics studies. Expert Rev Proteomics. 2009; 6:631–45.

58. Park S, Lee M-R, and Shin I. Carbohydrate microarrays as powerful tools in studies of carbohydrate-mediated biological processes. Chem Commun. 2008; (37):4389–99.

59. Park S, Gildersleeve JC, Blixt O, and Shin I. Carbohydrate microarrays. Chem Soc Rev. 2013; 42:4310–26.

60. Wang D, Liu S, Trummer BJ, Deng C, and Wang A. Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells. Nat Biotechnol. 2002; 20:275–81.

61. Wang Y, Yu G, Han Z, Yang B, Hu Y, Zhao X, Wu J, Lv Y, and Chai W. Specificities of *Ricinus communis* agglutinin 120 interaction with sulfated galactose. FEBS Lett. 2011; 585:3927–34.

62. Palma AS, Liu Y, Childs RA, Herbert C, Wang D, Chai W, and Feizi T. The human epithelial carcinoma antigen recognized by monoclonal antibody AE3 is expressed on a sulfoglycolipid in addition to neoplastic mucins. Biochem Biophys Res Commun. 2011; 408:548–52.

63. Liu Y, Feizi T, Campanero-Rhodes MA, Childs RA, Zhang Y, Mulloy B, Evans PG, Osborn HMI, Otto D, Crocker PR and Chai W. Neoglycolipid Probes Prepared via Oxime Ligation for Microarray Analysis of Oligosaccharide-Protein Interactions. Chem Biol. 2007; 14:847–59.

64. Liu Y, Childs RA, Palma AS, Campanero-Rhodes MA, Stoll MS, Chai W, and Feizi T. Neoglycolipid-based oligosaccharide microarray system: preparation of NGLs and their noncovalent immobilization on nitrocellulose-coated glass slides for microarray analyses. Methods Mol Biol Clifton NJ. 2012; 808:117–36.

65. Leteux C, Stoll MS, Childs RA, Chai W, Vorozhaikina M, and Feizi T. Influence of oligosaccharide presentation on the interactions of carbohydrate sequence-specific antibodies and the selectins: Observations with biotinylated oligosaccharides. J Immunol Methods. 1999; 227:109–19.

66. Uzawa H, Ito H, Izumi M, Tokuhisa H, Taguchi K, and Minoura N. Synthesis of polyanionic glycopolymers for the facile assembly of glycosyl arrays. Tetrahedron. 2005; 61:5895–905.

67. Ko K-S, Jaipuri FA, and Pohl NL. Fluorous-Based Carbohydrate Microarrays. J Am Chem Soc. 2005; 127:13162–3.

68. Mamidyala SK, Ko K-S, Jaipuri FA, Park G, and Pohl NL. Noncovalent fluorous interactions for the synthesis of carbohydrate microarrays. J Fluor Chem. 2006; 127:571–9.

69. Chen G-S, and Pohl NL. Synthesis of Fluorous Tags for Incorporation of Reducing Sugars into a Quantitative Microarray Platform. Org Lett. 2008; 10:785–8.

70. Park S, Lee M, Pyo S-J, and Shin I. Carbohydrate Chips for Studying High-Throughput Carbohydrate−Protein Interactions. J Am Chem Soc. 2004; 126:4812–9.

71. de Boer AR, Hokke CH, Deelder AM, and Wuhrer M. General Microarray Technique for Immobilization and Screening of Natural Glycans. Anal Chem. 2007; 79:8107–13.

72. Andersen MCF, Kračun SK, Rydahl MG, Willats WGT, and Clausen MH. Synthesis of β-1,4-Linked Galactan Side-Chains of Rhamnogalacturonan I. Chem – Eur J. 2016; 22:11543–8.

73. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, Die IV, Burton DR, Wilson IA, Cummings R, Bovin N, Wong C-H,

and Paulson C. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. Proc Natl Acad Sci U S A. 2004; 101:17033–8.

74. Yang J, Moraillon A, Siriwardena A, Boukherroub R, Ozanam F, Gouget-Laemmel AC, and Szunerits S. Carbohydrate Microarray for the Detection of Glycan–Protein Interactions Using Metal-Enhanced Fluorescence. Anal Chem. 2015; 87:3721–8.

75. Lee M, and Shin I. Facile Preparation of Carbohydrate Microarrays by Site-Specific, Covalent Immobilization of Unmodified Carbohydrates on Hydrazide-Coated Glass Slides. Org Lett. 2005; 7:4269–72.

76. Angeloni S, Ridet JL, Kusy N, Gao H, Crevoisier F, Guinchard S, Kochhar S, Sigrist H, and Sprenger N. Glycoprofiling with micro-arrays of glycoconjugates and lectins. Glycobiology. 2005; 15:31–41.

77. Tong Q, Wang X, Wang H, Kubo T, and Yan M. Fabrication of Glyconanoparticle Microarrays. Anal Chem. 2012; 84:3049–52.

78. Sundhoro M, Wang H, Boiko ST, Chen X, Jayawardena HSN, Park J, and Yan M. Fabrication of carbohydrate microarrays on a poly(2-hydroxyethyl methacrylate)-based photoactive substrate. Org Biomol Chem. 2016; 14:1124–30.

79. Wang H, Zhang Y, Yuan X, Chen Y, and Yan M. A Universal Protocol for Photochemical Covalent Immobilization of Intact Carbohydrates for the Preparation of Carbohydrate Microarrays. Bioconjug Chem. 2011; 22:26–32.

80. Feizi T. Carbohydrate recognition in the immune system: contributions of neoglycolipid-based microarrays to carbohydrate ligand discovery. Ann N Y Acad Sci. 2013; 1292:33–44.

81. Lis H, and Sharon N. Lectins: Carbohydrate-Specific Proteins That Mediate Cellular Recognition. Chem Rev. 1998; 98:637–74.

82. Knox JP. Revealing the structural and functional diversity of plant cell walls. Curr Opin Plant Biol. 2008; 11:308–13.

83. Drickamer K, and Taylor ME. Evolving views of protein glycosylation. Trends Biochem Sci. 1998; 23:321–4.

84. Rüdiger H, and Gabius H-J. Plant lectins: Occurrence, biochemistry, functions and applications. Glycoconj J. 2001; 18:589–613.

85. Goldstein IJ, and Hayes CE. The Lectins: Carbohydrate-Binding Proteins of Plants and Animals. Adv Carbohydr Chem Biochem. 1978; 35:127–340.

86. Yu Y, Mishra S, Song X, Lasanajak Y, Bradley KC, Tappert MM, Air GM, Steinhauer DA, Halder S, Cotmore S, Tattersall P, Agbandje-McKEnna M, Cummings RD, and Smith DF. Functional Glycomic Analysis of Human Milk Glycans Reveals the Presence of Virus Receptors and Embryonic Stem Cell Biomarkers. J Biol Chem. 2012; 287:44784–99.

87. Yu Y, Lasanajak Y, Song X, Hu L, Ramani S, Mickum ML, Ashline DJ, Prasad BVV, Estes MK, Reinhold VN, Cummings RD, and Smith DF. Human Milk Contains Novel Glycans That Are Potential Decoy Receptors for Neonatal Rotaviruses. Mol Cell Proteomics. 2014; 13:2944–60.

88. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, Die IV, Burton DR, Wilson IA, Cummings R, Bovin N, Wong C-H, and Paulson C. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. Proc Natl Acad Sci U S A. 2004; 101:17033–8.

89. Wang L, Cummings RD, Smith DF, Huflejt M, Campbell CT, Gildersleeve JC, Gerlach JQ, Kilcoyne M, Joshi L, Serna S, Reichardt N-C, Pera NP, Pieters RJ, Eng W, and Mahal LK. Cross-platform comparison of glycan microarray formats. Glycobiology. 2014; 24:507–17.

90. Baenziger JU, and Fiete D. Structural determinants of *Ricinus communis* agglutinin and toxin specificity for oligosaccharides. J Biol Chem. 1979; 254:9795–9.

91. Palma AS, Liu Y, Zhang H, Zhang Y, McCleary BV, Yu G, Huang Q, Guidolin LS, Ciocchini AE, Torosantucci A, Wang D, Carvalho AL, Fontes CMGA, Mulloy B, Childs RA, Feizi T, and Chai W.

Unravelling Glucan Recognition Systems by Glycome Microarrays Using the Designer Approach and Mass Spectrometry. Mol Cell Proteomics. 2015; 14:974–88.

92. McGreal EP, Rosas M, Brown GD, Zamze S, Wong SYC, Gordon S, MArtinez-Pomares L, and Taylor PR. The carbohydrate-recognition domain of Dectin-2 is a C-type lectin with specificity for high mannose. Glycobiology. 2006; 16:422–30.

93. Holla A, and Skerra A. Comparative analysis reveals selective recognition of glycans by the dendritic cell receptors DC-SIGN and Langerin. Protein Eng Des Sel. 2011; 24:659–69.

94. Tateno H, Ohnishi K, Yabe R, Hayatsu N, Sato T, Takeya M, Narimatsu H, and Hirabayashi J. Dual Specificity of Langerin to Sulfated and Mannosylated Glycans via a Single C-type Carbohydrate Recognition Domain. J Biol Chem. 2010; 285:6390–400.

95. Panagos CG, Thomson DS, Moss C, Hughes AD, Kelly MS, Liu Y, Chai W, Venkatasamy R, Spina D, Page CP, Hogwood J, Woods RJ, Mulloy B, Bavington CD, and Uhrín D. Fucosylated Chondroitin Sulfates from the Body Wall of the Sea Cucumber *Holothuria forskali*. J Biol Chem. 2014; 289:28284–98.

96. Imperial College London. Part I: Structures and functions of animal lectins [Web page]. London: Division of Molecular Biosciences; 2014 [update 2017; cited 2017 8 Jun]. Available from: http://www.imperial.ac.uk/research/animallectins/ctld/lectins.html#Location

97. Noll AJ, Gourdine J-P, Yu Y, Lasanajak Y, Smith DF, and Cummings RD. Galectins are human milk glycan receptors. Glycobiology. 2016; 26:655–69.

98. Knox JP. The use of antibodies to study the architecture and developmental regulation of plant cell walls. Int Rev Cytol. 1997; 171:79–120.

99. Moller I, Sørensen I, Bernal AJ, Blaukopf C, Lee K, Øbro J, Pettolino F, Roberts A, Mikkelsen JD, Knox JP, Bacic A, and Willatas WGT. High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. Plant J. 2007; 50:1118–28.

100. Pedersen HL, Fangel JU, McCleary B, Ruzanski C, Rydahl MG, Ralet M-C, Farkas V, Schantz LV, Marcus SE, Andersen MCF, Field R, Ohlin M, Knox JP, Clausen MH, and Willats WGT. Versatile High Resolution Oligosaccharide Microarrays for Plant Glycobiology and Cell Wall Research. J Biol Chem. 2012; 287:39429–38.

101. Moore JP, Nguema-Ona E, Fangel JU, Willats WGT, Hugo A, and Vivier MA. Profiling the main cell wall polysaccharides of grapevine leaves using high-throughput and fractionation methods. Carbohydr Polym. 2014; 99:190–8.

102. Meikle PJ, Hoogenraad NJ, Bonig I, Clarke AE, and Stone BA. A (1→3,1→4)-β-glucan-specific monoclonal antibody and its use in the quantitation and immunocytochemical location of (1→3,1→4)-β-glucans. Plant J. 1994; 5:1–9.

103. Gao C, Liu Y, Zhang H, Zhang Y, Fukuda MN, Palma AS, Kozak RP, Childs RA, Nonaka M, Li Z, Siegel DL, Hanfland P, Peehl DM, Chai W, Greene MI, and Feizi T. Carbohydrate Sequence of the Prostate Cancer-associated Antigen F77 Assigned by a Mucin O-Glycome Designer Array. J Biol Chem. 2014; 289:16462–77.

104. Boraston AB, Bolam DN, Gilbert HJ, and Davies GJ. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J. 2004; 382:769–81.

105. Gilbert HJ. The Biochemistry and Structural Biology of Plant Cell Wall Deconstruction. Plant Physiol. 2010; 153:444–55.

106. Gilbert HJ, Knox JP, and Boraston AB. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. Curr Opin Struct Biol. 2013; 23:669–77.

107. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, and Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014; 42:D490–5.

108. Zhang M, Chekan JR, Dodd D, Hong P-Y, Radlinski L, Revindran V, Nair SK, Macjie RI, and Cann I. Xylan utilization in human gut commensal bacteria is orchestrated by unique modular organization of polysaccharide-degrading enzymes. Proc Natl Acad Sci. 2014; 111:E3708–17.

109. Consortium THMP. Structure, function and diversity of the healthy human microbiome. Nature. 2012; 486:207–14.

110. Koropatkin NM, and Smith TJ. SusG: A Unique Cell-Membrane-Associated α-Amylase from a Prominent Human Gut Symbiont Targets Complex Starch Molecules. Structure. 2010; 18:200–15.

111. Foley MH, Cockburn DW, and Koropatkin NM. The Sus operon: a model system for starch uptake by the human gut Bacteroidetes. Cell Mol Life Sci. 2016; 73:2603–17.

112. Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP, Abbot DW, Henrissat B, Gilbert HJ, Bolam DN, and Gordon JI. Recognition and Degradation of Plant Cell Wall Polysaccharides by Two Human Gut Symbionts. PLOS Biol. 2011; 9:e1001221.

113. Ursell LK, Metcalf JL, Parfrey LW, and Knight R. Defining the Human Microbiome. Nutr Rev. 2012; 70:S38–44.

114. Ndeh D, Rogowski A, Cartmell A, Luis AS, Baslé A, Gray J, Venditto I, Briggs J, Zhang X, Labourel A, Terrapon N, Buffeto F, Nepogodiev S, Xiao Y, Field RA, Zhu Y, O'Neill MA, Urbanowicz BR, York WS, Davies GJ, Abbot DW, Ralet M-C, Martens EC, Henrissat B, and Gilbert HJ. Complex pectin metabolism by gut bacteria reveals novel catalytic functions. Nature. 2017; 544:65–70.

115. Wang J, Zhang Q, Zhang Z, and Li Z. Antioxidant activity of sulfated polysaccharide fractions extracted from *Laminaria japonica*. Int J Biol Macromol. 2008; 42:127–32.

116. Lourenço SO, Barbarino E, Lavín PL, Marquez UML, and Aidar E. Distribution of intracellular nitrogen in marine microalgae: Calculation of new nitrogen-to-protein conversion factors. Eur J Phycol. 2004; 39:17–32.

117. Selvendran RR, March JF, and Ring SG. Determination of aldoses and uronic acid content of vegetable fiber. Anal Biochem. 1979; 96:282–92.

118. Coimbra MA, Delgadillo I, Waldron KW, and Selvendran RR. Isolation and Analysis of Cell Wall Polymers from Olive Pulp. Modern Methods of Plant Analysis. 1996; 17:19–44.

119. Dodgson KS, and Price RG. A note on the determination of the ester sulphate content of sulphated polysaccharides. Biochem J. 1962; 84:106–10.

120. Ciucanu I, and Kerek F. A simple and rapid method for the permethylation of carbohydrates. Carbohydr Res. 1984; 131:209–17.

121. Ciucanu I. Per-*O*-methylation reaction for structural analysis of carbohydrates by mass spectrometry. Anal Chim Acta. 2006; 576:147–55.

122. Miller IJ, and Blunt JW. Desulfation of algal galactans. Carbohydr Res. 1998; 309:39-43.

123. Block H, Maertens B, Spriestersbach A, Brinker N, Kubicek J, Fabis R, Labahn J, and Schäfer F. Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC). Methods Enzymol. 2009; 463:439–73.

124. ISBio training course, Microarray sceening analysis supporting material. 2015; Caparica, Portugal.

125. Marcus SE, Blake AW, Benians TAS, Lee KJD, Poyser C, Donaldson L, Leroux O, Rogowski A, Petersen HL, Boraston A, Gilbert HJ, Willats WGT, and Knox JP. Restricted access of proteins to mannan polysaccharides in intact plant cell walls. Plant J. 2010; 64:191–203.

126. Wu H, and Miao X. Biodiesel quality and biochemical changes of microalgae *Chlorella pyrenoidosa* and *Scenedesmus obliquus* in response to nitrate levels. Bioresour Technol. 2014; 170:421–7.

127. Iwamoto K, and Shiraiwa Y. Salt-Regulated Mannitol Metabolism in Algae. Mar Biotechnol. 2005; 7:407–15.

128. Pal D, Khozin-Goldberg I, Didi-Cohen S, Solovchenko A, Batushansky A, Kaye Y, Sikron N, Samani T, Fait A, and Boussiba S. Growth, lipid production and metabolic adjustments in the euryhaline eustigmatophyte *Nannochloropsis oceanica* CCALA 804 in response to osmotic downshift. Appl Microbiol Biotechnol. 2013; 97:8291–306.

129. Obata T, Schoenefeld S, Krahnert I, Bergmann S, Scheffel A, and Fernie AR. Gas-Chromatography Mass-Spectrometry (GC-MS) Based Metabolite Profiling Reveals Mannitol as a Major Storage Carbohydrate in the Coccolithophorid Alga *Emiliania huxleyi*. Metabolites. 2013; 3:168–84.

130. Davison IR, and Reed RH. The physiological significance of mannitol accumulation in brown algae: the role of mannitol as a compatible cytoplasmic solute. Phycologia. 1985; 24:449–57.

131. Souto M, Saavedra M, Pousão-Ferreira P, and Herrero C. Riboflavin enrichment throughout the food chain from the marine microalga *Tetraselmis suecica* to the rotifer *Brachionus plicatilis* and to White Sea Bream (*Diplodus sargus*) and Gilthead Sea bream (*Sparus aurata*) larvae. Aquaculture. 2008; 283:128–33.

132. Brown MR, and Farmer CL. Riboflavin content of six species of microalgae used in mariculture. J Appl Phycol. 1994; 6:61–5.

133. Brown MR, Mular M, Miller I, Farmer C, and Trenerry C. The vitamin content of microalgae used in aquaculture. J Appl Phycol. 1999; 11:247–55.

134. Belitz H-D, Grosch W, and Shieberle P. Food Chemistry. 4th ed. Springer; 2009.

135. Li H, Mao W, Chen Y, Ren S, Qi X, Chen Y, Zhao C, Li N, Wang C, Lin C. Yan M, and Shan J. Sequence analysis of the sulfated rhamno-oligosaccharides derived from a sulfated rhamnan. Carbohydr Polym. 2012; 90:1299–304.

136. Ropellato J, Carvalho MM, Ferreira LG, Noseda MD, Zuconelli CR, Gonçalves AG, Ducatti DRB, Kenski JCN, Nasato PL, Winnischofer SMB, and Duarte MER. Sulfated heterorhamnans from the green seaweed *Gayralia oxysperma*: partial depolymerization, chemical structure and antitumor activity. Carbohydr Polym. 2015; 117:476–85.

137. Jerez CG, Malapascua JR, Sergejevová M, Figueroa FL, and Masojídek J. Effect of Nutrient Starvation under High Irradiance on Lipid and Starch Accumulation in *Chlorella fusca* (Chlorophyta). Mar Biotechnol. 2016; 18:24–36.

138. Geun Goo B, Baek G, Jin Choi D, Il Park Y, Synytsya A, Bleha R, Seong DH, Lee C-G, and Park JK. Characterization of a renewable extracellular polysaccharide from defatted microalgae *Dunaliella tertiolecta*. Bioresour Technol. 2013; 129:343–50.

139. Schulze C, Wetzel M, Reinhardt J, Schmidt M, Felten L, and Mundt S. Screening of microalgae for primary metabolites including β-glucans and the influence of nitrate starvation and irradiance on β-glucan production. J Appl Phycol. 2016; 28:2719–25.

140. Popper ZA, and Tuohy MG. Beyond the Green: Understanding the Evolutionary Puzzle of Plant and Algal Cell Walls. Plant Physiol. 2010; 153:373–83.

141. Salmeán AA, Duffieux D, Harholt J, Qin F, Michel G, Czjzek M, Willats WGT, and Hervé C. Insoluble (1 → 3), (1 → 4)-β-D-glucan is a component of cell walls in brown algae (Phaeophyceae) and is masked by alginates in tissues. Sci Rep. 2017; 7:2880.

142. Eder M, Tenhaken R, Driouich A, and Lütz-Meindl U. Occurrence and Characterization of Arabinogalactan-Like Proteins and Hemicelluloses in *Micrasterias* (streptophyta). J Phycol. 2008; 44:1221–34.

143. Yamagishi T, Müller DG, and Kawai H. Comparative transcriptome analysis of *Discosporangium mesarthrocarpum* (Phaeophyceae), *Schizocladia ischiensis* (Schizocladiophyceae), and *Phaeothamnion confervicola* (Phaeothamniophyceae), with special reference to cell wall-related genes. J Phycol. 2014; 50:543–51.

144. Lin Y-C, Campbell T, Chung C-C, Gong G-C, Chiang K-P, and Worden AZ. Distribution Patterns and Phylogeny of Marine Stramenopiles in the North Pacific Ocean. Appl Environ Microbiol. 2012; 78:3387–99.

145. Burton RA, and Fincher GB. (1,3;1,4)-beta-D-glucans in cell walls of the poaceae, lower plants, and fungi: a tale of two linkages. Mol Plant. 2009; 2:873–82.

146. Scholz MJ, Weiss TL, Jinkerson RE, Jing J, Roth R, Goodenough U, Posewitz MC, and Gerken HG. Ultrastructure and Composition of the *Nannochloropsis gaditana* Cell Wall. Eukaryot Cell. 2014; 13:1450–64.

147. Tsubaki S, Hiraoka M, Hadano S, Nishimura H, Kashimura K, and Mitani T. Functional group dependent dielectric properties of sulfated hydrocolloids extracted from green macroalgal biomass. Carbohydr Polym. 2014; 107:192–7.

148. Prasanna V, Prabha TN, and Tharanathan RN. Fruit ripening phenomena - an overview. Crit Rev Food Sci Nutr. 2007; 47:1–19.

149. Henshaw JL, Bolam DN, Pires VMR, Czjzek M, Henrissat B, Ferreira LMA, Fontes CMGA, and Gilbert HJ. The Family 6 Carbohydrate Binding Module *Cm*CBM6-2 Contains Two Ligand-binding Sites with Distinct Specificities. J Biol Chem. 2004; 279:21552–9.

150. Abbott DW, Ficko-Blean E, van Bueren AL, Rogowski A, Cartmell A, Coutinho PM, Henrissat B, Gilbert HJ, and Boraston AB. Analysis of the structural and functional diversity of plant cell wall specific family 6 carbohydrate binding modules. Biochemistry (Mosc). 2009 ; 48:10395–404.

151. Kent M, Welladsen HM, Mangott A, and Li Y. Nutritional Evaluation of Australian Microalgae as Potential Human Health Supplements. PLoS ONE. 2015; 10:e0118985.

152. Gu N, Lin Q, Li G, Tan Y, Huang L, and Lin J. Effect of salinity on growth, biochemical composition, and lipid productivity of *Nannochloropsis oculata* CS 179. Eng Life Sci. 2012; 12:631–7.

153. Smith DF, Song X, and Cummings RD. Use of Glycan Microarrays to Explore Specificity of Glycan-Binding Proteins. Methods Enzymol. 2010; 480:417–44.

154. Fei Y, Sun Y-S, Li Y, Lau K, Yu H, Chokhawala HA, Huang S, Landry JP, Chen X, and Zhu X. Fluorescent labeling agents change binding profiles of glycan-binding proteins. Mol BioSyst. 2011; 7:3343–52.

Websites:

http://www.cazy.org/

http://www.sigmaaldrich.com

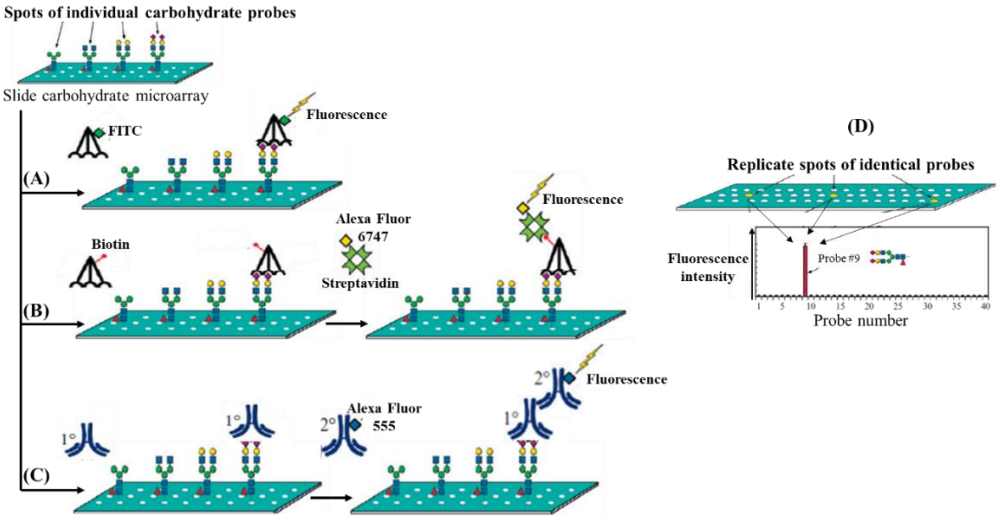
http://web.expasy.org/protparam

Supplementary Material

Supplementary information 1

Detection and analysis of carbohydrate-protein interactions through fluorescence-measurements

Fluorescence-based measurements are the most common approaches used to detect and analyse carbohydrate-protein interactions formed during the binding assay (57,153). This method relies mainly on incubation of carbohydrate microarrays, after printing of carbohydrate probes in each spot, with fluorescent markers coupled to the proteins in analysis or to other reagents in order to monitor molecular interactions. The fluorescence intensities of each spot derived from the interaction between the carbohydrate probe printed in the spot and the protein in analysis is quantified using fluorescence detectors, such as microarray scanners. The fluorescence intensities are directly proportional to the amount of bound carbohydrates on the surface (supplementary figure 1). Various fluorophore can be used to read out, such as Alexa-Fluor (AF) (74,92,100), fluorescein isothiocyanate (FITC) (67,68,79), and Cyanine (Cy)3 or Cy5 (63,70). In particular, Cy3 may be included during the printing step together with the probes, working as marker for spot location and array quality control (80).



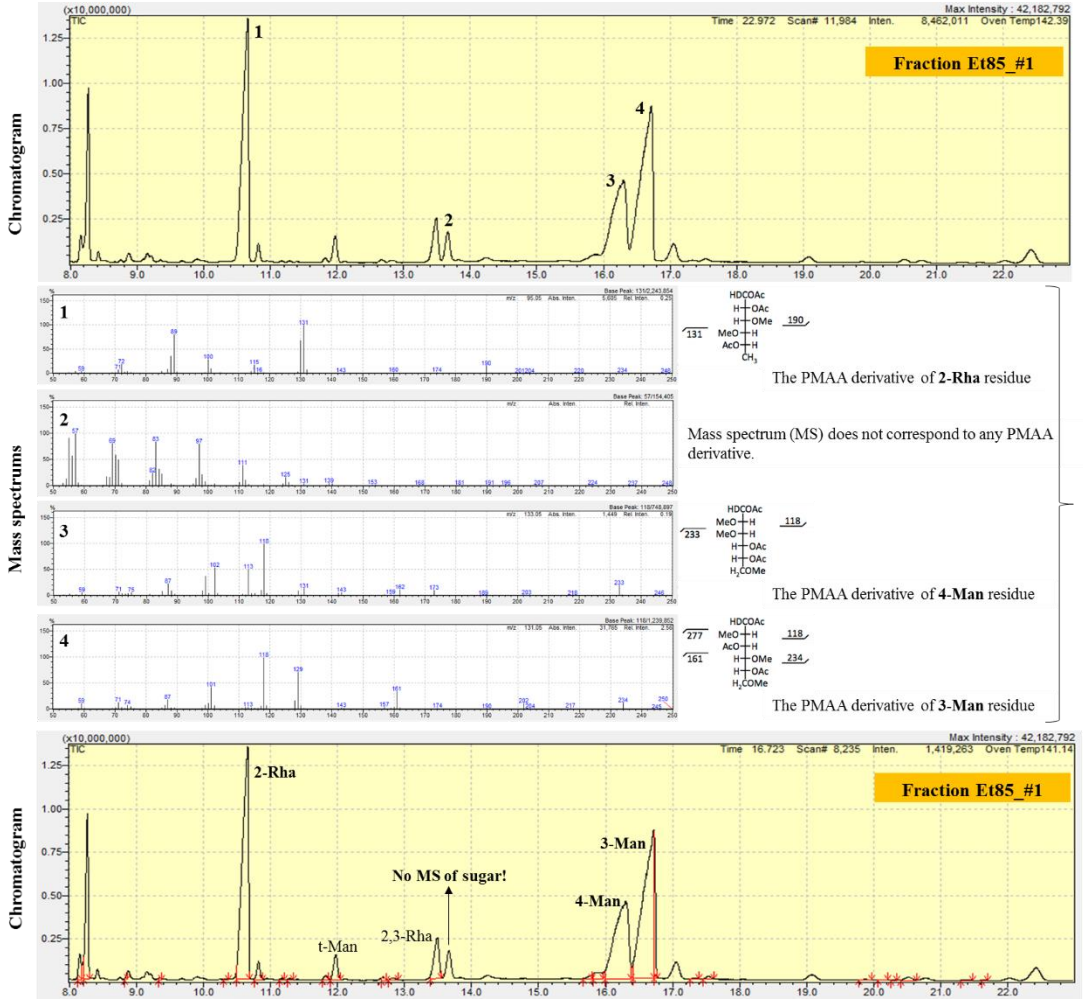
Supplementary Figure 1. Approaches to detect molecular interactions through fluorescence-based methods. **(A)** Direct fluorophore (FITC) labelling of the protein, **(B)** use of fluorophore (Alexa Fluor 6747)-labelled secondary reagent (streptavidin) that binds to biotinylated protein, and **(C)** use of fluorescent secondary antibody that binds to anti-glycan antibody. **(D)** The average relative fluorescence units generated during the process of fluorescence scanning of replicate spots are calculated and the data are presented as histograms of fluorescence intensity. Image adapted from Smith and Cummings (153).

Various approaches can be used to detect molecular interactions through fluorescence-based methods (153). These include: **1)** the direct fluorophore labelling of the protein in analysis (67,68), such as FITC-labelled ConA (67) (sup. Fig. 1 A); **2)** the use of fluorophore-labelled secondary reagent that binds a tag (*e.g.* biotin or histidine tags) present on the protein (61,62,92), such as using fluorophore-labelled streptavidin (*e.g.* AF-6747-labelled streptavidin) that binds a biotinylated protein, since streptavidin is highly specific to recognise the biotin (sup. Fig. 1 B) (62); and **3)** the use of a fluorophore-labelled secondary reagent that binds the protein (72,100), for instance, anti-glycan antibodies normally are detected with fluorescent secondary antibodies like the anti-rat secondary antibody conjugated to AF 555 (100) (sup. Fig. 1 C).

Although the fluorescence approaches used to detect the carbohydrate-protein interactions are satisfactory applied in the most cases, they can exhibit some limitations. As example, the direct labelling of proteins with fluorophores may cause protein denaturation or interfere with their binding properties with OS ligands (154). Despite using secondary reagents may be considered as an alternative strategy to avoid protein denaturation, this procedure also reveals limitations. For example, the lack of secondary reagents for many proteins or the possibility of these reagents possesses their own OS-binding properties.

Supplementary information 2

A representative scheme of the GC-MS analysis of the PMAAs derived from fraction Et85_#1 to determine the glycosidic linkages present in the PS of this fraction is shown below.



1º Separation of PMAAs derived from methylation analysis through gas chromatography (GC).

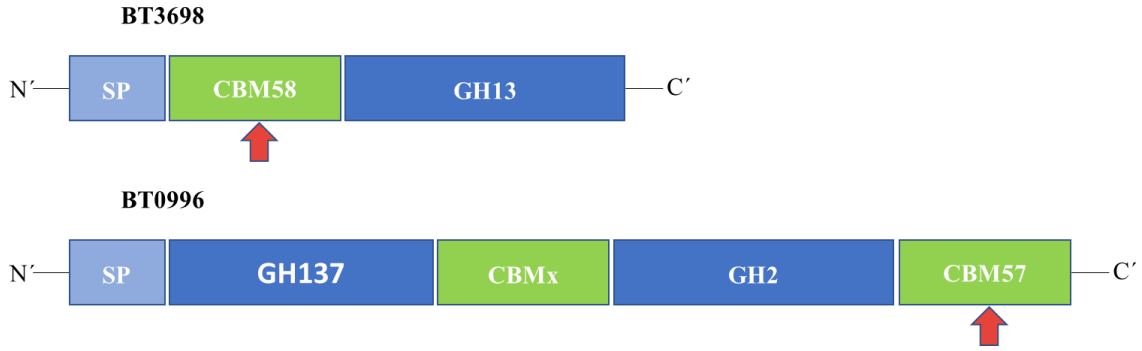
2º Identification of chromatogram peaks of the PMAAs derivatives and correspondent sugar residue through mass spectrometry (MS)

3º Integration of the peaks corresponding to sugar residues.

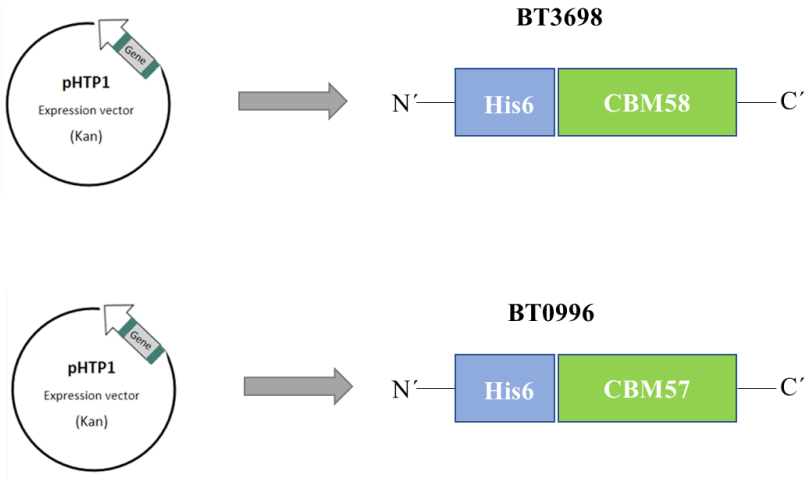
Supplementary Figure 2. Representative scheme of the steps performed to identify the glycosidic linkages present in the PS of fraction Et85_#1. **1º** Separation of the PMAAs by gas chromatography (GC); **2º** Identification of each peak in the chromatogram using mass spectrometry (MS); and **3º** Integration of the area of each peak on the chromatogram corresponding to sugars and posterior relative quantification.

Supplementary figures

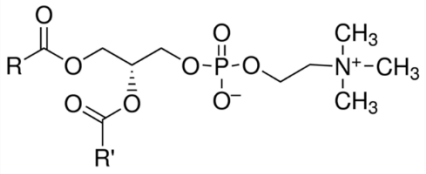
Molecular architecture of the native proteins BT3698 and BT0996 expressed by *B. thetaiotaomicron* are described in supplementary figure 3. The recombinant proteins overexpressed in *Escherichia coli* competent cells, BT3698 (CBM58) and BT0996 (CBM57), are shown in supplementary figure 4. The chemical structures of the carrier lipids used to form the liposome probes that were employed in the carbohydrate microarray experiments comprising the lipid fraction dialysed (GL) extracted from *N. oculata* are shown in supplementary figure 5.



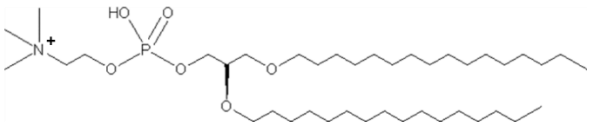
Supplementary Figure 3. Native proteins expressed by *Bacteroides thetaiotaomicron*, BT3698 and BT0996. Red arrows are the selected domains used for cloning. Key: SP, signal peptide; GH, glycoside hydrolase; CBM, carbohydrate-binding module; N', N-terminal; C', C-terminal; GH and CBM alternative numbers indicate different families according to CAZy classification (http://www.cazy.org/, (107) and Correia, V. G. *et al.* unpublished).



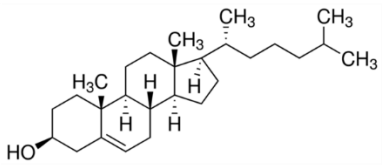
Supplementary Figure 4. Recombinant proteins overexpressed in *Escherichia coli* BL21 (DE3) and TUNER (DE3) expression strains, BT3698 and BT0996, respectively. A pHTP1 vector with kanamycin (Kan) resistance was used for cloning the domains CBM58 and CBM57. The recombinant proteins BT3698 and BT0996 contained an N-terminal (N') His-tag (His6) comprising the following sequence: MGSSHHHHHHSSGPQQGLR, and the respective CBM module (Correia, V. G. *et al.* unpublished).



1,2-Diacyl-*sn*-glycero-3-phosphocholine (PC1)
(Lecithin from egg yolk)



1,2-Di-*O*-hexadecyl-*sn*-glycero-3-phosphocholine (PC2)



Cholesterol (C)

Supplementary Figure 5. Chemical structures of the carrier lipids used to produce the liposome probes. **1,2-Diacyl-*sn*-glycero-3-phosphocholine** (PC1) from egg yolk, R and R'= Fatty acid chains. Typical egg yolk phosphatidylcholine have fatty acid contents of approximately 33 % 16:0 (palmitic), 13 % 18:0 (stearic), 31 % 18:1 (oleic), and 15 % 18:2 (linoleic) (other fatty acids being in minor contributors). **1,2-Di-*O*-hexadecyl-*sn*-glycero-3-phosphocholine** (PC2) and **Cholesterol** (C). Images adapted from http://www.sigmaaldrich.com.

Supplementary tables

The parameters of the proteins BT3698 and BT0996 produced by recombinant expression are described in supplementary table 1. In supplementary tables 2 and 3 are depicted the conditions used during binding assay for each protein analysed in microarrays developed with PS and liposome probes, respectively.

Supplementary Table 1. Carbohydrate-binding modules (CBMs) expressed by *Bacteroides thetaiotaomicron* produced through recombinant expression, CBM57 and CBM58. The protein parameters of each protein are depicted, which include the length in number of amino acids (aa), the molecular weight (*Mw*), the extinction coefficient and the isoelectric point (parameters were calculated using the tool online bioninformatic tool http://web.expasy.org/protparam).

Organism	**Molecular Architecture**	**Cloned Domain**	**Protein ID**	**Length** (aa)	**Protein *Mw*** (kDa)	**Ext. Coef.** ($M^{-1}.cm^{-1}$)	**pI**
B. thetaiotaomicron	CBM58-GH13	CBM58	BT3698	138	15.34	26 930	6.02
B. thetaiotaomicron	GH137-CBMx-GH2-CBM57	CBM57	BT0996	141	15.89	16 960	9.43

Supplementary Table 2. Conditions of the proteins analysed in the carbohydrate microarrays with Microalgae PS set 1. BI, biotin.

	BI-Concanavalin A (Con A)	**BI-*Ricinus Communis* Agglutinin I (RCA_{120})**
Batch	Vector Lab, B-1005	Vector Lab, B-1085 (Lot X0411)
Concentration	2 µg/mL	2 µg/mL
Precomplex	-	-
Antibody	-	-
Biodetector	-	-
Blocker/Diluent	3% BSA, 5 mM $CaCl_2$	1% BSA, 0.02% Casein in HBS, 5mM $CaCl_2$
	Anti-Heteromannan (Rat IgM) LM21	**Anti-(β1→3, β1→4)-D-glucan (Mouse IgG)**
Batch	Plant Probes	Biosupplies
Concentration	Dilution 1/10	10 µg/mL
Precomplex	-	-
Antibody	BI anti-Rat IgM 3 µg/mL	BI anti-mouse IgG 3 µg/mL
Biodetector	-	-
Blocker/Diluent	1% BSA, 0.02% Casein in HBS, 5mM $CaCl_2$	1% BSA, 0.02% Casein in HBS, 5mM $CaCl_2$

	His-*Cm*CBM6-2	**A33-BT3698-His**	**A53-BT0996-C-His**
Batch	Harry Gilbert	Viviana & Carolina (FTC-NOVA University, Lisboa)	Frederico & Catarina & Carolina (FTC-NOVA University, Lisboa)
Concentration	10 µg/mL	10 µg/mL	25 µg/mL
Precomplex	1:3:3	1:3:3	1:3:3
Antibody	Mouse anti-His 30 µg/mL	Mouse anti-His 30 µg/mL	Mouse anti-His 75 µg/mL
Biodetector	BI anti-mouse IgG 30 µg/mL	BI anti-mouse IgG 30 µg/mL	BI anti-mouse IgG 75 µg/mL
Blocker/Diluent	1% BSA, 0.02% Casein in HBS, 5mM $CaCl_2$	3%(Block)/ 1% BSA in HBS, 5 mM $CaCl_2$	1% BSA, 0.02% Casein in HBS, 5mM $CaCl_2$

Supplementary Table 3. Conditions of the proteins analysed in the carbohydrate microarrays with Microalgae GL set 2. BI, biotin.

	BI-*Ricinus Communis* Agglutinin I (RCA_{120})	**hGalectin-3**
Batch	Vector Lab, B-1085 (Lot X0411, 5 mg/mL)	abcam ab89487 (Lot GR243053-7; 0.5mg/mL, 10 µL aliq.)
Concentration	5 µg/mL	50 µg/mL
Pre-complex	-	(antibodies 1:1 – 10 µg)
Antibody	-	Manti-His (Sigma, 2.2 mg/mL)
Biodetector	-	BI anti-Mouse IgG (Sigma B7264, 0.6 mg/mL)
Blocker/Diluent	3 % BSA in HBS, 5mM $CaCl_2$	0.5 % Casein in HBS, 5 mM $CaCl_2$

Supplementary information 3

The composition of the medium cultures used for recombinant expression of the proteins BT3698 and BT0996 is described below:

- **Luria-Bertani medium culture (1L):**

- 10 g Tryptone (Sigma-Aldrich®);
- 10 g NaCl (Panreac®);
- 5 g Yeast (NZYeTch®);
- Required volume of distilled water.

Autoclave for 20 min at 121 ºC.

- **Auto-induction medium from NZYTech® (1L):**

- 50 g NZY Auto-induction LB medium (powder);
- 10 mL Glycerol;
- Required volume of distilled water.

Autoclave for 15 min at 121 ºC.

Supplementary information 4

In this supplementary section is briefly described the information about the carbohydrate composition of five commercial polysaccharides (PS): Polygalacturonic acid (PGA) apple (purchased from Sigma-Aldrich), PGA citrus (purchased from Megazyme), High Methylated Galacturonate (HMG) and Low Methylated Galacturonate (LMG) purchased from OligoTech, and Oat β-glucan (purchased from Megazyme). The carbohydrate composition was determined at the University of Aveiro.

Characterization of PS:

The PS were analysed in terms of neutral sugars (NS) and uronic acids (UA), as well as in composition in glycosidic linkages, as described in section 2.1.2.3. The composition of the five commercial PS analysed is summarized in supplementary table 4 and figure 5. The main glycosidic linkages present in each PS sample are described in supplementary table 5.

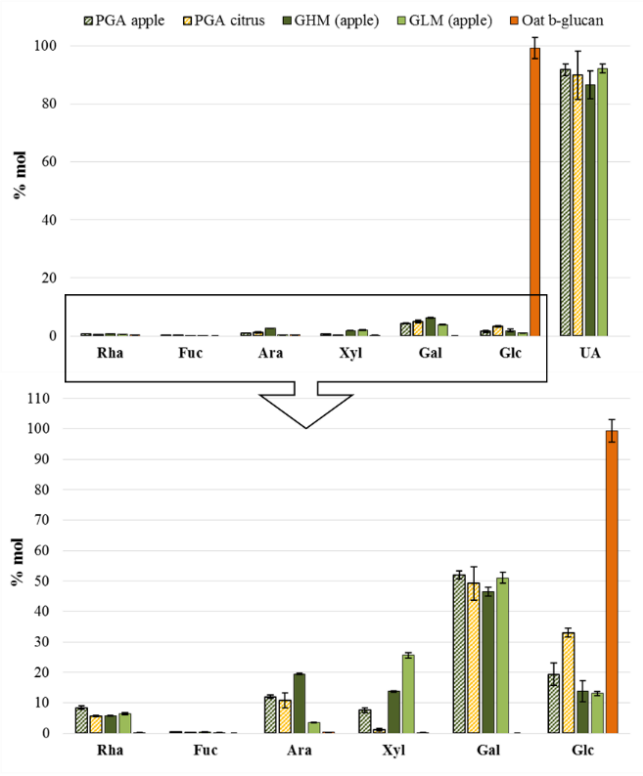
UA were found as major component in PGA apple, PGA citrus, HMG and LMG (sup. table 4). In these PS were found that the main NS was Gal, as expected for pectins, being found other sugars in lower content, namely Glc, Xyl, Ara and Rha (sup. Table 4 and Fig. 6). Both PGAs showed an identical sugar profile (sup. Fig. 6). Regarding to galacturonates (HM and LM), they also contained an identical sugar profile but Ara was found in higher amount in HMG than in LMG, whereas Xyl was in higher amount in LMG (sup. Fig. 6). Relatively to composition in glycosidic linkages (sup. Table 5), Gal was found mainly as terminal and (1→4)-linked Gal and GalA residues, as expected for pectins. In addition, it was found that Ara occurred mainly as (1→4)-linked Ara*p* or (1→5)-linked Ara*f*, and Rha as (1→2)-linked Rha.

The sugar analysis of the Oat β-glucan showed that this PS is composed essentially of Glc (sup. Table 4 and Fig. 6), which occurred mainly as (1→4)-linked Glc (68 % molar of the total NS) and (1→3)-linked Glc (22 %) (sup. Table 5), suggesting the presence of mixed (1→3, 1→4)-glucans, as expected.

Supplementary Table 4. Composition in neutral sugars (NS) and content in acidic sugars (UA) present in the following commercial polysaccharides: PGA apple (purchased from SIGMA), PGA citrus (purchased from Megazyme), High Methylated Galacturonate (HMG) and Low Methylated Galacturonate (LMG) purchased from OligoTech, and Oat β-glucan (purchased from Megazyme). Results are expressed in weight percentage (% w/w).

	(mean ± SD) % (w/w)*								
	Rha	Fuc	Ara	Xyl	Gal	Glc	**Total NS**	**UA**	**Total (NS +UA)**
PGA apple	0.67 ± 0.04	0.04 ± 0.01	0.86 ± 0.04	0.54 ± 0.05	4.56 ± 0.12	1.70 ± 0.32	8.37 ± 0.59	106.80 ± 2.40	115.17 ± 2.99
PGA citrus	0.47 ± 0.02	0.01 ± 0.01	0.81 ± 0.18	0.09 ± 0.04	4.51 ± 0.50	3.03 ± 0.14	8.90 ± 0.89	88.60 ± 8.20	97.50 ± 9.09
HMG	0.67 ± 0.02	0.04 ± 0.03	2.03 ± 0.02	1.43 ± 0.02	5.92 ± 0.19	1.77 ± 0.45	11.88 ± 0.75	89.70 ± 5.00	101.58 ± 5.75
LMG	0.42 ± 0.02	0.01 ± 0.01	0.21 ± 0.01	1.50 ± 0.06	3.67 ± 0.13	0.94 ± 0.05	6.75 ± 0.27	92.60 ± 1.50	99.35 ± 1.77
Oat β-glucan	0.25 ± 0.04	-	0.25 ± 0.04	0.14 ± 0.02	-	98.70 ± 3.67	99.34 ± 3.77	-	99.34 ± 3.77

***-** Data represent the mean values ± SD in weight percentage (% w/w) from triplicate (n=3) and n=4 analyses for NS and UA, respectively.



Supplementary Figure 6. Top graph: Composition in neutral and acidic sugars (UA) of the five commercial polysaccharides (PS) (PGA apple, PGA citrus, GHM, GLM, and Oat β-glucan). Results are expressed in molar percentage (mol %) of the total content in sugars (NS + UA). **Bottom graph:** Composition in NS of the same commercial PS, results are expressed in mol % of the total content in NS.

Supplementary Table 5. Glycosyl linkage composition (mol % of the total sugars) of the five commercial polysaccharides (PS) (PGA apple, PGA citrus, HMG, LMG, and Oat β-glucan). Glycosidic linkages were determined by methylation analysis with reduction of the carboxyl groups in PGA apple, PGA citrus, HMG and LMG, and in Oat β-glucan by methylation analysis without reduction of carboxyl groups.

	% mol				
	Methylation with reduction of carboxyl groups				
Glycosidic Linkage	**PGA apple**	**PGA citrus**	**HMG**	**LMG***	**Oat β-glucan**
2-Rha	10.1	5.9	10.7	-	-
t-Ara*f*	-	1.5	5.1	-	-
4-Ara*p*=5-Ara*f*	5.0	1.7	3.0	-	-
t-Xyl	-	-	7.3	-	-
4-Xyl	-	-	4.3	-	-
t-Gal	7.1	2.6	6.3	-	-
4-Gal	6.3	5.1	10.3	-	-
4-GalA	4.6	2.3	8.2	-	-
t-Glc	-	-	-	-	1.1
3-Glc	-	-	-	-	22.0
4-Glc	-	-	-	-	68.5

*- LMG methylation analysis did not allow to identify any sugar residue.
**- Other sugar residues were found in minor amounts.