



Università  
Ca'Foscari  
Venezia

Scuola Dottorale di Ateneo  
Graduate School

Dottorato di ricerca  
in Scienze Chimiche  
Ciclo XXVII  
Anno di discussione 2014

***Methodological developments based on mass  
spectrometry for the analysis of glycoproteins and  
polysaccharides of plant gums: an application to cultural  
heritage samples***

SETTORE SCIENTIFICO DISCIPLINARE DI AFFERENZA: CHIM/01  
Tesi di Dottorato di Clara Granzotto, matricola 820280

*Coordinatore del Dottorato*  
Prof. Maurizio Selva

*Tutore del Dottorando*  
Prof. Carlo Barbante  
Dr. Caroline Tokarski



*Alla mia super mamma*



# Acknowledgments

Here it is, the end of three long years that have been marked by a lot of ups and downs. I cannot finish this journey without giving thanks to all the wonderful people that I met along this, let's say, adventure spent between Venice and Lille.

First of all I would like to thank my two tutors. Prof. Carlo Barbante from Ca' Foscari University of Venice, thanks for this opportunity and for all your suggestions. Dr. Caroline Tokarski from Université de Lille1, a huge thank for welcoming me in France, for your support and for believing in me and in my project, even when I was the first one doubting. It was a great honor to work with you and your teams.

I sincerely thank Dr. Philippe Walter, UPMC, UMR CNRS 8220 LAMS, and Prof. Peter Vandenabeele, Ghent University, for their availability in judging my thesis in quality of referees.

A sincere thanks to Julie Arslanoglu (The Metropolitan Museum of Art, Department of Scientific Research) and Rachel Mustalish (The Metropolitan Museum of Art, Department of Paper Conservation) for providing the 50 years old gum arabic and the old watercolor samples. and to Patrick Ravines (Buffalo State, The state University of New York, Art Conservation Department) for providing several gum arabic samples collected from different parts of the world.

During these three years I had the chance to work in two amazing groups. First I want to thank Francesca Guidi for her precious help with all the bureaucracy and her availability in answering to the hundreds email I sent to her. Thanks to all the labmates of "laboratorio Cescon" in Venice. Warren, la Crotti, il Roman, thanks for your huge support and suggestions, you have been a great example. Silviuz, Chiara, Fabio, Elena B., Stefano, Enrico, thanks for all the funny and crazy moments in our office, you have been the best cure in times of difficulties. And a huge thanks to all the others for their smiles, the coffee times shared together and for being part of a great group: Luisette, Elisa S., Elena G., Elena A., Marco V., Francesca, Elisa M., Marta, Sarah, Fabiana, Torben, Piero, Andrea, Michela, Jacopo, Natalie, Prof. Gambaro, Prof. Piazza, Roberta, Patrizia, Giulio e Clara.

Now a huge thank to the MSAP group in Lille. Mr. Rolando, Fabrice, Stéphanie, Christophe, Nicolas, Chloé, thanks for your support, presence and suggestions, for the great help with my beloved MALDI and the thousands of gels I've prepared. Thanks to Charlène, I wish you all the best in Munich, and among all thanks to Sophie, a super colleague but, among all, a super friend. Thanks for your help, for listening to me all the time my analysis were not working, thanks for the tea time and the tons of chocolate we had to cheer me up. Tu vas me manquer Mme Dallongeville.

Deby, a fantastic friend, Nicolaus, Ema e Michelozzo, also known as "noialtri zingue", thanks for being always ready to meet up when I was coming back from Lille, even if for few days. Our dinners have always been a breath of fresh air for me. Francesca, Eric and Sacha, thanks for your help with the informatics tools, but most of all thanks for your friendship, your "let's do it!" and for the Sunday chicken. Sophie and Romain, my super "colocs" in Lille, thanks for all the great

moments we had together. Your presence and laughs have been essential while I was writing my dissertation, merci!.

An enormous thanks to three important persons that live far away but that I've always felt right close to me in the worst moments. Anna (aka Duke of Bristol), and Hae Young (aka Dr. Lee), thanks for your support, your emails and words. No matter how busy you were, you have always found time for me. Fede (aka Fred), I have not enough words to thank you for all our chats, for listening to me and for cheering me up. I wish you all the best and I hope one day we're going to be in the same, at least, continent!

I'm really glad that all of you crossed my life, you taught me a lot with your example, life and experience, helping me to grow up.

Finally, I want to thanks my mom, dad and sister. Even if I was far from home, you have always been there every time things were not going well. Thanks for your support, for believing in me, for sharing my happy and sad moments and for making these three years adventure more liveable. I felt your love and presence every single day and I knew I could always count on you. An immense thanks goes to my little sister for our weekends together around Europe, eating macarons, "frites" and gaufres. It meant a lot to me, grazie Stracchia. You three are a great example for me. Vi voglio bene.

# Table of contents

INTRODUCTION .....	xiii
--------------------	------

## **CHAPTER I**

### **Overview of plant gums: plant source, chemical structure and applications in artworks**

1. DEFINITION, PROPERTIES AND GENERAL APPLICATIONS .....	19
1.1. Gum arabic .....	20
1.2. Gum tragacanth .....	20
1.3. Fruit tree gums (Prunus species) .....	20
1.4. Gum karaya .....	21
1.5. Gum ghatti.....	21
1.6. Guar and locust bean gum.....	21
2. PLANT GUMS IN CULTURAL HERITAGE SAMPLES: APPLICATIONS AND THEIR CHEMICAL CHARACTERIZATION.....	22
2.1. The organic binding media.....	22
2.2. Employment of plant gums in art and archaeology.....	23
2.3. Study of polysaccharide based materials in art samples: state of art .....	24
3. THE PROTEINACEOUS COMPONENT OF PLANT GUMS .....	28
3.1. Gum arabic .....	28
3.2. Features of other plant gums protein fraction .....	32
1.2.1. Protein content.....	32
1.2.2. Amino acid composition .....	32
1.2.3. Molecular weight.....	34
4. THE POLYSACCHARIDE COMPONENT OF PLANT GUMS.....	35
4.1. Gum Arabic .....	35
4.2. Gum Tragacanth.....	38
4.3. Fruit tree gums (Prunus species) .....	39
4.4. Guar and locust bean gum.....	40
4.5. Ghatti gum.....	41
4.6. Karaya gum .....	43
5. CONCLUSIONS .....	45
Annex .....	46
REFERENCES .....	47

## **CHAPTER II**

### **Investigation of plant gums (glyco)protein profile by chromatographic and electrophoretic methods**

1. AIM OF THE RESEARCH .....	63
2. INVESTIGATION OF PLANT GUM PROTEIN COMPONENT BY SIZE EXCLUSION CHROMATOGRAPHY.....	64
2.1. Size exclusion chromatography.....	64
2.2. Sample preparation & materials .....	65
2.3. Optimization of HPLC-SEC and column calibration.....	66
2.4. SEC of several plant gums and molecular weight evaluation .....	68
2.4.1. Gum arabic .....	68
2.4.2. Tragacanth gum.....	69
2.4.3. Fruit tree gums.....	70
2.4.4. Guar, locust bean, ghatti and karaya gums.....	72
2.5. Conclusions of SEC analysis.....	73
3. INVESTIGATION OF PLANT GUMS PROTEIN PROFILE USING GEL ELECTROPHORESIS .....	74
3.1. Gel electrophoresis .....	74
3.1.1. Polyacrylamide gel electrophoresis (PAGE).....	74
3.1.2. Sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) .....	75
3.1.3. Gel staining.....	76
3.2. Optimization of electrophoretic experiments .....	77
3.3. SDS-PAGE profile of proteins in several plant gums .....	82
3.4. PAGE under non-denaturing conditions .....	87
3.6. Conclusions of SDS-PAGE analysis.....	89
REFERENCES .....	90

## **CHAPTER III**

### **A new strategy for plant gums identification by stepwise enzymatic digestion and mass spectrometry**

1. AIM OF THE RESEARCH .....	101
2. ANALYTICAL TECHNIQUE AND ENZYMATIC APPROACH.....	102
2.1. Matrix assisted laser desorption ionization mass spectrometry.....	102
(MALDI-MS) .....	102
2.1. MALDI source .....	102
2.2. Time of flight detector.....	102
2.3. MALDI-MS analysis of polysaccharides .....	103
2.4. Tandem mass spectrometry .....	104
2.2. Stepwise enzymatic degradation of polysaccharides.....	106



2.2.1. Enzymes acting on Type II arabinogalactans.....	107
2.2.2. Enzymes acting on galactomannans.....	109
2.2.3. Enzymes acting on rhamnogalacturonan.....	109
3. METHODOLOGICAL DEVELOPMENT ON STANDARD POLYSACCHARIDES .....	111
3.1. Experimental procedure optimization .....	111
3.1.1 Standard polysaccharides description and preparation.....	111
3.1.2. Enzymes and enzymatic hydrolysis general procedure.....	111
3.1.3. MALDI-TOF matrix .....	114
3.1.4. Oligosaccharides derivatization by 3-aminoquinoline .....	116
2.1.5. Oligosaccharides purification.....	119
3.2. Optimization of larch arabinogalactan enzymatic digestion .....	120
3.2.1. Evaluation of enzyme amount.....	120
3.2.2. Evaluation of enzymes combination and digestion time.....	122
3.2.3. MALDI-TOF profile of the polysaccharide .....	125
3.3. Optimization of carob galactomannan enzymatic digestion .....	128
4. APPLICATION AND METHODOLOGICAL DEVELOPMENT ON PLANT GUMS SAMPLES.....	131
4.1. Plant gum samples preparation .....	131
4.2. Gum arabic .....	132
4.2.1. Evaluation of enzymatic digestion protocol.....	132
4.2.2. Identification of the released oligosaccharides .....	137
4.2.3. Analysis of different gum arabic samples from Acacia Senegal var. Senegal .....	145
4.3. Tragacanth gum.....	147
4.3.1. Evaluation of enzymatic digestion protocol and oligosaccharides identification .....	148
4.3.2. Analysis of tragacanth samples from different brands .....	151
4.4 Cherry gum.....	153
3.4.1. Evaluation of enzymatic digestion protocol.....	153
4.4.2. Analysis of different cherry gum samples.....	155
4.5. Locust bean and guar gum .....	157
4.5.1. Digestion and identification of oligosaccharides released from LBG.....	157
4.5.3. Digestion and identification of oligosaccharides released from guar gum .....	161
5. DEVELOPMENT OF AN ENZYME COCKTAIL.....	164
5.1. Cocktail formulation and digestion procedure .....	164
5.2. Results and discussions .....	164
5.2.1. Gum arabic .....	165
5.2.2. Tragacanth gum.....	165
5.2.3. Cherry gum.....	166
5.2.4. Locust bean gum .....	167
5.2.5. Guar gum.....	167
6. APPLICATION ON REAL SAMPLES .....	169
6.1. A 50 years old gum arabic sample .....	169

6.1.1. Composition and sample preparation .....	169
6.1.2. Results and discussion.....	169
6.2. Fresh watercolor samples .....	171
6.2.1. Composition and sample preparation .....	171
6.2.2. Results and discussion.....	171
6.3. Old watercolor sample.....	173
6.3.1. Composition and sample preparation .....	173
6.3.2. Results and discussion.....	173
7. CONCLUSIONS of MALDI-TOF ANALYSIS .....	176
REFERENCES.....	177
<b>FINAL CONCLUSIONS.....</b>	<b>181</b>
<b>Materials and Methods.....</b>	<b>183</b>
1. PROTEIN QUANTITATION AND SEPARATION.....	187
1.1. Protein quantitation by bicinchoninic acid (BCA).....	187
1.2. Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis.....	187
1.3. SDS agarose-acrylamide composite gel electrophoresis.....	188
1.4. Size Exclusion Chromatography (SEC).....	188
2. PURIFICATION PROCEDURE.....	188
3. ENZYMATIC DIGESTION.....	189
4. MALDI-MS AND MS/MS ANALYSIS.....	189

# Abbreviations

ACN	: Acetonitrile
AFs	: Arabinofuranosidases
AG	: Arabinogalactan
AGE	: Agarose Gel Electrophoresis
AGP	: Arabinogalactan proteins / Arabinogalactan protein complex
Ala	: Alanine
APs	: Arabinopyranosidases
APS	: Ammonium persulfate
Ara	: Arabinose
Arg	: Arginine
Asp	: Aspartic acid
BSTFA	: <i>N,O</i> -bistrimethylsilyltrifluoroacetamide
CE	: Capillary Electrophoresis
Cis	: Cystine
Da	: Dalton
DHB	: Dehydroabiatic acid
dHex	: Deoxyhexose
DP	: Degree of polymerization
DTT	: Dithiothreitol
EC	: Enzyme Commission number
ELISA	: Enzyme-Linked Immunosorbent Assay
ESI-MS	: Electrospray-ionization mass spectrometry
FA	: Formic acid
FAO	: Food and Agricultural Organization
FTIR	: Fourier Transform Infrared Spectroscopy
Fuc	: Fucose
Gal	: Galactose
GalNAc	: N-acetylgalactosamine
GALs	: Galactanases (enzyme family)
GC	: Gas chromatography
Glc	: Glucose
GlcA	: Glucuronic acid
Glu	: Glutamic acid
Gly	: Glycine
GMs	: Glucuronomannans
GP	: Glycoprotein
h	: hour
HAC	: Hydrophobic Affinity Chromatography
HCCA	: $\alpha$ -Cyano-4-hydroxycinnamic acid
Hex	: Hexose
HexA	: Hexuronic acid
HexNAc	: N-Acetylamine hexose
His	: Histidine
HMDS	: Hexamethyldisilazane
HPAEC	: High-Performance Anion Exchange Chromatography
HPLC	: High Performance Liquid Chromatography
Hyp	: Hydroxyproline
JECFA	: Joint Expert Committee for Food Additives
Ile	: Isoleucine
LBG	: Locust bean gum
Leu	: Leucine
LOD	: Limit of detection

Lys	: Lysine
MALDI	: Matrix assisted laser desorption ionization
MALLS	: Multi angle laser light scattering
Man	: Mannose
Met	: Methionine
MeOH	: Methanol
MS	: Mass spectrometry
MS <sup>n</sup>	: Tandem mass spectrometry
m/z	: mass-to-charge ratio
MW	: Molecular weight
[M+H] <sup>+</sup>	: Proton adduct ion
[M+K] <sup>+</sup>	: Potassium adduct ion
[M+Na] <sup>+</sup>	: Sodium adduct ion
NeuAc	: N-Acetylneuraminic acid
NMR	: Nuclear Magnetic Resonance
o/n	: over night
PCS	: Photon correlation microscopy
Pent	: Pentose
Phe	: Phenylalanine
Pro	: Proline
Py-GC/MS	: Pyrolysis Gas Chromatography
RGs	: Rhamnogalacturonans
Rha	: Rhamnose
RI	: Refractive index
SDS	: Sodium Dodecyl Sulfate
SDS-PAGE	: Sodium Dodecyl Sulfate - PolyAcrylamide Gel Electrophoresis
SEC	: Size exclusion chromatography
Ser	: Serine
TEM	: Transmission Electrons Microscopy
TFA	: Trifluoroacetic acid
TFE	: 2,2,2-Trifluoroethanol
THAP	: 2',4',6'-Trihydroxyacetatophenone monohydrate
Thr	: Threonine
TLC	: Thin-layer chromatography
TMAH	: Tetramethylammonium Hydroxide
TMCS	: Trimethylchlorosilane
TEMED	: N, N, N, N-tetramethylethylenediamine
Tris	: Tris(hydroxymethyl)aminomethane
Tris-HCl	: Tris(hydroxymethyl)aminomethane hydrochloride
Tyr	: Tyrosine
Val	: Valine
v/v	: volume/volume
w/v	: weight/volume
Xyl	: Xylose

# Introduction

Art is all around and is a tangible legacy left by artists, craftsmen and designers for us and future generations. Complementary to art historical research, the ability to identify chemical components in complex art materials, has not only provided important information on the making of the works of art, but has also been crucial in addressing questions of authenticity and the development of appropriate conservation strategies.

In the past decades, great advances have been made in the identification of natural organic compounds in museum artefacts. Materials such as oils and fats, waxes, proteinaceous compounds, plant resins, and polysaccharides, have been used as binders, adhesives and coatings since ancient times. In the cultural heritage field, several analytical techniques have been widely adopted for the characterization of these compounds. However, while it is often straightforward to determine the class of material (for example using infrared spectroscopy), discriminating specific materials within a class (e.g. egg protein from collagen, linseed from sunflower oil) has proved to be more challenging, especially for aged and complex samples. In addition, the objects under investigation are of high value thus the sample amount available for analysis is usually small (< 1 mg). The organic binder, which is usually degraded, represents a minimal portion of the sample (< 10%), and the presence of inorganic compounds, such as pigments, can interfere with the analytical technique.

Currently the category of plant gums is less investigated than other organic materials, largely because of their chemically complex nature. Plant gums are natural products derived from several plant species and/or extracted from the endosperm of some seeds. From a chemical point of view, gums are complex mixtures of mainly polysaccharides and polypeptide chains. Because of their adhesive properties, their application in the cultural heritage field was already known in early cultures. The most widely known and traditionally used gums are gum arabic, tragacanth and fruit tree gums (e.g. cherry gum). Locust bean gum, guar, ghatti and karaya are also believed to have been used since ancient times. In the cultural heritage field, gum identification is traditionally performed by complete acid hydrolysis of the analyte, followed by gas chromatography mass spectrometry (GCMS) analysis. However, inorganic or/and organic compounds present in the sample, can alter the gum chromatographic profile thus preventing the correct identification.

Since one of the main question of conservators and art historians concerns the identification of the art materials, the present research aims to develop and optimize a proper analytical approach for plant gums discrimination based on their chromatographic, electrophoretic and mass spectrometric profile.

In the first chapter, a brief overview of plant gums application in cultural heritage is presented and the current analytical techniques employed for gums identification in samples from works of art is provided. Furthermore, the potential structure and composition of the polysaccharide and protein component of plant gums proposed in the literature are described.

In regard to the methodological development, the thesis is divided in two main chapters.

The second chapter is focused on the identification of plant gums according to their protein component profile by means of chromatographic and electrophoretic techniques. Plant gums

samples were analyzed by both size exclusion chromatography (SEC), coupled to a UV detector, and polyacrylamide gel electrophoresis (PAGE). These two protein separation techniques also allowed to better understand the material under investigation, for example in terms of molecular weight, and to decomplexify the analytes.

In the third chapter, a novel method involving partial enzymatic digestion of plant gums, followed by analysis of the released oligosaccharides by matrix assisted laser desorption ionization mass spectrometry (MALDI-MS), is presented as a possible tool suitable for gum identification in works of art. The strategy was applied to gum arabic, tragacanth, cherry, locust bean and guar gum, and resulted in the possibility to distinguish the gums by their characteristic mass spectrometric profile or “saccharidic mass fingerprint”. In addition the method was tested on an aged gum arabic sample and on fresh and old watercolors.

# Chapter I

## Overview of plant gums: plant source, applications in artworks and chemical structure

### Abstract

This first chapter is intended to provide a quick overview of plant gums, their plant source and their use in the cultural heritage field as adhesive and binders. A short review of the current analytical techniques used for plant gums identification in samples from works of art is given and pros and cons of each methodology are highlighted. In addition, the potential chemical structure and composition of several plant gums is described according to the information reported in the literature. Both the proteinaceous and the polysaccharide components are examined.





## TABLE OF CONTENTS

1. DEFINITION, PROPERTIES AND GENERAL APPLICATIONS .....	19
1.1. Gum arabic .....	20
1.2. Gum tragacanth .....	20
1.3. Fruit tree gums (Prunus species) .....	20
1.4. Gum karaya .....	21
1.5. Gum ghatti.....	21
1.6. Guar and locust bean gum.....	21
2. PLANT GUMS IN CULTURAL HERITAGE SAMPLES: APPLICATIONS AND THEIR CHEMICAL CHARACTERIZATION.....	22
2.1. The organic binding media.....	22
2.2. Employment of plant gums in art and archaeology.....	23
2.3. Study of polysaccharide based materials in art samples: state of art .....	24
3. THE PROTEINACEOUS COMPONENT OF PLANT GUMS .....	28
3.1. Gum arabic .....	28
3.2. Features of other plant gums protein fraction .....	32
1.2.1. Protein content.....	32
1.2.2. Amino acid composition .....	32
1.2.3. Molecular weight.....	34
4. THE POLYSACCHARIDE COMPONENT OF PLANT GUMS.....	35
4.1. Gum Arabic .....	35
4.2. Gum Tragacanth.....	38
4.3. Fruit tree gums (Prunus species) .....	39
4.4. Guar and locust bean gum.....	40
4.5. Ghatti gum.....	41
4.6. Karaya gum .....	43
5. CONCLUSIONS .....	45
Annex .....	46
REFERENCES.....	47

## List of figures

Fig. 1. Schematic representation of the different arabinogalactan-protein complex tertiary structure (A) the “wattle blossom” model; (B) the “twisted hairy rope model. ....	30
Fig. 2. Schematic representation of the AGP fraction proposed by Mahendran <i>et al.</i> (2008).....	31
Fig. 3. Schematic representation of gum ghatti structure. ....	34
Fig. 4 One of the first proposed structure of gum arabic polysaccharide component. ....	35
Fig. 5. (A) Schematic structure of gum arabic arabinogalactan protein; (B) zoom of the polypeptide chain with the O-linked arabinogalactan polysaccharide; (C) proposed structure of gum arabic ( <i>A. Senegal</i> species) polysaccharide core. ....	36
Fig. 6. Scheme of the possible gum tragacanth fractions. ....	38
Fig. 7. Main structure of tragacanthic acid suggested by Aspinall <i>et al.</i> 1963.....	38
Fig. 8. Three suggested components of cherry gum polysaccharide moiety.....	39
Fig. 9. Proposed chemical structure of: (A) guar gum and (B) locust bean gum polysaccharide moiety. ....	41
Fig. 10. Proposed structural features of gum ghatti. ....	42
Fig. 11. Proposed structure of gum ghatti: (a) fraction F80; (b) soluble fraction.....	42
Fig. 12. Proposed structures of the two main regions (I and II) of the polysaccharide chain of Karaya gum. ....	43

## List of tables

Tab. 1. List of plant gums and their botanical sources. ....	19
Tab. 2. Classification of the main natural organic binders according to their composition .....	22
Tab. 3. Comparison of the monosaccharide profile (%) of raw gums.....	25
Tab. 4. Amino acid composition (nmol/mg gum dry wt) of <i>Acacia Senegal</i> .....	29
Tab. 5. Protein content of several plant gums and corresponding references. ....	32
Tab. 6. Amino acid composition (%) of raw gums exuded from <i>Prunus</i> species .....	33
Tab. 7. Amino acid profile of gum karaya by GC-MS .....	33
Tab. 8. Summary of plant gums molecular weight and corresponding analytical technique used for its identification. UV (ultraviolet); MALLS (Multi angle laser light scattering); DRI (differential refractive index), RI (refractive index); LALLS (low-angle laser light scattering); DP (differential-bridge viscosity). ....	34
Tab. 9 Polysaccharide types detected in several plant gums. ....	44

## 1. DEFINITION, PROPERTIES AND GENERAL APPLICATIONS

The term “gum” is generally used to describe a group of naturally occurring non-starch polysaccharides coming from various sources as plant cell walls, seeds, trees or shrubs exudates and seaweed extracts [1-3]. Gums belong to the more general category of hydrocolloids [4] and are usually referred as water-soluble gums since they have the property to form viscous adhesives, jellies or pastes by absorption or dispersion in water. Plant gums are precisely water-dispersible and not always water soluble in the scientific sense, but the designation “water-soluble gums” is used in order to distinguish these plant products from other resinous substances, rich in terpenoids, which are erroneously commonly called “gums” (e.g. gum dammar, gum copal, gum Congo, Manila gum etc.) [5]. Plant gums are mainly polysaccharides but, as described in the following chapters, the presence of a small but significant protein component associated to the polysaccharide was reported in the literature [6,7]. Gum structure and composition, such as molecular weight, amino acid composition, monosaccharide composition and sequence, configuration and glycosidic linkages position, have a direct influence on the gum rheological and physical properties [8]. Due to their adhesive properties and their ability to form gels, stabilize emulsion, foams and dispersions, water-soluble gums have been used for various applications: in the food industry as thickening and gelling agents [9], dietary fiber and modifier of texture and organoleptic properties of food [3], in cosmetic, textile and biomedical/pharmaceutical field as binders and for drug delivery [10,11].

Among the numerous plant gums which have been defined/classified by J.N. BeMiller [2] and recently reviewed by Mirhosseini et al. [12], attention is focused on the ones extracted from the endosperm of plant seeds and obtained as exudates of trees trunk and shrubs (Tab. 1). In the last case, a large number of plant species are capable of producing gums as a result of the protection mechanism against mechanical or microbial injury. Gums are produced by degradation of certain cells or plant tissues. Due to gum swelling, the cortex of the plant may be broken and an aqueous gum solution is exuded through the split. This solution, in contact with air and sunlight, dries to form a glass-like nodule of hard gum [13].

**Tab. 1.** List of plant gums and their botanical sources.

<b>Gum</b>	<b>Genus</b>	<b>Plant source</b>
Arabic	<i>Acacia senegal</i>	Tree or shrub exudates
Tragacanth	<i>Astragalus</i>	Shrub exudate
Karaya	<i>Sterculia urens</i>	Tree or shrub exudates
Ghatti	<i>Anogeissus latifolia</i>	Bark exudates
Guar	<i>Cyamopsis tetragonolobus</i>	Seeds endosperm
Locust bean (carob)	<i>Ceratonia siliqua</i>	Fruit pod endosperm
Fruit gums	<i>Prunus</i>	Tree or shrub exudate

## 1.1. GUM ARABIC

Gum arabic, also known as acacia gum, has been defined by the Joint FAO/WHO (Food and Agricultural Organization of the United Nations/World Health Organization) Expert Committee for Food Additives (JECFA) as: “a dried exudate obtained from the stems and branches of *Acacia Senegal* (L.) Willdenow or *Acacia Seyal* (Fam. Leguminosae)” [14]. Among the almost 900 species of *Acacia*, traditionally the gum is obtained from the species of *Acacia Senegal*, but the commercial one is usually collected from other *Acacia* species such as *A. Seyal* and *A. Polyacantha*. *Acacia* trees are found in Australia, India and America but they mainly grow across the so called “gum belt of Africa” at the border of Sahara desert, from the western part of Senegal to Somalia [15]. Gum arabic exudes from the tree as a sticky substance after tapping the trunk during dry season. The liquid dries to form hard nodules which are picked up manually, cleaned and graded. After a previous manual cleaning, gum can be subjected to multiple processing steps which involve: mechanical treatments, purification by dry or in-solution methods and agglomeration [16].

Gum arabic has the highest solubility among all the other gums and for this reason it has multiple applications such as in food (e.g. confectionery, flavour encapsulation, emulsifier) and pharmaceutical industry, cosmetics, lithography, textile plus other applications as in ink and pigment manufacture, ceramics and polishes [1,17-19]. As described later (section 2.2), gum arabic is maybe the most ancient and well known gum and its application goes back to about 5000 years ago during the Egyptians era [20]. Since that time, after its introduction in Europe by some Arabian ports, gum arabic has become a material of great interest in the market because of its broaden properties [21-24].

## 1.2. GUM TRAGACANTH

Gum tragacanth was defined by JECFA as “a dried exudation obtained from the stems and branches of *Astragalus gummifer* Labillardière and other Asiatic species of *Astragalus* (Fam. Leguminosae)” [25]. Shrubs of *Astragalus* species are native of semi-desert regions of western Asia and southeastern Europe, and specifically in the highlands and deserts of Turkey, Iran, Iraq, Syria, Lebanon, Afghanistan, Pakistan and Russia. Even if tragacanth gum exudes spontaneously, the process is accelerated by making incisions on shrubs lower stems, roots and branches from which the gum exudes and hardens as curled ribbons or flakes. After collection in the hottest summer months, the gum is sorted in different grades and then enters the market [5,18].

Gum tragacanth is not completely soluble. Due to its properties [26-28] such as high acid stability, the gum is widely used in the food industry in acidic conditions as stabilizer for sauces and dressings, it helps lowering the interfacial tension, it acts as a thickener and provides texture to different food products. Gum tragacanth is also used in pharmaceutical field as suspending and stabilizer agent, in the glazing process of ceramic, as stabilizer in insecticides, in paint as binder and in the textile industry [18,29,30]. Gum tragacanth is maybe, after gum arabic, one of the most important commercial water-soluble gums.

## 1.3. FRUIT TREE GUMS (PRUNUS SPECIES)

Cherry, peach, plum, apricot and almond gums are all exudates of several *Prunus* species which belong to the *Rosaceae* family. The gums are respectively exuded from: *P. cerasus* (*P. avium* and *P. virginiana*), *P. persica*, *P. spinosa*, *P. armeniaca* and *P. dulcis*. The gum exudation from these plants trunk and fruits is caused by a disease (gummosis) which takes place mainly after mechanical injury followed by microbial attack of *Polyporus* sp. [31].

#### 1.4. GUM KARAYA

Gum Karaya, also named sterculia gum, has been defined as "a dried exudation from the stems and branches of *Sterculia urens* Roxburgh and other species of *Sterculia* (Fam. Sterculiaceae) or from *Cochlospermum gossypium* A.P. De Candolle or other species of *Cochlospermum* (Fam. Bixaceae)" [32]. Trees of *S. urens* are indigenous to central and northern India but other species such as *S. villosa* and *S. setigera*, coming respectively from Pakistan and some African countries (Senegal, Mali), are employed. Tree trunks are incised or tapped; once dried the exuded irregular tears are collected mainly from April to June before the monsoon season, cleaned and sorted for the world market [5,18]. Karaya gum is actually the least soluble among the gums; it doesn't dissolve in water but it simply swallows. Due to its rheological properties [33-35], karaya gum has been employed in food industry as stabilizer and modifier of food texture, colour and appearance; in the pharmaceutical industry as adhesive, as thickening agents in textile printing techniques and in paper industry [18,29,30,36]. Differently from gum arabic, the first application of karaya gum is pretty recent and it goes back to around 100 years ago when it began to be used as adulterant for tragacanth gum since their properties similarity.

#### 1.5. GUM GHATTI

Gum ghatti, or Indian gum, is a dried exudate of the bark of *Anogeissus latifolia* (Fam. Combretaceae), a tree native of India and Sri Lanka. The gum appears as glassy nodules and, after collection, it is generally subjected to different processes that can lead to a final refined powder called GATIFOLIA SD, which is completely soluble [37]. Gum ghatti was first introduced around 1900 as a substitute for gum arabic, but due to its emulsifying properties [38], its application is increasing mainly in the food and pharmaceutical industry, especially in regard to the refined product [39-41].

#### 1.6. GUAR AND LOCUST BEAN GUM

Guar gum is recovered from the endosperm of *Cyamopsis tetragonolobus* (Fam. Leguminosae) seeds following different process steps. This plant has been grown for years in the Indian subcontinent where the final product was used for human and animal nutrition. Nowadays guar gum is widely used in food, pharmaceutical, cosmetic and textile industry [11,42].

Locust bean gum (LBG), or carob gum, is obtained from the ground endosperm of *Ceratonia Siliqua* seed (Fam. Fabaceae). These trees grow semi arid areas in the Mediterranean basin and the pods are harvested once a year. Locust bean gum forms very viscous aqueous solutions [43] and it is widely applied in food products as binding, gelling, thickening and suspending agent [11,42].

## 2. PLANT GUMS IN CULTURAL HERITAGE SAMPLES: APPLICATIONS AND THEIR CHEMICAL CHARACTERIZATION

### 2.1. THE ORGANIC BINDING MEDIA

Besides the distinct category of organic colorants, natural organic materials such as drying oils, waxes, resins, polysaccharide and proteinaceous compounds have been used in objects manufacture since ancient time as binding media, adhesives and coatings. Among the different applications, all these materials have been used as binders.

When considering a painting, its structure generally consists of different layers which include: the support (e.g. canvas, panel, plaster), a ground layer to create a smooth surface that can be painted upon (e.g. gypsum/chalk and glue) [44], one or more paint layers and, eventually, final varnish/coating layer/s for either aesthetic or protective purposes.[45]. The two main components of a paint layer are the pigments, which provide the color, and the organic binding media. Since pigments are by their nature insoluble in the medium, it is the role of the organic material to bind the pigment particles together and adhere them to the support. Therefore the binder must be characterized by specific physical-chemical proprieties such as: stability, adhesiveness, viscosity and other optical properties as transparency and absence of coloration [46]. Besides the synthetic polymers, which are the binding media of choice for most of paints since the 20<sup>th</sup> century, a number of naturally occurring materials have been used as both binders and, some of them, coatings. The organic binders of the tradition can be classified in four main groups according to the organic compounds they consist of. These were the only media available from antiquity and they continue to be used by contemporary artists, thus revealing their durability and versatility [47]. A List of the four categories of binders is reported in Tab. 2 [48-50].

**Tab. 2.** Classification of the main natural organic binders according to their composition

Binder category	Materials
Proteinaceous media	This category includes: egg, animal glues (e.g. rabbit, fish) and casein. Egg consists of different proteins such as ovalbumin, lysozyme and phosphorylated proteins. Both egg yolk and egg white (glair) were used. However, egg yolk was preferred because the presence of ~20% lipids allowed the formation of an hydrophobic and therefore more resistant film. Glues are obtained by boiling animal skin, bones and connective tissues and they mainly consist of collagen, a fibrous protein. Casein is a phosphorylated protein complex obtained by acidification of milk. It was mainly used in mural paintings.
Polysaccharide media	This group includes a wide number of plant gums which are exuded from several plant species or extracted from seeds. The most known and used are gum arabic, tragacanth gum and gums exuded from fruit trees (cherry, apricot, almond and plum gum). Besides these material also honey, plant and fruit juice and, in more circumscribed application, starch (e.g. Japanese prints), were used.
Fatty acid containing	<u>Drying oils</u> Drying oils are glycerol esters of fatty acids. These esters are

media	<p>characterized by an high amount of unsaturated fatty acids which ensure the oil to dry by polymerization processes. The most frequently used oils were: linseed oil (pressed from the seeds of flax plant), poppy-seed oil, walnut oil and sunflower oil.</p> <p><u>Waxes</u></p> <p>Waxes are a complex mixture of esters of fatty acids with monohydroxy high molecular mass alcohols, free fatty acids and hydrocarbons. Natural waxes can be of animal (beeswax), vegetable (carnuba wax), or mineral (montan wax) origin. Beeswax was the most used mainly as waterproofing agent.</p>
Resinous media	<p>Resins are water-insoluble substance that exude from several tree species and they are mainly composed of terpenoid compounds. The most know resins are sandarac, Venice turpentine, amber, copal, mastic and dammar. They were mainly used as picture varnish and in glazes in oil paintings.</p>

## 2.2. EMPLOYMENT OF PLANT GUMS IN ART AND ARCHAEOLOGY

Beyond the wide application in food, pharmaceutical, lithography and cosmetic industry as thickening agent or emulsion stabilizers, some plant gums have been widely used in the field of cultural heritage since the third millennium B.C., time of ancient Egyptians, throughout the middle ages and they are still highly used by contemporary artists.

Use of gums as food traces back in the Stone Age [51], but the wider application of polysaccharide materials is dated to ca. 2600 B.C. during the Fourth Dynasty of Egyptian time [20]. Ancient Egyptian inscriptions refer to plant gums, and precisely to gum arabic, as “kami”. The polysaccharide material was reported to be used as a binder and adhesive for mineral pigments on cartonnage and in the practices of mummification [20,52]. During the last centuries, arabic, tragacanth and fruit trees gums such as peach, apricot, plum, cherry and almond, have been found in mural paintings in Greece [53], Asian wall paintings [54], paintings on silk and in manuscript illumination from the middle ages [50]. Less frequently they were employed for metallo-gallic inks preparation [55] and, on occasion, as sizing materials and as varnish/coating [56]. However, due to their high water solubility and swelling, plant gums have been mostly employed as binding agents and the major area of application lies in watercolors.

The historical investigation confirmed that the primary gum employed in watercolor painting has been, and continues to be, gum arabic. Ancient treaties reported how extensively plant gums have been used in Europe. In the *Schedula diversarum artium* of Theophilus, early XII century, there is a note on the use of gums in place of sundried oil. In addition Cennino Cennini in its “The craftsman’s handbook”, early XV century, spoke about the use of a egg white/gum mixture as a medium for “mosaic gold” [57], which also appeared in the Naples Manuscript, late XIV century. Documents showed that gums were frequently mixed with other materials in order to modify their physical/chemical properties and this procedure has been extensively used in Europe for centuries [45]. Historical treatises reported the use of different materials to be added to gum-based paints: proteinaceous materials (animal glues, isinglass, egg and casein), linseed oil, honey, candy sugar

and various preservatives (e.g. essential oils). Honey and sugar, in its different forms (brown, white etc.), have been added to gum-based watercolors as humectants, plasticizers to prevent cracking, and to impart strength and transparency. In modern practice glycerine took the place of honey adding [49]. Gum mixtures were also suggested (e.g. same proportion of gum arabic and tragacanth), mostly in order to decrease surface tension of gum arabic and to reduce its viscosity [58].

Among all the natural gums available, gum arabic is the best known and used one. The gum found its way into Europe through various Arabian ports and it was called gum arabic because of its place of origin [18]. Ormsby *et al.* [58] reported an interesting investigation about the use of natural gums in british watercolors preparation from the XVIIIth to the early XXth century. Treatises revealed that arabic gums of different geographical origin (e.g. from Iran and Senegal) were used for distinct purposes. They have been employed for lighter and darker colors according to their natural coloration, and they were usually made up into 1:2 (strong gum) or 1:3 (weak gum) w/v solutions in water. The use of gum arabic as a temporary varnish, particularly for darker colors, has also been mentioned [58].

Gum tragacanth is the second most used gum in art and it was described by Theophrastus in the third century B.C. Its physical properties, such as film strength and partial insolubility, made it available for certain applications. Tragacanth gum has been mainly used in the manufacture of pastels, as a medium for painting on linen and it was the sole binder for “scene painting” [59]. It was also mentioned as a good medium for miniatures, even better after addition of borax (sodium borate). Gum tragacanth solutions are usually 5-10% w/v in water and it was reported to be the gum preferred by some artists as it was more sticky when dried [58].

Cherry, apricot, almond, peach and plum gums have found use in Europe for centuries since they come from indigenous trees. They have been mainly used in medieval manuscript illumination and mixture of cherry gum with tragacanth and isinglass has been reported, indicating that cherry may have been used for tempera paints [59].

Karaya gum has been used as a cheaper substitute for gum tragacanth and, apart those mentioned above, also ghatti, guar, locust bean gums were used and they were important materials mainly in the Indian subcontinent [50].

### **2.3. STUDY OF POLYSACCHARIDE BASED MATERIALS IN ART SAMPLES: STATE OF ART**

As already mentioned in the first chapter, plant gums are commonly used in food industry, cosmetic, textile, biomedical/pharmaceutical field and they have been used in art and archaeological objects as binders, adhesives and coatings. Though the analytical methods used for their identification are common in the different fields, specific challenges are encountered due to the different type of substrates (e.g. food, drugs, art samples). Therefore, the focus of this section is to provide an overview of the analytical techniques used for plant gums identification in samples from works of art. The different methods of analysis range from non-destructive techniques to the most know chromatographic methods, from capillary electrophoresis to enzyme-linked immunosorbent assay techniques (ELISA). The following paragraphs intend to highlight the information that is possible to obtain with these techniques, as well as their limitations.



Fourier Transform Infrared Spectroscopy (FTIR) [60,61] and Raman microscopy [59] are invasive non-destructing techniques and they have been used for the study of polysaccharide based materials. The advantages of these methods rely on the fact that: the sample is not destroyed, thus allowing its use for additional analytical investigations, and direct analysis can be done on certain objects without sampling (e.g. illuminated manuscripts). Spectroscopic techniques allowed the distinction of plant gums from the other organic binding media (drying oils, waxes, proteins and resins) [61], and of gums from starch [59]. However the major downside of these methods is that different plant gums (e.g. arabic, tragacanth and cherry gum), analyzed by Raman and FTIR, give similar spectra, thus hindering specific discrimination among them.

The most frequently employed technique in the analysis of polysaccharide based materials is chromatography [49,62,63]. This method evolved from thin-layer chromatography (TLC) [64] to the most known gas chromatographic techniques (GC) coupled with mass spectrometry (MS). A conventional analytical approach consists of three main steps: acid hydrolysis is performed in order to break the glycosidic linkages, followed by derivatized of released monosaccharides and GC-MS analysis [54,65-67]. Plant recognition is then based on quantitative assessment of the relative proportions of monosaccharides, or on the qualitative determination of the absence/presence of characteristic monosaccharides. The steps can be altered in certain cases to obtain better results. Acid hydrolysis can be replaced by methanolysis, followed by silylation derivatization, for best results in recovery of aldose and uronic acids [68-70]. However, the obtained chromatograms are highly complex because of the presence of multiple peaks for each monosaccharide, corresponding to the different isomers. This complexity implies a loss in sensitivity and therefore monomers quantification cannot be reliable. Due to these limitations, new GC-MS derivatization procedures were developed to eliminate monosaccharide multiple peaks and increase chromatographic separation. Pitthard *et al.* [71] developed a method to transform the monosaccharides into oximes. Specifically, methyloxyme acetates derivatives were obtained by subsequent derivatization with acetic anhydride. The method was successfully applied for the analysis of polysaccharides materials in different art objects [72,73]. Another method was proposed by Bonaduce *et al.* [74] who introduced a two steps procedure. After microwave assisted acid hydrolysis, monosaccharides and uronic acids were initially transformed into the corresponding diethyl-dithioacetals and lactones, and then acidic and alcoholic moieties were trimethylsilylated. Each gum was characterized by a specific monosaccharide composition (see Table 3, data obtained from [74] and (\*) [75]). Therefore, if only one type of gum was present in an unknown sample, it was possible to identify it according to the absence/presence of specific monosaccharides. For example karaya gum does not contain arabinose; fucose is present only in tragacanth (7.4%) so it can be considered a marker of this gum, and xylose is specific of gums extracted from fruit trees.

**Tab. 3.** Comparison of the monosaccharide profile (%) of raw gums.

	Ara	Xyl	Gal	Rha	Man	Glc	GlcA
Ghatti	47.0	-	36.5	3.5	2.5		11.0
Karaya	-	-	64.0	25.5	-		4.0
LBG	1.5	-	17.5	-	81.0		-
Guar	2.5	-	34.5	-	63.0		-
Arabic	37.2	-	44.7	11.2	-		7.0
Tragacanth	34.7	14.9	10.8	2.8	-	11.5	3.5
Cherry	35.1	6.6	37.0	2.8	5.5	-	13.0

Peach	34.0	6.1	38.0	2.8	5.1	-	14.0
Plum	36.0	6.9	36.0	2.8	5.3	-	13.0
Almond*	46.8	10.9	35.5	0.8	0.8	0.2	6.0

The method was found to be highly reproducible since only one peak was obtained for each monosaccharide. However, the procedure was time consuming because a second derivatization step was necessary. In addition, if a mixture of gums is present, results can lead to misinterpretations as the sugar profile will be altered by the sugar contribution of each gum.

When dealing with samples from works of art, another aspect has to be considered. Gums are rarely found alone but usually they are mixed with many different materials (organic and inorganic). Therefore additional procedure was needed that would allow gum identification even when present in a complex mixture. A first step toward this goal has been recently made. A new analytical procedure based on GC-MS for the determination of glycerolipids, natural waxes, terpenoid resins, proteinaceous and polysaccharide materials in the same paint micro sample was developed [76]. The method involved (i) microwave acidic hydrolysis, (ii) cleanup step to remove inorganic materials and (iii) derivatization step consisting of a mercaptalation followed by a double silylation. The same author also proposed a new model based on GC-MS analysis for the reliable identification of saccharide binders in paintings based on the evaluation of markers that are stable to ageing and are unaffected by pigments interference [77]. The effect of the simultaneous presence of inorganic materials, proteinaceous binders and environmental contaminants was investigated and results showed how these compounds were responsible for modifications in plant gums monosaccharide profile. This result highlighted how discrimination of different gums based on their monosaccharide composition can be influenced by factors such as the age and the presence of other materials the gums are found with.

Due to the mentioned limitations encountered with GC-MS analysis, an alternative approach based on pyrolysis coupled to GC-MS (Py-GC/MS) was developed and applied to different real samples of historical inks, watercolors and wall painting samples [52,55,78-83]. This technique consists in the thermal decomposition of the compound under investigation and, on the contrary of the traditional GC-MS analysis, it is based on online derivatization. Py-GC/MS requires a small amount of sample and it is less time consuming since any preliminary treatment, except the addition of the derivatizing reagent prior to pyrolysis, is necessary. At the beginning of pyrolysis development for sugar analysis, experiments were carried out without sample derivatization [78,84,85]. Though the procedure was fast, results showed how obtained pyrograms were not informative enough for gums identification. For this reason, as for GC-MS, attention was focused on derivatization of the sugars to improve the pyrolysis profile.

In the literature two different approaches for derivatization of the hydroxyl groups are reported. The OH groups can be converted either into methylated derivatives by the use of tetramethylammonium hydroxide (TMAH) [86-89], or into silylated derivatives. In this last case both hexamethyldisilazane (HMDS) or *N,O*-bistrimethylsilyltrifluoroacetamide (BSTFA) can be used with or without trimethylchlorosilane (TMCS) [84,85]. Both derivatization methods presented some disadvantages, thus preventing the correct gum classification. The missing detection of uronic acids with methylation [86,90], and the underivatization of some products with silylating mixtures [85], led to a partial loss of structural information.

Among the chromatographic techniques, methods based on liquid chromatography (HPLC) represent a more direct way for plant gums analysis since any derivatization step is required. However HPLC have been mainly employed in biochemistry research and food industry analysis than in the conservation science field. An application in the cultural heritage was proposed by Colombini *et al.* [53] who introduced a procedure based on the classic acid hydrolysis, assisted by microwaves, followed by cation-exchange clean-up procedure, analysis by high-performance anion exchange chromatography (HPAEC) and final evaluation of analytical results by chemometric tools such as Principal Components Analysis (PCA). The method was successfully applied to raw gums, watercolors, painting mockups and real painting samples and allowed plant gum identification at a level of a few micrograms. However, to our knowledge, this approach was not further investigated.

Besides the chromatographic techniques, a complete different methodology based on capillary electrophoresis (CE) was introduced. This technique is suitable for the analysis of polysaccharides since they possess weakly basic OH groups and therefore have a different mobility when an electric field is applied. As in chromatography, gums were completely hydrolyzed in the corresponding monosaccharides by a strong acid, (e.g. trifluoroacetic acid) followed by detection of the monomers. Once the sugar profile was obtained, gum identification was performed, as with GC-MS analysis, according to a “decision scheme” based on the presence/absence of certain monosaccharides. Even if the technique had the advantage that any derivatization step was required, results showed that the electrophoretic method was not sufficiently sensitive and resolute [91,92].

The above mentioned techniques allowed the identification of the main different types of gums (arabic, tragacanth and cherry gum) and they are all based on the evaluation of the monosaccharide composition. The main problem with these techniques is that the obtained results do not allow specific identification if a mixture of polysaccharide materials is present, as the sugar composition changes depending on the relative proportion of each gum. In order to overcome this limitation, other techniques were developed such as enzyme-linked immunosorbent assay (ELISA) which is an immunological based technique. Even if further studies are necessary, ELISA was shown to offer great potential in identifying different proteinaceous binding media and it has also been employed for gums identification in art samples using polyclonal primary antibodies [93].

### 3. THE PROTEINACEOUS COMPONENT OF PLANT GUMS

#### 3.1. GUM ARABIC

Many studies have been undertaken on the chemical composition and structure of this gum, since its critical importance mainly in the food industry. In the past, gum arabic was thought to be composed entirely of carbohydrates, thus first experiments focused on the elucidation of the carbohydrate structure and mode of linkage [94]. However, the presence of nitrogen, a possible indication of proteins, was reported in the 60s [95], and more in-depth analysis began to be run in the 80s. Akiyama *et al.* [96] reported the amino acid composition of the intact gum, performed by complete hydrolysis, and results showed that the main amino acids were Hydroxyproline (Hyp), Serine (Ser) and Aspartic acid (Asp). Furthermore, they discovered that gum arabic precipitated with Yariv reagent, a group of chemical compounds that specifically bind and precipitate plant arabinogalactan proteins (AGPs) [97], demonstrating for the first time that the gum resembled the AGP composition/structure. Due to the similarity of the gum to this group of plant proteoglycans, the author further investigated the gum arabic carbohydrate-protein linkage. In order to achieve this goal, gum arabic was hydrolyzed with barium hydroxide and sodium borohydrate, followed by subsequent fractionation by gel filtration chromatography. Analysis of the amino acid and monosaccharide composition, before and after hydrolysis, were indicative of both hydroxyproline- and serine-carbohydrate bond.

Following this study and the discovery of about 2% polypeptide in acacia gum samples [96,98], in the last 30 years research of plant gums has made great strides on the investigation of the proteinaceous component and its influence on the structural and functional properties of the gum [99]. Anderson and McDougall [100,101], working on Smith degradation, a three steps procedure that allow to selectively degrade the polysaccharide to a smaller polysaccharide or to oligosaccharides, and amino acids analysis, deduced that gum arabic structure consisted of a protein core, with hydroxyproline as major amino acid. Vandeveld and Fenyo [102] studied the gum by applying size exclusion chromatography. The presence of two distinct fractions was proposed, thus demonstrating gum heterogeneity. Results revealed a low molecular weight protein-deficient fraction, which constituted ~70% of the total, and an high molecular weight protein-rich fraction which was described as an arabinogalactan protein complex. Connolly *et al.* [103,104] and Duvall *et al.* [105] performed a proteolytic digestion of gum arabic using pronase, followed by gel permeation chromatography (GPC) analysis. Results showed a decrease in the molecular weight of the gum, while no structural modification was observed. It was therefore concluded that the change in MW was due to the hydrolysis of the polypeptide fraction. This provided evidence that the gum had a “wattle blossom structure”, which had already been suggested as a model for describing AGP tertiary structure [106]. As showed in Fig. 1A (from Tan *et al.* [107]), this model consists of a proteinaceous core surrounded by several carbohydrate blocks. It was concluded that the gum had a main polypeptide chain with covalently linked carbohydrate blocks of approximately  $2 \times 10^5$  molecular mass. Thus the final structure was kind of a spheroidal macromolecule where polysaccharide chains were folded around the protein core.

Further studies on gum arabic structure were undertaken by Randall *et al.* [108]. The author fractionated the gum by hydrophobic affinity chromatography (HAC) and isolated three main fractions with significant differences in their size and protein content. In general the major fraction

(~90% of the total weight of the gum) was described by the author as an highly branched block-type structure arabinogalactan (AG) with a MW of  $\sim 2.5 \times 10^5$  and small amount of protein (< 1%). A second component, which represented ~10% of the total, was described as an AGP complex with a MW of  $1-2 \times 10^6$  and a significant 11.8% content of protein. The third and newly reported minor fraction (~1% of the total) had a molecular weight of  $2 \times 10^5$  and the higher protein content (~50% w/w), thus consisting of one, or two, glycoproteins (GP). While HPLC analysis of the ratio of the monosaccharides released by acid hydrolysis (sulfuric acid) was reported to be fairly the same among the different fractions, significant differences in their relative amino acid composition were observed. While fractions I and II had a similar composition with hydroxyproline and serine being the most abundant, the third fraction had a very different profile with aspartic acid predominating (see Tab. 4 [108], analysis performed by HPLC after hydrochloric acid hydrolysis).

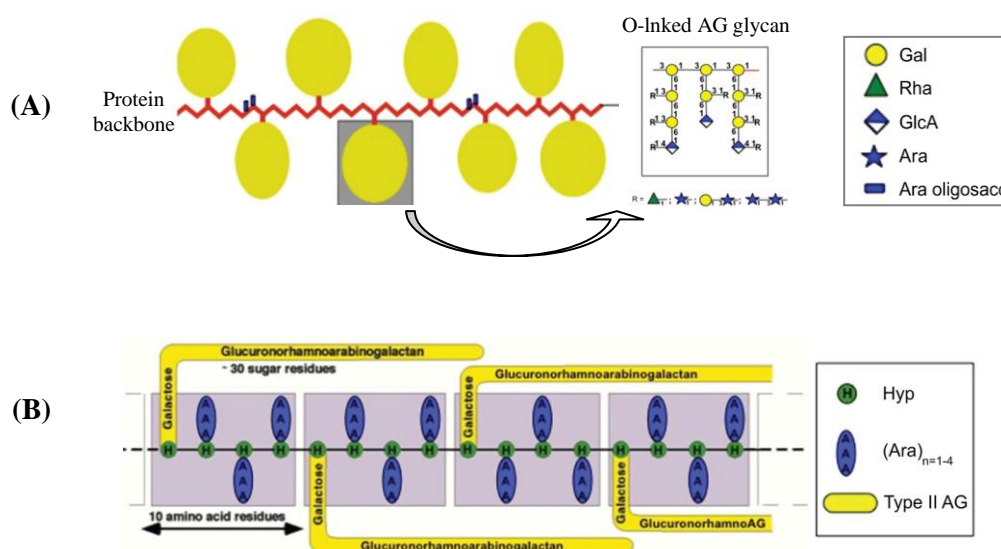
**Tab. 4.** Amino acid composition (nmol/mg gum dry wt) of Acacia Senegal sample and its fractions.

	Whole gum	Fraction I -AG-	Fraction II -AGP-	Fraction III -GP-
Hydroxyproline	41.0	5.8	291.0	228.0
Aspartic acid	12.0	2.1	30.0	432.0
Threonine	12.0	2.6	67.0	168.0
Serine	22.0	5.4	115.0	324.0
Glutamic acid	6.2	3.1	7.2	324.0
Proline	13.0	2.9	55.0	216.0
Glycine	9.3	2.7	27.0	312.0
Alanine	4.9	1.9	7.8	192.0
Cystine	/	0.4	/	11.9
Valine	5.9	1.4	8.7	300.0
Methionine	0.3	0.08	/	6.0
Isoleucine	2.0	0.7	2.6	84.0
Leucine	12.0	2.0	51.0	300.0
Tyrosine	2.8	0.42	4.6	83.0
Phenylalanine	6.0	0.37	10.0	372.0
Histidine	8.8	1.2	51.0	132.0
Lysine	4.3	0.8	9.9	168.0
Arginine	1.5	2.2	1.2	47.0

Similar results were obtained by other authors in later works [23,108-118]. Enzymatic attack of AGP, AG and GP fractions of gum arabic by pronase, followed by GPC analysis of the single fractions before and after digestion, showed various degrees of degradation, thus demonstrating differences in their structure. According to the chromatograms, only the high molecular weight fraction (AGP), which contains about 10% protein, underwent degradation. The MW decreased from 1500000 to 250000. This suggested the presence of 5 or 6 carbohydrate blocks linked to the polypeptide chain, thus confirming the “wattle blossom-type” structure. [109]. On the contrary, both the arabinogalactan and the fraction containing the highest amount of protein, the so called GP, did not show any degradation effect.

An alternative model to the “wattle blossom”, for arabinogalactan-protein complex, was introduced by Qi et al. [119]. The author isolated the AGP fraction by preparative gel permeation chromatography and subjected it to chemical deglycosylation by hydrofluoric acid. Results showed that the fraction consisted of a Hyp/Pro-rich polypeptide backbone of around 400 amino acids. A

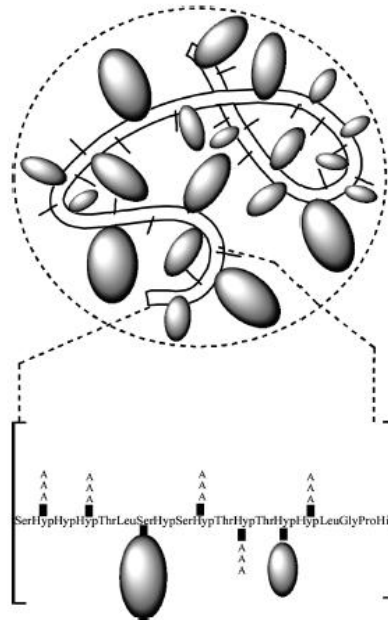
10 to 12 residue peptide motif [Hyp<sub>4</sub>Ser<sub>2</sub>ThrProGlyLeuHis] of around 7 kDa was also predicted as the basic repetitive unit of the AGP polypeptide backbone. Alkaline hydrolysis performed by Ba(OH)<sub>2</sub> allowed to conclude that the carbohydrates were attached to the polypeptide chain through hydroxyproline and a further partial acid hydrolysis of this fraction, separated by GPC, showed that carbohydrates were O-linked. By means of GPC and hydroxyproline glycosidic chromatography with a specific hydroxyproline assay analysis, performed after gum alkaline hydrolysis, the author suggested that 12.1% of the hydroxyprolines were non glycosylated, the 63.5% were linked to short-chain oligomers (~3 sugar units) and the 24.3% to carbohydrate blocks (~30 sugar residues). Based on the results, gum arabic glycoprotein shared several characteristic with arabinogalactan proteins [120], thus confirming the hypothesis of Akiyama *et al.* [96]. Investigation of the gum by transmission electron microscopy (TEM) showed that the polypeptide chain was 150 nm long and was rod shaped. Therefore it was suggested that the gum structure resembled a “twisted hairy rope”. The proposed model is reported in figure 1B (from Qi. *et al.* [119]).



**Fig. 1.** Schematic representation of the different arabinogalactan-protein complex tertiary structure (A) the “wattle blossom” model; (B) the “twisted hairy rope model”.

Following the analytical approach reported by Qi *et al.*, Goodrum *et al.* [121] suggested for the AGP fraction a 19 residue motif [Ser-Hyp<sub>3</sub>-Thr-Leu-Ser-Hyp-Ser-Hyp-Thr-Hyp-Thr-Hyp-Hyp-Leu-Gly-Pro-His]. They also deduced the glycosylation sites on the base of Hyp contiguity hypothesis [122]: six Hyp-arabinosides on contiguous Hyp residues and two polysaccharide attachment sites in clustered non-contiguous Hyp residues.

The AGP structure was also investigated by small-angle X-ray and neutron scattering combined with cryo TEM [123]. Measurements revealed a complex shape composed of many spheroidal aggregates, assigned to the polysaccharide, with a small amount of larger coils. Mahendran *et al.* [114] further investigated the structure of AGP. The innovative aspect concerned the hypothesis that the shape of the macromolecule would be a kind of spheroidal random coil. The model consisted of a folded polypeptide, carrying short arabinose side chain and carbohydrate blocks with a possible thin oblate ellipsoid structure (Figure 2, [114]).



**Fig. 2.** Schematic representation of the AGP fraction proposed by Mahendran *et al.* (2008).

Al-Assaf *et al.* [115] treated gum arabic by maturation techniques and enzymatic digestion by pronase, followed by evaluation of the changes in molecular weight by gel permeation chromatography. Results allowed the author to suggest the hypothesis that the AGP fraction was basically an aggregated fraction characterized by several AG units which are stabilized through hydrophobic associations by low molecular weight highly proteinaceous components. Therefore, the presence of just a single polypeptide chain was excluded. According to this new hypothesis, the so-called AGP fraction was described as a molecular association and it was in a middle way between the “wattle blossom” and the “hairy rope” models. The existence of assembled AG as part of the AGP component was also confirmed by Renard *et al.* [116] who analyzed the AGP fraction, obtained by hydrophobic interaction chromatography, by transmission electron microscopy (TEM) and size exclusion chromatography. Arabinogalactan-protein was supposed to be made up of two populations characterized respectively by a long-chain branching and globular shape population, the first, and short-chain branching and elongated shape population, the second. The AGP fraction was therefore described as an heterogeneous population of all possible molecular combinations of AGP, AG and GP.

In conclusion, gum arabic contains a high molecular weight fraction that resembles the arabinogalactan protein structure. It consists of a polypeptide chain to which highly branched carbohydrates are *O*-linked through hydroxyproline and possibly serine. The polypeptide chain cannot be completely degraded by proteolytic enzymes [109] and the tertiary structure was suggested to consist of a compact spherical structure.

### 3.2. FEATURES OF OTHER PLANT GUMS PROTEIN FRACTION

#### 1.2.1. Protein content

The protein content of plant gums has been calculated by means of different techniques. Colorimetric methods allowed to estimate a protein content lower than 1% for peach gum [31], while to our knowledge no values are reported in the literature for the other exudates of *Prunus* species trees (cherry, plum, apricot and almond trees). Several authors evaluated the protein content by the Kjeldahl procedure according to Anderson [124]. The general procedure implies the acid digestion of the sample and the protein content is then calculated from the nitrogen concentration, using a proper conversion factor (*F*). An amount of around 5-6% was measured for locust bean gum [125-128] but the value was demonstrated to be variable according to the area of origin of the plant [129]. Same method was used by Al-Assaf *et al.* to investigate gum ghatti [130]. It was reported that the gum had an higher content of protein (~3%) [130,131] if compared to the 2% of gum arabic [98]. A protein content of  $3.46 \pm 0.06\%$  was reported for guar gum [132]. Finally a value of ~3-4% was measured for tragacanth gum [133] and karaya gum was showed to have  $1.2 \pm 0.6\%$  protein content [134,135], which is lower than the othe gums. A summary of the plant gum protein content is reported in Tab. 5.

**Tab. 5.** Protein content of several plant gums and corresponding references.

Plant gum	Protein content %	Ref.
Arabic	2%	[96]
Tragacanth	3-4%	[133]
Peach	< 1%	[31]
Guar	3.5%	[132]
LBG	5-6%	[125-128]
Ghatti	3%	[130]
Karaya	1%	[134,135]

#### 1.2.2. Amino acid composition

The most common method for identifying the amino acid composition consists in the complete hydrolysis of the polypeptide chain in order to break the peptidic linkages and obtain the single amino acids. Analysis of the monosaccharide percentage can be then performed by HPLC or GC-MS.

A complete study of fruit tree gums was carried out by Shilling *et al.* [136]. The author performed acid hydrolysis of the plant gums, followed by GC-MS analysis of the corresponding ethyl chloroformate derivatives. The results obtained for cherry, almond and apricot gums are reported in Tab. 6.



**Tab. 6.** Amino acid composition (%) of raw gums exuded from *Prunus* species obtained by GC-MS analysis of the ethyl chloroformate derivatives .

Amino acids	Cherry	Almond	Apricot
Ala	10.9	7.7	9.9
Val	3.7	4.8	4.2
Ile	1.7	3.5	2.0
Leu	7.4	6.9	4.1
Gly	15.2	9.5	13.9
Pro	5.8	8.3	4.5
Asp	17.5	15.3	18.6
Thr	4.8	6.0	5.7
Met	-	1.9	0.6
Ser	12.0	9.7	12.9
Glu	9.9	12.8	10.4
Phe	3.3	2.8	0.8
Hyp	3.7	4.7	0.6
Lys	4.1	5.8	11.8

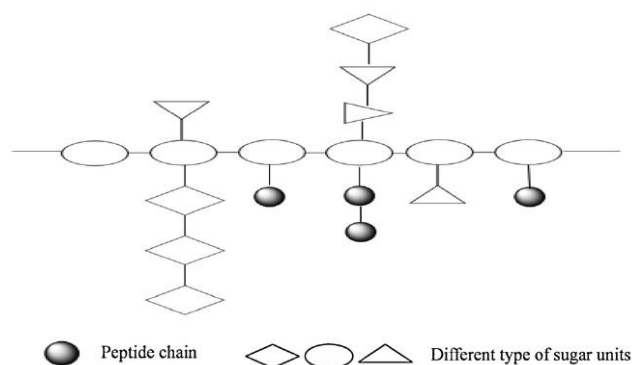
In tragacanth gum Asp, Hyp, Ser, Pro, Val and His were reported to be the major amino acids, whom proportion varies among different *Astragalus* species [133,137,138]. Acid hydrolysis followed by GC-MS analysis of karaya gum showed the major amino acids to be aspartic acid and glutamic acid (Tab. 7, from [139]), thus resulting in a significantly different amino acid composition from gum arabic for which hydroxyproline, serine and proline are present in higher percentage [140]. However, the proportion of amino acids in the gum karaya proteinaceous component is reported by some authors to vary significantly according to the botanical source and the age of the tree [35,134].

**Tab. 7.** Amino acid profile of gum karaya by GC-MS .

Amino acids	Content [ $\mu\text{g/g}$ ]
Alanine	ND
Glycine	$4.8 \pm 0.45$
Valine	ND
Leucine	$3.9 \pm 0.28$
Proline	$30.5 \pm 1.86$
Methionine	ND
Aspartic acid	$64.2 \pm 2.44$
Threonine	$25.2 \pm 1.06$
Glutamic acid	$34.2 \pm 1.44$
Tyrosine	ND
Tryptophan	ND

In regard of gum ghatti, Kang et al. [141] reported an high percentage of Asp, Gly, Lys and low percentage of Phe, Pro and Ser, indicating that gum ghatti has a different protein structure compared to gum arabic. Furthermore, the gum was partially digested by a mild acid method and a non-defined protease, followed by HPSEC. The strategy allowed to obtain several fractions with different molecular weights that were further analyzed by matrix assisted laser desorption ionization (MALDI) and nuclear magnetic resonance (NMR). These two techniques allowed the

author to better understand the linkage between proteinaceous and polysaccharide moieties. It was supposed that most of the protein was covalently linked to the backbone of the polysaccharide, instead of the side chains, and the proteins were described as N-linked through GlcNAc. Therefore a new model for gum ghatti structure was proposed. It consists of a 1,6-linked galactose backbone with several sugar side chains and proteins directly attached to the polysaccharide core (see Fig. 3, from [141]).



**Fig. 3.** Schematic representation of gum ghatti structure.

### 1.2.3. Molecular weight

The molecular weight measured for all the plant gums under investigations are summarized in Tab. 8. Data might be referred to the proteinaceous fraction with the eventual addition of polysaccharide blocks since, unfortunately, no distinction is made in the literature.

**Tab. 8.** Summary of plant gums molecular weight and corresponding analytical technique used for its identification. UV (ultraviolet); MALLS (Multi angle laser light scattering); DRI (differential refractive index), RI (refractive index); LALLS (low-angle laser light scattering); DP (differential-bridge viscosity).

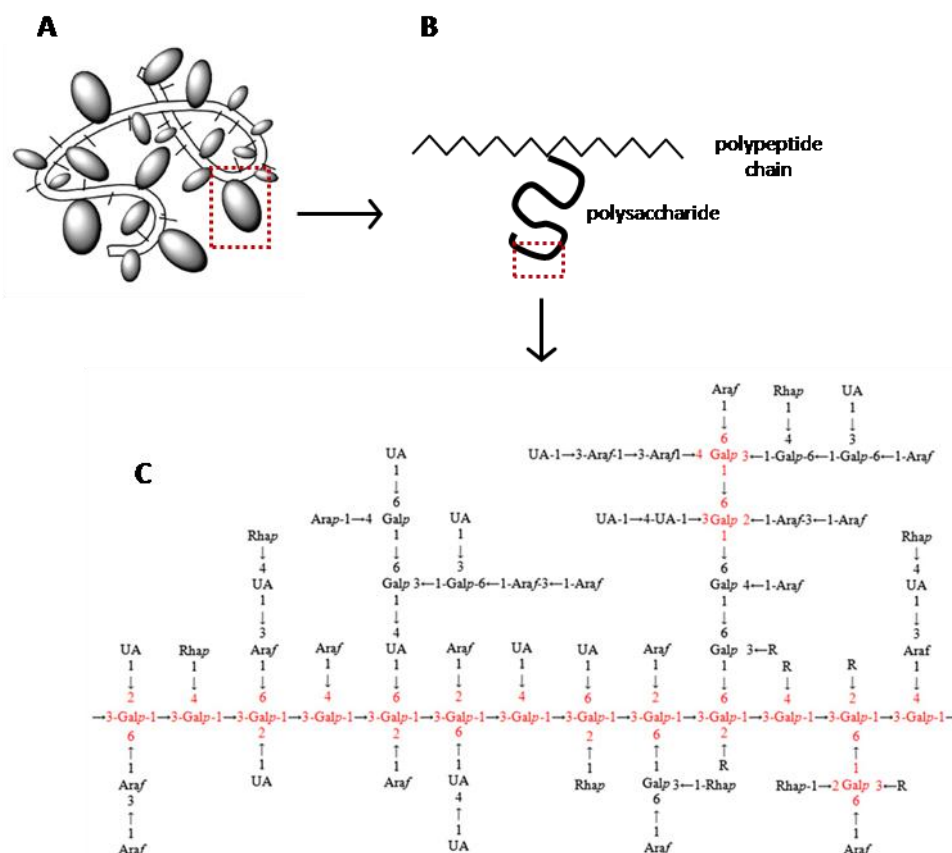
Gum	MW	Analytical technique	Ref.
Tragacanth <sup>1</sup>	$1.6 \times 10^6$	GPC-MALLS	[26]
Cherry	$1.3 \times 10^6$	GPC-DRI	[142]
Peach	$5.61 \times 10^6$	HPSEC-RI	[31]
Guar	$2.2 \times 10^6$	SEC-LALLS	[143]
LBG	$\sim 1 \times 10^6$	SEC-LS/RI/DP; colorimetric methods	[144,145]
Ghatti	$8.7 \times 10^5$	SEC-RI; UV; MALLS	[39]
Karaya	$16 \times 10^6$	SEC-RI	[35]

<sup>1</sup>: data refer to the gum tragacanth soluble fraction (tragacanthin).



Churms *et al.* [150] performed methylation analysis and molecular weight estimation of the gum products obtained after two smith degradation. Results confirmed the presence of uniform sub-units of 1,3-linked DGalp having a molecular weight of 8000 Dalton (Da), like other arabinogalactans of simpler structure. These first investigations revealed how the polysaccharide component of gum arabic could be related to the family of Type II arabinogalactans [151]. Further studies by 1D-NMR, 2D-NMR and electrospray-ionization mass spectrometry (ESI-MS) of the oligosaccharide fractions released from the gum by subsequent smith degradations, mainly based on NMR, showed how gum arabic structure was much more complicated than a classic Type II arabinogalactan since it was characterized by a variety of different substitutes [152]. The combination of the three techniques allowed the author to detect multiple disaccharides (e.g.  $\alpha$ -D-Galp-(1,3)- $\alpha$ -L-Araf), trisaccharides (e.g.  $\alpha$ -L-Araf-(1,4)- $\beta$ -D-Galp-(1,6)- $\alpha$ -D-Galp), a branched pentasaccharide ( $\alpha$ -D-Galp-(1,3)- $\alpha$ -L-Araf-(1,3)- $\beta$ -D-Galp-(1,6)- $\alpha$ -D-Galp substituted with  $\alpha$ -L-Araf) and a doubly branched heptasaccharide. These oligosaccharides were considered to represent the side-chains of the polysaccharide structure of gum arabic.

More recently Nie *et al.* investigated and compared the polysaccharide core of both *Acacia seyal* [153] and *A. senegal*, species [154]. Besides demonstrating for the first time the presence of galacturonic acid (13.66%) in the *seyal* species, they mainly revised the molecular structure of the gum arabic polysaccharide moiety by means of several NMR spectral analysis. The proposed structure is reported in Figure. 5C (from Nie *et al.* [154]).



**Fig. 5.** (A) Schematic structure of gum arabic arabinogalactan protein; (B) zoom of the polypeptide chain with the O-linked arabinogalactan polysaccharide; (C) proposed structure of gum arabic (*A. Senegal* species) polysaccharide core.

R represents one of these following residues: T-Rhap1→, T-L-Araf 1→, T-L-Arap 1→, T-UA1→, T-UA1→, T-L-Araf 1→3-L-Araf 1→, T-UA1 →4-UA1→. All the galactose and uronic acid units (UA) are in β-D form while the arabinose and rhamnose units are in α-L form. The authors reported for the first time the substitution of the main chain with residues of →2,3,6-β-D-Galp1→, →3,4-Galp1→, →3,4,6-Galp1→, in O-2 and O-4 position, besides the classical O-6. This last study demonstrated the high complexity of the gum arabic polysaccharide. The proposed structure was the one considered as a possible model for the study performed in this research.

As previously reported, gum arabic polysaccharide component was described to resemble the structure of arabinogalactan polysaccharides. Arabinogalactans are described as plant cell branched polysaccharides containing mainly arabinose and galactose. They can be classified in two groups according to their structure: type I and type II arabinogalactans. A third group, consisting of polysaccharides with arabinogalactan side chains, has been also suggested [151]. Type I AGs usually bind to pectins, which consist of a homogalacturonan and rhamnogalacturonan-I backbones, partially esterified with methanol and/or acetate galacturonic acid [155]. Differently, type II AGs link to proteins and the combination of the two is referred as arabinogalactan proteins (AGPs). AGPs are a specific protein family rich in Hyp, Ser, Ala, and Gly in which the polysaccharide represents the 90% of the molecular weight [120].

Since type II AGs polysaccharides have been found in some plant exudate gums (e.g. gum arabic), particular attention will be given on the structure of this polysaccharide family. The general structure of type II AGs consists of a β-1,3-galactan backbone with short side chains. About half of the main chain is substituted with β-1,3-galactopyranose (Galp) dimers and about a quarter with single galactopyranose units. The other side chains contain three or more residues, which might include different monosaccharides such as L-arabinofuranose (Araf), L-arabinopyranose (Arap), D-glucuronic acid (GlcA), 4-O-methyl-D-glucuronic acid, L-rhamnose (Rha) and L-fucose (Fuc) [156,157]. This structure has been frequently detected in type II AGs from larch heartwood samples of different species [157-159]. Furthermore it has also been observed in AGs from other sources such as *Coffea Arabica* powder [160], Chinese Wolfberry [161], Norway spruce and Scots pine heartwood [162], even if in the last three cases the unusual presence of glucuronic acid in the side chains was detected. However, variations in the carbohydrate structure have been encountered and dependence to development stage and plant tissue was discussed [163]. 1,6-linkage of galactose monomers were detected in the main chain of type II AGs coming from the pollen of timothy grass [164] and Tryfona *et al.* described the presence of long side chains, up to 20 Gal monomers and highly substituted, in wheat flour [165] and Arabidopsis leaves arabinogalactans [166]. Moreover, as described above, the type II AGs found in exudates gums from Acacia species, appear to have a complex branched structure [154]. Therefore the study of type II arabinogalactans is very challenging since, even if the main chain composition is well known, the side structures are not well defined and structural heterogeneity/complexity represents a big obstacle to their analysis.

## 4.2. GUM TRAGACANTH

Gum tragacanth is described as a complex, highly branched, heterogeneous hydrophilic polysaccharide. It is supposed to consist of two major fractions which are present in different ratios according to the *Astragalus* species [167]: the so-called *tragacanthin*, water soluble, and *bassorin* which is insoluble but it swells in water. Tragacanthin can be further divided in two other different fractions: an ethanol soluble minor fraction and an ethanol insoluble fraction [168]. The tragacanth gum possible fractionation reported by Aspinall and Baillie is reproduced in Fig. 6 [168,169].

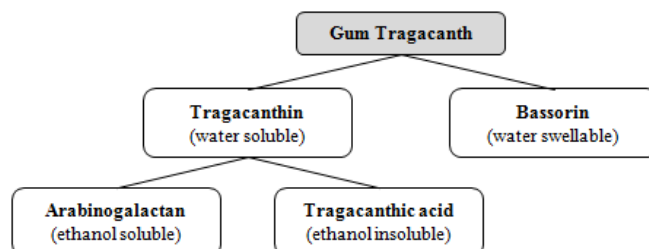


Fig. 6. Scheme of the possible gum tragacanth fractions.

The structure of the gum tragacanth polysaccharide has been extensively studied in the past by several authors [170-173] and more in detail by James and Smith [174,175] and Aspinall and Baillie [168,169]. The water-swellaable fraction main chain is probably formed by 1,4-linked  $\alpha$ -D-galacturonic acid residues with side chains attached 1,3 of D-xylose units, sometimes ending with D-Gal or L-Fuc [4] (Fig. 7, [168]). As described by the authors mentioned above, the water-soluble tragacanthin fraction displays a structure possibly similar to arabinogalactans Type II and it probably consists of a core composed of 1,6- and 1,3-linked D-galactose with attached chains of 1,2-, 1,3- and 1,5-linked arabinose as mono- or oligosaccharides. However, the classification of the water-soluble part as an arabinogalactan and the insoluble part as tragacanthic acid has been recently questioned. Investigation by NMR and ESI-MS of the two sub-fractions of tragacanthin component, further fractionated by subsequent smith degradations, showed how the structure was much more complex than the one reported in previous studies since it contained arabinogalactan, tragacanthic acid (galacturonans) and starch (0.6%) [176]. Free reducing oligosaccharides have been detected such as:  $\alpha$ -Araf-(1,2)- $\alpha$ -Araf-(1,4)-Ara and  $\alpha$ -Araf-(1,2)- $\alpha$ -Araf-(1,5)-Ara (typical of arabinogalactans),  $\beta$ -Galp-(1,4)- $\beta$ -Galp-(1,4)- $\beta$ -Galp-(1,4)-Gal and mixture of  $\beta$ -Galp-(1,4)- $\beta$ -Galp-(1,4)-Gal and  $\beta$ -GlcP-(1,4)- $\beta$ -Galp-(1,4)- $\beta$ -Galp-(1,4)-Gal which are not characteristic of AGs side chains. Furthermore, the study of Balaghi *et al.* [27] on the monosaccharide composition of tragacanth gum samples obtained from six different *Astragalus* species showed that, because of the high amount of xylose and fucose, some tragacanth gums might be referred as xylogalacturonan or fucogalacturonan. For this reason, classifying this gum as an arabinogalacturonan might not be correct.

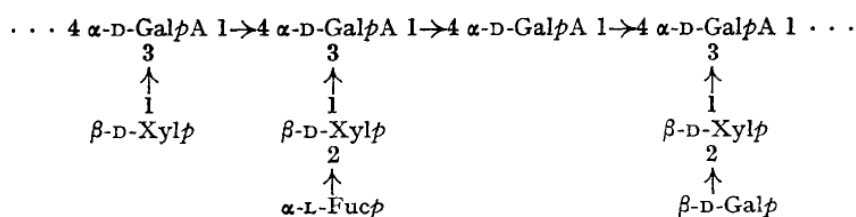


Fig. 7. Main structure of tragacanthic acid suggested by Aspinall *et al.* 1963.



### PEACH GUM

The polysaccharide component of the gum exuded from peach tree was first studied in 1950 [183], but it was deeply investigated only in 2008 by Simas *et al.* using GC-MS, HPSEC-RI, <sup>13</sup>C NMR and ESI-MS [31]. In regard to the monosaccharide composition, the high amount of arabinose (36-37%) and galactose (42%) suggested the hypothesis of the polysaccharide to be an arabinogalactan. The suggested structure consisted in a highly branched polysaccharide with a backbone mainly composed of (1→6)-linked Galp units with some (1→3) glycosidic linkages. Galactose units of the main chain were substituted by side chains containing  $\alpha$ -l-Araf (non-reducing end, 3-*O*-, 5-*O*-, and 2,5-di-*O*-subst.),  $\beta$ -l-Arap (4-*O*- and 2,4-di-*O*-subst.),  $\beta$ -d-Galp (3-*O*-, 2,3-di-*O*-, 3,4-di-*O*-, 3,6-di-*O*- and 3,4,6-tri-*O*-subst.),  $\alpha$ -and/or  $\beta$ -D-Xylp non-reducing end units and 4-*O*-Me- $\alpha$ -D-GlcpA units. This structure was partially in contrast with the classical arabinogalactans such as gum from *A. senegal* which have (1→3)-linked  $\beta$ -Galp main-chain. However a recent study, which also focused on gum rheological properties, confirmed the acidic arabinogalactan structure of the polysaccharide fraction [184]. No significant differences were found in the structure of the polysaccharide from nectarine gum compared to peach tree gum [185].

### APRICOT GUM

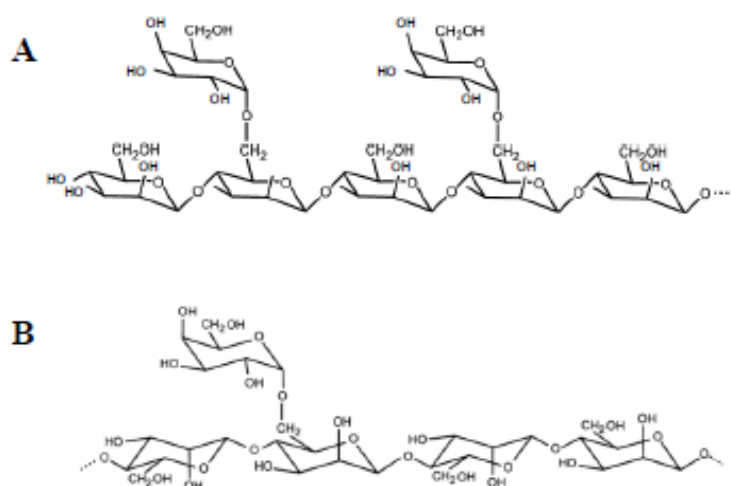
Besides the recent investigation of the polysaccharide structure of apricot seeds [186], just few information are known in relation to the structure of apricot-tree gum [187,188]. The core is a 1,6 and 1,3-linked D-Galp chain with alternating residues of 4-*O*-substituted D-GlcA and 2-*O*-substituted D-Man. Peripheral sugar units of Ara, Rha and Xyl were identified. Therefore gum from apricot tree can be classified as substituted arabinogalactan.

## **4.4. GUAR AND LOCUST BEAN GUM**

Guar gum and locust bean gum (LBG) are supposed to consist of mainly galactopyranosyl and mannopyranosyl residues. The two gums are described to resemble the typical structure of plant galactomannans which is characterized by a  $\beta$ -D-(1,4) linked Man backbone with single  $\alpha$ -D-Gal linked in 6-*O*-position as side chains [189]. The only difference between the two gums is the M/G ratio: about 2 for guar gum and 4 for locust bean gum, which therefore contain half as many Gal units compared to guar gum [190]. Methylation analysis suggested, for guar gum, the presence of repeating units of three hexose composed of two mannoses and one galactose as side chain [191]. In the structure of LBG the single Gal as side chain is linked to one Man out of three. The proposed structure for the two gums is reported in Fig. 9 (images taken from Moreira et Filho [192]).

Recently the presence of pentose residues and acetyl groups in LBG, as substituents on the main mannose backbone, was highlighted. The (1,4)-linked mannan backbone of locust bean gum was subjected to partial degradation using the enzyme endo- $\beta$ -mannanase. The released oligosaccharides were fractionated by size exclusion chromatography and later analyzed by both ESI-MS<sup>n</sup> (tandem mass spectrometry) and GC-qMS after derivatization [193].





**Fig. 9.** Proposed chemical structure of: (A) guar gum and (B) locust bean gum polysaccharide moiety.

Besides the presence of different substituents (e.g. pentose residues), the main polysaccharide structure of both guar and locust bean gum resembled the family of the galactomannans.

Galactomannans are mainly present in the seeds of Leguminosae. Together with linear mannan, glucomannans and galactoglucomannans, they belong to the mannan polysaccharides, an important constituent group of hemicelluloses [194]. They are characterized by a main chain of  $\beta$ -D-(1,4) linked mannose (Man) to which  $\alpha$ -D-1,6-Gal units are attached. What differentiates the galactomannans from the other mannan subfamilies, in terms of definition, is the presence of galactose units whose amount, according to Aspinall, has to be higher than 5% [195].

Some non-classical structures have been identified in the literature and, recently, they have been reviewed by both Srivastava et Kapoor [190] and Moreira et Filho [192]. It has been shown how some galactomannans are highly branched (up to 91%). Side chains containing both Gal and Man units with more than one residue for chain sometimes linked 1,3- and 1,2- were described in the mannan backbone. At the same time the presence of arabinose, glucose residues and widespread hexose acetylation were revealed in galactomannans obtained from non conventional sources [196] and green and roasted coffee infusions [197].

#### 4.5. GHATTI GUM

The structure of gum ghatti polysaccharide core was widely studied in the past by Aspinall et al. [198-201]. Later in 2002, Tischer et al. carried out a research aimed to highlight the nature of the side chains simplifying the gum structure by Smith degradation and performing  $^{13}\text{C}$  NMR analysis on the gum products [202]. The gum was shown to have a backbone of alternating 4-*O*- and 2-*O*-substituted  $\beta$ -D-glucuronic acid and  $\alpha$ -D-mannopyranose units to which L-arabinofuranose residues (single or in the oligosaccharide form) are linked 1,6 to the mannosyl residues (Fig. 10, from Deshmukh et al. [203]). The Man<sub>p</sub> residues are also substituted with 1,6-Galp residues of different length which can end by 1,6-GluA residue. Other substituents and side chains were discovered such as 6% of rhamnose linked as  $\alpha$ -Rhap-(1,4) to the  $\beta$ -Glc<sub>p</sub>A- chain. This linkage is common with the linkage found in most of plant gums.





The polysaccharide backbone of karaya gum was demonstrated to be an interrupted sequence of rhamnose and mannose, thus resulting in a rhamnogalacturonoglycan.

Rhamnogalacturonans (RGs), together with homogalacturonans, apiogalacturonan and xylogalacturonan, represent one of the main domains of pectins [155]. RGs comprise two subfamilies: the rhamnogalacturonans I (RG-I) and the rhamnogalacturonans II (RG-II). RG-I structure and functions have been recently reviewed by B.M. Yapo [218]. Even if variations in the structure may be encountered from source to source, these polysaccharides are all characterized by a backbone of  $[\rightarrow 4)\text{-}\alpha\text{-D-GalpA-(1,2)-}\alpha\text{-L-Rhap-(1}\rightarrow\text{)]}_n$  repeating units where some of the GalA residues are *O*-acetylated [219]. RGs-I structure can be more or less branched with many different side chains, of varying degrees of polymerization (DP), linked to rhamnose in *O*-4 or *O*-3 position. The most frequent residues are (1,4)-linked  $\alpha\text{-D-galactan}$  and (1,5)-linked  $\alpha\text{-L-arabinan}$  [155]. Furthermore terminal units of  $\beta\text{-D-Xylp}$ ,  $\alpha\text{-L-Fucp}$ ,  $\alpha\text{-DFucp}$ ,  $\alpha\text{-D-Galp}$ , 3-*O*-methyl- $\beta\text{-D-Galp}$ ,  $\alpha\text{-D-Glcp}$ ,  $\beta\text{-D-GlcpA}$  and its 4-*O*-methyl (ether) derivative may be present in the side chains [218]. The structure of RG II is much more complex and it is characterized by the presence of 12 possible different sugar residues [155].

As described above, different techniques have been employed to elucidate the structure of the polysaccharide component of plant gums. Results showed the significant complexity of these polysaccharides due to the presence of branched side chains composed of several substituents. In the literature, different potential structures have been suggested for the polysaccharide component of plant gums. Gum arabic was found to have a carbohydrate component similar to Type II arabinogalactans. This structure was also revealed in some fractions of gum tragacanth and gums from fruit trees. LBG and guar resembled the characteristic structure of galactomannans, while gum ghatti and karaya showed a complex composition which had respectively some similarities with the glucuronomannan and rhamnogalacturonan polysaccharides. Therefore, in the light of these consideration, it is possible to propose a classification of the plant gums under investigation according to different polysaccharide categories (Tab. 9).

**Tab. 9.** Polysaccharide types detected in several plant gums.

Plant gum	Polysaccharide type	Ref.
Arabic	Type II Arabinogalactan	[146,171,220]
Tragacanth	Mixture of type II arabinogalactan and galacturonan	[221]
Fruit tree	Substituted arabinogalactan	[182]
Guar	Galactomannan	[222]
Locust bean	Galactomannan	[222]
Ghatti	Substituted glucuronomannan	[223]
Karaya	Substituted rhamnogalacturonan	[224]

## 5. CONCLUSIONS

This first chapter aimed to summarize the information available in the literature about the chemical structure and composition of several plant gums. Furthermore, special emphasis was placed on their use in the cultural heritage field and on the analytical techniques that are commonly used for gums identification. The three described sections allowed to draw the following conclusions.

Most of the different analytical techniques employed for the study of plant gums in samples from artworks, suffer from some limitations. Non-destructive techniques, such as FTIR and Raman, can be useful in performing a initial survey on the sample to check the presence or not of polysaccharides. However, differentiation of gums cannot be performed. The most employed chromatographic techniques (GC-MS, Py-GC/MS, HPAEC) are all based on complete hydrolysis of the polysaccharide in the corresponding monomers. Gum identification is based on quantitative assessment of the relative proportions of monosaccharides, or on the qualitative determination of the absence/presence of characteristic monosaccharides. Though these techniques allowed the discrimination among the most known plant gums (e.g. gum arabic, tragacanth gum and cherry gum), the presence of inorganic and organic materials such as proteinaceous binders, environmental contaminants, as well as other plant gums, can be responsible of modifications in the monosaccharide profile of the gum, thus resulting in a not always reliable gum identification.

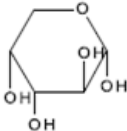
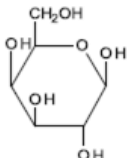
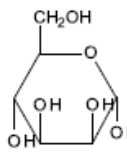
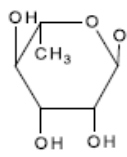
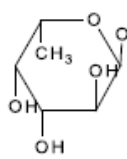
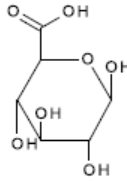
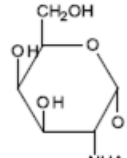
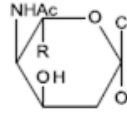
These limitations highlight the need for greater research efforts in the study of plant gums in cultural heritage materials by means of different strategies and analytical techniques.

In addition, the study of the literature showed how plant gums are characterized by some significant differences concerning both their proteinaceous (e.g. protein amount, amino acid composition and molecular weight) and polysaccharide nature. In regard of this last point, gum arabic and a fraction of tragacanth and fruit tree gums. were found to resemble the Type II arabinogalactans structure; locust bean gum and guar gum were described to have a carbohydrate structure similar to galactomannans, while gum ghatti and karaya showed a composition which had some similarities with respectively the glucuronomannan and rhamnogalacturonan polysaccharides.

These considerations represent the keystone of this thesis research since they reveal the possibility to set up alternative strategies for plant gums identification in samples from art works according to their differences in the proteinaceous and polysaccharide component.

**ANNEX:**

List of monosaccharides, their abbreviation and the corresponding structure.

Monosaccharide (abbreviation)	Example (abbreviation)	Haworth projection*
Pentose (Pent)	Arabinose (Ara)	
Hexose (Hex)	Galactose (Gal)	
	Mannose (Man)	
Deoxyhexose (dHex)	Rhamnose (Rha)	
	Fucose (Fuc)	
Hexuronic acid (HexA)	Glucuronic acid (GlcA)	
N-Acetylamine hexose (HexNAc)	N-acetylgalactosamine (GalNAc)	
N-Acetylneuraminic acid (NeuAc)	Sialic acid (SA)	

(\*): L-fucose;  $\alpha$ -L-rhamnose;  $\alpha$ -D-arabinopyranose;  $\beta$ -D-galactopyranose;  $\alpha$ -D-mannopyranose

## REFERENCES

1. BeMiller JN (2000) Gums. Kirk-Othmer Encyclopedia of Chemical Technology: John Wiley & Sons, Inc.
2. BeMiller JN (2001) Plant Gums. eLS: John Wiley & Sons, Ltd.
3. Williams PA, Phillips GO (2009) Introduction to food hydrocolloids. Handbook of hydrocolloids. In G.O. Phillips, P.A. Williams (Eds.) ed. New York, NY: CRC Press. pp. 1-19.
4. Phillips GO, Williams PA (2009) Handbook of hydrocolloids. New York, NY: CRC Press.
5. Mantell CL (1947) The Water-Soluble Gums. New York: Reinhold Publishing Corp.
6. Stephen AM, Churms SC, Vogt DC (1990) Exudates gums. In: (Ed.) PMD, editor. Methods in plant biochemistry: Academic Press. pp. 483-522.
7. Anderson DMW, Howlett JF, McNab CGA (1986) The hydroxyproline content of gum exudates from several plant genera. *Phytochemistry* 26: 309-311.
8. Cui SW (2005) Structural Analysis of Polysaccharides. Food Carbohydrates: CRC Press.
9. Saha D, Bhattacharya S (2010) Hydrocolloids as thickening and gelling agents in food: a critical review. *Journal of Food Science and Technology* 47: 587-597.
10. Rana V, Rai P, Tiwary AK, Singh RS, Kennedy JF, et al. (2011) Modified gums: Approaches and applications in drug delivery. *Carbohydrate Polymers* 83: 1031-1047.
11. Avachat AM, Dash RR, Shrotriya SN (2011) Recent Investigations of Plant Based Natural Gums, Mucilages and Resins in Novel Drug Delivery Systems. *Ind J Pharm Edu Res* 45: 86-99.
12. Mirhosseini H, Amid BT (2012) A review study on chemical composition and molecular structure of newly plant gum exudates and seed gums. *Food Research International* 46: 387-398.
13. Spiridon I, Popa VI (2005) Hemicelluloses: Structure and Properties. In: S. Dumitriu MD, editor. Polysaccharides Structural Diversity and Functional Versatility. New York: CRC Press. pp. 475-489.
14. FAO (1999) Gum Arabic. Rome: (Food and nutrition paper 52, addendum 7).
15. Imeson AP (1992) Exudate gums. In: Imeson A, editor. Thickening and Gelling Agents for Food: Springer US. pp. 66-97.
16. Kennedy JF, Phillips GO, Williams PA (2012) Gum arabic. Cambridge: The Royal Society of Chemistry.
17. Thevenet F (2009) Acacia Gum (Gum Arabic). Food Stabilisers, Thickeners and Gelling Agents: Wiley-Blackwell. pp. 11-30.
18. Verbeken D, Dierckx S, Dewettinck K (2003) Exudate gums: occurrence, production, and applications. *Applied Microbiology and Biotechnology* 63: 10-21.
19. Williams PA, Phillips GO (2000) Gum arabic. Handbook of hydrocolloids. In G.O. Phillips, P.A. Williams (Eds.) ed. New York, NY: CRC Press. pp. 252-273.
20. Lucas A (1962) Ancient Egyptian Materials and Industries; Arnold, editor. London.
21. Alfrén J, Penarrieta JM, Bergenstahl B, Nilsson L (2012) Comparison of molecular and emulsifying properties of gum arabic and mesquite gum using asymmetrical flow field-flow fractionation. *Food Hydrocolloids* 26: 54-62.
22. Al-Assaf S, Phillips GO, Aoki H, Sasaki Y (2007) Characterization and properties of Acacia senegal (L.) Willd. var. senegal with enhanced properties (Acacia (sen) SUPER GUM (TM)): Part 1 - Controlled maturation of Acacia senegal var. senegal to increase viscoelasticity, produce a hydrogel form and convert a poor into a good emulsifier. *Food Hydrocolloids* 21: 319-328.
23. Fauconnier ML, Blecker C, Groyne J, Razafindralambo H, Vanzeveren E, et al. (2000) Characterization of two Acacia gums and their fractions using a Langmuir film balance. *Journal of Agricultural and Food Chemistry* 48: 2709-2712.
24. Wang Q, Burchard W, Cui SW, Huang X, Phillips GO (2008) Solution properties of conventional gum arabic and a matured gum arabic (Acacia (sen) SUPER GUM). *Biomacromolecules* 9: 1163-1169.

25. FAO (1992) Tragacanth gum. Rome: (Food and nutrition paper 53).
26. Mohammadifar MA, Musavi SM, Kiumarsi A, Williams PA (2006) Solution properties of targacanthin (water-soluble part of gum tragacanth exudate from *Astragalus gossypinus*). *International Journal of Biological Macromolecules* 38: 31-39.
27. Balaghi S, Mohammadifar MA, Zargaraan A, Gavlighi HA, Mohammadi M (2011) Compositional analysis and rheological characterization of gum tragacanth exudates from six species of Iranian *Astragalus*. *Food Hydrocolloids* 25: 1775-1784.
28. Gavlighi AH, Meyer AS, Zaidel DNA, Mohammadifar MA, Mikkelsen JD (2013) Stabilization of emulsions by gum tragacanth (*Astragalus* spp.) correlates to the galacturonic acid content and methoxylation degree of the gum. *Food Hydrocolloids* 31: 5-14.
29. Mayes JM (2009) Gum Tragacanth and Karaya. *Food Stabilisers, Thickeners and Gelling Agents*: Wiley-Blackwell. pp. 167-179.
30. Lópex-Franco Y, Higuera-Ciapara I (2009) Other exudates: tragacanth, karaya, mesquite gum and larchwood arabinogalactan. In: Williams GOPaPA, editor. *Handbook of hydrocolloids*. Cambridge: CRC Press LCC. pp. 495-534.
31. Simas FF, Gorin PAJ, Wagner R, Sasaki GL, Bonkerner A, et al. (2008) Comparison of structure of gum exudate polysaccharides from the trunk and fruit of the peach tree (*Prunus persica*). *Carbohydrate Polymers* 71: 218-228.
32. FAO (1992) Karaya gum. Rome: (Food and nutrition paper 52).
33. Brito ACF, Silva DA, de Paula RCM, Feitosa JPA (2004) *Sterculia striata* exudate polysaccharide: characterization, rheological properties and comparison with *Sterculia urens* (karaya) polysaccharide. *Polymer International* 53: 1025-1032.
34. Galla NR, Dubasi GR (2010) Chemical and functional characterization of Gum karaya (*Sterculia urens* L.) seed meal. *Food Hydrocolloids* 24: 479-485.
35. Le Cerf D, Irinei F, Muller G (1990) Solution properties of gum exudates from *Sterculia-Urens* (karaya gum). *Carbohydrate Polymers* 13: 375-386.
36. Ibrahim NA, Abo-Shosha MH, Allam EA, El-Zairy EM (2010) New thickening agents based on tamarind seed gum and karaya gum polysaccharides. *Carbohydrate Polymers* 81: 402-408.
37. Sakai E, Katayama T, Ogasawara T, Mizuno M (2013) Identification of *Anogeissus latifolia* Wallich and analysis of refined gum ghatti. *J Nat Med* 67: 276-280.
38. Ido T, Ogasawara T, Katayama T, Sasaki Y, Al-Assaf S P (2008) Emulsification properties of GATIFOLIA (gum ghatti) used for emulsions in food products. *Foods Food Ingrid J Jpn* 213: 365-371.
39. Al-Assaf S, Phillips GO, Amar V (2009) Gum ghatti. *Handbook of hydrocolloids*. In G.O. Phillips, P.A. Williams (Eds.) ed. New York, NY: CRC Press. pp. 477-494.
40. Deshmukh AS, Setty CM, Badiger AM, Muralikrishna KS (2012) Gum ghatti: A promising polysaccharide for pharmaceutical applications. *Carbohydrate Polymers* 87: 980-986.
41. Katayama T, Ido T, Sasaki Y, Ogasawara T, Al-Assaf S, et al. (2008) Characteristics of the adsorbed component of gum ghatti responsible for its oil-water interface advantages. *Foods Food Ingrid J Jpn* 213: 372-376.
42. Wielinga WC (2009) Galactomannans. *Handbook of hydrocolloids*. In G.O. Phillips, P.A. Williams (Eds.) ed. New York, NY: CRC Press. pp. 229-251.
43. Lazaridou A, Biliaderis CG, Izydorczyk MS (2001) Structural characteristics and rheological properties of locust bean galactomannans: a comparison of samples from different carob tree populations. *Journal of the Science of Food and Agriculture* 81: 68-75.
44. Stols-Witlox M, Ormsby BA, Gottsegen M (2012) Grounds, 1400-1900. Including: Twentieth-century grounds. *Conservation of Easel Paintings*. Oxon.
45. Gettens RJ, Stout GL (1966) *Painting materials: a short encyclopedia*; Ed. D, editor.
46. Matteini M, Moles A (2003) *La chimica nel restauro. I materiali dell'arte pittorica*; Nardini, editor.
47. Kühn H (1986) *Conservation and restoration of works of art and antiquities*. London.
48. Doménech-Carbó MT (2008) Novel analytical methods for characterising binding media and protective coatings in artworks. *Analytica Chimica Acta* 621: 109-139.



49. Zadrozna I, Polec-Pawlak K, Gluch I, Ackacha MA, Mojski M, et al. (2003) Old master paintings - A fruitful field of activity for analysts: Targets, methods, outlook. *Journal of Separation Science* 26.
50. Mills JS, White R (1984) The organic chemistry of museum objects.
51. Chevalier A (1924) Sur la production de la gomme arabique en Afrique Occidentale. *Revue de Botanique Appliquée & d'Agriculture Coloniale* 4: 256-263.
52. Wright MM, Wheals BB (1987) Pyrolysis-mass spectrometry of natural gums, resins, and waxes and its use for detecting such materials in ancient Egyptian mummy cases (cartonnages). *Journal of Analytical and Applied Pyrolysis* 11: 195-211.
53. Colombini MP, Ceccarini A, Carmignani A (2002) Ion chromatography characterization of polysaccharides in ancient wall paintings. *Journal of Chromatography A* 968: 79-88.
54. Birstein VJ (1975) On the technology of Central Asia wall paintings: the problem of binding media. *Studies in Conservation* 20: 8-19.
55. Keheyani Y, Giulianelli L (2006) Identification of historic ink ingredients using pyrolysis-GC-MS: a model study. *e-Preservation Science*. pp. 5-10.
56. Broekman-Bokstijn M, van Asperen de Boer JRJ, Serck-Dewaide M (1970) The scientific examination of the polychromed sculpture in the Herlin altarpiece. *Studies in Conservation* 15: 370-399.
57. Frezzato F (2004) *Il libro dell'arte di Cennino Cennini*; Pozza N, editor. Vicenza.
58. Ormsby BA, Townsend JH, Singer BW, Dean JR (2005) British watercolour cakes from the eighteenth to the early twentieth century. *Studies in Conservation* 50: 45-66.
59. Vandenabeele P, Wehling B, Moens L, Edwards H, De Reu M, et al. (2000) Analysis with micro-Raman spectroscopy of natural organic binding media and varnishes used in art. *Analytica Chimica Acta* 407: 261-274.
60. Morris WW, Salmon W (1963) Preparation of Films for the Infrared Spectrophotometric Analysis of Plant Gums. *Anal Chem* 35: 600.
61. Jennings KR (1968) Collision-induced decompositions of aromatic molecular ions. *International Journal of Mass Spectrometry and Ion Physics* 1: 227-235.
62. Vallance SL (1997) Applications of chromatography in art conservation: Techniques used for the analysis and identification of proteinaceous and gum binding media. *Analyst* 122: R75-R81.
63. Vallance SL, Singer BW, Hitchen SM, Townsend JH (1998) The development and initial application of a gas chromatographic method for the characterization of gum media. *Journal of the American Institute for Conservation* 37: 294-311.
64. Kharbade BV, Joshi GP (1995) Thin layer chromatography and hydrolysis methods for the identification of plant gums in art-objects. *Studies in Conservation* 40: 93-102.
65. Birstein VJ, Tul'chinskii VM (1976) An investigation and identification of polysaccharides isolated from archaeological specimens. *Khim Prirod Soedin* 1.
66. Twilley JW (1984) The analysis of exudate plant gums in their artistic applications: an interim report. *Archaeological chemistry III*. pp. 357-394.
67. Schneider U, Kenndler E (2001) Identification of plant and animal glues in museum objects by GC-MS, after catalytic hydrolysis of the proteins by the use of a cation exchanger, with simultaneous separation from the carbohydrates. *Fresenius Journal of Analytical Chemistry* 371.
68. Bleton J, Mejanelle P, Sansoulet J, Goursaud S, Tchaplà A (1996) Characterization of neutral sugars and uronic acids after methanolysis and trimethylsilylation for recognition of plant gums. *Journal of Chromatography A* 720: 27-49.
69. Marinach C, Papillon MC, Pepe C (2004) Identification of binding media in works of art by gas chromatography-mass spectrometry. *Journal of Cultural Heritage* 5: 231-240.
70. Bleton J, Mejanelle P, Goursaud S, Tchaplà A (1996) Significance of GC/MS coupling demonstrated by characterization of 4-O-methylglucuronic acid in certain plant gums. *Analisis* 24: M14-M16.
71. Pitthard V, Finch P (2001) GC-MS analysis of monosaccharide mixtures as their diethylthioacetal derivatives: Application to plant gums used in art works. *Chromatographia* 53: S317-S321.

72. Pitthard V, Griesser M, Stanek S (2006) Methodology and application of GC-MS to study altered organic binding media from objects of the Kunsthistorisches Museum, Vienna. *Annali Di Chimica* 96: 561-573.
73. Pitthard V, Griesser M, Stanek S, Bayerova T (2006) Study of complex organic binding media systems on artworks applying GC-MS analysis: Selected examples from the Kunsthistorisches Museum, Vienna. *Macromolecular Symposia* 238: 37-45.
74. Bonaduce I, Brecoulaki H, Colombini MP, Lluveras A, Restivo V, et al. (2007) Gas chromatographic-mass spectrometric characterisation of plant gums in samples from painted works of art. *Journal of Chromatography A* 1175: 275-282.
75. Mahfoudhi N, Chouaibi M, Donsi F, Ferrari G, Hamdi S (2012) Chemical composition and functional properties of gum exudates from the trunk of the almond tree (*Prunus dulcis*). *Food Sci Technol Int* 18: 241-250.
76. Lluveras A, Bonaduce I, Andreotti A, Colombini MP (2010) GC/MS analytical procedure for the characterization of glycerolipids, natural waxes, terpenoid resins, proteinaceous and polysaccharide materials in the same paint microsample avoiding interferences from inorganic media. *Analytical chemistry* 82: 376-386.
77. Lluveras-Tenorio A, Mazurek J, Restivo A, Colombini MP, Bonaduce I (2012) The development of a new analytical model for the identification of saccharide binders in paint samples. *PloS one* 7: E49383.
78. Derrick MR, Stulik DC (1987) Identification of natural gums in works of art using pyrolysis-gas chromatography. *ICOM Committee for the Conservation, 9th Triennial Meeting, Dresden, 26-31 August 1990* 1: 9-13.
79. Stevanato R, Rovea M, Carbini M, Favretto D, Traldi P (1997) Curie-point pyrolysis gas chromatography mass spectrometry in the art field .3. The characterization of some non-proteinaceous binders. *Rapid Communications in Mass Spectrometry* 11: 286-294.
80. Keheyani Y, Eliazyan G, Engel P, Rittmeier B (2009) Py/GC/MS characterisation of naturally and artificially aged inks and papers. *Journal of Analytical and Applied Pyrolysis* 86: 192-199.
81. Chiavari G, Montalbani S, Prati S, Keheyani Y, Baroni S (2007) Application of analytical pyrolysis for the characterisation of old inks. *Journal of Analytical and Applied Pyrolysis* 80: 400-405.
82. Brecoulaki H, Andreotti A, Bonaduce I, Colombini MP, Lluveras A (2012) Characterization of organic media in the wall-paintings of the "Palace of Nestor" at Pylos, Greece: evidence for a secco painting techniques in the Bronze Age. *Journal of Archaeological Science* 39: 2866-2876.
83. He L, Nie M, Chiavari G, Mazzeo R (2007) Analytical characterization of binding medium used in ancient Chinese artworks by pyrolysis-gas chromatography/mass spectrometry. *Microchemical Journal* 85: 347-353.
84. Scalarone D, Chiantore O, Riedo C (2008) Gas chromatographic/mass spectrometric analysis of on-line pyrolysis-silylation products of monosaccharides. *Journal of Analytical and Applied Pyrolysis* 83: 157-164.
85. Chiantore O, Riedo C, Scalarone D (2009) Gas chromatography-mass spectrometric analysis of products from on-line pyrolysis/silylation of plant gums used as binding media. *International Journal of Mass Spectrometry* 284: 35-41.
86. Riedo C, Scalarone D, Chiantore O (2010) Advances in identification of plant gums in cultural heritage by thermally assisted hydrolysis and methylation. *Analytical and Bioanalytical Chemistry* 396: 1559-1569.
87. Challinor JM (2001) Review: the development and applications of thermally assisted hydrolysis and methylation reactions. *Journal of Analytical and Applied Pyrolysis* 61: 3-34.
88. Fabbri D, Helleur R (1999) Characterization of the tetramethylammonium hydroxide thermochemolysis products of carbohydrates. *Journal of Analytical and Applied Pyrolysis* 49: 277-293.
89. Riedo C, Scalarone D, Chiantore O (2013) Multivariate analysis of pyrolysis-GC/MS data for identification of polysaccharide binding media. *Anal Methods* 5: 4060-4067.
90. Knill CJ, Kennedy JF (2003) Degradation of cellulose under alkaline conditions. *Carbohydrate Polymers* 51: 281-300.

91. Grossl M, Harrison S, Kaml I, Kenndler E (2005) Characterisation of natural polysaccharides (plant gums) used as binding media for artistic and historic works by capillary zone electrophoresis. *Journal of Chromatography A* 1077: 80-89.
92. Kaml I, Kenndler E (2007) Characterization of natural organic binding media in museum objects by capillary electrophoresis. *Current Analytical Chemistry* 3: 33-40.
93. Arslanoglu J, Schultz J, Loike J, Peterson K (2010) Immunology and art: Using antibody-based techniques to identify proteins and gums in artworks. *Journal of Biosciences* 35: 3-10.
94. Smith F (1940) 191. The constitution of arabic acid. Part V. Methylated arabic acid. *Journal of the Chemical Society (Resumed)*: 1035-1051.
95. Anderson DMW, Stoddard JF (1966) Studies on uronic acid materials. *Carbohydrate Res* 2.
96. Akiyama Y, Eda S, Kato K (1984) Gum arabic is a kind of arabinogalactan-protein. *Agric Biol Chem* 48: 235-237.
97. Yariv J, Lis H, Katchalski E (1967) Precipitation of arabic acid and some seed polysaccharides by glycosylphenylazo dyes. *The Biochemical journal* 105.
98. Anderson DMW, Bridgeman MME, Farquhar JGK, McNab CGA (1983) The chemical characterization of the test article used in toxicological studies of gum Arabic (*Acacia Senegal* (L) Willd). *International Tree Crops Journal* 2: 245-254.
99. Islam AM, Phillips GO, Sljivo A, Snowden MJ, Williams PA (1997) A review of recent developments on the regulatory, structural and functional aspects of gum arabic. *Food Hydrocolloids* 11.
100. Anderson DMW, McDougall FJ (1987) Degradative studies of gum-arabic (*Acacia-Senegal* (L) Willd) with special reference to the fate of the amino-acids present. *Food Additives and Contaminants* 4: 247-255.
101. Anderson DMW, McDougall FJ (1987) The amino-acid-composition and quantitative sugar-amino acid relationships in sequential smith-degradation products from gum-arabic (*acacia-senegal* (l) willd). *Food Additives and Contaminants* 4: 125-132.
102. Vandeveld MC, Fenyo JC (1985) Macromolecular distribution of acacia-senegal gum (gum arabic) by size-exclusion chromatography. *Carbohydrate Polymers* 5: 251-273.
103. Connolly S, Fenyo J-C, Vandeveld M-C (1987) Heterogeneity and homogeneity of an arabinogalactan-protein: *Acacia senegal* gum. *Food Hydrocolloids* 1.
104. Connolly S, Fenyo JC, Vandeveld MC (1988) Effect of proteinase on the macromolecular distribution of acacia-senegal gum. *Carbohydrate Polymers* 8: 23-32.
105. Duvallat S, Fenyo JC, Vandeveld MC (1989) Meaning of molecular-weight measurements of gum arabic. *Polymer Bulletin* 21: 517-521.
106. Fincher GB, Stone BA, Clarke AE (1983) Arabinogalactan-proteins: structure, biosynthesis and function. *Annu Rev Plant Physiol* 34: 47-70.
107. Tan L, Showalter AM, Egelund J, Hernandez-Sanchez A, Doblin MS, et al. (2012) Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. *Frontiers in plant science* 3.
108. Randall RC, Phillips GO, Williams PA (1989) Fractionation and characterization of gum from *Acacia senegal*. *Food Hydrocolloids* 3: 65-75.
109. Osman ME, Menzies AR, Williams PA, Phillips GO, Baldwin TC (1993) The molecular characterization of the polysaccharide gum from acacia-senegal. *Carbohydrate Research* 246: 303-318.
110. Renard D, Lavenant-Gourgeon L, Ralet M-C, Sanchez C (2006) *Acacia senegal* gum: Continuum of molecular species differing by their protein to sugar ratio, molecular weight, and charges. *Biomacromolecules* 7: 2637-2649.
111. Siddig NE, Osman ME, Al-Assaf S, Phillips GO, Williams PA (2005) Studies on acacia exudate gums, part IV. Distribution of molecular components in *Acacia seyal* in relation to *Acacia senegal*. *Food Hydrocolloids* 19.
112. Osman ME, Menzies AR, Williams PA, Phillips GO (1994) Fractionation and characterization of gum-arabic samples from various african countries. *Food Hydrocolloids* 8: 233-242.
113. Idris OHM, Williams PA, Phillips GO (1998) Characterisation of gum from *Acacia senegal* trees of different age and location using multidetection gel permeation chromatography. *Food Hydrocolloids* 12: 379-388.

114. Mahendran T, Williams PA, Phillips GO, Al-Assaf S, Baldwin TC (2008) New insights into the structural characteristics of the arabinogalactan - Protein (AGP) fraction of gum Arabic. *Journal of Agricultural and Food Chemistry* 56: 9269-9276.
115. Al-Assaf S, Sakata M, McKenna C, Aoki H, Phillips GO (2009) Molecular associations in acacia gums. *Structural Chemistry* 20: 325-336.
116. Renard D, Garnier C, Lapp A, Schmitt C, Sanchez C (2012) Structure of arabinogalactan-protein from Acacia gum: From porous ellipsoids to supramolecular architectures. *Carbohydrate Polymers* 90: 322-332.
117. Ray AK, Bird PB, Iacobucci GA, Clark BC (1995) Functionality of gum-arabic: fractionation, characterization and evaluation of gum fractions in citrus oil emulsions and model beverages. *Food Hydrocolloids* 9: 123-131.
118. Osman ME, Menzies AR, Martin BA, Williams PA, Phillips GO, et al. (1995) Characterization of gum-arabic fractions obtained by anion exchange chromatography. *Phytochemistry* 38: 409-417.
119. Qi W, Fong C, Lamport DTA (1991) Gum-arabic glycoprotein is a twisted hairy rope - a new model based on *O*-galactosylhydroxyprolines as the polysaccharide attachment site. *Plant Physiology* 96: 848-855.
120. Showalter AM (2001) Arabinogalactan-proteins: structure, expression and function. *Cellular and Molecular Life Sciences* 58: 1399-1417.
121. Goodrum LJ, Patel A, Leykam JF, Kieliszewski MJ (2000) Gum arabic glycoprotein contains glycomodules of both extensin and arabinogalactan-glycoproteins. *Phytochemistry* 54: 99-106.
122. Kieliszewski MJ, Lamport DT (1994) Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. *Plant J* 5: 157-172.
123. Dror Y, Cohen Y, Yerushalmi-Rozen R (2006) Structure of gum arabic in aqueous solution. *Journal of Polymer Science Part B-Polymer Physics* 44: 3265-3271.
124. Anderson DMW (1986) Nitrogen conversion factors for the proteinaceous content of gums permitted as food additives. *Food Additives and Contaminants* 3: 231-234.
125. Glicksman M (1969) *Gum Technology in the Food Industry*. New York: Academic Press.
126. Koek MS (2007) A comparative study on the compositions of crude and refined locust bean gum: In relation to rheological properties. *Carbohydrate Polymers* 70: 68-76.
127. Wielinga WC (2009) Galactomannans. In: Williams GOPaPA, editor. *Handbook of hydrocolloids*: CRC Press. pp. 228-297.
128. Dakia PA, Blecker C, Robert C, Wathelet B, Paquot M (2008) Composition and physicochemical extracted from whole seeds by acid properties of locust bean gum or water dehulling pre-treatment. *Food Hydrocolloids* 22.
129. Bouzouita N, Khaldi A, Zgoulli S, Chebil L, Chekki R, et al. (2007) The analysis of crude and purified locust bean gum: A comparison of samples from different carob tree populations in Tunisia. *Food Chemistry* 101.
130. Al-Assaf S, Amar V, Phillips GO (2008) Characterization of gum ghatti and comparison with gum Arabic. In: P.A. Williams aGOPE, editor. *Gums and stabilisers for the food industry* 14. Cambridge: RSC. pp. 281-290.
131. Kaur L, Singh J, Singh H (2009) Characterization of Gum Ghatti (*Anogeissus latifolia*): A Structural and Rheological Approach. *Journal of Food Science* 74: E328-E332.
132. Wu Y, Cui W, Eskin NAM, Goff HD (2009) An investigation of four commercial galactomannans on their emulsion and rheological properties. *Food Research International* 42: 1141-1146.
133. Anderson DMW, Howlett JF, McNab CGA (1985) Studies of uronic-acid materials .81. The amino-acid composition of the proteinaceous component of gum tragacanth (*asiatic astragalus* spp). *Food Additives and Contaminants* 2: 231-235.
134. Anderson DMW, Howlett JF, McNab CGA (1985) Studies of uronic acid materials. 74. The amino-acid composition of the proteinaceous component of gum karaya (*sterculia* spp). *Food Additives and Contaminants* 2: 153-157.
135. Janaki B, Sashidhar RB (2000) Subchronic (90-day) toxicity study in rats fed gum kondagogu (*Cochlospermum gossypium*). *Food and Chemical Toxicology* 38.

136. Schilling MR, Khanjian HP, Souza LAC (1996) Gas chromatographic analysis of amino acids as ethyl chloroformate derivatives. 1. Composition of proteins associated with art objects and monuments. . *Journal of the American Institute for Conservation* 35.
137. Whistler RL (1993) Exudate gums. In: R. L. Whistler aJNBE, editor. *Industrial gums, polysaccharides and their derivatives*. San Diego: Academic Press. pp. 309-339.
138. Anderson DMW, Grant DAD (1988) The chemical characterization of some *Astragalus* gum exudates. *Food Hydrocolloids* 2: 417-423.
139. Vinod VTP, Sashidhar RB, Sarma VUM, Raju SS (2010) Comparative amino acid and fatty acid compositions of edible gums kondagogu (*Cochlospermum gossypium*) and karaya (*Sterculia urens*). *Food Chemistry* 123: 57-62.
140. Anderson DMW, Howlett JF, McNab CGA (1985) Studies of uronic-acid materials .75. The amino-acid composition of the proteinaceous component of gum arabic (*acacia-senegal* (l) willd). *Food Additives and Contaminants* 2: 159-164.
141. Kang J, Cui SW, Guo Q, Chen J, Wang Q, et al. (2012) Structural investigation of a glycoprotein from gum ghatti. *Carbohydrate Polymers* 89: 749-758.
142. Amarioarei G, Lungu M, Sorin C (2012) Molar mass characterization of cherry tree exudate gums of different seasons. *Cellulose Chem Technol* 46: 583-588.
143. Vijayendran BR, Bone T (1984) Absolute molecular weight and molecular weight distribution of guar by size exclusion chromatography and low-angle light scattering. *Carbohydrate Polymers* 4: 299-313.
144. Pollard MA, Kelly R, Fischer PA, Windhab EJ, Eder B, et al. (2008) Investigation of molecular weight distribution of LBG galactomannan for flours prepared from individual seeds, mixtures, and commercial samples. *Food Hydrocolloids* 22: 1596-1606.
145. da Silva JAL, Goncalves MP (1990) Studies on a purification method for locust bean gum by precipitation with isopropanol. *Food Hydrocolloids* 4: 277-287.
146. Street CA, Anderson DMW (1983) Refinement of structures previously proposed for gum arabic and other acacia gum exudates. *Talanta* 30.
147. Defaye J, Wong E (1986) Structural studies of gum-arabic, the exudate polysaccharide from *acacia-senegal*. *Carbohydrate Research* 150: 221-231.
148. Aspinall GO, Knebl MC (1986) The location of alpha-d-galactopyranose residues in gum-arabic. *Carbohydrate Research* 157: 257-260.
149. McIntyre DD, Ceri H, Vogel HJ (1996) Nuclear Magnetic Resonance Studies of the Heteropolysaccharides Alginate, Gum arabic and Gum Xanthan. *Starch - Stärke* 48: 285-291.
150. Churms SC, Merrifield EM, Stephen AM (1983) Some new aspects of the molecular structure of *Acacia senegal* gum (Gum arabic). *Carbohydrate Res* 123: 267-279.
151. Clarke AE, Anderson RL, Stone BA (1979) Form and function of arabinogalactans and arabinogalactan-proteins. *Phytochemistry* 18: 521-540.
152. Tischer CA, Gorin PAJ, Iacomini M (2002) The free reducing oligosaccharides of gum arabic: aids for structural assignments in the polysaccharide. *Carbohydrate Polymers* 47: 151-158.
153. Nie S-P, Wang C, Cui SW, Wang Q, Xie M-Y, et al. (2013) The core carbohydrate structure of *Acacia seyal* var. *seyal* (Gum arabic). *Food Hydrocolloids* 32: 221-227.
154. Nie S-P, Wang C, Cui SW, Wang Q, Xie M-Y, et al. (2013) A further amendment to the classical core structure of gum arabic (*Acacia senegal*). *Food Hydrocolloids* 31: 42-48.
155. Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydrate Research* 344: 1879-1900.
156. Fincher GB, Stone BA, Clarke AE (1983) Arabinogalactan-Proteins: Structure, Biosynthesis, and Function. *Annual Review of Plant Physiology* 34: 47-70.
157. Ponder GR, Richards GN (1997) Arabinogalactan from Western larch, Part III: alkaline degradation revisited, with novel conclusions on molecular structure. *Carbohydrate Polymers* 34: 251-261.
158. Goellner EM, Utermoehlen J, Kramer R, Classen B (2011) Structure of arabinogalactan from *Larix laricina* and its reactivity with antibodies directed against type-II-arabinogalactans. *Carbohydrate Polymers* 86: 1739-1744.
159. Karácsonyi Š, Kováčik V, Alföldi J, Kubačková M (1984) Chemical and <sup>13</sup>C-n.m.r. studies of an arabinogalactan from *Larix sibirica* L. *Carbohydrate Research* 134: 265-274.

160. Matulová M, Capek P, Kaneko S, Navarini L, Liverani FS (2011) Structure of arabinogalactan oligosaccharides derived from arabinogalactan-protein of *Coffea arabica* instant coffee powder. *Carbohydrate Research* 346: 1029-1036.
161. Redgwell RJ, Curti D, Wang J, Dobruchowska JM, Gerwig GJ, et al. (2011) Cell wall polysaccharides of Chinese Wolfberry (*Lycium barbarum*): Part 1. Characterisation of soluble and insoluble polymer fractions. *Carbohydrate Polymers* 84: 1344-1349.
162. Willför S, Sjöholm R, Laine C, Holmbom B (2002) Structural features of water-soluble arabinogalactans from Norway spruce and Scots pine heartwood. *Wood Science and Technology* 36: 101-110.
163. Tsumuraya Y, Ogura K, Hashimoto Y, Mukoyama H, Yamamoto S (1988) Arabinogalactan-Proteins from Primary and Mature Roots of Radish (*Raphanus sativus* L.). *Plant Physiology* 86: 155-160.
164. Brecker L, Wicklein D, Moll H, Fuchs EC, Becker W-M, et al. (2005) Structural and immunological properties of arabinogalactan polysaccharides from pollen of timothy grass (*Phleum pratense* L.). *Carbohydr Res* 340: 657-663.
165. Tryfona T, Liang H-C, Kotake T, Kaneko S, Marsh J, et al. (2010) Carbohydrate structural analysis of wheat flour arabinogalactan protein. *Carbohydrate Research* 345: 2648-2656.
166. Tryfona T, Liang H-C, Kotake T, Tsumuraya Y, Stephens E, et al. (2012) Structural characterization of Arabidopsis leaf arabinogalactan polysaccharides. *Plant Physiol* 160: 653-666.
167. Balaghi S, Mohammadifar MA, Zargaraan A (2010) Physicochemical and Rheological Characterization of Gum Tragacanth Exudates from Six Species of Iranian Astragalus. *Food Biophysics* 5.
168. Aspinall GO, Baillie J (1963) Gum tragacanth. Part I. Fractionation of the gum and the structure of tragacanthin acid. *Journal of the Chemical Society*. pp. 1702-1714.
169. Aspinall GO, Baillie J (1963) Gum tragacanth. Part II. The arabinogalactan. *Journal of the Chemical Society*. pp. 1714-1721.
170. Jones JKN, Smith F (1949) Plant gums and Mucilages. *Adv Carbohydr Chem* 4: 243-291.
171. Smith F, Montgomery R (1959) The chemistry of plant gums and mucilages. New York: Reinhold Publishing.
172. Al-Hazmi MI, Stauffer KR (1986) *J Food Sci* 51: 1091.
173. Mantell CL, Reinhold BB (1948) The water-soluble gums. *Journal of Polymer Science* 3: 302-302.
174. James SP, Smith F (1945) The chemistry of gum tragacanth. Part I. Tragacanthic acid. *Journal of the Chemical Society* 193: 739-746.
175. James SP, Smith F (1945) The chemistry of gum tragacanth. Part III. *Journal of the Chemical Society* 195: 749-751.
176. Tischer CA, Iacomini M, Gorin PAJ (2002) Structure of the arabinogalactan from gum tragacanth (*Astragalus gummifer*). *Carbohydrate Research* 337: 1647-1655.
177. Gavlighi HA, Michalak M, Meyer AS, Mikkelsen JD (2013) Enzymatic Depolymerization of Gum Tragacanth: Bifidogenic Potential of Low Molecular Weight Oligosaccharides. *Journal of Agricultural and Food Chemistry* 61: 1272-1278.
178. Stephen AM (1983) Other plant polysaccharides. In: (Ed.) GOA, editor. *Polysaccharides*. New York, NY: Academic Press. pp. 97-180.
179. Shcherbukhin VD, Rosik J, Kubala J (1968) Polysaccharide from the cherry-tree gum (*Prunus avium* var *duracina*). *Chemical Papers* 22: 248-256.
180. Rosik J, Zitko V, Bauer S, Kubala J (1966) A polysaccharide from the wild cherry tree gum (*Prunus avium* L. subsp. *avium*). *Collect Czech Chem Commun* 31: 3353-3361.
181. Rosik J, Zitko V, Kubala J (1966) Structure of sour cherry-tree gum (*Prunus cerasus* L.). *Collect Czech Chem Commun* 31: 1569-1577.
182. Aspinall GO, Chaudhari AS, Whitehead CC (1976) Methylated cherry gum. *Carbohydrate research* 47: 119-127.
183. Jones JKN (1950) The structure of peach gum. Part I. The sugars produced on partial hydrolysis of the gum. *Journal of the Chemical Society*: 534-537.

184. Simas-Tosin FF, Barraza RR, Petkowicz CLO, Silveira JLM, Sasaki GL, et al. (2010) Rheological and structural characteristics of peach tree gum exudate. *Food Hydrocolloids* 24: 486-493.
185. Simas-Tosin FF, Wagner R, Santos EMR, Sasaki GL, Gorin PAJ, et al. (2009) Polysaccharide of nectarine gum exudate: Comparison with that of peach gum. *Carbohydrate Polymers* 76: 485-487.
186. Dass C (2007) *Fundamentals of contemporary mass spectrometry*; N.M. DDMN, editor: Jhon Wiley & Sons, Inc.
187. Stephen AM, Churms SC (1986) Smith degradation of gums from prunus species - observations on the core structure of prunus-armeniaca (apricot-tree) gum. *South African Journal of Chemistry* 39: 7-14.
188. Zitko V., Rosík R., Bruteničová M., J. K (1965) *Collect Czech Chem Commun* 30: 3501.
189. McCleary BV, Clark AH, Dea ICM, Rees DA (1985) The fine structure of carob and guar galactomannans. *Carbohydrate Research* 139: 237-260.
190. Srivastava M, Kapoor VP (2005) Seed Galactomannans: An Overview. *Chemistry & Biodiversity* 2: 295-317.
191. Ahmed ZF, Whistler RL (1950) The Structure of Guaran1. *Journal of the American Chemical Society* 72: 2524-2525.
192. Moreira LRS, Filho EXF (2008) An overview of mannan structure and mannan-degrading enzyme systems. *Applied Microbiology and Biotechnology* 79: 165-178.
193. Simoes J, Nunes FM, Rosario Domingues M, Coimbra MA (2011) Demonstration of the presence of acetylation and arabinose branching as structural features of locust bean gum galactomannans. *Carbohydrate Polymers* 86: 1476-1483.
194. Dey PM (1978) Biochemistry of plant galactomannans. *Adv Carbohydr Chem Biochem* 35: 341-376.
195. Aspinall GO (1959) *Adv Carbohydr Chem* 14: 429.
196. Cerqueira MA, Souza BWS, Simões J, Teixeira JA, Domingues MRM, et al. (2011) Structural and thermal characterization of galactomannans from non-conventional sources. *Carbohydrate Polymers* 83: 179-185.
197. Nunes FM, Domingues MR, Coimbra MA (2005) Arabinosyl and glucosyl residues as structural features of acetylated galactomannans from green and roasted coffee infusions. *Carbohydrate Research* 340: 1689-1698.
198. Aspinall GO, Hirst EL, Wickstrøm A (1955) Gum ghatti (Indian gum). The composition of the gum and the structure of the two aldo biouronic acids derived from it. *Journal of the Chemical Society*: 1160.
199. Aspinall GO, Auret BJ, Hirst EL (1958) Gum ghatti (Indian gum). Part II. The hydrolysis products obtained from methylated degraded gum and the methylated gum. *Journal of the Chemical Society*: 4408-4418.
200. Aspinall GO, Auret BJ, Hirst EL (1958) Gum ghatti (Indian gum). Part III. Neutral polysaccharide formed on partial acid hydrolysis of the gum. *Journal of the Chemical Society*: 4408-4418.
201. Aspinall GO, Bhavanandan VP, Christensen TB (1965) Gum ghatti (Indian gum). Part V. Degradation of the periodate-oxidised gum. *Journal of the Chemical Society*: 2677.
202. Tischer CA, Iacomini M, Wagner R, Gorin PAJ (2002) New structural features of the polysaccharide from gum ghatti (*Anogeissus latifolia*). *Carbohydrate Research* 337.
203. Deshmukh AS, Setty CM, Badiger AM, Muralikrishna KS (2012) Gum ghatti: A promising polysaccharide for pharmaceutical applications. *Carbohydrate Polymers* 87.
204. Kang J, Cui SW, Chen J, Phillips GO, Wu Y, et al. (2011) New studies on gum ghatti (*Anogeissus latifolia*) part I. Fractionation, chemical and physical characterization of the gum. *Food Hydrocolloids* 25: 1984-1990.
205. Kang J, Cui SW, Phillips GO, Chen J, Guo Q, et al. (2011) New studies on gum ghatti (*Anogeissus latifolia*) part II. Structure characterization of an arabinogalactan from the gum by 1D, 2D NMR spectroscopy and methylation analysis. *Food Hydrocolloids* 25: 1991-1998.
206. Kang J, Cui SW, Phillips GO, Chen J, Guo Q, et al. (2011) New studies on gum ghatti (*Anogeissus latifolia*) Part III: Structure characterization of a globular polysaccharide

- fraction by 1D, 2D NMR spectroscopy and methylation analysis. *Food Hydrocolloids* 25: 1999-2007.
207. Fry SC (2010) Cell Wall Polysaccharide Composition and Covalent Crosslinking. *Annual Plant Reviews: Wiley-Blackwell*. pp. 1-42.
  208. Redgwell RJ, O'Neill MA, Selvendran RR, Parsley KJ (1986) Structural features of the mucilage from the stem pith of kiwifruit (*actinidia deliciosa*: part II, structure of the neutral oligosaccharides. *Carbohydrate Research* 153: 107-118.
  209. Honda Y, Inaoka H, Takei A, Sugimura Y, Otsuji K (1996) Extracellular polysaccharides produced by tuberose callus. *Phytochemistry* 41: 1517-1521.
  210. Aspinall GO, Nasir-ud-din (1965) Plant gums of the genus *Sterculia*. Part I. The main structural features of *sterculia urens* gum. *J Chem Soc*: 2710-2720.
  211. Aspinall GO, Fraser RN (1965) Plant gums of the genus *sterculia*. Part II. *Sterculia caudata* gum. *J Chem Soc*: 4318-4325.
  212. Aspinall GO, Fraser RN, Sanderson GR (1965) Plant gums of the genus *Sterculia*. Part III. *Sterculia setigera* and *Cochlospermum gossypium* gums. *J Chem Soc*: 4325-4329.
  213. Aspinall GO, Sanderson GR (1970) Plant gums of the Genus *Sterculia*. Part V. Degradation of carboxy-reduced *Sterculia urens* gum. *J Chem Soc*: 2259-2264.
  214. Aspinall GO, Khondo L, Williams BA (1987) The hex-5-enose degradation - cleavage of glycosiduronic acid linkages in modified methylated *sterculia* gums. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* 65.
  215. Stephen AM, Churms SC (1995) Gums and mucilages. In: (ed.) SAM, editor. *Food Polysaccharides and Their Applications*. New York: Marcel Dekker. pp. 377-440.
  216. Meer K Gum Karaya. In: (Ed.) RLD, editor. *Handbook of water-soluble gums and resins*. New York: McGraw-Hill. pp. 10-11.
  217. Silva DA, Brito ACF, de Paula RCM, Feitosa JPA, Paula HCB (2003) Effect of mono and divalent salts on gelation of native, Na and deacetylated *Sterculia striata* and *Sterculia urens* polysaccharide gels. *Carbohydrate Polymers* 54: 229-236.
  218. Yapo BM (2011) Rhamnogalacturonan-I: A Structurally Puzzling and Functionally Versatile Polysaccharide from Plant Cell Walls and Mucilages. *Polymer Reviews* 51: 391-413.
  219. Komalavilas P, Mort AJ (1989) The acetylation of O-3 of galacturonic acid in the rhamnose-rich portion of pectins. *Carbohydrate Research* 189: 261-272.
  220. Glicksman M, Sand RE (1973) Gum Arabic. In: R. L. Whistler AP, editor. *Industrial Gums*. 2nd edn ed. New York. pp. 197-263.
  221. Stauffer KR (1980) Gum Tragacanth. In: R. L. Davidson e, editor. *Handbook of Water-Soluble Gums and Resins*. New York.
  222. McCleary BV, Clark AH, Dea ICM, Rees DA (1985) The fine structures of carob and guar galactomannans. *Carbohydrate Research* 139: 237-260.
  223. Glicksman M (1983) Gum ghatti (Indian gum). In: M. Glicksman e, editor. *Food Hydrocolloids*. Boca Raton, FL: CRC Press. pp. 31.
  224. Glicksman M (1982) Gum karaya. In: (Ed.) MG, editor. *Food Hydrocolloids Vol II*. Boca Raton, FL: CRC Press. pp. 39-48.



# Chapter II

## Investigation of plant gums glycoprotein profile by chromatographic and electrophoretic methods

### Abstract

Plant gums are complex proteoglycans which contain a small but significant amount of protein. In this chapter the possibility to identify plant gums according to their proteinaceous profile by size exclusion chromatography (SEC) and polyacrylamide gel electrophoresis (PAGE) is evaluated. These techniques allowed to decomplexify the gum samples and to estimate their molecular weight (MW) distribution. Analytical parameters have been optimized for the analysis of arabic, tragacanth, karaya, guar, ghatti, locust bean and fruit tree gums raw samples. While SEC allowed to further investigate the protein component showing the presence of protein fractions with MWs over 1 million Dalton, the more resolutive power of PAGE resulted in the ability to differentiate some plant gums according to their electrophoretic profile.



## TABLE OF CONTENTS

1. AIM OF THE RESEARCH .....	63
2. INVESTIGATION OF PLANT GUM PROTEIN COMPONENT BY SIZE EXCLUSION CHROMATOGRAPHY .....	64
2.1. Size exclusion chromatography .....	64
2.2. Sample preparation & materials .....	65
2.3. Optimization of HPLC-SEC and column calibration.....	66
2.4. SEC of several plant gums and molecular weight evaluation .....	68
2.4.1. Gum arabic .....	68
2.4.2. Tragacanth gum.....	69
2.4.3. Fruit tree gums .....	70
2.4.4. Guar, locust bean, ghatti and karaya gums.....	72
2.5. Conclusions of SEC analysis.....	73
3. INVESTIGATION OF PLANT GUMS PROTEIN PROFILE USING GEL ELECTROPHORESIS .....	74
3.1. Gel electrophoresis.....	74
3.1.1. Polyacrylamide gel eletrophoresis (PAGE) .....	74
3.1.2 Sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) .....	75
3.1.3. Gel staining .....	76
3.2. Optimization of electrophoretic experiments .....	77
3.3. SDS-PAGE profile of proteins in several plant gums .....	82
3.4. PAGE under non-denaturing conditions .....	87
3.6. Conclusions of SDS-PAGE analysis.....	89
REFERENCES.....	90



## List of figures

Fig. 1. Representation of the internal volumes in a SEC column and of the elution principle and profile. ....	64
Fig. 2. Calibration curve for (A) BioSuite HR (linear regression $y = -0.2134x + 4.5493$ , $R^2 0.9676$ ), and (B) Biosep S2000 (the linear regression $y = -0.9803x + 4.2658$ , $R^2 0.9167$ ).....	67
Fig. 3. SEC chromatogram of gum arabic (Zecchi, Italy) monitored at 214 nm: (A) MW range 100-7000 kDa; (B) MWs range 1-100 kDa. On the right: list of elution volume and estimated molecular weight of each observed fraction. ....	68
Fig. 4. SEC chromatogram of gum tragacanth (Zecchi, Italy) monitored at 214 nm: (A) MW range 100-7000 kDa; (B) MWs range 1-100 kDa. On the right: list of elution volume and estimated molecular weight of each observed fraction. ....	70
Fig. 5. SEC chromatogram monitored at 214 nm of apricot gum, almond gum and cherry gum. (A) MW range 100-7000 kDa; (B) MWs range 1-100 kDa. On the right: list of elution volume and estimated molecular weight of fractions observed for each gum. ....	71
Fig. 6. SEC chromatogram of ghatti gum and karaya monitored at 214 nm. (A) MW range 100-7000 kDa; (B) MWs range 1-100 kDa. Below each graph: list of elution volume and estimated molecular weight of each observed fraction. ....	72
Fig. 7. Representation of acrylamide chains linked together by bis molecules. ....	74
Fig. 8. SDS-PAGE 10% acrylamide of gum arabic (Zecchi). Different ratios gum/Laemmli (v/v) tested: (1) MW standards; (2) 1:7; (3) 1:3; (4) 1:2; (5) 1:1; (6) 1:7 gum heated at 45°C for 60 minutes; (7) 1:2 gum heated at 45°C for 60 minutes; (8) 1:1 gum heated at 45°C for 60 minutes. ....	78
Fig. 9. SDS-PAGE pattern of (1) MW standards; (2) gum arabic. (A) 4.7% polyacrylamide gel and (B) 7% polyacrylamide gel.....	79
Fig. 10. SDS-PAGE (4.7% acrylamide + 2000 kDa dextran) pattern of proteins from : (1) MW standards; (2) gum arabic. ....	80
Fig. 11. DS-AgPAGE (2% acrylamide + 0.5% agarose) pattern of proteins from : (1) MW standards; (2) gum arabic. ....	81
Fig. 12. SDS-PAGE pattern of (1) MW standards; (2) gum arabic. (A) 10% polyacrylamide gel and (B) 12% polyacrylamide gel.....	82
Fig. 13. SDS-PAGE (4.7% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.....	83
Fig. 14. SDS-PAGE (7% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum. ....	84
Fig. 15. SDS-PAGE (10% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.....	85
Fig. 16. SDS-PAGE (12% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.....	86
Fig. 17. SDS-PAGE (15% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.....	87
Fig. 18. Non-denaturing-PAGE (10% acrylamide) pattern of proteins from : (1) MW standards; (2) ghatti; (3) gum arabic; (4) tragacanth gum; (5) LBG; (6) guar gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.....	88

## List of tables

Tab. 1. Elution volume of the standard proteins used for respectively the two SEC column BioSuite HR and Biosep 2000. ( <sup>a</sup> ) corresponds to the volume of total inclusion ( $V_t$ ), ( <sup>b</sup> ) represents the totally excluded..	67
Tab. 2. Range of protein MWs resolved by gel with a different percentage of acrylamide .....	75
Tab. 3. Estimation of the protein amount loaded on the gel.....	78
Tab. 4. Summary of the gums analyzed by SDS-PAGE, corresponding amount of gum solubilized in Laemmli buffer and brands. ....	83

## 1. AIM OF THE RESEARCH

As described in the Chapter I section 3, gums are characterized by a small but significant amount of protein and differences were observed in terms of content, amino acid composition and molecular weights. These features represent the keystone of this research, whose aim is to differentiate several plant gums according to their proteinaceous component.

Size exclusion chromatography coupled to an ultraviolet (UV) detector and Sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) are two of the most used analytical techniques for separating and determining the molecular mass of macromolecules such as proteins. Furthermore, any specific sample treatment is required thus resulting in fairly simple and straightforward methods. At the light of these considerations, SEC and SDS-PAGE were considered the most suitable techniques to achieve the research goal.

The chapter is divided in two sections corresponding respectively to the chromatographic and electrophoretic approaches.

In the first part, the principles of protein separation in a size exclusion column are outlined (2.1). Sample preparation, column calibration and chromatography optimization will be discussed (2.2 and 2.3). In the further section (2.4) the chromatograms of gum arabic, tragacanth, karaya, guar, ghatti, locust bean and fruit tree gums (cherry, apricot, almond and plum gums) are showed and the obtained results related to the molecular weight discussed.

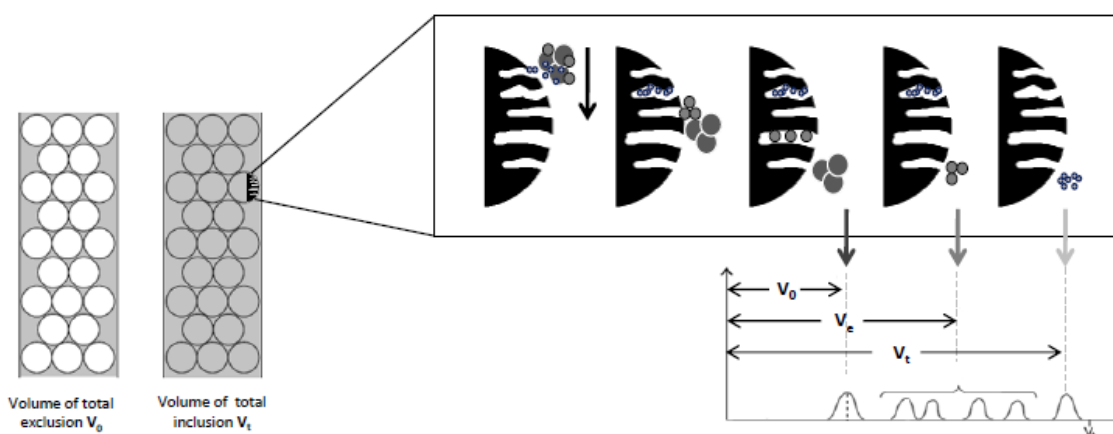
In the second part, besides an overview on polyacrylamide gel electrophoresis (3.1), the optimization of SDS-PAGE, in terms of both sample handling and matrix composition, is described (3.2). The protein component of gum arabic samples from different *Acacia* species has already been evaluated by polyacrylamide gel electrophoresis [1]. However, the developments performed in this research are focused on the application of SDS-PAGE to several different gums. The protein electrophoretic profiles, obtained for each plant gum using different percentages of polyacrylamide, are showed and discussed in section 3.3. Polyacrylamide gel electrophoresis was also performed without the addition of any denaturant detergent (e.g. SDS). The obtained gels are introduced and compared with the results obtained by classical SDS-PAGE (section 3.4).

## 2. INVESTIGATION OF PLANT GUM PROTEIN COMPONENT BY SIZE EXCLUSION CHROMATOGRAPHY

### 2.1. SIZE EXCLUSION CHROMATOGRAPHY

Size exclusion chromatography (SEC), also referred as gel filtration chromatography (GFC), is a liquid chromatographic technique that allows molecule separation according to their size. The first experiment on separation of macromolecules according to their size was performed in the 50s by Porath and Flodin [2], who separated various water-soluble macromolecules by a column packed with a cross-linked polydextran gel, swollen in aqueous media. In the 60s, molecular size fractionation of synthetic polymers in organic solvents was performed by Moore [3] using cross-linked polystyrene beads as stationary phase. Nowadays, different packing materials are available that are compatible with several polymers (water-soluble or hydrophobic) and mobile phases, and capable to separate from monomers and oligomers, to polymers, therefore from thousands to millions Dalton. One of the widest application of size exclusion chromatography concerns the protein separation [4].

In size exclusion chromatography the sample is loaded in a column filled with a porous material and it is carried along the column by a specific mobile phase. The separation in SEC depends on the hydrodynamic diameter of the analyzed molecules and on the pore size of the stationary phase. Large molecules, whose diameter is bigger than the average pore diameter of the stationary phase, will not enter the internal volume of the pores. Molecules with a hydrodynamic diameter comparable to the average pore size, will have access to some internal volumes, but not all, thus resulting in a fractionation of the analytes. Molecules whose diameter is smaller than the pores of the stationary phase beads, will access all the internal volumes. Therefore, as showed in Fig. 1, the order of elution of a mixture of molecules will be the inverse of their hydrodynamic radius: large molecules will elute at the beginning of the chromatogram since they are totally excluded from the pores, while smaller molecules will elute at the end of the chromatogram since their pathway along the column is longer. The molecules with an intermediate diameter will be separated along the chromatographic run according to their radius and they will represent the linear range of the column.



**Fig. 1.** Representation of the internal volumes in a SEC column and of the elution principle and profile.



Since in size exclusion chromatography no interaction is supposed to occur between the stationary phase and the analyte, the parameter that is measured in the chromatogram corresponds to the volume of elution of the molecule ( $V_e$ ). This value is related to other volume values, characteristic of the column, by the following equation (Eq. 1):

$$K_D = \frac{V_e - V_0}{V_t - V_0} \quad (1)$$

$K_D$  is defined as the partition coefficient of the analyzed molecules.  $V_0$  is the so called “volume of total exclusion” and corresponds to the external volume of the beads in the stationary phase;  $V_t$  is the “volume of total inclusion” and represents the sum of the external and internal volume within the beads. These two parameters can be experimentally determined by measuring the volume of elution of a molecule that is, respectively, totally excluded from the pores and totally included.  $V_e$  is the volume of elution of the analyte and it is obtained by multiplying the retention time for the flow. The  $K_D$  value can be determined for some standard proteins with a known molecular weight. This process consist in the calibration of the column. Therefore, the molecular weight of an unknown protein can be obtained by simply interpolate the empirical  $K_D$  in a graph of the logarithm of the molecular weight of standard proteins as function of  $K_D$ . Usually  $K_D$  is preferred to  $V_e$  since its value is column-independent.

It is well known that the size of a macromolecule (such as proteins) might not be always related to its molecular weight due to multiple possible shapes. Therefore in size exclusion chromatography the elution behavior of proteins is better related to their Stokes radii (RS) instead of their MW [5]. This value, also known as hydrodynamic radii (RH), is defined as the radius of a sphere that diffuses in the column as the molecule under investigation. RS is known for most of the standard proteins and it can be therefore obtained for an unknown protein by performing column calibration. The reason why all the SEC column usually refer to a MW range of separation is that, when size exclusion chromatography is used for a general evaluation of the protein mass, information about its molecular weight is much more informative than its stock radius.

## 2.2. SAMPLE PREPARATION & MATERIALS

Plant gum samples were prepared at 1% w/v in the same buffer used for chromatographic separation. The solution was mixed over night at room temperature for complete gum solubilization. The soluble part was then recovered after centrifugation at 13400 rpm for 30 minutes and 100  $\mu$ L were injected in the SEC column. Two different SEC columns were used in order to cover the widest molecular weight range (from 1 kDa to 7 MDa):

- BioSuite HR SEC (Waters), 7.8 x 300 mm, particle size 8  $\mu$ m and pore size 450 Å. The linear range declared by the producer is between 20 kDa – 7 MDa.
- Biosep S2000 (Phenomenex), 4.6 x 300 mm, particle size 5  $\mu$ m and pore size 145 Å. The linear range declared by the producer is between 1-300 kDa.

The eluted proteins were monitored using a Diode Array Detector (HPLC-DAD system, Agilent 1200 Series) operating at  $\lambda = 280$  and 214 nm (BW,4 nm). Both wavelengths are specific for protein detection:  $\lambda = 280$  nm is used to detect the absorbance of protein aromatic groups (present

in amino acids such as tryptophan, tyrosine and phenylalanine), while the  $\lambda=214\text{nm}$  is used to detect the absorbance of peptide bonds.

### 2.3. OPTIMIZATION OF HPLC-SEC AND COLUMN CALIBRATION

In size exclusion chromatography, protein separation occurs according to the size of the molecule and, on the contrary of other chromatographic techniques (e.g. reversed phase chromatography), there is no interaction between the protein and the stationary phase. However, the presence of some functional groups on the protein surface, may leads to non-specific interactions between the protein and the stationary phase, thus resulting in a decrease of protein recovery and alteration of the retention time [6]. Therefore, one of the most important parameter to consider when optimizing the protein separation by SEC is the choice of the mobile phase.

In the present work several buffers have been tested and the protein recovery was evaluated using two standard proteins. Since, as described in chapter I, Tab. 8, plant gums appear to have a significant high molecular weight, proteins such as catalase (MW = 232 kDa) and tyroglobulin (MW = 670 kDa) were selected in order to resemble the analytes size. The following mobile phases have tested with and without the addition of NaCl 50 mM:

- NaCl (200, 100, 50, 25 mM);
- Ammonium phosphate buffer (200, 100, 50, 25 mM), pH = 6;
- Sodium phosphate buffer, pH = 6.5;
- Tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) 50 mM, pH = 7.

Among the different buffers, Tris-HCl 50 mM, with the addition of NaCl 50 mM, gave the best recovery (97%) and it was therefore selected as mobile phase. Once the mobile phase was selected, the second step consisted on the calibration of the chromatographic column.

The determination of the molecular weight of an unknown protein by size exclusion chromatography is based on the comparison of its elution time with those of standard proteins whose molecular weights are known. Therefore the column has to be first calibrated with a set of standard proteins that cover the column MW range. The molecular weight of the unknown protein is therefore estimated by a graph of the logarithm of the molecular weight as a function of the elution volume, which is obtained by the SEC data of standard proteins [7].

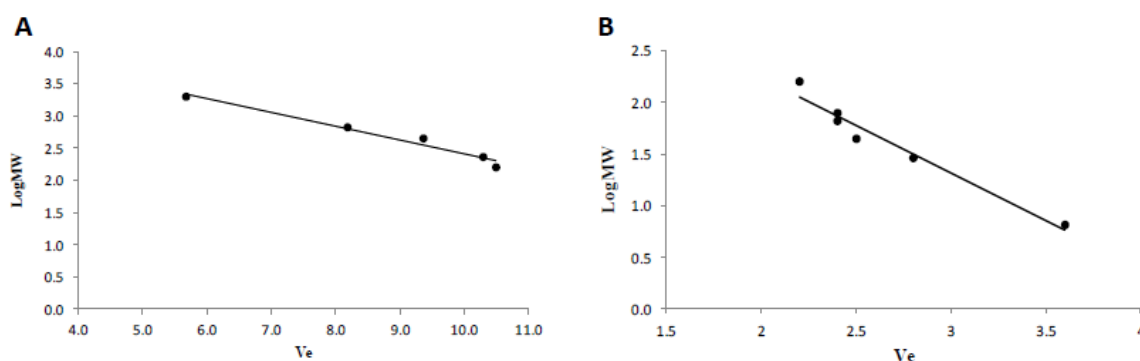
In the present work the two chromatographic columns were calibrated using the standards proteins reported in Tab. 1. Proteins with different molecular weights were specifically selected for each column in order to span the full range of sizes that could be analyzed by the respective stationary phases. 100  $\mu\text{L}$  of each standard protein, at the appropriate concentration ( $\sim 10^{-6}$  M), were injected and the obtained retention time was used to calculate the corresponding elution volume ( $V_e$ ).  $V_e$  is calculated by multiplying the retention time for the flow rate (0.5 mL/min for the column BioSuite HR and 0.2 mL/min for the column Biosep S200). The volume of total inclusion ( $V_i$ ) was evaluated by injection of milliQ water. The selected buffer has an absorbance cut-off at 205 nm of wavelength. Therefore, once water is injected, a negative peak can be observed in the chromatograms due to buffer dilution. The corresponding retention time was used to estimate the

$V_t$  value. In regard to the total excluded volume ( $V_0$ ), blue dextran is usually employed since its absorbance can be measured at  $\lambda=280$ , due to the presence of reactive blue 2 dyes, and it has an high molecular weight (2MDa) [8]. However, in the case of the BioSuite HR column, the MW range goes up to 7MDa, therefore the  $V_0$  could not be evaluated.

**Tab. 1.** Elution volume of the standard proteins used for respectively the two SEC column BioSuite HR and Biosep 2000. <sup>(a)</sup> corresponds to the volume of total inclusion ( $V_t$ ), <sup>(b)</sup> represents the totally excluded volume ( $V_0$ ) and “/” indicates that the protein was not used for column calibration. For each standard protein the corresponding molecular weight (MW) is reported.

Protein	MW [kDa]	LogMW	Ve [mL]	
			BioSuite	Biosep
Water	0.018	-1.74	12.4 <sup>a</sup>	
Cobalamin	1.3	0.13	/	4.2 <sup>a</sup>
Aprotinin	6.5	0.81	/	3.6
Carbonic Anhydrase	29	1.46	/	2.8
Ovalbumin	44.5	1.65	10.8	2.5
Bovine Serum Albumin	66.3	1.82	10.6	2.4
Apo-trasferrin	79	1.90	/	2.4
Aldolase	160	2.20	10.5	2.2
Catalase	232	2.37	10.3	/
Ferritin	450	2.65	9.4	/
Thyroglobulin	670	2.85	8.2	1.2 <sup>b</sup>
Dextran blue	2000	3.30	5.7	/

The calibration curve corresponding to the two SEC columns are reported respectively in Fig. 2A and B. As described in the previous chapter, for a correct MW determination, the x-axis of the graph usually corresponds to the  $K_D$  value, which takes into consideration the retention factor of proteins on the SEC column stationary phase. However, since no  $V_0$  could be measured for the BioSuite column, it was decided to plot the logMW values with the volume of elution, instead of  $K_D$ . In addition, for the Biosuite HR column, the values obtained for the standard proteins ovalbumin and bovine serum albumin were not plotted since they got distant from linearity.



**Fig. 2.** Calibration curve for (A) BioSuite HR (linear regression  $y = -0.2134x + 4.5493$ ,  $R^2 0.9676$ ), and (B) Biosep S2000 (the linear regression  $y = -0.9803x + 4.2658$ ,  $R^2 0.9167$ ).

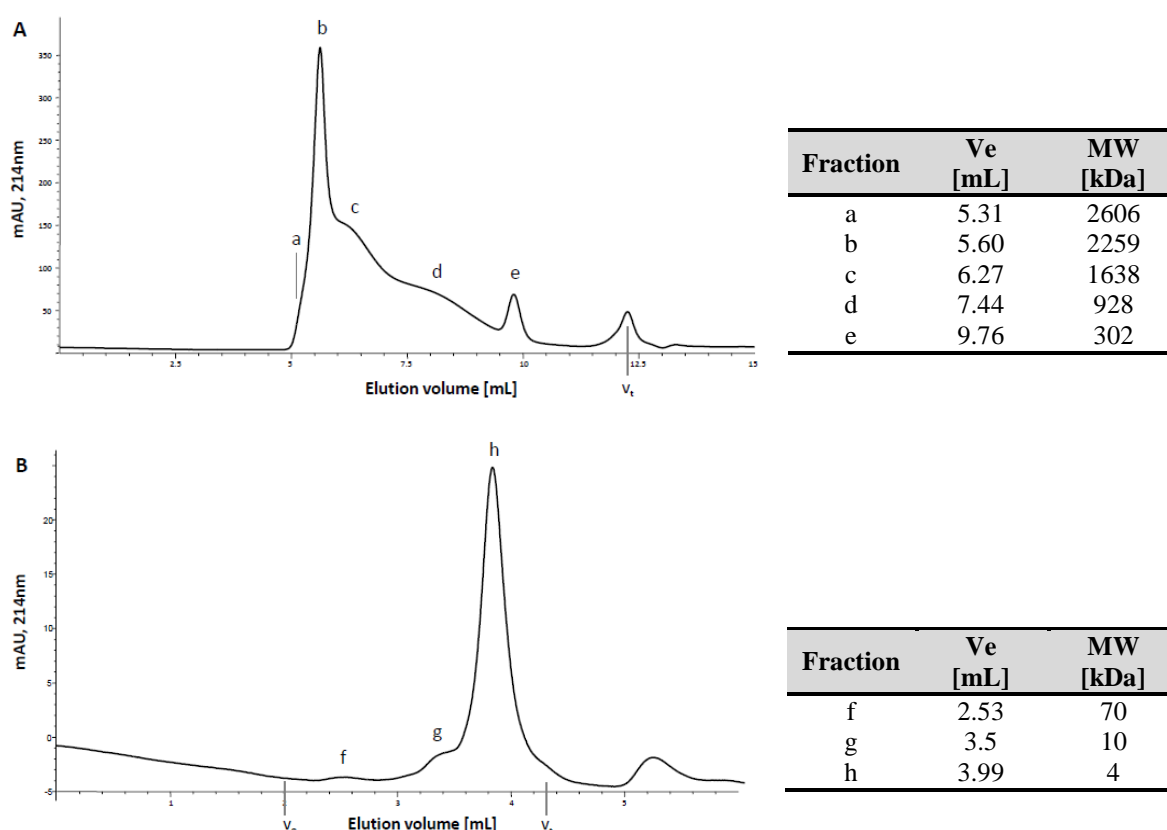
After columns calibration, the plant gums solubilized in Tris-HCl 50 mM, with the addition of NaCl 50 mM, were injected and the molecular weights of the obtained fractions were evaluated according to the calibration curves of the respective columns. Data related to each plant gum are showed in the following section.

## 2.4. SEC OF SEVERAL PLANT GUMS AND MOLECULAR WEIGHT EVALUATION

The evaluation of the molecular weight of the plant gums' proteinaceous component was performed in two steps. First the solubilized gums were injected in the column BioSuite HR SEC in order to evaluate the MW range 100 kDa – 7 MDa. The fractions that eluted over the volume of total inclusion of the column were collected, concentrated and re-injected in the second column (Biosep S2000) to evaluate lower molecular weights components. For each gum the two chromatograms, corresponding to the high MWs range and low MWs range will be reported, together with a table summarizing the estimated molecular weights. These values have been determined by interpolation of the measured  $V_e$  of each fraction with the calibration curves showed in Fig. 2.

### 2.4.1. Gum arabic

The SEC chromatogram showing the molecular mass distribution of gum arabic (Zecchi, Italy) is reported in Fig. 3. The fraction eluted from the first column over the volume of total inclusion ( $V_t = 12.4$ ) (see Fig. 3A), was recovered and injected in the second column. The corresponding chromatogram is reported in Fig. 3B.



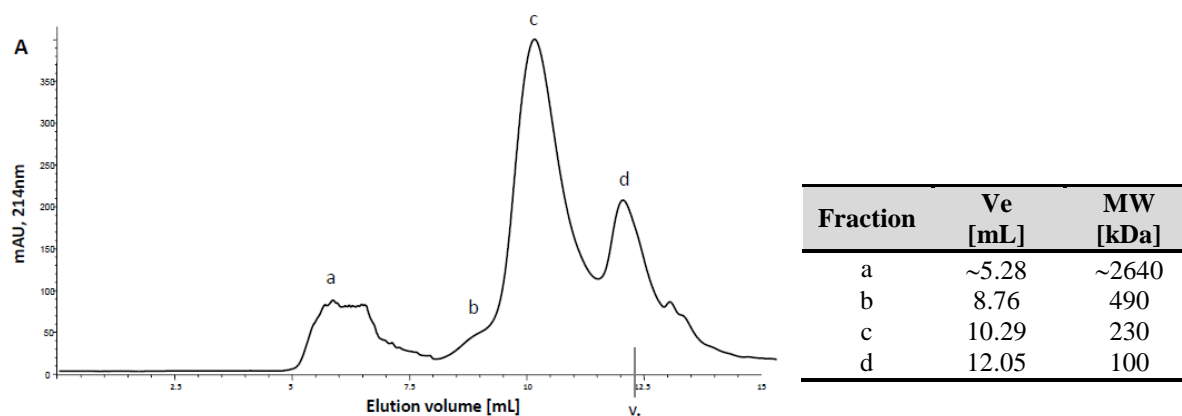
**Fig. 3.** SEC chromatogram of gum arabic (Zecchi, Italy) monitored at 214 nm: (A) MW range 100-7000 kDa; (B) MWs range 1-100 kDa. On the right: list of elution volume and estimated molecular weight of each observed fraction.

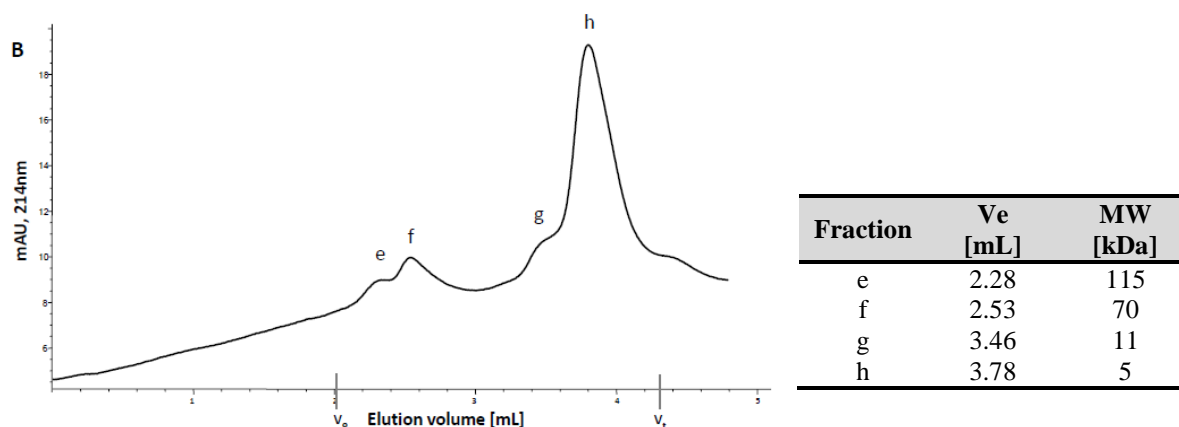
The chromatogram reported in Fig. 3A clearly showed the poor resolution of this technique [9], mainly referring to the peaks *a*, *b*, *c* and *d*. In addition a shoulder (peak *a*), was identified in correspondence to the peak *b*. Therefore, in order to better evaluate the volume of elution of each fraction and therefore the estimated MW, each fraction was collected and re-injected in the same column. In this way, narrower peaks were obtained and the MW could be evaluated. The SEC chromatogram confirmed the polydispersity of the gum, which is characterized by a certain molecular mass distribution. When considering the chromatogram obtained with the BioSep S2000 column (Fig. 3B), the presence of three protein fractions with a molecular weight included between 4 and 70 kDa was observed. More information could be obtained by the MW evaluation of the chromatogram reported in Fig. 3A. The sample was found to consist of four peaks. A major peak (peak *b*) with a shoulder (peak *a*), for which a molecular mass of ~2.3 million Da was estimated, and three minor peaks (*c*, *d* and *e*) whose MW was estimated to be respectively 1600, 900 and 300 kDa. Since UV response at 214 nm is indicative of the protein component within the gum, the chromatogram demonstrated how the observed fractions had a different protein content. Peak *b* was the most intense, thus revealing the higher amount of protein, while the other fractions had a decreasing intensity. The SEC chromatogram of gum arabic allowed to assign the observed peaks to the three fractions reported by Randall et al. [10]: AGP (peak *b*), AG (possibly peaks *c* and *d*) and GP (peak *e*) (Fig. 3A). In the literature three peaks are usually described for gum arabic, while the obtained SEC chromatogram showed the presence of two possible AG components corresponding to peak *c* and *d*. The obtained molecular weights were consistent with the ones reported in the literature, in terms of order of magnitude, although the values are higher [11]. However, it must be taken into consideration that the obtained MWs are an estimation of the real ones since the standard proteins used for calibration are usually globular.

In conclusion, it was demonstrated that gum arabic was characterized by a certain polydispersity. The presence of high molecular weight fractions, up to 2-3 million Dalton, was revealed and differences in intensities showed how these fractions are characterized by a different protein content.

#### 2.4.2. *Tragacanth gum*

The molecular mass distribution obtained by SEC of gum Tragacanth (Zecchi, Italy) is reported in Fig. 4. The high molecular weight fractions are showed in Fig. 4A, while the chromatogram related to the low molecular weight components is reported in Fig. 4B.





**Fig. 4.** SEC chromatogram of gum tragacanth (Zecchi, Italy) monitored at 214 nm: (A) MW range 100-7000 kDa; (B) MWs range 1-100 kDa. On the right: list of elution volume and estimated molecular weight of each observed fraction.

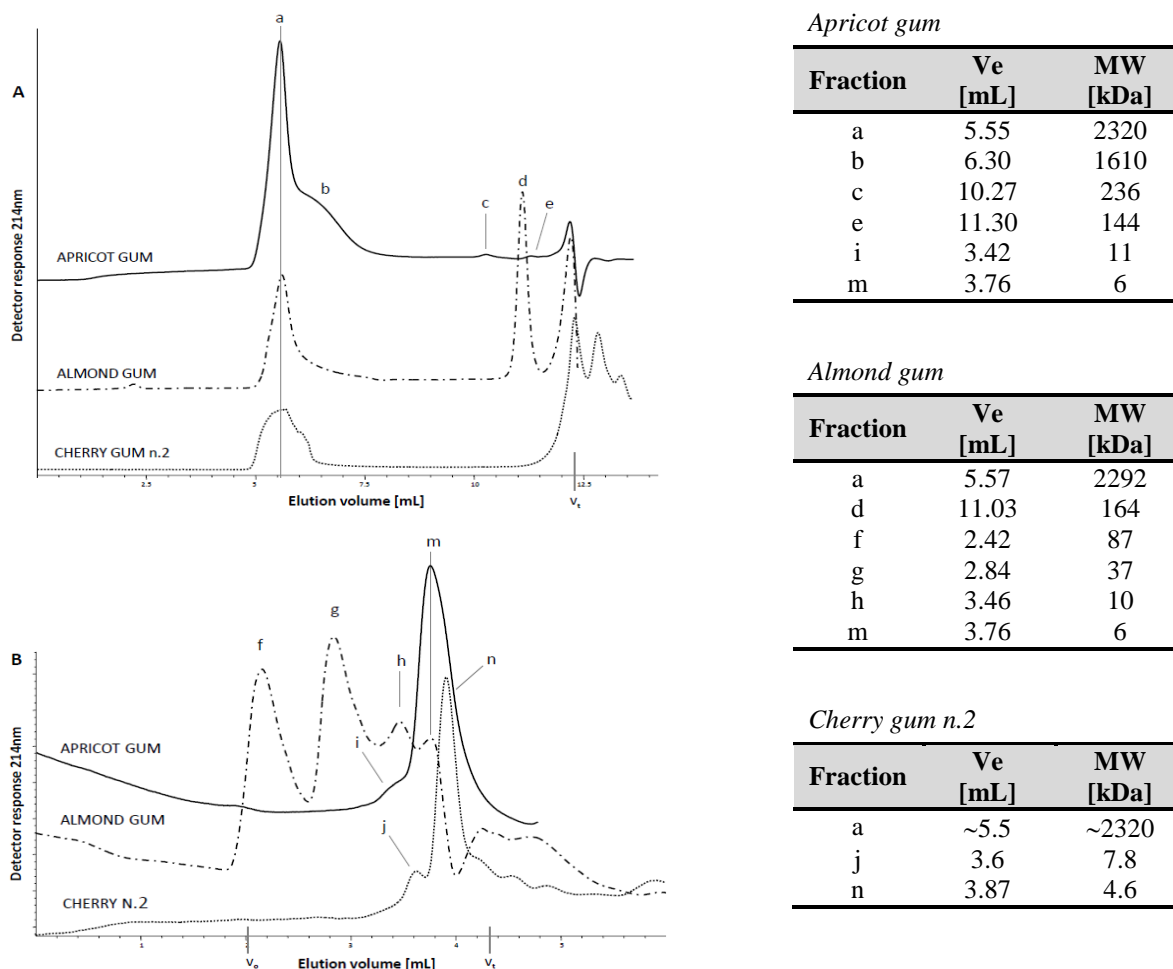
The SEC chromatogram obtained eluting the gum with the BioSuite HR column (Fig. 5A) showed the presence of a broad non-resolved peak named peak *a* ( $10.4 < R_t < 13.6$  minutes). Several fractions were collected along the elution of peak *a* and re-injected to evaluate the presence of hidden peaks. It was observed that the peak consisted of a unique fraction. The fact that it was not well resolved was maybe due to non-specific interaction between the macromolecules and the column stationary phase. However, the fraction was supposed to have a molecular weight over 2 million Da. Peak *c* (estimated MW = 230 kDa) had the highest intensity, thus showing that an higher content of protein was associated to this fraction. The last peak *d* eluted in proximity to the  $V_t$  of the column. Therefore its molecular weight was approximated to  $\sim 100$  kDa. The SEC chromatogram showed a certain gum polydispersity, while in the literature a single fraction with a MW =  $1.6 \times 10^6$  Da was identified [12]. In regard to the lower MW range (Fig. 4B), four peaks were observed with an intense one corresponding to peak *h* (estimated MW = 5 kDa). Peaks *e*, *f* and *g* were not well resolved and of low intensity.

#### 2.4.3. Fruit tree gums

The molecular mass distribution obtained by SEC of fruit tree gums is reported respectively in Fig. 5A, for the MW range 100-7000 kDa, and Fig. 5B for the mass range 1-100 kDa. Apricot, almond and cherry gums were collected in the summer 2012 in Veneto region (Italy).

The chromatogram reported in Fig. 5A showed some similarities and differences among the three analyzed gums. Apricot gum was characterized by 3 peaks: the first and with highest intensity (peak *a*) to which a molecular weight of 2320 kDa was attributed, peak *b* that was not well resolved but showed a significant intensity (MW = 1610 kDa) and two low intense peaks (*c* and *e*) that could contain a fairly limited amount of protein (MW = 236 and 144 kDa). The molecular mass distribution of apricot gum appeared to resemble the SEC chromatogram of gum arabic (Fig. 3A), both in terms of relative intensities of the peaks and corresponding estimated molecular weights. Therefore, it was supposed that the two gums could have a similar proteinaceous component. Regarding almond and cherry gums, a major peak with a molecular weight of  $\sim 2.3$  million Da was observed. The broad shaped peak observed in the chromatogram of cherry gum (peak *a*) could be

caused to both the presence of several fractions containing proteins with a similar MW, or to the effect of non-specific interaction between the analyte and the stationary phase. For almond gum, a well resolved peak (peak *d*) eluting at 22 minutes was observed (MW = 164 kDa).



**Fig. 5.** SEC chromatogram monitored at 214 nm of apricot gum, almond gum and cherry gum. (A) MW range 100-7000 kDa; (B) MWs range 1-100 kDa. On the right: list of elution volume and estimated molecular weight of fractions observed for each gum.

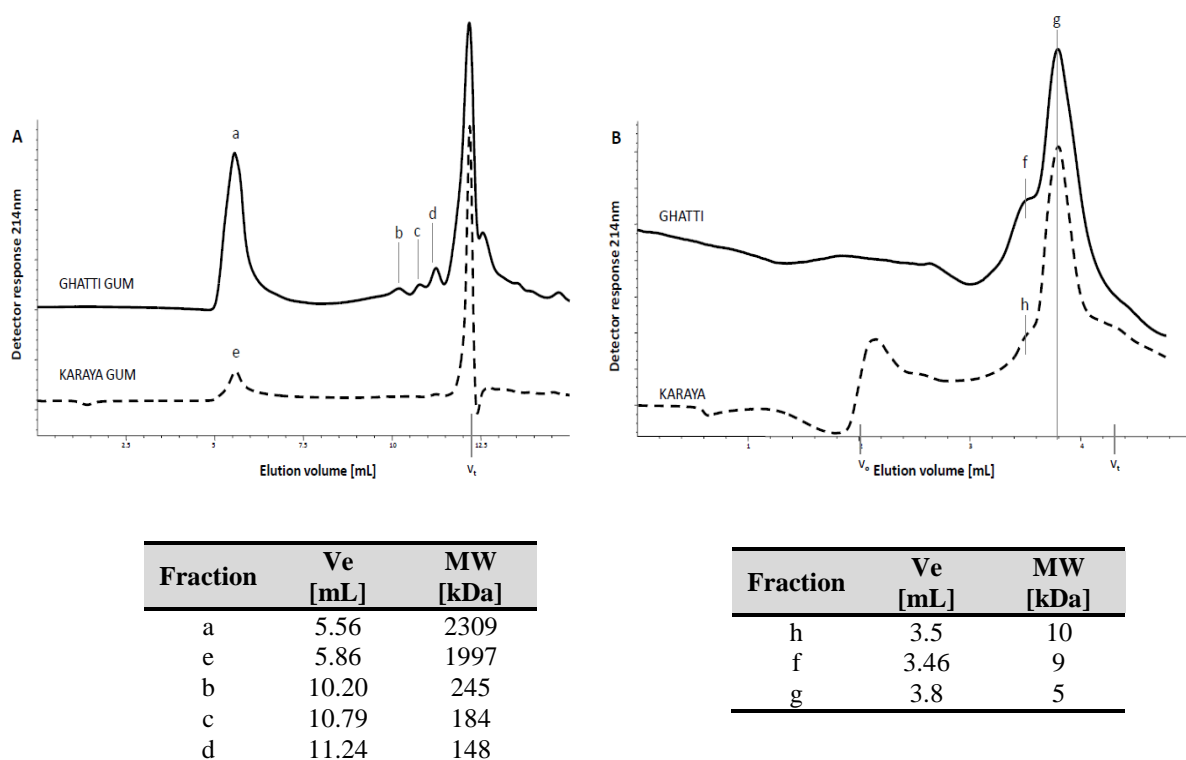
The chromatograms obtained eluting the gums with the Biosep S2000 column (Fig. 5B), showed the presence of several protein fractions with a molecular weight lower than 100 kDa. In regard of almond gum, four fractions with similar intensity and estimated molecular weights of 87, 37, 10 and 6 kDa were observed. A fraction with a comparable MW (around 6 kDa) was also identified in apricot gum. The profile showed how this fraction (peak *m*) could have the highest amount of protein if compared to the lower intense peak *i*. For cherry gum only two fractions were observed with an estimated molecular weight of 7.8 and 4.6 kDa.

In conclusion, the analyzed fruit tree gums were all characterized by an high molecular weight fraction of ~2.3 million Dalton. The intensity of the peak suggested a significant protein content in the corresponding fraction. The mass distribution at lower molecular weight showed instead some differences among the gums exuded from *Prunus* specie trees.

## 2.4.4. Guar, locust bean, ghatti and karaya gums

Among the other gums studied in this research, no chromatogram could be obtained for guar and locust bean gum. One of the main problems encountered concerned the sample preparation. Gum solutions had to be diluted several times in order to perform the injection. Therefore it was thought that the SEC technique was not sensitive enough to detect the proteinaceous component in a too diluted gum samples.

The molecular mass distribution obtained by SEC of ghatti and karaya gum are reported respectively in Fig. 6A, for the MW range 100-7000 kDa, and Fig. 6B for the mass range 1-100 kDa.



**Fig. 6.** SEC chromatogram of ghatti gum and karaya monitored at 214 nm. (A) MW range 100-7000 kDa; (B) MWs range 1-100 kDa. Below each graph: list of elution volume and estimated molecular weight of each observed fraction.

The chromatogram reported in Fig. 6A showed, for both gums, the presence of a fraction with an estimated molecular weight of around 2 million Da (peaks *a* and *e*) and characterized by a significant protein content. Three more minor fractions were observed for ghatti gum (peaks *b*, *c* and *d*). The intense peak in correspondence to the volume of total inclusion of the column indicated the presence of proteinaceous component with a MW lower than 100 kDa. This fraction was collected and injected in the second BiosepS2000 column. The corresponding chromatograms are showed in Fig. 6B. Ghatti and karaya gum showed a similar molecular mass distribution in the low MW range. The profile was characterized by an intense peak, to which a molecular weight of 5 kDa was assigned, and a non-resolved peak that appeared as a shoulder (MW ~ 10 kDa).



## 2.5. CONCLUSIONS OF SEC ANALYSIS

The capability of SEC columns to handle a significant wide range of molecular weights, from few thousand up to several million Dalton, allowed to further investigate the proteinaceous component of several plant gums. Analysis of gums is challenging due to the presence of glycoproteins and polysaccharides characterized by high molecular weights. While this feature represents a limit for most of the other analytical techniques, size exclusion chromatography resulted in a suitable method for gums investigation, even if it is characterized by a low resolution.

Different profiles were obtained for several plant gums. In the specific, the SEC chromatogram of gum arabic showed the presence of the characteristic AGP, AG and GP fractions reported in the literature. In addition, the presence of two possible AG components was observed. High molecular weight protein fractions, up to 2-3 million Dalton, were revealed and differences in intensities showed how these fractions are characterized by a different protein content. Tragacanth gum was characterized by an intense peak corresponding to a protein fraction with an estimated MW of 230 kDa. Fruit tree gums (cherry, apricot and almond gums) resulted to be all characterized by a common high molecular weight fraction of ~2.3 million Dalton and the high intensity of the peak suggested a significant protein content. However, investigation of the corresponding gums at lower molecular weights (1-100 kDa) showed some differences mainly related to almond gum that resulted to be characterized by four intense protein fractions.

In conclusion, different plant gums showed different chromatographic profiles characterized by (glyco)protein components with a molecular weight up to few million Dalton. However, these profiles did not result in a characteristic pattern suitable for plant gums identification. Therefore other analytical techniques will be employed in the following sections.

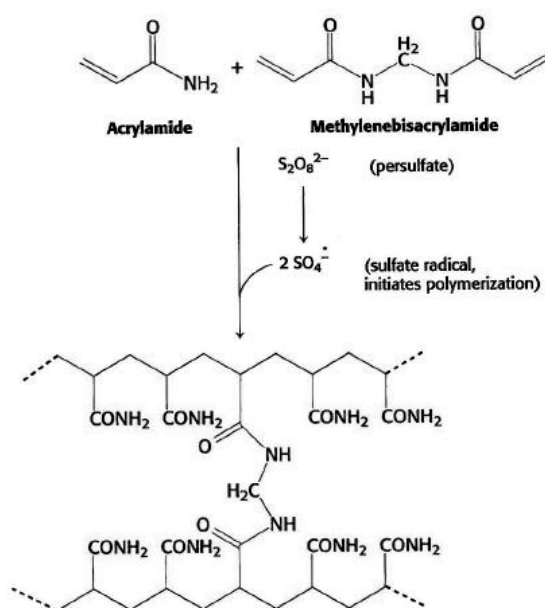
### 3. INVESTIGATION OF PLANT GUMS PROTEIN PROFILE USING GEL ELECTROPHORESIS

#### 3.1. GEL ELECTROPHORESIS

Gel Electrophoresis includes a number of techniques which are used for molecules separation such as DNA, RNA and proteins according to size, charge or shape. Electrophoresis is based on the principle according to which charged molecules can migrate when an electric field is applied, and since molecules with different physical characteristics will move in the matrix with a different speed, the separation is achieved. Gel electrophoresis is usually employed as a preparative technique before applying other methods such as mass spectrometry, mainly for protein analysis. As suggested by the name, gel electrophoresis is performed on a gel matrix which does not chemically react with the biomolecules and acts as a molecular sieve. Composition and concentration of the matrix depend on the type of molecules analyzed. For protein separation the most used are agarose, for molecular weights higher than 200 kDa, and acrylamide gels for smaller proteins (MWs up to 200 kDa).

##### 3.1.1. Polyacrylamide gel eletrophoresis (PAGE)

The use of acrylamide for gel preparation was first introduced by Shapiro *et al.* [13] and it is based on the cross linked polymerization of acrylamide and N,N'-methylenebisacrylamide. The reaction is a free radical catalyzed reaction and usually ammonium persulfate (APS) and the quaternary amine N,N,N',N'-tetramethylethylenediamine (TEMED) are used respectively as reaction initiator and catalyst (Fig. 7 [14]). The concentration of the initiator and the catalyst is of significant importance since it can influence the gel polymerization. A too rapid (less than 10 minutes) or too low polymerization (higher than 60 minutes) can result in a non homogeneous gel.



**Fig. 7.** Representation of acrylamide chains linked together by bis molecules.

If prepared correctly, polyacrylamide gels have a uniform porosity which can however be adapted according to the molecular weight of the proteins under investigation. The size of the pores is determined by two parameters: the weight percentage of total monomer (%T), which includes both acrylamide and N,N'-methylenebisacrylamide) and the ratio of cross-linker to acrylamide monomers (%C). If the %C value is kept fixed, the higher is %T, the higher is the number of chains in the gel, so the smaller is the pore size. Thus low percentage gels (e.g. 7%T) are better to resolve high molecular weight molecules, while gels with high %T (e.g. 15%T) are employed to resolve smaller proteins [15] (see Tab. 2). In conclusion, when using a homogeneous gel, a low reticulation allow to better separate molecules with an high molecular weight, while an higher reticulation is preferred when separation of low molecular weight analytes is required. A different type of gel is the so-called "gradient gel" which is characterized by a low acrylamide percentage at the top and an higher at the bottom. These gels are used for the separation of a wider range of molecular weights [16].

**Tab. 2.** Range of protein MWs resolved by gel with a different percentage of acrylamide

Acrylamide %	Protein MW range [kDa]
15	~1-100
12	~20-150
10	~30-250
7	~50-300
4	~60-400

### 3.1.2 Sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) allows protein separation according to their molecular weight. This method, also referred to monodimensional electrophoresis, implies the use of SDS, an anionic detergent that forms complexes with the proteins, conferring them an homogeneous net negative charge. Therefore, when an electric field is applied, the negatively charged proteins will migrate along the polyacrylamide gel and will be separated just according their molecular weight. The polyacrylamide gel will work as a sieve, thus separating the proteins according to their size. Therefore, the high molecular weight proteins will get stuck at the beginning of the gel, while the proteins with a lower molecular weight will be found at the end of the gel, both in the form of bands. As described before, polyacrylamide gel reticulation can be modified, thus resulting in the possibility to separate proteins in a different MW range. In addition, if standard proteins with a known MW are run simultaneously with unknown samples, the MWs of these last ones can be estimated.

An electrophoretic approach that combines the isoelectric focusing (IEF) [17] and SDS-PAGE methods was developed in 1975 [18] and consisted in the two-dimensional gel electrophoresis (2-DE). In the first dimension proteins are separated according to their isoelectric point, while in the orthogonal second dimension proteins are separated according to their molecular weight as in a classic SDS-PAGE.

### 3.1.3. Gel staining

Once the electrophoresis run is ended, the following step is the staining of the gel which is necessary to visualize the proteins position in the gel matrix. Proteins are usually fixed on the gel with a proper fixing solution and then stained. The most used staining solutions are: coomassie blue, silver and SYPRO ruby staining.

#### COOMASSIE BLUE

Coomassie blue R250 binds non-specifically and approximately stoichiometrically to almost all proteins and protein coloration occurs by interaction of the organic colorant with the basic amino acids of the protein. Although this staining method is less sensitive if compared to other reagents (sensitivity 10-300 ng), it is commonly used due to its fast colorimetric response and convenience. The gel is simply soaked in a dye solution and, after some washing steps of the gel, proteins are detected as blue bands while the not bound dye diffuses through the gel and it can be washed away [19]. A modification of the classical coomassie blue consists in the so-called colloidal coomassie. The method, which implies the use of a modified coomassie dye (G-250), was developed in the 80s [20] and was further optimized to improve sensitivity [21]. The main advantages of this method concern the improved sensitivity (around 10 ng) and reproducibility since any de-staining of the gel has to be performed. This is due to the fact that, on the contrary of the classical coomassie dye, the colloidal particles do not enter the gel pores. One of the main disadvantage of both staining methods based on blue coomassie is the narrow linear dynamic range (10-30).

#### SILVER STAINING

The silver staining procedure is based on the property of proteins to bind silver ions which, under specific conditions, can be reduced to  $Ag^0$  to form a brown insoluble precipitate of metallic silver, thus making the protein visible. Even if several silver staining protocols are proposed in the literature, the procedure is based on some defined steps: (1) fixation, in order to remove possible interfering compounds; (2) sensitization; (3) impregnation of the gel with a silver solution; (4) development to obtain the metal silver by formaldehyde and (5) rinse to end the development [22]. Several variations of each step have been suggested in the literature but all the different protocols can be grouped in two main families according to the type of silver solution that is used to impregnate the gel: silver nitrate or silver-ammonia complex. While the silver nitrate solution is used in an alkaline solution, the silver ammonia complex is associated to an acid solution. This staining method is really sensitive (LOD from 100 pg to few ng of proteins) but it is time consuming and susceptible to different factors, thus its reproducibility is poor.

#### SYPRO RUBY

Besides the above described staining techniques which allow to visualize the proteins in the visible spectra, there are other techniques which are based on the fluorescence. Sypro staining belongs to this category and, in addition to the possibility to detect amount of proteins in the order 1-2 ng, this method allows also to stain compounds such as glycoproteins, lipoproteins, fibrillar proteins and others which are usually difficult to be stained [23]. Furthermore, the linear dynamic range covers three orders of magnitude [24]. Different fluorophores have been commercialized (e.g. SYPRO Red, SYPRO Orange) and among these the SYPRO Ruby was also introduced [25]. Sypro Ruby

staining process is based on a metal chelate stain of ruthenium as part of an organic complex which interacts non-covalently with proteins. After a first step of staining and a quick washing, the gel can be scanned under blue light (maximum excitation at 302 and 470 nm) and so the protein position can be recorded. Less than one ng of protein can be detected by this staining method [23]. This type of stain does not interfere with later applications such as in-gel protein extraction [26].

### 3.2. OPTIMIZATION OF ELECTROPHORETIC EXPERIMENTS

The first step to evaluate during the optimization of an analytical method consists in the sample preparation, which has to be appropriate for the type of analytes and analysis that will be performed. In the present research, two main parameters had to be considered: the gum amount and the proper buffer/solvent for gum solubilization. Regarding the first point, as already pointed out in Chapter I, section 3, since the protein content in plant gums is usually  $< 5\%$  w/w, the amount of gum loaded in the gel has to be enough to get a detectable amount of protein. Secondly, plant gums have different physical/chemical properties that can influence the selection of the buffer/solvent. Since gums are mainly carbohydrates, the best choice fell on water. However, only gum arabic is completely soluble in water. All the other gums are only partially soluble (e.g. the soluble component of gum tragacanth represents the 30% of the total weight [12]). Therefore, if the concentration is higher than 5% w/v, a gel-like consistence is obtained, thus preventing any further sample handling.

Optimization of sample preparation was performed working with gum arabic (Zecchi, Italy). As first experiment, the gum was solubilized in milliQ water. Several concentrations were tested and a value of 3% w/v (30 mg/mL) was observed to be the maximum amount of gum that allowed to get a handling sample. Solution was mixed overnight at room temperature and different aliquots were loaded on the polyacrylamide gel.

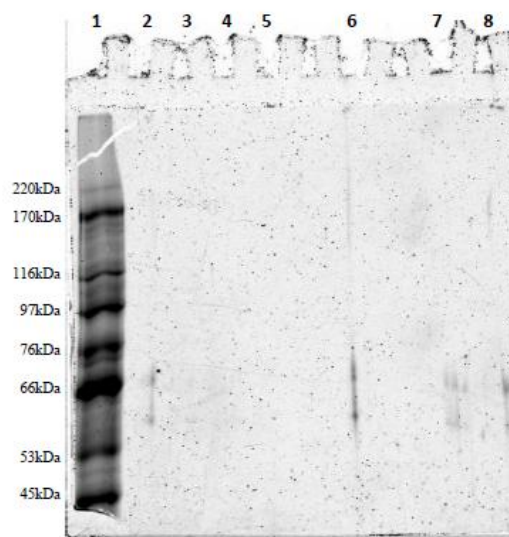
In order to have an idea of the amount of protein loaded, quantitation of gum arabic protein content was performed by BCA Protein Assay Kit (Thermo Scientific Pierce®). The method is based on the capability of proteins to reduce  $\text{Cu}^{2+}$  to  $\text{Cu}^{+}$  in alkaline solution. The used reagent contains bicinchoninic acid (BCA) [27]. Two molecules of BCA chelate one  $\text{Cu}^{+}$  ion resulting in water-soluble complexes that have a strong absorbance at 562 nm. The absorbance is linear in a protein range concentration of 20 - 2000  $\mu\text{g/mL}$ . Therefore, the unknown protein concentration of a sample can be evaluated with reference to the concentration curve of a specific standard. All procedural details are described in the chapter “Materials and Methods”. A protein content of 1% was estimated ( $R^2 = 0.998$ ).

Since the analyzed samples were not pure proteins, SDS-PAGE was performed by testing different proportions sample/Laemmli buffer [28] (for the gel preparation procedure see the chapter “Materials and Methods”). The different proportions and the corresponding estimated protein content of each loaded sample are reported in Tab. 3. The scanned image of the final polyacrylamide gel is reported in Fig. 8.

**Tab. 3.** Estimation of the protein amount loaded on the gel.

Gum/Laemmli	Gum volume [ $\mu\text{L}$ ]*	Protein content [ $\mu\text{g}$ ]
1/7	3.8	1.1
1/3	7.5	2.2
1/2	10	3
1/1	15	4.5

\* maximum total volume is 30 $\mu\text{L}$



**Fig. 8.** SDS-PAGE 10% acrylamide of gum arabic (Zecchi). Different ratios gum/Laemmli (v/v) tested: (1) MW standards; (2) 1:7; (3) 1:3; (4) 1:2; (5) 1:1; (6) 1:7 gum heated at 45°C for 60 minutes; (7) 1:2 gum heated at 45°C for 60 minutes; (8) 1:1 gum heated at 45°C for 60 minutes.

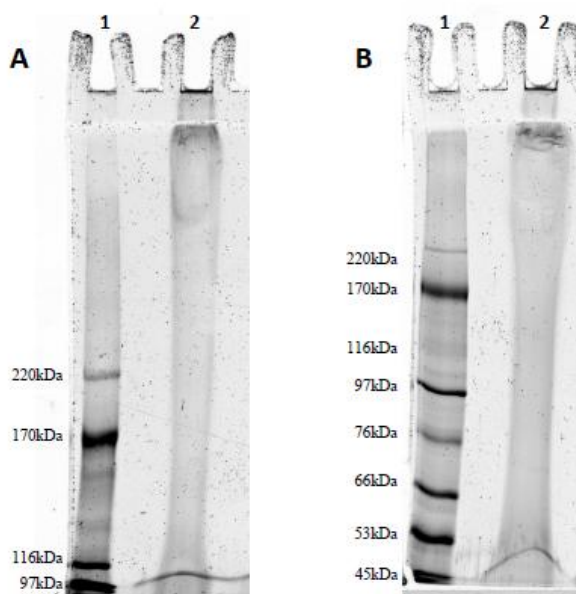
Two faint bands between 50 and 66 kDa were observed for gum arabic sample heated at 45°C for 60 minutes (lanes 7 and 8). However, besides these ones, no other protein bands were observed in the gel. It was concluded that either the gum proteins were not completely solubilized, or the amount of gum loaded on the gel was not enough to get a sufficient protein content that could be detected. However, the concentration of the gum in solution could not be increased, due to the limitations explained above related to the solution viscosity. In addition, even the volume of sample loaded could not be increased to get more protein, since the maximum volume that could be loaded on the gel was of 30  $\mu\text{L}$ . Therefore, a test consisting in increasing the dimension of the wells in the stacking gel was performed. Although the experiment allowed to load a double volume of sample (~60  $\mu\text{L}$ ), no protein bands were observed.

In order to overcome these limitations, a new strategy consisting in dissolving the gum directly in the Laemmli denaturant buffer was tested. As described in the chapter “Materials and Methods”, Laemmli solution contains denaturing, detergent and reducing agents that should promote gum/protein solubilization. Furthermore, dissolving the gum directly in Laemmli buffer, and not mixing a certain aliquot, would allow to increase the amount of gum loaded in the gel.

Referring to gum arabic, this method allowed to improve significantly the amount of protein loaded in the gel, that was estimated to be around 9  $\mu\text{g}$  (in a total volume of 30  $\mu\text{L}$ ). According to the

results obtained with size exclusion chromatography, proteinaceous component with an high molecular weight (over 300 kDa) are present in gum arabic. Therefore, first experiments were focused on the optimization of polyacrylamide gels in order to resolve these high molecular weight fractions. In order to achieve this goal, gels with low reticulation (%T) were prepared and gum arabic sample was analyzed. Both 7% and 4.7% gels were used and the corresponding scanned images are reported respectively in Fig. 9A and B.

If compared to Fig. 8, it was immediately possible to conclude that the solubilization of gum in Laemmli buffer was essential in order to visualize the gum proteinaceous component in the gel. Observing lane 2, corresponding to gum arabic, it was possible to conclude that gels with a low reticulation allowed high molecular weight proteins to have access to the separation gel. In the 7% gel (Fig. 9B) the proteins with higher MW appeared to be blocked in the stacking gel, along it and at the right beginning of the separation gel. The stacking gel has a 4.7% of acrylamide, thus resulting in large pores which should facilitate the access to the separation gel of high molecular weight molecules (around 400 kDa). It was therefore confirmed that the proteins that could not enter the gel had a molecular mass over 300 kDa. In the 4.7% gel (Fig. 9A) the high molecular weight proteins of gum arabic seemed to better enter the separation gel.

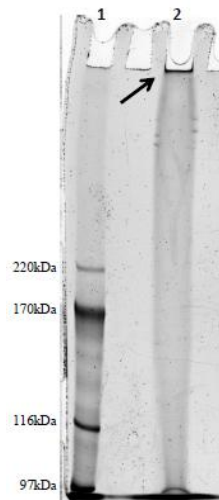


**Fig. 9.** SDS-PAGE pattern of (1) MW standards; (2) gum arabic. (A) 4.7% polyacrylamide gel and (B) 7% polyacrylamide gel.

However, the 7% and 4.7% polyacrylamide gels did not result to be enough resolutive since any protein band could be observed.

A further experiment was performed consisting in the preparation of a 4.7% acrylamide gel with a higher molecular weight dextran. As reported in the chapter “Materials and Methods”, around 0.1% w/v of dextran 500 (MW = 500 kDa) is added to the acrylamide/bisacrylamide solution in order to improve the gel resolution. A 2 million Dalton MW dextran was instead added to the acrylamide solution and gum arabic sample was loaded on the gel. As showed in Fig. 10, the high molecular weight proteins could better enter the separation gel. However, a black line due to Sypro staining (lane 2) still indicated the presence of some macromolecules that could not migrate along the gel,

thus showing how the use of a 2 million Dalton MW dextran was not suitable for the separation of these macromolecules.



**Fig. 10.** SDS-PAGE (4.7% acrylamide + 2000 kDa dextran) pattern of proteins from : (1) MW standards; (2) gum arabic.

In the light of the previous experiments, in order to further investigate and improve the separation of the high molecular weight protein components, a lower percentage of acrylamide, which correspond to a minor gel reticulation, is required. However, the low the acrylamide percentage is, the more complex the handling of the gel is. Though 3% is the lowest percentage of acrylamide that allows gel polymerization, even 4.7% acrylamide gels are difficult to work with. In addition, at this low percentage (see Fig. 9A), the PAGE gel matrix was too reticulated to allow the complete entrance of large molecules. Therefore, in order to achieve the objective, a different type of gel had to be developed.

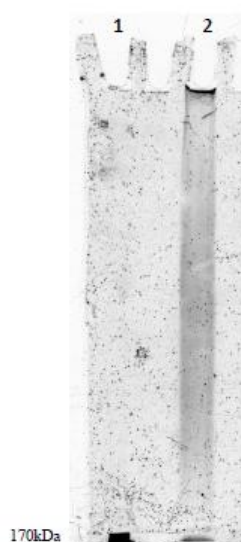
As described in the literature, gum arabic contains an arabinogalactan component that is characterized by a polypeptide chain to which polysaccharide blocks are *O*-linked. The result is an heavily glycosylated protein with high molecular weight that resembles mucins structure [29]. PAGE has proved to be unsuitable for the separation of these heavily glycosylated proteins with molecular weight higher than 500-600 kDa [30,31]. In order to access the characterization of these macromolecules, agarose gel electrophoresis (AGE) was reported to be an appropriate technique for their analysis [32-34]. However, as it can be observed in the gels reported in the literature, agarose gel electrophoresis has a low resolving power and they are characterized by not well defined bands [35,36].

A solution to the limitations associated to both polyacrylamide and agarose gel matrix, for the analysis of high molecular weight proteins, resulted in a combination of the useful properties from the two of them. Uriel J. in 1966 [37] described a new method for protein electrophoresis based on a gel composed of both acrylamide and agarose. However, an acrylamide amount lower than 3% could not be achieved. In 1968, Peacock and Dingman developed a composite gel of 2% acrylamide and 0.5% agarose for separation of ribonucleic acids [38]. A similar gel was optimized by Tatsumi and Hattori for SDS-denaturing electrophoresis of giant myofibrillar proteins [39]. Since the first developments, agarose-acrylamide composite gels (Ag-PAGE) have been recently applied for separation of high molecular weight mucins [40] and some modifications, as the



addition of urea to simplify gel casting procedure, were proposed [41]. The inclusion of agarose in Ag-PAGE composite gel was demonstrated to allow the preparation of low percentage acrylamide gels (< 4%). The new mixed matrix gel is therefore suitable for separation of high molecular weight proteins, while still retaining enough strength to be handled and the good resolution properties of classical acrylamide gels.

At the light of these considerations, gum arabic was investigated by agarose-acrylamide composite gel electrophoresis in denaturing conditions (SDS-AgPAGE). A 2% acrylamide/0.5% agarose gel was prepared following the procedure reported in the chapter “Materials and Methods” [39]. Gel was stained by Sypro and the corresponding scanned image is reported in Fig. 11.

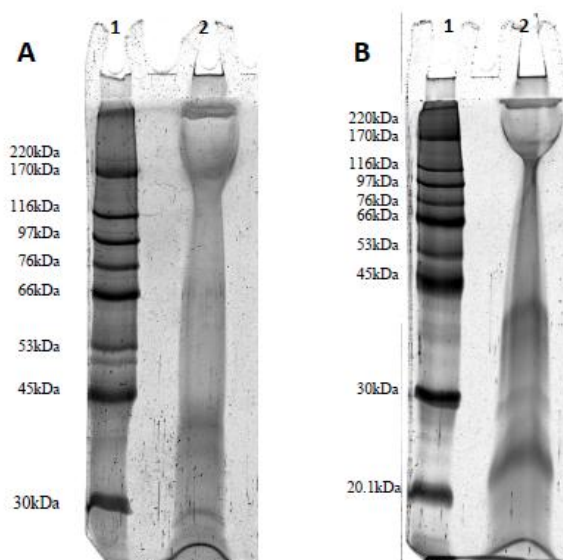


**Fig. 11.** DS-AgPAGE (2% acrylamide + 0.5% agarose) pattern of proteins from : (1) MW standards; (2) gum arabic.

The composite agarose-acrylamide gel allowed the proteins with higher molecular weight to better migrate along the gel. Any alteration of the gel could be observed in correspondence of gum arabic (lane 2), thus demonstrating the access of the proteins into the separation gel. However, Sypro staining revealed how some other protein components were still blocked at the beginning of the stacking gel (black line), while the proteins that could migrate were not well resolved.

An additional agarose-acrylamide composite gel was tested and acrylamide percentage was decreased to 1%. Unfortunately, no results could be achieved since the gel did not polymerized. It was therefore concluded that 2% acrylamide was the lowest percentage suitable for complete polymerization.

In conclusion, the four tested gels (4.7%, 7%, 4.7% with the addition of 2 million Da dextran and the composite SDS-AgPAGE) demonstrated how the high molecular weight proteins could access the low reticulated gel, but no electrophoretic profile could be obtained for gum arabic because of the low resolutive power. Therefore, in order to achieve this goal, it was decided to reduce the gel porosity and analyze the low molecular weight protein components by 10% and 12% polyacrylamide gels. The corresponding scanned images are reported in Fig. 12A and B.



**Fig. 12.** SDS-PAGE pattern of (1) MW standards; (2) gum arabic. (A) 10% polyacrylamide gel and (B) 12% polyacrylamide gel.

In the entering area of the 10% and 12% separation gel (over 220 kDa), strong protein bands were observed. These proteins were even blocked in the 4.7% stacking gel and could not be resolved by SDS-PAGE because of their high molecular weight.

However, the two gel showed a first electrophoretic profile of gum arabic. In the 10% gel (Fig. 12A), three fairly intense and large bands were observed: one at 170 kDa and two between 45 kDa and 30 kDa. Other two faint bands could be identified at ~60 and 70 kDa, and more proteins were observed to have a molecular weight lower than 30 kDa. The 12% (Fig. 12B) acrylamide gel allowed to better resolve some protein bands that in the 10% acrylamide gel were fairly visible at the end of the gel. A rather intense band at 25 kDa was observed in gum arabic.

In conclusion, optimization of classical SDS-PAGE for the analysis of plant gum samples, resulted in the necessity to solubilize the gum directly in the denaturing Laemmli buffer. Furthermore, although polyacrylamide gels with a low reticulation (4.7% and 7%) allowed the entrance of high molecular weight proteins, they resulted to be not enough resolutive. On the contrary, more reticulated gels (10% and 12%) showed the possibility to get an electrophoretic profile of gum arabic.

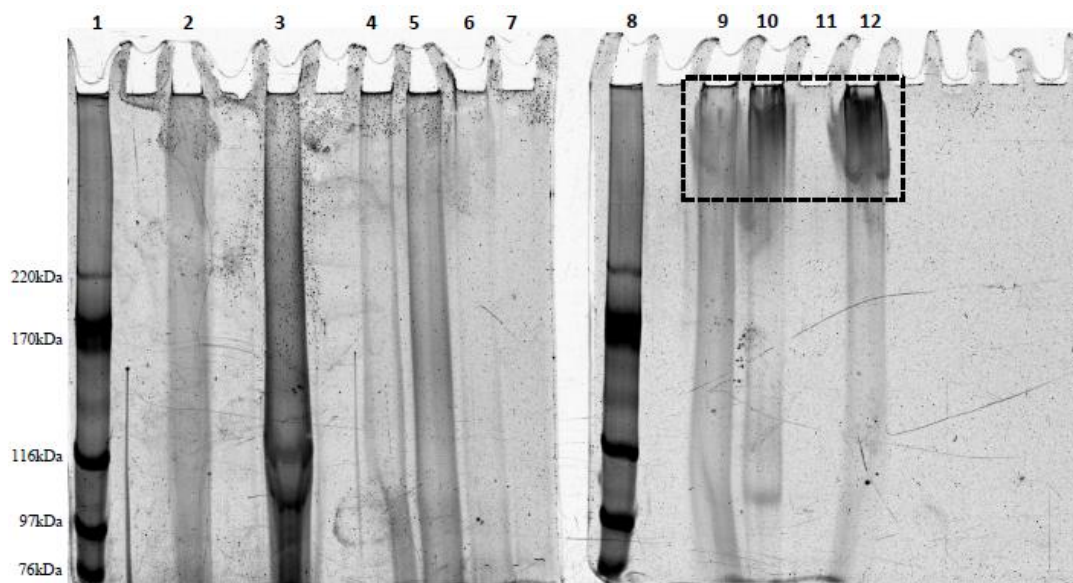
### 3.3. SDS-PAGE PROFILE OF PROTEINS IN SEVERAL PLANT GUMS

Several plant gums such as tragacanth gum, ghatti, guar, karaya, locust bean and a series of fruit tree gums (cherry, apricot, almond and plum gums) were analyzed by polyacrylamide gel electrophoresis and different percentage of acrylamide were tested in order to evaluate the widest range of molecular weights and to identify a possible electrophoretic profile for each gum. All different gums were solubilized in Laemmli at a specific concentrations reported in Tab. 4. Solutions were mixed overnight at room temperature and the soluble part was then recovered after centrifugation at 13400 rpm for 30 minute. 30  $\mu$ L solution were collected and loaded on the gel.

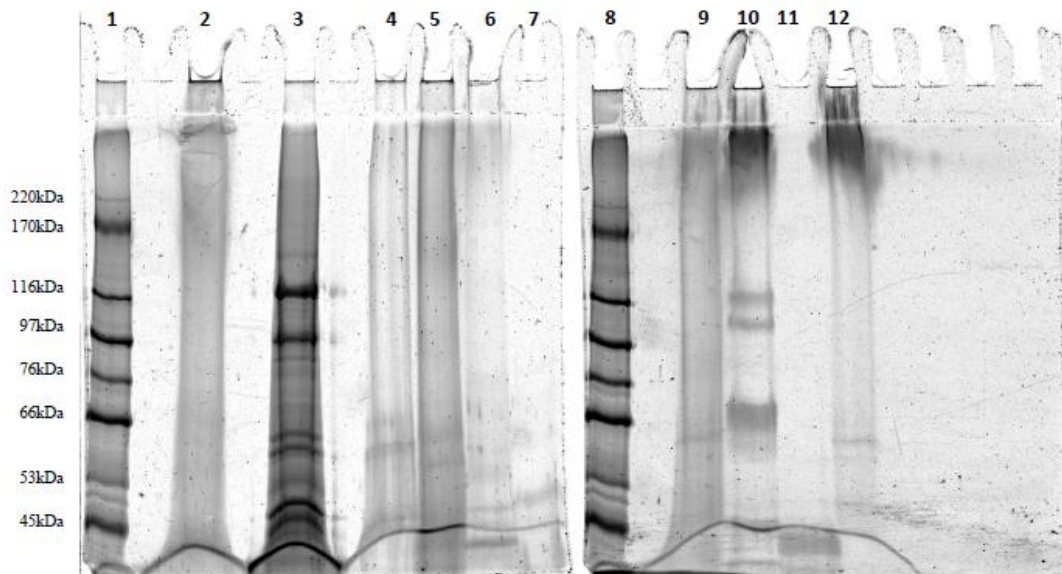
**Tab. 4.** Summary of the gums analyzed by SDS-PAGE, corresponding amount of gum solubilized in Laemmli buffer and brands.

Gum	% [w/v]	Brand
Arabic	3%	Zecchi (Italy)
Tragacanth	1%	Bresciani (Italy)
Fruit gums	2%	Collected in Veneto Region (Italy)
LBG	1%	Sigma Aldrich (Germany)
Guar	0.8%	Sigma Aldrich (Germany)
Ghatti	1%	Sigma Aldrich (Germany)
Karaya	1%	Sigma Aldrich (Germany)

Although investigation of gum arabic by polyacrylamide gels with a low reticulation (e.g. < 7%) did not allow to get an electrophoretic profile of the gum, several other plant gums were loaded on these gel in order to evaluate if any difference occurred concerning the high molecular weight proteinaceous components. Both 4.7% and 7% gels were used and the corresponding scanned images are reported respectively in Fig. 13 and 14.



**Fig. 13.** SDS-PAGE (4.7% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.



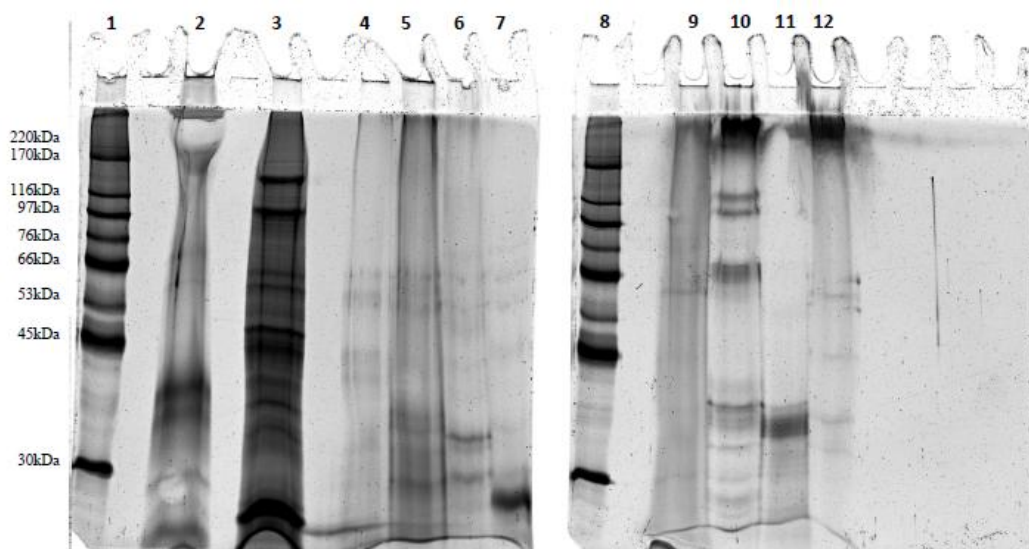
**Fig. 14.** SDS-PAGE (7% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.

Analysis of other plant gums by a 4.7% polyacrylamide gel (Fig. 13) showed some similarities and difference with the already investigated gum arabic (Fig. 9A). For guar gum (lane 4), tragacanth gum (lane 5), ghatti (lane 6) and karaya (lane 7), any electrophoretic profile could be obtained. However, as already observed by size exclusion chromatography, the presence of high molecular weight protein components, that appeared stuck at the beginning of the gel, was confirmed. On the contrary, two intense bands at around 120 kDa and 100 kDa were observed for locust bean gum (lane 3). Interesting results were also obtained for the fruit tree gums. Besides an intense black area (frame in Fig. 13) indicating an high concentration of proteins with a MW higher than 250 kDa, one band could be observed for apricot gum (lane 10) between 116 and 97 kDa.

More interesting information were obtained by increasing the percentage of acrylamide from 4.7% to 7% (Fig. 14). Several bands, more or less intense, with a molecular weight lower than 66 kDa were observed for LBG (lane 3), guar gum (lane 4), tragacanth gum (lane 5), ghatti (lane 6) and karaya gum (lane 7). Concerning the fruit tree gums, a possible profile of apricot gum (lane 10) was revealed with some protein bands below 116 kDa. Several faint protein bands in correspondence of a MW lower than 66 kDa were also obtained for the other gums from *Prunus* tree species.

These promising results demonstrated the possibility to get an electrophoretic profile by investigating the proteinaceous components with a molecular weight lower than 120 kDa. Therefore, in order to achieve this objective, the plant gums under investigation were analyzed by polyacrylamide gels with an increasing reticulation: 10%, 12% and 15%.

At the beginning a 10% acrylamide gel was prepared in order to have a broad view of protein fractions in a range between 250 and 30 kDa. The Sypro stained gel of all gum samples and MW markers is shown in Fig. 15.



**Fig. 15.** SDS-PAGE (10% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.

The two obtained 10% acrylamide gels are characterized by a number of protein bands spanning the whole range of molecular weights, thus showing a specific electrophoretic profile for each plant gum.

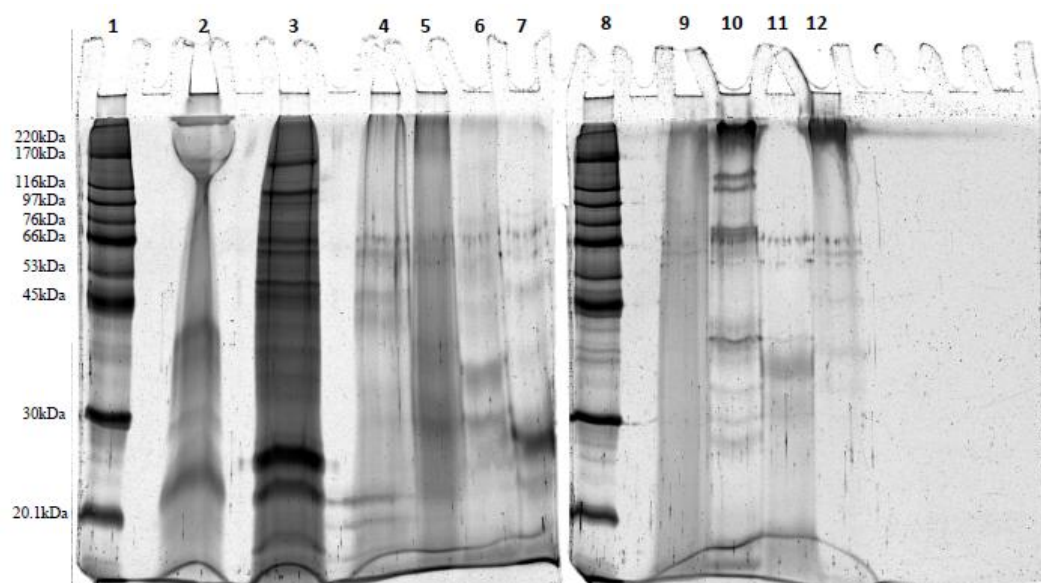
The first gel on the left represents the SDS-PAGE protein patterns of gum arabic (lane 2), locust bean gum (lane 3), guar (lane 4), tragacanth (lane 5), ghatti (lane 6) and karaya gum (lane 7). As already observed, gum arabic was characterized by proteins with a molecular mass higher than 220 kDa. These macromolecules could not enter the gel and they resulted blocked in the 4.7% stacking gel, along it and at the right beginning of the 10% separation gel. In addition, three fairly intense and large bands were observed at 170 kDa and below 45 kDa. Other two faint bands could be identified at ~60 and 70 kDa, and more proteins were observed to have a molecular weight lower than 30 kDa.

Several bands were observed for locust bean gum (lane 3). The gum seemed to contain proteins with a molecular weight lower than 220 kDa. Bands resulted to be intense and well resolved, with several faint bands that were supposed to correspond to some denaturation products since the electrophoresis was performed in denaturing conditions (Laemmli). The higher amount of protein bands was observed in the mass range between 45 and 30 kDa, and below 30 kDa. Therefore, LBG proteins was further investigated with higher percentage acrylamide gels (12% gel in Fig. 16 and 15% gel in Fig. 17). Guar gum (lane 4) was characterized by few weak bands at 60 and 40 kDa. A similar pattern profile was observed for ghatti (lane 6) and karaya gum (lane 7). However, the SDS-PAGE pattern of these two gums showed the presence of some more intense and resolved bands: three protein bands between 30 and 40 kDa for gum ghatti, and a significant intense band of around 30 kDa for karaya gum. Finally, for gum tragacanth (lane 5) the SDS-PAGE analysis showed the presence of high molecular weight proteins that got blocked along the lane. Besides three faint bands between 66 and 53 kDa, several protein bands were observed in the mass range 45-30 kDa, revealing the presence of proteins with low molecular mass.

Gums collected from different fruit trees were analyzed and results are showed in the gel on the right of Fig. 15. Among them, almond gum (lane 11) resulted to be the only gum not containing high molecular mass proteins. The stacking gel and the initial part of the separation gel were not stained by Sypro, while an intense band could be observed at around 35 kDa. Differently, the lanes corresponding to cherry (lane 9) and mainly apricot (lane 10) and plum (lane 12) gums, showed how these fruit tree gums contain some proteins with a molecular weight higher than 220 kDa. The protein components of cherry gum were not resolved by SDS-PAGE and only a faint band was observed at 60 kDa. The 10% acrylamide gel of apricot gum (lane 10) clearly indicated that the gum contain several proteins with a molecular mass spanning the whole range of MWs. Two strong and resolved bands were observed at ~ 100 kDa, followed by a large one at 66 kDa and other faint ones with a molecular mass lower than 45 kDa. Plum gum (lane 12) showed instead a strong band corresponding to a high molecular weight proteinaceous fraction (over 220 kDa), and few weak bands at 60, 45 and ~ 40 kDa.

In conclusion, the 10% polyacrylamide gel displayed how different plant gums have a different electrophoretic profile, which is characterized by protein bands at diverse molecular weights. Gum arabic, locust bean gum and apricot gum showed some intense and well resolved bands, thus indicating the possibility to distinguish them by a simple analytical method such as gel electrophoresis.

Due to the presence of protein components with a molecular weight lower than 45 kDa, plant gums were analyzed by SDS-PAGE using gels with an higher percentage of acrylamide. Both 12% and 15% gels were used and the corresponding scanned images are reported in Fig. 16 and 17.



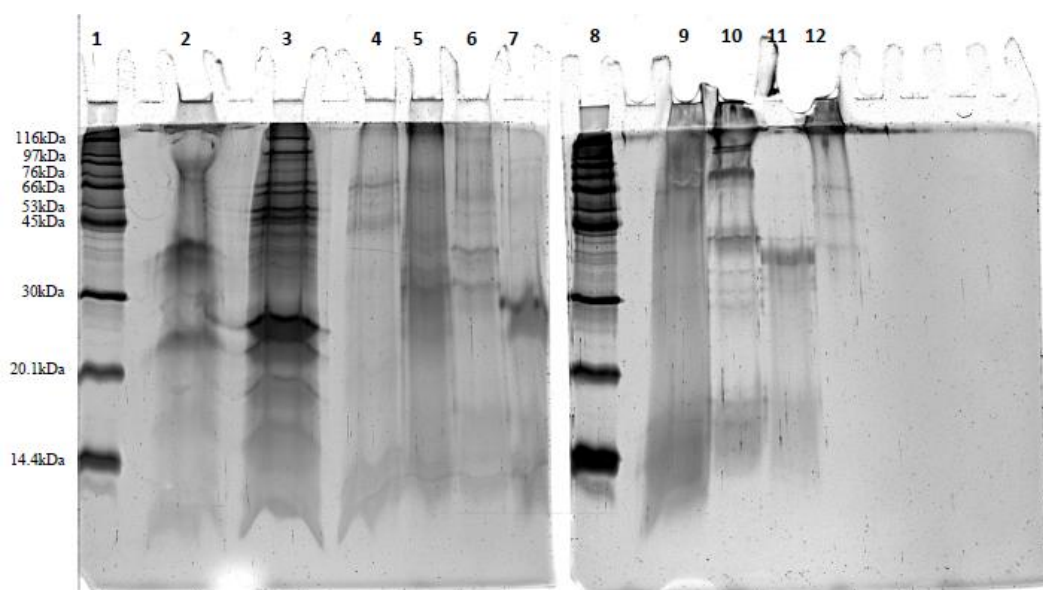
**Fig. 16.** SDS-PAGE (12% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.

The 12% acrylamide gel allowed to examine in depth the proteins with molecular weights between 30 and 20 kDa, while the 15 % acrylamide gel was selected in order to highlight the MWs range below 20 kDa. In both cases, high molecular weight fractions could not be resolved since high

values of %T implied low pores size, resulting in the impossibility of high MW proteins to enter the gel and be separated. However the protein profile of the gums could be further investigated.

For both gum arabic and LBG, the 12% acrylamide gel allowed to better resolve some protein bands that in the 10% acrylamide gel (Fig. 15) were fairly visible at the end of the gel. A rather intense band at 25 kDa was observed in gum arabic (lane 2), while two significant intense bands between 30 and 20 kDa, with a faint one below 20 kDa, were revealed for locust bean gum (lane 3). No further information were obtained for tragacanth (lane 5) and ghatti gum (lane 6), while two intense bands at ~ 20 kDa were revealed for guar gum. Furthermore, a new faint protein band with a molecular weight of around 25 kDa was observed for karaya gum (lane 7). Its low intensity suggested that the band might corresponded to a product of denaturation of the more intense protein band at 30 kDa. Regarding fruit tree gums (gel on the right of Fig. 16), no more protein fractions were observed under 30 kDa, except for a weak band observed for apricot gum (lane 10) close to the end of the gel.

The 15% acrylamide gel reported in Fig. 17, showed how proteins with a molecular weight below 20 kDa could not be resolved. For all gums the corresponding area of the gel resulted stained by Sypro Ruby but no bands could be observed.

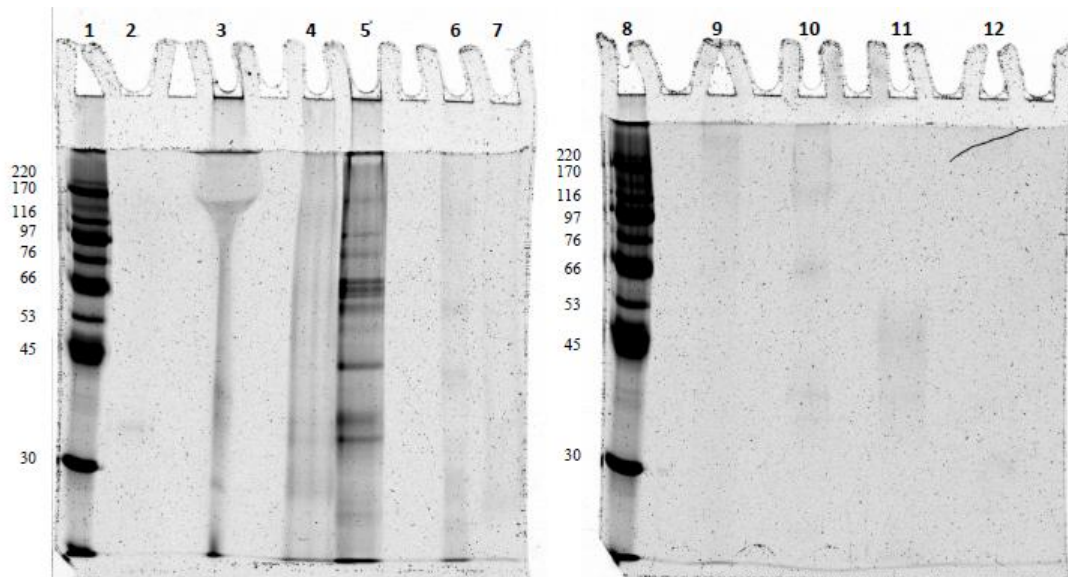


**Fig. 17.** SDS-PAGE (15% acrylamide) pattern of proteins from : (1) MW standards; (2) gum arabic; (3) LBG; (4) guar gum; (5) tragacanth gum; (6) ghatti gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.

### 3.4. PAGE UNDER NON-DENATURING CONDITIONS

Polyacrylamide gel electrophoresis under non-denaturing conditions allow to analyze proteins in their native form since no denaturing compound (e.g. SDS) or reducing agent (e.g. DTT) are used [42,43]. Therefore, plant gums were analyzed in non-denaturing conditions and results were compared to the SDS-PAGE gels in order to investigate the proteinaceous fractions in their native

conformation. Gum samples were solubilized in Laemmli buffer (see chapter “Materials and Methods”) which was prepared without the addition of SDS and DTT. Samples were analyzed by a standard 10% acrylamide gel and the results are reported in Fig. 18.



**Fig. 18.** Non-denaturing-PAGE (10% acrylamide) pattern of proteins from : (1) MW standards; (2) ghatti; (3) gum arabic; (4) tragacanth gum; (5) LBG; (6) guar gum; (7) karaya gum; (8) MW standards; (9) cherry gum; (10) apricot gum; (11) almond gum; (12) plum gum.

Results obtained with PAGE in non-denaturing conditions showed some differences if compared to the corresponding 10% acrylamide SDS-PAGE gels (Fig. 15). The most evident dissimilarity concerned the fruit tree gums (Fig. 18, gel on the right). Except for some faint bands observed for cherry, apricot and almond gums (lanes 9, 10 and 11), on the contrary of the SDS-PAGE gel any high molecular weight protein fraction appeared to be blocked in the stacking gel. Therefore it was assumed that SDS and DTT could facilitate the solubilization of these high MW and heavily glycosylated proteins. Tragacanth, guar and karaya gums showed a similar protein pattern, except for the absence of the band at 30 kDa that was clearly observed for karaya gum by SDS-PAGE. The three bands in the 35-30 kDa range observed for ghatti gum in denaturing conditions (Fig. 15, lane 6) were maybe the products of denaturation of some higher MW proteins since, in non-denaturing conditions, only one faint band at ~35 kDa was observed (Fig. 18, lane 2). In regard to gum arabic (lane 3), the gel reticulation appeared to be altered and the same could be said for the electrophoretic lane. Furthermore protein bands appeared to be less resolved. Finally, analysis of locust bean gum in non-denaturing conditions (lane 5) allowed to conclude that, most of the protein bands observed below 45 kDa in denaturing conditions (Fig. 15, lane 3), could correspond to some products of denaturation of higher mass proteins.



### 3.6. CONCLUSIONS OF SDS-PAGE ANALYSIS

In conclusion, SDS-PAGE with Sypro staining allowed to successfully establish the electrophoretic profile of the protein component of some plant gums in a molecular weight range from 20 to 220 kDa.

Among the several polyacrylamide gels tested, gels with a low reticulation (4.7%) showed for the majority of the samples, particularly gum arabic and fruit tree gums, the presence of proteins with a molecular weight by far higher than 220 kDa. These fractions appeared blocked in the stacking gel and could not migrate along the separation gel due to its increased reticulation. Addition of 2 million Da molecular weight dextran to the classical acrylamide/bisacrylamide gel and optimization of agarose-acrylamide composite gel, did not result to be enough resolutive and any protein band could be observed.

Focusing on the MW range below 220 kDa, an electrophoretic profile of the protein component of several plant gums could be obtained by increasing the percentage of acrylamide. 10% and 12% polyacrylamide gels resulted in the best choice. Locust bean gum was showed to be characterized by a different profile composed of several proteins with MWs spanning the whole range of masses. Intense and well resolved bands were obtained also for apricot gum that showed a SDS-PAGE protein pattern significantly different from the other fruit tree gums (cherry, almond and plum gums). The 10%, and mainly the 12%, polyacrylamide gels highlighted the presence of some bands in ghatti, tragacanth, guar and karaya gums. Guar was characterized by two bands around 45 kDa and other two at 20 kDa; tragacanth showed three large bands at respectively 66, 40 and 30 kDa; a large and intense protein band was observed for ghatti gum at around 40 kDa and, finally, karaya showed one large and intense band between 30 and 25 kDa. However their proteinaceous electrophoretic profile resulted, in general, fairly similar with the presence of few faint bands with molecular weights mainly between 80 and 20 kDa.

Therefore, SDS-PAGE resulted in a high resolutive technique that allowed to reveal how several plant gums are characterized by different electrophoretic profiles that reflect the differences related to the (glycol)protein component.

## REFERENCES

1. Motlagh S, Ravines P, Karamallah KA, Ma QF (2006) The analysis of Acacia gums using electrophoresis. *Food Hydrocolloids* 20: 848-854.
2. Porath J, P. F (1959) Gel filtration: a method for desalting and group separation. *Nature* 183: 1657- 1659.
3. Moore J (1964) Gel permeation chromatography. I. A new method for molecular weight distribution of high polymers. *J Polym Sci A* 2: 835.
4. Goetz H, Kuschel M, Wulff T, Sauber C, Miller C, et al. (2004) Comparison of selected analytical techniques for protein sizing, quantitation and molecular weight determination. *Journal of Biochemical and Biophysical Methods* 60: 281-293.
5. Le Maire M, Aggerbeck LP, Monteilhet C, Andersen JP, Møller JV (1986) The use of high-performance liquid chromatography for the determination of size and molecular weight of proteins: A caution and a list of membrane proteins suitable as standards. *Analytical Biochemistry* 154: 525-535.
6. Arakawa T, Ejima D, Li T, Philo JS (2010) The critical role of mobile phase composition in size exclusion chromatography of protein pharmaceuticals. *Journal of Pharmaceutical Sciences* 99: 1674-1692.
7. Kostanski LK, Keller DM, Hamielec AE (2004) Size-exclusion chromatography—a review of calibration methodologies. *Journal of Biochemical and Biophysical Methods* 58: 159-186.
8. Rögner M (1999) Chapter 2 Size Exclusion Chromatography. In: Michael K, editor. *Journal of Chromatography Library*: Elsevier. pp. 89-145.
9. Stellwagen E (2009) Chapter 23 Gel Filtration. In: Richard RB, Murray PD, editors. *Methods in Enzymology*: Academic Press. pp. 373-385.
10. Randall RC, Phillips GO, Williams PA (1989) Fractionation and characterization of gum from *Acacia senegal*. *Food Hydrocolloids* 3: 65-75.
11. Idris OHM, Williams PA, Phillips GO (1998) Characterisation of gum from *Acacia senegal* trees of different age and location using multidetection gel permeation chromatography. *Food Hydrocolloids* 12: 379-388.
12. Mohammadifar MA, Musavi SM, Kiumarsi A, Williams PA (2006) Solution properties of targacanthin (water-soluble part of gum tragacanth exudate from *Astragalus gossypinus*). *International Journal of Biological Macromolecules* 38: 31-39.
13. Shapiro AL, Viñuela E, V. Maizel Jr J (1967) Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochemical and Biophysical Research Communications* 28: 815-820.
14. Kavooosi G, Ardestani SK (2012) Gel Electrophoresis of Protein - From Basic Science to Practical Approach. *Gel Electrophoresis - Principles and Basics*: Dr. Sameh Magdeldin (Ed.).
15. Chiari M, Righetti PG (1995) New types of separation matrices for electrophoresis. *Electrophoresis* 16: 1815-1829.
16. Walker J (2002) Gradient SDS Polyacrylamide Gel Electrophoresis of Proteins. In: Walker J, editor. *The Protein Protocols Handbook*: Humana Press. pp. 69-72.
17. Righetti PG, Chillemi F (1978) Isoelectric focusing of peptides. *Journal of Chromatography A* 157: 243-251.
18. O'Farrell P (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250: 4007-4021.
19. Smejkal GB (2004) The Coomassie chronicles: past, present and future perspectives in polyacrylamide gel staining. *Expert Review of Proteomics* 1: 381-387.
20. Neuhoﬀ V, Arold N, Taube D, Ehrhardt W (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9: 255-262.
21. Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri GM, et al. (2004) Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* 25: 1327-1333.

22. Chevallet M, Luche S, Rabilloud T (2006) Silver staining of proteins in polyacrylamide gels. *Nat Protocols* 1: 1852-1858.
23. Berggren K, Chernokalskaya E, Steinberg TH, Kemper C, Lopez MF, et al. (2000) Background-free, high sensitivity staining of proteins in one- and two-dimensional sodium dodecyl sulfate-polyacrylamide gels using a luminescent ruthenium complex. *Electrophoresis* 21: 2509-2521.
24. Nishihara JC, Champion KM (2002) Quantitative evaluation of proteins in one- and two-dimensional polyacrylamide gels using a fluorescent stain. *Electrophoresis* 23: 2203-2215.
25. Berggren K, Steinberg TH, Lauber WM, Carroll JA, Lopez MF, et al. (1999) A Luminescent Ruthenium Complex for Ultrasensitive Detection of Proteins Immobilized on Membrane Supports. *Analytical Biochemistry* 276: 129-143.
26. Lopez MF, Berggren K, Chernokalskaya E, Lazarev A, Robinson M, et al. (2000) A comparison of silver stain and SYPRO Ruby Protein Gel Stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling. *Electrophoresis* 21: 3673-3683.
27. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, et al. (1985) Measurement of protein using bicinchoninic acid. *Analytical Biochemistry* 150: 76-85.
28. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London, U K)* 227: 680-685.
29. Perez-Vilar J, Hill RL (1999) The Structure and Assembly of Secreted Mucins. *Journal of Biological Chemistry* 274: 31751-31754.
30. Aksoy N, Corfield AP, Sheehan JK (2000) Preliminary study pointing out a significant alteration in the biochemical composition of MUC2 in colorectal mucinous carcinoma. *Clinical Biochemistry* 33: 167-173.
31. Aksoy N, Thornton DJ, Corfield A, Paraskeva C, Sheehan JK (1999) A study of the intracellular and secreted forms of the MUC2 mucin from the PC/AA intestinal cell line. *Glycobiology* 9: 739-746.
32. Aksoy N, Unlu S (2003) Increased resolution of macromolecules with agarose gel electrophoresis compared with polyacrylamide gel electrophoresis. *Macromol Biosci* 3: 482-486.
33. Thornton DJ, Howard M, Devine PL, Sheehan JK (1995) Methods for Separation and Deglycosylation of Mucin Subunits. *Analytical Biochemistry* 227: 162-167.
34. Thornton DJ, Carlstedt I, Howard M, Devine PL, Price MR, et al. (1996) Respiratory mucins: identification of core proteins and glycoforms. *Biochem J* 316: 967-975.
35. Stellwagen NC (2009) Electrophoresis of DNA in agarose gels, polyacrylamide gels and in free solution. *Electrophoresis* 30 Suppl 1: S188-195.
36. Barril P, Nates S (2012) Introduction to Agarose and Polyacrylamide Gel Electrophoresis Matrices with Respect to Their Detection Sensitivities. In: Sameh M, editor. *Gel Electrophoresis- Principles and Basics*. Croatia. pp. 3-14.
37. Uriel J (1966) Method of electrophoresis in acrylamide-agarose gels. *Bull Soc Chim Biol (Paris)* 48: 969-982.
38. Peacock AC, Dingman CW (1968) Molecular weight estimation and separation of ribonucleic acid by electrophoresis in agarose-acrylamide composite gels. *Biochemistry* 7: 668-674.
39. Tatsumi R, Hattori A (1995) Detection of Giant Myofibrillar Proteins Connectin and Nebulin by Electrophoresis in 2% Polyacrylamide Slab Gels Strengthened with Agarose. *Analytical Biochemistry* 224: 28-31.
40. Andersch-Bjoerkman Y, Thomsson KA, Larsson JMH, Ekerhovd E, Hansson GC (2007) Large scale identification of proteins, mucins, and their O-glycosylation in the endocervical mucus during the menstrual cycle. *Mol Cell Proteomics* 6: 708-716.
41. Issa SMA, Schulz BL, Packer NH, Karlsson NG (2011) Analysis of mucosal mucins separated by SDS-urea agarose polyacrylamide composite gel electrophoresis. *Electrophoresis* 32: 3554-3563.
42. Davis BJ (1964) Disc electrophoresis - method and application to human serum proteins. *Annals of the New York Academy of Sciences* 121: 404-427.
43. Ornstein L (1964) Disc electrophoresis - background and theory. *Annals of the New York Academy of Sciences* 121: 321-349.



# Chapter III

## **A new strategy for plant gums identification by stepwise enzymatic digestion and mass spectrometry**

### **Abstract**

In the cultural heritage field, characterization of gums is usually based on the monosaccharide composition, as determined by gas chromatography mass spectrometry (GCMS) after complete acid hydrolysis of the gum. However the gum chromatographic profile can be altered by the presence of other organic/inorganic compounds. Therefore, in this chapter, a novel method involving partial enzymatic digestion of plant gums, followed by analysis of the released oligosaccharides by MALDI-MS is proposed. Due to the different polysaccharide structure of the gums, the obtained MALDI mass spectrum represents the unique profile or “saccharide mass fingerprint” of the gum. The developed strategy allows the discrimination of gum arabic, tragacanth, cherry, guar and locust bean gum. Mass profile reproducibility was verified and a unique enzyme cocktail, suitable for the digestion of all of the most common plant gums, was developed. Analysis of a watercolor sample from 1870 revealed the mass profile of gum arabic, thus demonstrating the promising applicability of the method to aged materials from works of art.



## TABLE OF CONTENTS

1. AIM OF THE RESEARCH .....	101
2. ANALYTICAL TECHNIQUE AND ENZYMATIC APPROACH .....	102
2.1. Matrix assisted laser desorption ionization mass spectrometry .....	102
(MALDI-MS) .....	102
2.1. MALDI source .....	102
2.2. Time of flight detector.....	102
2.3. MALDI-MS analysis of polysaccharides .....	103
2.4. Tandem mass spectrometry .....	104
2.2. Stepwise enzymatic degradation of polysaccharides .....	106
2.2.1. Enzymes acting on Type II arabinogalactans.....	107
2.2.2. Enzymes acting on galactomannans.....	109
2.2.3. Enzymes acting on rhamnogalacturonan.....	109
3. METHODOLOGICAL DEVELOPMENT ON STANDARD POLYSACCHARIDES .....	111
3.1. Experimental procedure optimization .....	111
3.1.1 Standard polysaccharides description and preparation.....	111
3.1.2. Enzymes and enzymatic hydrolysis general procedure.....	111
3.1.3. MALDI-TOF matrix .....	114
3.1.4. Oligosaccharides derivatization by 3-aminoquinoline .....	116
3.1.5. Oligosaccharides purification.....	119
3.2. Optimization of larch arabinogalactan enzymatic digestion .....	120
3.2.1. Evaluation of enzyme amount.....	120
3.2.2. Evaluation of enzymes combination and digestion time.....	122
3.2.3. MALDI-TOF profile of the polysaccharide .....	125
3.3. Optimization of carob galactomannan enzymatic digestion .....	128
4. APPLICATION AND METHODOLOGICAL DEVELOPMENT ON PLANT GUMS SAMPLES.....	131
4.1. Plant gum samples preparation .....	131
4.2. Gum arabic .....	132
4.2.1. Evaluation of enzymatic digestion protocol.....	132
4.2.2. Identification of the released oligosaccharides .....	137
4.2.3. Analysis of different gum arabic samples from Acacia Senegal var. Senegal .....	145
4.3. Tragacanth gum.....	147
4.3.1. Evaluation of enzymatic digestion protocol and oligosaccharides identification .....	148
4.3.2. Analysis of tragacanth samples from different brands .....	151
4.4 Cherry gum.....	153
3.4.1. Evaluation of enzymatic digestion protocol.....	153
4.4.2. Analysis of different cherry gum samples.....	155
4.5. Locust bean and guar gum .....	157
4.5.1. Digestion and identification of oligosaccharides released from LBG.....	157

4.5.3. Digestion and identification of oligosaccharides released from guar gum.....	161
5. DEVELOPMENT OF AN ENZYME COCKTAIL.....	164
5.1. Cocktail formulation and digestion procedure .....	164
5.2. Results and discussions .....	164
5.2.1. Gum arabic .....	165
5.2.2. Tragacanth gum.....	165
5.2.3. Cherry gum.....	166
5.2.4. Locust bean gum .....	167
5.2.5. Guar gum.....	167
6. APPLICATION ON REAL SAMPLES.....	169
6.1. A 50 years old gum arabic sample .....	169
6.1.1. Composition and sample preparation .....	169
6.1.2. Results and discussion.....	169
6.2. Fresh watercolor samples .....	171
6.2.1. Composition and sample preparation .....	171
6.2.2. Results and discussion.....	171
6.3. Old watercolor sample.....	173
6.3.1. Composition and sample preparation .....	173
6.3.2. Results and discussion.....	173
7. CONCLUSIONS of MALDI-TOF ANALYSIS .....	176
REFERENCES.....	177



## List of figures

Fig. 1. Schematic representation of a reflectron TOF mass analyzer.....	103
Fig. 2. Scheme of sugar derivatization by reductive amination .....	104
Fig. 3. Scheme of the two fragmentation pathways and corresponding nomenclature of carbohydrates MS/MS analysis.....	105
Fig. 4. Schematic representation of a complex type II AGs with the site of attack of various enzymes.....	108
Fig. 5. Supposed enzymatic attack on galactomannans. ....	109
Fig. 6. Mass spectrum in linear mode (10-60kDa) of larch AG digested and spotted with DHB in MeOH/H <sub>2</sub> O 50/50 0.1% TFA. Mono and double charged ions of exo- $\beta$ -1,3-GAL enzyme. ....	115
Fig. 7. Mass spectra comparison of the matrix (above) and the larch arabinogalactan digested with $\alpha$ -L-arabinofuranosidase followed by exo- $\beta$ -1,3-galactanase for 24 hours (below).....	116
Fig. 8. Scheme of oligosaccharide derivatization by reductive amination.....	117
Fig. 9. Mass spectrum in positive reflector mode of .....	117
Fig. 10. MALDI-TOF mass spectrum of derivatized oligosaccharides of larch arabinogalactan after enzymatic digestion (zoom 600-100Da). ....	118
Fig. 11. Mass spectra (and corresponding enlarged image of the mass range 600-1000Da) of digested larch AG after fractionation by hypercarb: (A) loading solution; (B) washing solution and (C) elution with ACN 10%.....	120
Fig. 12. MALDI-TOF enlarged spectra (1000-1500 Da) of larch arabinogalactan digested with 25, 100 and 1000 U of each enzyme. ....	121
Fig. 13. MALDI-TOF enlarged spectrum (600-1400Da) of larch arabinogalactan after digestion with 4 enzymes ( $\alpha$ -L-arabinofuranosidase + $\beta$ -glucuronidase + $\alpha$ -L-rhamnosidase + $\beta$ -galactosidase). (*) described in the following section.....	123
Fig. 14. Mass spectra comparison of larch arabinogalactan digested with five enzymes for 3, 5 and 93 hours.....	124
Fig. 15. MALDI-TOF spectra of larch arabinogalactan digested with exo- $\beta$ -1,3-galactanase for (1) 5 minutes; (2) 1 hour and (3) 120 hours. The enlargement of the mass range 1000 - 1400Da is showed respectively in the mass spectra A, B and C. ....	125
Fig. 16. MALDI-CID spectrum of 3-AQ derivatized oligosaccharide released by exo- $\beta$ -1,3-galactanase (m/z 1027.56).....	126
Fig. 17. MALDI-TOF spectra of carob galactomannan digested with $\alpha$ -galactosidase (24 hours) and endo-1,4- $\beta$ -mannanase for: (A) 5minute; (B) 1h; (C) 3h; (D) 8h; (E) 24 hours and (F) 6 days. ....	128
Fig. 18. MALDI-TOF spectrum of carob galactomannan digested with endo-1,4- $\beta$ -mannanase for 8h.....	129
Fig. 19. Illustration of various enzymes possibly involved in the degradation of gum arabic polysaccharide component. <i>Galp</i> galactopyranose, <i>Araf</i> arabinofuranose, <i>Arap</i> .....	132
Fig. 20. Mass spectra (range 600-1800 Da) of gum arabic sample: (A) non digested (diluted 1/10); (B) digested with $\alpha$ -L-arabinofuranosidase; (C) after adding $\beta$ -glucuronidase; (D) after adding $\alpha$ -L-rhamnosidase; (E) after adding $\beta$ -galactosidase (93h); (F) after adding exo- $\beta$ -1,3-galactanase; (F corner) zoom 1300-1900 Da; (G) enlarged image (1050-1300Da) of Fig. 20F. ....	134
Fig. 21. MALDI-TOF spectra (600-1300 Da) of gum arabic digested with exo- $\beta$ -1,3-galactanase for (A) 5 minutes; (B) 5 h; (C) 96 h and (D) 2 weeks. On the right a zoom of the mass range 1300-1700 Da is showed.....	136
Fig. 22. MS spectrum (range 600-2300 Da) of gum arabic sample digested for 5 hours with exo- $\beta$ -1,3-galactanase. Significant mass differences between peaks are reported. ....	138
Fig. 23. MS spectra of the gum arabic digested with: (A) exo- $\beta$ -1,3-galactanase and (B) exo- $\beta$ -1,3-galactanase and $\alpha$ -L-arabinofuranosidase. On the right corner the corresponding zoom of mass range 1300-1500 Da. ....	139
Fig. 24. MALDI-CID spectrum of 3-AQ derivatized oligosaccharide released by exo- $\beta$ -1,3-galactanase (m/z 1055.46).....	140

Fig. 25. MALDI-CID spectrum of 3-AQ derivatized oligosaccharide released by $\text{exo-}\beta\text{-1,3-galactanase}$ (m/z 1187.43).....	141
Fig. 26. MALDI-TOF mass spectrum of gum arabic (Zecchi - Italy) digested with $\text{exo-}\beta\text{-1,3-galactanase}$ for 5h with the respective assigned oligosaccharides (all derivatized with 3-aminoquinoline).....	144
Fig. 27. <i>Acacia senegal</i> var. <i>senegal</i> gum nodules from: (1) Kordufan (1994), (2) Damazene (1999) and (3) Damazene (1994). .....	145
Fig. 28. Comparison of MALDI-TOF spectra of standard gum arabic (Zecchi) with gum arabic nodules ( <i>Acacia senegal</i> var. <i>senegal</i> ) digested for 5h with $\text{exo-}\beta\text{-1,3-galactanase}$ . .....	146
Fig. 29. Illustration of various enzymes possibly involved in the degradation of tragacanth gum arabinogalactan component. ....	148
Fig. 30. Mass spectra (range 800-1200 Da) of tragacanth gum sample: (A) non digested (diluted 1/100); (B) digested with $\alpha\text{-L-arabinofuranosidase}$ ; (C) digested with $\alpha\text{-L-arabinofuranosidase}$ + $\text{endo-}\beta\text{-1,4-galactanase}$ + $\text{exo-}\beta\text{-1,3-galactanase}$ ; (D) zoom in signal/noise of Fig. 30B (frame). .....	149
Fig. 31. MALDI-TOF mass spectrum of tragacanth gum (Bresciani - Italy) digested with $\text{exo-}\beta\text{-1,3-galactanase}$ for 5 h with the assigned oligosaccharides (all derivatized with 3-aminoquinoline).....	150
Fig. 32. MALDI-TOF spectra of tragacanth gum digested for 5 h with $\text{exo-}\beta\text{-1,3-galactanase}$ : (A) Bresciani; (B) Zecchi; (C) Okhra and (D) Sigma. ....	152
Fig. 33. Enlargement of the MS spectrum of digested tragacanth gum from Okhra (1200-2300 Da). .....	152
Fig. 34. Illustration of various enzymes possibly involved in the degradation of cherry gum. ....	153
Fig. 35. MALDI-MS spectrum of (A) cherry gum (n.2) digested with $\text{exo-}\beta\text{-1,3-galactanase}$ for 5h and (B) sample not digested. ....	155
Fig. 36. MALDI-TOF mass spectra of cherry gum samples digested with $\text{exo-}\beta\text{-1,3-GAL}$ for 5h: (A) cherry n.2; (B) Kremer; (C) cherry n.1 and (D) cherry n.3. Marked region indicates the ions of interest. ....	156
Fig. 37. MALDI-TOF spectra of LBG (A) not digested and digested with $\text{endo-1,4-}\beta\text{-mannanase}$ for (B) 5minute; (C) 1 hour; (D) 5h; (E) 8h; (F) 24 hours and (G) 48h. On the right corner of each spectrum the enlarged mass range 2000-3000Da is shown.....	158
Fig. 38. MALDI-T MALDI-TOF spectrum of LBG digested with $\text{endo-1,4-}\beta\text{-mannanase}$ for 8 hours. On the right corner: zoom of the range 900-1100 Da with the assigned oligosaccharides. ....	159
Fig. 39. MALDI-TOF spectra of guar gum (A) not digested and digested with $\text{endo-1,4-}\beta\text{-mannanase}$ for (B) 5 minutes; (C) 5h. On the right corner the corresponding enlarged mass range 1200-3000Da is shown. ....	161
Fig. 40. MALDI-TOF spectrum of guar gum digested with $\text{endo-1,4-}\beta\text{-mannanase}$ for 5 hours (mass range 600-1450 Da) .....	162
Fig. 41. Comparison of MALDI-TOF spectra of gum arabic (Zecchi) digested for 24 hours with: (A) $\text{exo-}\beta\text{-1,3-galactanase}$ and (B) enzyme cocktail. ....	165
Fig. 42. Comparison of MALDI-TOF spectra of tragacanth gum (Bresciani) digested for 24 hours with: (A) $\text{exo-}\beta\text{-1,3-galactanase}$ and (B) enzyme cocktail (with enlarged image of mass range 1200-3000 Da).....	165
Fig. 43. Comparison of MALDI-TOF spectra of cherry gum n.2 digested for 24 hours with: (A) $\text{exo-}\beta\text{-1,3-galactanase}$ and (B) enzyme cocktail (with enlarged image of mass range 2200-3000Da).....	166
Fig. 44. Comparison of MALDI-TOF spectra of LBG (Sigma) digested for 24 hours with: (A) $\text{endo-}\beta\text{-1,4-mannanase}$ and (B) enzyme cocktail. ....	167
Fig. 45. Comparison of MALDI-TOF spectra of guar gum (Sigma) digested for 24 hours with: (A) $\text{endo-}\beta\text{-1,4-mannanase}$ and (B) enzyme cocktail. ....	168
Fig. 46. Mass spectra comparison of gum arabic (Zecchi) and a 50 years old gum arabic sample (Metropolitan Museum of Art, NY) digested with $\text{exo-}\beta\text{-1,3-galactanase}$ for 24 hours. ....	170
Fig. 47. Mass spectra comparison of gum arabic (Zecchi), watercolor n. 278 (Maimeri) and watercolor n. 588 (Daler-Rowney) digested with $\text{exo-}\beta\text{-1,3-galactanase}$ for 24 hours.....	172

Fig. 48. Mass spectra comparison of gum arabic (Zecchi) and watercolor n. 588 (Daler-Rowney) digested with enzyme cocktail (exo- $\beta$ -1,3-galactanase and endo-1,4- $\beta$ -mannanase) for 24 hours.....	173
Fig. 49. Mass spectra comparison of gum arabic (Zecchi) and watercolor sample from 1870, digested with exo- $\beta$ -1,3-galactanase for 24 hours.....	174

## List of tables

Tab. 1. Monosaccharides, abbreviation and respectively monoisotopic mass ( <a href="http://web.expasy.org/glycomod/glycomod_masses.html">http://web.expasy.org/glycomod/glycomod_masses.html</a> ).....	106
Tab. 2. Enzymes classification.....	106
Tab. 3 Summary of exoglycosidases and endoglycosidases involved in type II arabinogalactans degradation. Information obtained by BRENDA website.....	108
Tab. 4. Information about purchased enzymes. “Working buffer” is the solution at the specific pH in which each enzyme was diluted and “T °C” is the temperature kept during digestion. The “Unit definition” related to enzyme activity are described in the technical resources for Megazyme ( <a href="http://www.megazyme.com/">http://www.megazyme.com/</a> ) and Nzytech ( <a href="https://www.nzytech.com/site/">https://www.nzytech.com/site/</a> )...	113
Tab. 5. List of the assigned oligosaccharides for digested larch arabinogalactan. ( <sup>a</sup> ) the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline, H <sup>+</sup> , Na <sup>+</sup> or K <sup>+</sup> ).....	127
Tab. 6. Summary of amount and solubilization protocol for each plant gums. ....	131
Tab. 7. List of the assigned oligosaccharides for digested gum arabic (Zecchi, Italy). ( <sup>a</sup> ) the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline, H <sup>+</sup> , Na <sup>+</sup> or K <sup>+</sup> ).....	142
Tab. 8. List of the analyzed gum arabic samples from <i>Acacia Senegal</i> var. <i>Senegal</i> , area and date of collection. ....	145
Tab. 9. Comparison of theoretical masses (gum arabic characteristic oligosaccharides, Tab.7) with the experimental masses obtained after digestion of three gum arabic samples <i>Acacia Senegal</i> var <i>Senegal</i> . The oligosaccharides which are not present in the mass spectra of the analyzed samples are indicated by ‘/’.....	146
Tab. 10. List of the assigned oligosaccharides for digested gum tragacanth. ( <sup>a</sup> ) the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline, H <sup>+</sup> , Na <sup>+</sup> or K <sup>+</sup> ). 151	
Tab. 11. List of the assigned oligosaccharides for digested LBG. ( <sup>a</sup> ) the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline, H <sup>+</sup> , Na <sup>+</sup> or K <sup>+</sup> ). ....	160
Tab. 12. List of the assigned oligosaccharides for digested guar gum. ( <sup>a</sup> ) the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline, H <sup>+</sup> , Na <sup>+</sup> or K <sup>+</sup> ). ....	162
Tab. 13. Description of the enzyme cocktail composition and corresponding operative conditions. ....	164
Tab. 14. List of the gum arabic oligosaccharides recognized in the digested 50 years old gum arabic sample. The oligosaccharides which are not present in the mass spectrum are indicated by ‘/’ . ....	170
Tab. 15. Purchased watercolors, reference number and composition as reported by the company. ....	171
Tab. 16. List of the gum arabic oligosaccharides recognized in the digested old watercolor sample. The oligosaccharides which are not present in the mass spectrum are indicated by ‘/’ . ....	174



## 1. AIM OF THE RESEARCH

Identification of polysaccharide materials, such as plant gums, is challenging and especially difficult in art samples where polysaccharides are found in mixture with complex organic (e.g. proteins) and/or inorganic (e.g. pigments) materials. Due to the limitations encountered in gum identification when using the classical analytical techniques (Chapter I, section 2.3), the aim of this chapter is to develop and optimize a new strategy based on mass spectrometry, and in particular on Matrix Assisted Laser Desorption Ionization Time of Flight mass spectrometry (MALDI-TOF), for the discrimination of different plant gums in samples from works of art and cultural artefacts.

MALDI-MS is routinely used for proteomic studies in the biomedical field and it has already been successfully applied to the study of proteins in samples from works of art. The method, called peptide mass fingerprint (PMF), relies on the fact that proteolytic enzymes cleave proteins at specific amino acid sites producing unique patterns of peptides characteristic of the protein. A similar analytical strategy for plant gum analysis, based on stepwise enzymatic degradation of the polysaccharide, followed by MS analysis, was adopted. This method is applicable to differentiate plant gums since they show specific polysaccharide structures in terms of glycosidic linkages, monosaccharide sequence, and presence of more or less branched side chains. As described in Chapter I section 3.1, gum arabic showed similarities to Type II arabinogalactans. This structure was also revealed in some fractions of gum tragacanth and gums from fruit trees. LBG and guar resembled the structure of galactomannans, while gum ghatti and karaya had a complex composition which showed some similarities with respectively the glucuronomannan and rhamnogalacturonan polysaccharide families. Enzymes that are able to cleave the polysaccharide adjacent to specific monosaccharides or at certain types of linkage, releasing unique and characteristic oligosaccharides, have been selected. The composition of the released oligosaccharides depends on the gum polysaccharide structure. Therefore, MALDI-MS analysis enabled to obtain a characteristic mass spectrum of the gum.

In the first part (section 2), a brief overview of the employed instrumentation (MALDI-TOF-TOF) and of the mechanism of action of enzymes on polysaccharide structure is presented. Regarding the methodological development, in section 3 MALDI-MS parameters were optimized by means of standard polysaccharides which resemble the structure of the two most significant plant gums polysaccharide families: type II arabinogalactans and galactomannans. After that, the most adapt enzymatic digestion protocol was developed and optimized for each pure plant gum (section 4). Oligosaccharide profiles were acquired for gum arabic, tragacanth, locust bean, guar and cherry gum. However, with the prospect to analyze real art/archaeological samples, whose composition is usually unknown, a second goal of this chapter was to develop an unique enzyme cocktail that could be suitable for the digestion of all of the most common plant gums, regardless of their polysaccharide structure (section 5). Finally, the applicability of the novel method on real samples was verified on more complex and aged materials such as fresh and old watercolors (section 6).

## 2. ANALYTICAL TECHNIQUE AND ENZYMATIC APPROACH

### 2.1. MATRIX ASSISTED LASER DESORPTION IONIZATION MASS SPECTROMETRY (MALDI-MS)

#### 2.1. MALDI source

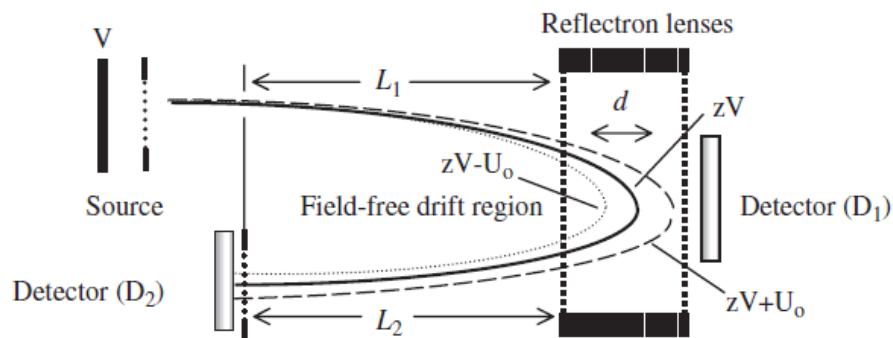
Matrix Assisted Laser Desorption Ionization (MALDI) is a soft ionization source which was introduced in the mid 80s for the ionization of biological molecules [1,2]. In desorption ionization the sample is analyzed by the impact of energetic particles, and in the case of MALDI it corresponds to a laser beam, which both desorbs and ionizes the analyte. A solid matrix is solubilized in a proper aqueous solution, usually containing a certain percentage of organic solvent, and mixed with the sample in a ratio of about 1/1000 sample/matrix. The mixture is then spot on the MALDI plate and let dry in order to get sample-matrix crystals. Several ways of mixing/depositing the sample-matrix mixture exist (e.g. dried droplet, sandwich method etc.) and more important different matrices are employed according to the type of sample analyzed. The most frequently used matrices are: sinapinic acid (SA) for proteins, 2,5-dihydroxybenzoic acid (DHB) for oligosaccharides and peptides, and  $\alpha$ -cyano-hydroxycinnamic acid (HCCA) for peptides. Once the plate is loaded, the co-crystallized sample-matrix is irradiated with a laser beam (commonly azote with a wavelength of  $\lambda = 337$  nm) which simultaneously desorbs and ionizes the sample and matrix molecules into the gas phase. Since a high energy input is used, the matrix absorbs large amount of energy at the wavelength of the laser radiation, it allows molecules desorption in gas-phase and ionization of the analytes by proton transfer, which is also facilitated by the presence of an acid in the matrix solution.

#### 2.2. Time of flight detector

The time of flight detector (TOF) consists in a field-free flight tube where ions are separated according to their velocity. Because of the relationship among velocity ( $v$ ), mass ( $m$ ) and charge ( $z$ ), the time of arrival of each ion depends on these parameters and thus a mass spectrum is obtained. Since TOF detectors need the ions to enter the flight tube at the same time, the pulsed laser beam used with MALDI was found to be the best ion source combination. Plus TOF guarantees the mass range requested. Since the ions obtained in source are accelerated with inherent dispersion in time, space and velocity, the TOF mass resolution was pretty limited until the introduction of *Delayed Extraction* (DE). Thanks to the DE, ions are delayed for few hundred nanoseconds before the potential is applied and they are accelerated into the flight tube. Therefore a discrete amount of ions of a particular mass but of different kinetic energy will reach the detector at the same time, thus increasing the resolution.

There are two main modes of analysis: linear and reflectron. Linear mode is mainly used for detection of molecules with medium/high molecular weight and resolution is low because of the time, spatial and kinetic energy dispersion of the ions. Therefore reflectron TOF instruments were introduced. It consists of grids and a series of ring electrodes with an increasing repelling potential (Fig. 1, from [3]). In this way all ions with the same  $m/z$  value arrive at the detector

simultaneously. This mode is usually employed for analysis of low molecular weight molecules (< 4000 Da).



**Fig. 1.** Schematic representation of a reflectron TOF mass analyzer.

### 2.3. MALDI-MS analysis of polysaccharides

Analysis of carbohydrates by MALDI was reported at the end of the 80s by Karas and Hillenkamp [4], followed by experiments on N-linked glycans by Mock *et al.* [5]. Since then, MALDI has become one of the most used techniques for glycan analysis and it was extensively reviewed by Harvey D.J. [6]. However, oligo- and polysaccharides ionization in mass spectrometry is a hard task due to their low surface activity, higher polarity, low stability and high variation in molecular mass (polydispersity [7]). In addition, while for proteins and peptides the nitrogen atom promotes the formation of protonated ions ( $[M+H]^+$ ), under MALDI conditions oligo- and polysaccharides commonly form sodium adduct ions ( $[M+Na]^+$ ) when analyzed in positive mode. Carbohydrates are neutral and usually there are no basic groups to enable protonation. However, the corresponding protonated ion and potassium adduct ( $[M+K]^+$ ) are also detected as minor ions in the mass spectrum, as well as possible adducts with the matrix ( $[M+matrix+Na/K]^+$ ) [8].

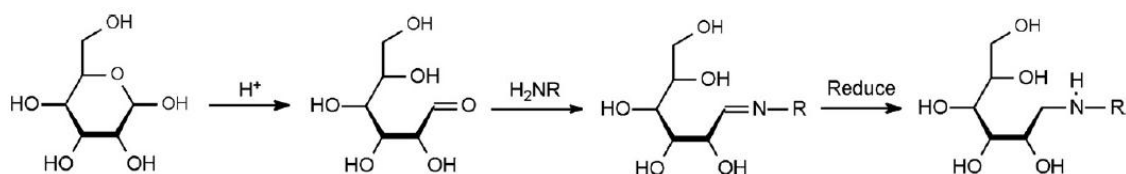
There are some important aspects to consider in order to obtain a good signal from carbohydrates: the nature of the matrix and the sample preparation procedure. These two steps are of critical importance since matrices that are effective for some compounds (e.g. HCCA for peptides), might be completely ineffective for others. Therefore, several different matrices have been tested for oligo- and polysaccharide analysis according to the nature of the carbohydrates (e.g. neutral, acidic and sulfated) [9]. However, the most used matrix for carbohydrate analysis is still nowadays 2,5-dihydroxybenzoic acid (2,5-DHB) [8], which have been in the years mixed with other compounds mainly to increase sensitivity or resolution (e.g. 2-hydroxy-5-methoxybenzoic acid [10]).

In addition, another common method employed to increase oligo- and polysaccharide ionization, consists in chemical modification, or derivatization, of the carbohydrate [11-13]. Several techniques are commonly used and, among the most common, two groups of techniques can be identified according to the carbohydrate group that is derivatized: (1) derivatization of the hydroxyl groups and (2) derivatization of the reducing terminal.

The most common derivatization method belonging to the first group is permethylation. Permethylation is used in mass spectrometry to improve the detection limit and to obtain information on the type of glycosidic linkages by fragmentation. The procedure can be difficult and purification steps before analysis are usually required. Therefore several procedural modifications

have been proposed in the literature during the past years [14]. However, compared to other methods such as alkylation and acylation, the mass increase due to derivatization (14 Da) is relatively small and this is an important aspect since the high number of hydroxyl groups in the carbohydrate structure.

Concerning the second group of derivatization methods, several procedure have been proposed such as the attachment of chromophores or fluorophores but, for MALDI-MS, the most popular method consists in the modification of the carbohydrate by reductive amination. Carbohydrates in their open form have a carbonyl group at their reducing terminus that can react with amines to form a Schiff base. These compounds are not stable but they can be reduced to the secondary amine by reductive amination (Fig. 2 from [12]).



**Fig. 2.** Scheme of sugar derivatization by reductive amination

Many amines have been used (e.g. 2-aminobenzimide, 2-aminopyridine) [12] and, among these, one was selected in this thesis work. Matrix preparation and effectiveness on the detection of the carbohydrates analyzed in this research will be described and explained in Chapter III, section 3.1.3.

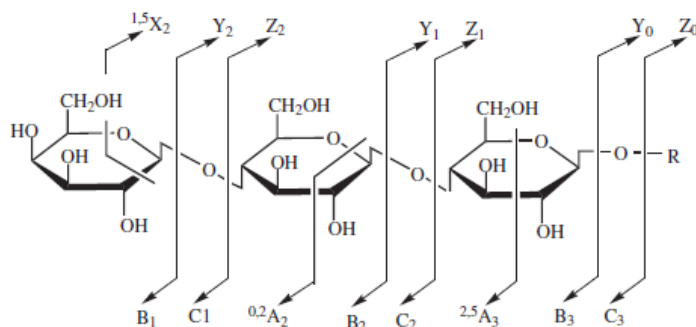
MALDI-MS is a valuable analytical technique to highlight the oligo- and polysaccharides nature. One important goal of MALDI-MS is the measure of the average molecular weight of polymers. However, problems are encountered for high polydisperse polysaccharides so MS results are usually compared to results obtained by other techniques such as size exclusion chromatography [7]. Regarding the last point, the most useful tool to gain insight into the carbohydrate sequence consists in tandem mass spectrometry.

#### 2.4. Tandem mass spectrometry

Fragmentation experiments are of significant importance since they allow to reconstruct the exact sequence of macromolecules such as peptides, according to their amino acids sequence, and oligosaccharides, according to the obtained monosaccharides. The analyte is selected, by its  $m/z$  value, and fragmented in the corresponding monomers. Several options are available to obtain a fragmentation spectrum by MALDI, and they mainly depend on the type of analyte and instrument. Usually fragmentation is performed in a collision cell, and this is the case of the instrument used in this thesis work (MALDI-TOF-TOF). This field-free region enables the fragmentation of the precursor ion by collision-induced dissociation (CID). Fragmentation occurs due to the collision of the target ion with an inert gas (e.g. Argon) or simply with air. Sometimes fragmentation can occur in the MALDI source by the action of the laser, for example by a Nd-YAG laser (laser-induced dissociation, LID). These ions are usually referred to in-source decay (ISD) ions, and they are distinguished from the ions formed after extraction for the source (post-source decay ions, PSD).



Fragmentation of carbohydrates, under CID conditions, occurs by two different pathways: glycosidic cleavage and cross-ring cleavage. The glycosidic cleavage involves the breaking of the bonds between the monosaccharide rings, the cross-ring cleavage consists on the break of the bonds comprising the sugar ring. While the first type of fragmentation leads to information about the carbohydrate sequence, the second pathways provides information on the type of glycosidic linkages inside the oligosaccharide. The MS/MS fragmentation of carbohydrates, with the corresponding nomenclature of the fragment ions introduced by Domon and Costello [15], is reported in Fig. 3.



**Fig. 3.** Scheme of the two fragmentation pathways and corresponding nomenclature of carbohydrates MS/MS analysis.

According to the reported scheme, the ion that retains the charge on the non-reducing end are labeled A, B and C, while the fragments retaining the charge on the reducing end are referred with the letters X, Y and Z. Among these, A and X fragments derive from cross-ring cleavage and they are labeled with superscript numbers related to the bonds cleaved. B, C, Y and Z ions result instead from glycosidic cleavage. All the fragments are in addition labeled with a subscript number related to the sugar residue retained. The type of fragments obtained after MS/MS analysis depends, besides from the oligosaccharide structure, by the state of the precursor ion. It is reported that, sodiated ions lead more easily to cross-ring cleavages, while the protonated ones provide more information about the glycosidic linkages [16].

If the MS/MS spectra is mainly characterized by B, C, Y and Z fragments, the oligosaccharide sequence can be easily reconstructed by calculating the mass difference between adjacent ions. This mass difference indicates the type of monosaccharide and, specifically, it refers to the monosaccharide residual mass that is obtained by subtracting the mass of a molecule of water from the intact monoisotopic mass. The list of the general monosaccharides and their specific mass residue is reported in Tab. 1.

**Tab. 1.** Monosaccharides, abbreviation and respectively monoisotopic mass ([http://web.expasy.org/glycomod/glycomod\\_masses.html](http://web.expasy.org/glycomod/glycomod_masses.html))

Monosaccharide (abbreviation)	Example (abbreviation)	Formula	Monoisotopic mass [Da]	
			Intact	Residue
Pentose (Pent)	Arabinose (Ara)	C <sub>5</sub> H <sub>10</sub> O <sub>5</sub>	150.0529	132.0423
Hexose (Hex)	Galactose (Gal)	C <sub>6</sub> H <sub>12</sub> O <sub>6</sub>	180.0634	162.0528
	Mannose (Man)			
	Glucose (Glu)			
Deoxyhexose (dHex)	Rhamnose (Rha)	C <sub>6</sub> H <sub>12</sub> O <sub>5</sub>	164.0685	146.0579
	Fucose (Fuc)			
Hexuronic acid (HexA)	Glucuronic acid (GlcA)	C <sub>6</sub> H <sub>10</sub> O <sub>7</sub>	194.0427	176.0321
N-Acetylamine hexose (HexNAc)	N-acetylgalactosamine (GalNAc)	C <sub>8</sub> H <sub>15</sub> NO <sub>6</sub>	221.0899	203.0794
N-Acetylneuraminic acid (NeuAc)	Sialic acid (SA)	C <sub>11</sub> H <sub>19</sub> NO <sub>9</sub>	309.2699	291.0954
	N-Glycolyl neuraminic acid (NeuGc)	C <sub>11</sub> H <sub>19</sub> NO <sub>10</sub>	325.2693	307.0903

## 2.2. STEPWISE ENZYMATIC DEGRADATION OF POLYSACCHARIDES

Enzymes are macromolecules produced by living organisms which act as catalysts in several metabolomic processes. Enzymes are mainly classified in six groups, according to the type of reaction they catalyze [17]:

**Tab. 2.** Enzymes classification

Group	EC code	Catalyzed reaction
Oxidoreductases	EC 1	Transfer of O, H atoms or electrons from one substrate to another
Transferases	EC 2	Transfer of a functional group
Hydrolases	EC 3	Hydrolysis of a substrate resulting in the formation of two products
Lyases	EC 4	Non hydrolytic removal/addition of a group
Isomerases	EC 5	Change of the molecular form
Ligases	EC 6	Synthesis of new bonds resulting in the joint of two molecules

Each group is characterized by a specific EC code which consists of a nomenclature system established by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB).

Among the different groups, in this chapter attention is focused on EC3 group and, specifically, on the glycoside hydrolases [18]. These enzymes are able to hydrolyze the glycosidic linkages of oligo- and polysaccharides. Their action on the carbohydrate can occur following two different pathways: exo- or endo-way. While exo-glycosidases hydrolyze terminal glycosidic bonds and release monosaccharides from the non-reducing end, the endo-glycosidases hydrolyze internal glycosidic linkages at specific or random position, releasing intact oligosaccharides [19,20].

Furthermore, glycoside hydrolases are characterized by an highly specificity in terms of glycosidic linkages, type of monosaccharide and anomeric configuration ( $\alpha$  and  $\beta$ ) [21].

Due to their ability to decomplexify polymeric structures, glycoside hydrolases are usually involved in sequential degradation of polysaccharides, followed by MS detection. This strategy is commonly employed for compositional and sequence analysis of the glycoproteins carbohydrate component [22,23]. Usually the method is based on the analysis by MALDI-MS of the carbohydrate before (undigested carbohydrate) and after each step digestion with well-defined exo-enzymes. After each incubation, the lost of the sugar residue is revealed in the mass spectrum by the mass difference corresponding to the specific number of monosaccharides released by each enzyme. Therefore, the monosaccharide sequence of the glycan chain can be reconstructed [24].

Sequential enzymatic digestion is an already established procedure for the investigation of fairly simple polysaccharides such as cellulose, hemicelluloses and pectins [25], and recently it has also been considered a valuable tool for degrading highly branched polysaccharides (e.g. arabinogalactans [26]). Therefore, this strategy was considered a potential method to highlight the plant polysaccharide structure.

As observed in the previous sections, naturally occurring polysaccharides have often and heterogeneous structure (hetero-polymers), characterized by several monosaccharides and variations in linkage patterns [27]. Therefore structural investigation, which is considered of significant importance to better understand the polysaccharide functions, is difficult to be achieved [28]. Due to the large size of most of natural polysaccharides (up to several million Da), enzymes that are capable of degrading the polymer into smaller oligosaccharides are a valuable tool for determining their structure. To achieve this objective, partial degradation by acid hydrolysis have also been employed. However, only with enzymes the polymer degradation occurs in a controlled manner and different information about the carbohydrate structure can be obtained.

As described above, plant gums represent a complex and branched substrate. Therefore, both exo- and endo-enzymes have to be used to decomplexify the polysaccharide structure. The focus of this section is to evaluate the possible enzymes that can be involved in the degradation of plant gums polysaccharide structure. Enzymes acting on arabinogalactans, galactomannans and rhamnogalacturonans will be described. General information on the enzyme action can be obtained from the website BRENDA (BRaunschweig ENzyme Database), an electronic resource that comprises information on enzymes which have been classified by the International Union of Biochemistry and Molecular Biology (IUBMB) (<http://www.brenda-enzymes.org/index.php4>). Several information on enzymes are reported in this database such as the kinetic constant for specific substrates, pH and temperature stability and other parameters.

### 2.2.1. Enzymes acting on Type II arabinogalactans

As described in depth in the review by Sakamoto et Ishimura [26], the enzymes which are involved in type II AGs degradation are: endo- $\beta$ -1,3-galactanases (GALs) (EC 3.2.1.181), exo- $\beta$ -1,3-GAL (EC 3.2.1.145), endo- $\beta$ -1,6-GAL (EC 3.2.1.164),  $\beta$ -galactosidases (EC 3.2.1.23),  $\alpha$ -L-arabinofuranosidases (AFs) (EC 3.2.1.55),  $\beta$ -L-arabinopyranosidases (APs) (EC 3.2.1.88),  $\beta$ -glucuronidase (EC 3.2.1.31) and  $\alpha$ -L-rhamnosidase (EC 3.2.1.40) (Fig. 4, from Sakamoto et al.

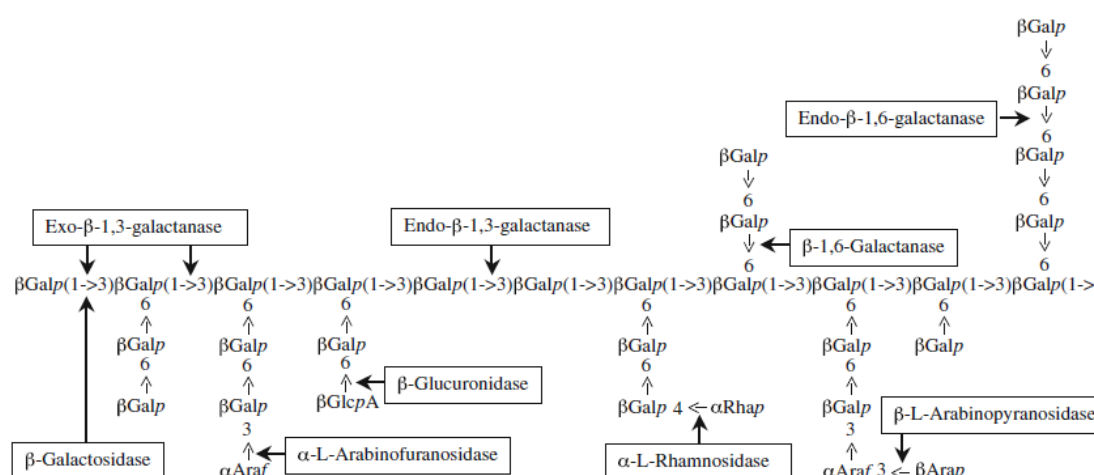
[26]). Most of the research is usually focused on the discovery, study and classification of new enzymes from various substrates. Their action on AGs might be fairly different according to the organism of origin, and the substrate. For this reason a table is introduced below in order to summarize the general activity of enzymes involved in the degradation of type II arabinogalactans and which were obtained, when possible, from a similar source of the purchased ones (Tab. 3).

**Tab. 3.** Summary of exoglycosidases and endoglycosidases involved in type II arabinogalactans degradation. Information obtained by BRENDA website.

Enzyme	Source	Linkage hydrolysis
Endo- $\beta$ -1,3-GAL <sup>a</sup>	Fungi	$\beta$ -1,3-galactan and $\beta$ -1,3-galactooligosaccharides
Exo- $\beta$ -1,3-GAL	Fungi/Bacteria	$\beta$ -1,3-galactan and $\beta$ -1,3-galactooligosaccharides and release monomeric Gal from their nonreducing ends
Endo- $\beta$ -1,6-GAL <sup>a</sup>	Fungi/Bacteria	$\beta$ -1,6-D-galactosidic linkages
$\beta$ -GALs	Fungi/Bacteria	$\beta$ -1,3- and $\beta$ -1,6-galactans
$\alpha$ -L-AFs	Bacteria	Terminal nonreducing residues from arabinose-containing polysaccharides
$\beta$ -L-APs <sup>a</sup>	Bacteria	Terminal nonreducing L-arabinose
$\beta$ -glucuronidase <sup>b</sup>	Fungi	$\beta$ -D-GluA located at the terminal nonreducing end of 1,6- $\beta$ -Gal side chains
$\alpha$ -L-rhamnosidase	Bacteria	Terminal nonreducing $\alpha$ -rhamnose

<sup>a</sup>: to our knowledge not available in the market

<sup>b</sup>: different source respect to the used one



**Fig. 4.** Schematic representation of a complex type II AGs with the site of attack of various enzymes.



- RGS backbone degrading enzymes.  
Rhamnogalacturonase-hydrolases (EC 3.2.1.171) and rhamnogalacturonase-lyases catalyze endo-depolymerisation of the rhamnogalacturonan backbone. RG-hydrolase cleaves the linkage GalA-(1,2)- $\alpha$ Rha releasing oligosaccharides with rhamnose at the non-reducing end, whereas RG-lyase cleaves the linkage Rha-(1,4)- $\alpha$ GalA [37,38].
- RGS side chains degrading enzymes.  
The galactan component is hydrolyzed by the employment of endo-galactanases and galactosidases, whose action have already been described for arabinogalactan polysaccharides; while the arabinan fraction is degraded by endo- $\alpha$ -1,5-arabinanase (EC 3.2.1.99) and  $\alpha$ -arabinofuranosidases [39,40].

### 3. METHODOLOGICAL DEVELOPMENT ON STANDARD POLYSACCHARIDES

The objective of this section is to evaluate and optimize all the parameters that are involved in (1) the detection of oligosaccharides by MALDI-MS and (2) the enzymatic digestion of polysaccharides. The study was carried out on standard polysaccharides which resemble the structure of the two most significant plant gums polysaccharide families: type II arabinogalactans (larch arabinogalactan) and galactomannans (carob galactomannan). In the first part (3.1) the materials that were used (standards and enzymes), followed by the evaluation of the best conditions for oligosaccharides detection by MALDI-TOF (matrix, derivatization and clean-up step) will be discussed. In the second part (3.2 and 3.3) the optimization of the enzymatic digestion protocol and, therefore, the choice of the enzyme amount, type of enzymes and incubation timing will be explained. Finally the profiles obtained for the two polysaccharides will be shown and discussed. The obtained results are necessary to establish an optimized protocol of analysis that will be later applied to the study of plant gum samples.

#### 3.1. EXPERIMENTAL PROCEDURE OPTIMIZATION

##### *3.1.1 Standard polysaccharides description and preparation*

Larch wood arabinogalactan and carob galactomannan (low viscosity) were purchased by Megazyme and they were selected since they respectively resemble gum arabic and LBG structure. Larch arabinogalactan (MW ~47 kDa obtained by SEC-MAALS) is prepared at 3% weight/volume (w/v) in milliQ water. The solution is stirred on a hot-plate magnetic stirrer at 80°C until it is clear (approximately 15 minute). It is then left cooling down, centrifuged for 15 minute at 13400 rpm in order to remove possible insoluble residues and then stored at 4°C.

Carob galactomannan (MW ~107 kDa obtained by SEC-MAALS) is prepared at 1% w/v in milliQ water. Before adding water the polysaccharide powder is wet with 95% MeOH (1  $\mu$ L/mg of gum). The solution is gently stirred for about 1 hour and then left over night at 4°C to enhance the polymer hydration. The solution is then stirred on a hot-plate magnetic stirrer at 120°C for about 20 minutes. Once cooled down, the solution is centrifuged for 15 minutes at 13400 rpm, the supernatant is kept and stored at 4°C.

##### *3.1.2. Enzymes and enzymatic hydrolysis general procedure*

Exo- $\beta$ -1,3-galactanase (EC 3.2.1.145),  $\alpha$ -L-arabinofuranosidase (EC 3.2.1.55),  $\alpha$ -galactosidase (EC 3.2.1.22) and endo-1,4- $\beta$ -mannanases (EC 3.2.1.78) were purchased from NZYtech while  $\beta$ -galactosidase (EC 3.2.1.23),  $\beta$ -glucuronidase (EC 3.2.1.31) and  $\alpha$ -L-rhamnosidase (EC 3.2.1.40) were purchased from Megazyme. The specific characteristics of each enzyme, as described from the suppliers, are reported in Tab. 4 together with the operating conditions (last three columns).

The enzymatic digestion protocol has been optimized specifically for each standard polysaccharide and it is described in the subsequent sections (3.2 and 3.3). However the general approach can be summarized in the following steps:

<b>Dehydration:</b>	1 mg of gum, previously solubilized, is dried by a vacuum centrifuge
<b>Re-solubilization:</b>	buffer solution with an appropriate pH, selected according to the enzyme, is added and gum re-suspended
<b>Enzyme addition:</b>	selected enzyme/s are added to the gum solution and mixed vigorously
<b>Digestion:</b>	sample is incubated at the precise digestion temperature
<b>Digestion quench:</b>	sample is heated at 100°C for 5-10 minutes



**Tab. 4.** Information about purchased enzymes. “Working buffer” is the solution at the specific pH in which each enzyme was diluted and “T °C” is the temperature kept during digestion. The “Unit definition” related to enzyme activity are described in the technical resources for Megazyme (<http://www.megazyme.com/>) and Nzytech (<https://www.nzytech.com/site/>).

Enzyme	Specific activity	MW [kDa]	Shipping buffer	pH stability	T <sub>max</sub> [°C]	Working buffer	pH	T [°C]
<b>Exo-β-1,3-galactanase</b> <i>Clostridium thermocellum</i>	20 U/mg	37.9	35 mM NaHepes pH 7.5, 750 mM NaCl, 200 mM imidazol, 3.5 mM CaCl <sub>2</sub> , 0.02% sodium azide, 25% glycerol	3-10	55	Phosphate buffer 50 mM	6	50
<b>α-L-Arabinofuranosidase</b> <i>Clostridium thermocellum</i>	125 U/mg	58.9	35 mM NaHepes pH 7.5, 750 mM NaCl, 200 mM imidazol, 3.5 mM CaCl <sub>2</sub> , 0.02% sodium azide, 25% glycerol	6-8	70	Phosphate buffer 50 mM	7	65
<b>Endo- β-1,4-mannanase</b> <i>Cellvibrio japonicus</i>	2200 U/mg	51.7	35 mM NaHepes pH 7.5, 750 mM NaCl, 200 mM imidazol, 3.5 mM CaCl <sub>2</sub> , 0.02% sodium azide, 25% glycerol.	6.5-8	50	Phosphate buffer 50 mM	7.5	45
<b>α-Galactosidase</b> <i>Cellvibrio mixtus</i>	150 U/mg	44.8	35 mM NaHepes pH 7.5, 750 mM NaCl, 200 mM imidazol, 3.5 mM CaCl <sub>2</sub> , 0.02% sodium azide, 25% glycerol.	8-9.5	40	Tris-HCl buffer 25 mM	8.5	37
<b>β-Galactosidase</b> <i>Aspergillus niger</i>	175 U/mg	/	3.2 M ammonium sulphate + 0.02% Na azide.	3-8	<70	Acetate buffer 50 mM	4.5	60
<b>β-Glucuronidase</b> <i>Escherichia Coli</i>	37 kU/mg	82.6	20 mM Tris-HCl pH 7.5, 50 mM NaCl, 0.1 mM EDTA plus 0.02% (w/v) sodium azide.	5-7.5	<50	Phosphate buffer 100 mM	6.8	37
<b>α-L-rhamnosidase</b> <i>Prokaryote</i>	190 U/mg	75.4	Ammonium sulphate suspension in 0.02 % (w/v) sodium azide.	4-9	<50	Phosphate buffer 20 mM	6.5	50
<b>Endo- β-1,4-galactanase</b> <i>Clostridium thermocellum</i>	10kU/mg	36.6	35 mM NaHepes pH 7.5, 750 mM NaCl, 200 mM imidazol, 3.5 mM CaCl <sub>2</sub> , 0.02% sodium azide, 25% glycerol	5-7.5	70	Phosphate buffer 50 mM	7	45

### 3.1.3. MALDI-TOF matrix

Optimization of mass spectrometric conditions was carried out on digested larch arabinogalactan standard polysaccharide. A general digestion was performed taking the cue from enzymatic digestion performed on general type II arabinogalactans [41]. All the available enzymes for type II AGs hydrolysis were used in order to remove first the side chains and then facilitate the access of further enzymes to the main backbone. The experiments related to the digestion procedure optimization will be described in the following section since it was first necessary to optimize the detection parameters.

The following enzymes have been used in the listed order:

1.  $\alpha$ -L-arabinofuranosidase;
2.  $\beta$ -glucuronidase;
3.  $\alpha$ -L-rhamnosidase;
4.  $\beta$ -galactosidase;
5. exo- $\beta$ -1,3-galactanase.

Enzymes  $\alpha$ -L-arabinofuranosidase,  $\beta$ -glucuronidase and  $\alpha$ -L-rhamnosidase were used at the beginning of the digestion in order to remove the terminus of the side chains and to debranch the larch arabinogalactan structure. Exo-galactanases enzymes can bypass the branching points of the main backbone of type II AGs (galactose monomers linked 1,3) and release intact side chains only if the structure is partially debranched. Usually restricted activity is described for native AGs [26,42,43].  $\alpha$ -L-arabinofuranosidase was also used as a first enzyme because it seems to help the action of  $\beta$ -glucuronidase [44]. Furthermore, besides the theoretical information on enzymes action, experimental evidences of application of the reported enzymes for type II arabinogalactans digestion is shown in the literature. Similar sequences of enzymes were used by Tryfona *et al.* for the investigation of the carbohydrate component of AG from wheat flour [41] and *Arabidopsis* leaf [45]; and by Huisman *et al.* for the AG component of soybean substances [46,47].

1 mg of larch arabinogalactan, previously solubilized in milliQ water (1% w/v), is dried, resuspended in the buffer and digested with 100 mU of each enzyme for 24 hours. Addition of a subsequent enzyme required head inactivation of the previous one by heating the sample at 100°C for 5-10 minutes. The solution is then dried, the sample dissolved in the proper new buffer and the following enzyme added. Buffer composition and digestion temperature are reported in Tab. 4. This process was performed for all 5 enzymes for a final 5 days procedure. At the end the sample was dried and resuspended in few microliters of milliQ water before being spotted on the MALDI plate, once mixed with the proper matrix.

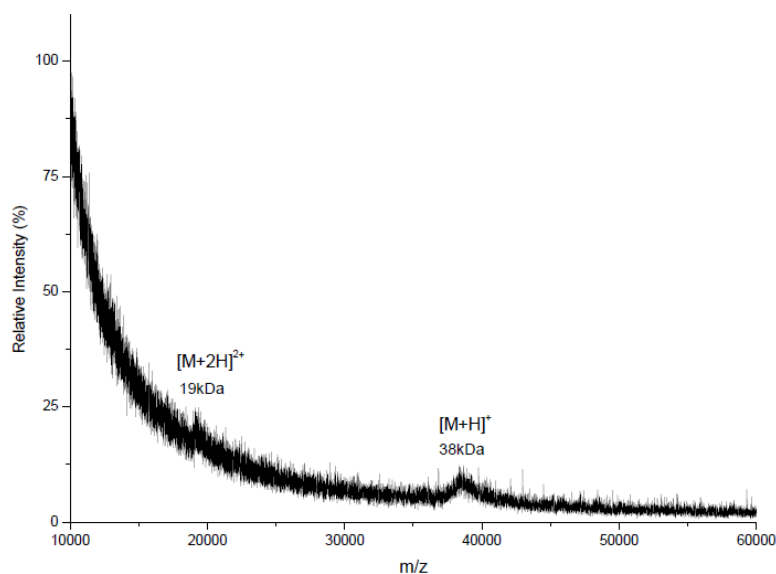
As reported by Stahl *et al.* [8], 2,5-DHB is so far the most used matrix for analysis of oligosaccharides. First tests were performed trying different solvents for DHB solubilization. For each matrix, different ratios sample/matrix (1/1; 1/4; 1/10; 1/20; 1/40) were used. Addition of inorganic salts (e.g. NaCl and LiCl) was also tested since they can enhance carbohydrate ionization [48]. The different matrices are listed below:

- DHB 10 mg/mL in 100% ACN + 0.1% TFA;
- DHB 10 mg/mL in ACN 50% in water + 0.1% TFA;

- DHB 10 mg/mL in ACN 75% in water + 0.1% TFA;
- DHB 10 mg/mL in MeOH 50% in water + 0.1% TFA;
- DHB 10 mg/mL in MeOH/Chloroform/water 1.5/3/0.25;
- DHB 10 mg/mL in MeOH + 2% NaCl (1 M in MeOH);
- DHB 10 mg/mL in MeOH + 2% LiCl (saturated solution in MeOH);

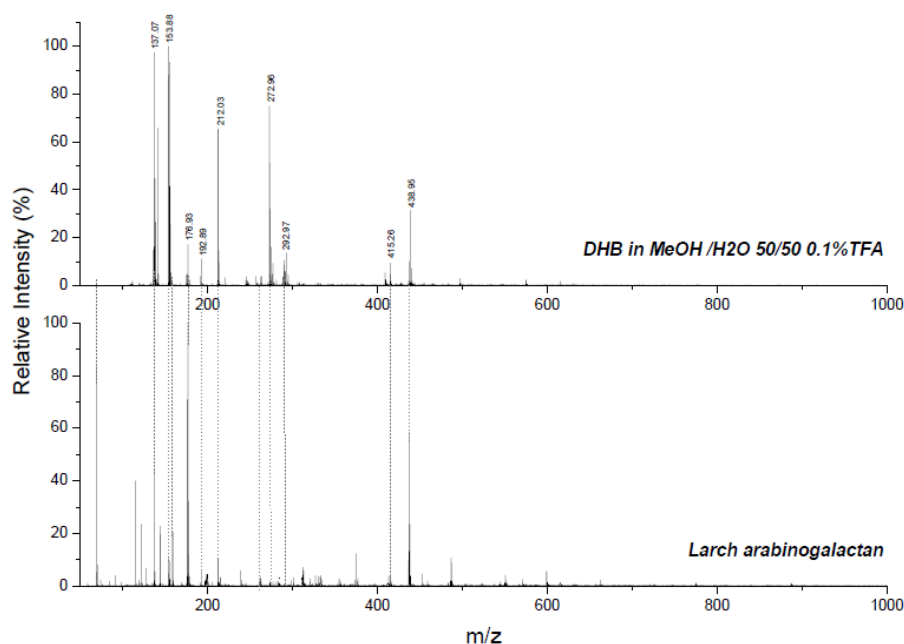
Besides the wide application of DHB, new matrices were introduced for the analysis of polysaccharides of over 3 kDa. Hsu *et al.* [49] applied the matrix 2',4',6'-trihydroxyacetophenone (THAP) for the first time to the study of sugars and they obtained good quality spectra and reproducibility for high MW linear polysaccharides, dextrans and glycoproteins. Therefore also THAP was tested and a solution 50 mM was prepared in 50% of acetonitrile in water + 0.1% TFA.

1  $\mu$ L of the mixture matrix/sample was spotted on the MALDI plate and analyzed using the parameters reported in the chapter "Materials and Methods". Even if the selected matrices are the classic ones used for the analysis of oligosaccharides, in the case of the analyzed samples no crystallization of the mixture matrix/sample occurred. Most of the spots appeared as a drop on the MALDI plate even after several dilution of the sample, thus no analysis could be performed. The only crystallization occurred when the sample was spotted with DHB dissolved in MeOH 50% in water + 0.1% TFA [41]. The typical needle-shaped crystals of DHB were formed in the periphery of the spot but any amorphous area in the center, usually containing the analytes plus salts and contaminants, was observed. Since the selected enzymes cut the polysaccharide in *exo*- mode, the presence of oligosaccharides with high molecular weight was investigated by analyzing the spot in linear mode (10-60 kDa) (Fig. 6).



**Fig. 6.** Mass spectrum in linear mode (10-60kDa) of larch AG digested and spotted with DHB in MeOH/H<sub>2</sub>O 50/50 0.1% TFA. Mono and double charged ions of *exo*-  $\beta$ -1,3-GAL enzyme.

In the selected mass range only the mono ( $m/z$  38000) and doubly charged ( $m/z$  19000) ions of the enzyme *exo*- $\beta$ -1,3-galactanase (MW = 37.9 kDa) were visible. Since no oligosaccharides with high MW were detected, analysis in reflector mode (50-3000 Da) was performed in order to check if smaller polysaccharide fragments were present.



**Fig. 7.** Mass spectra comparison of the matrix (above) and the larch arabinogalactan digested with  $\alpha$ -L-arabinofuranosidase followed by  $\text{exo-}\beta$ -1,3-galactanase for 24 hours (below).

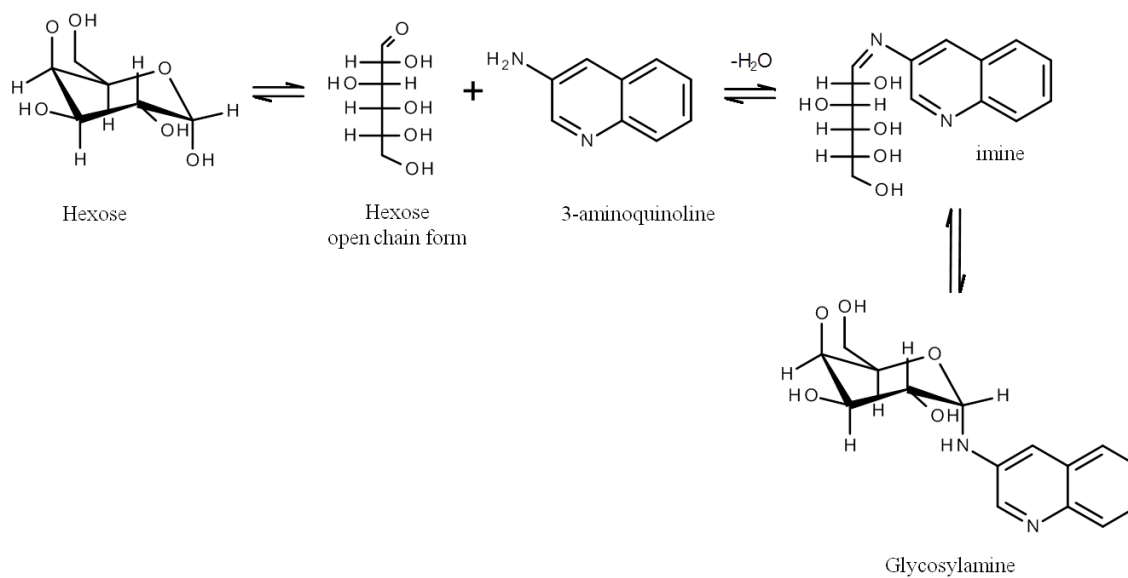
As shown in Fig. 7, the main peaks observed for the digested larch arabinogalactan corresponded to the matrix ions ( $m/z$  137.07; 153.88; 176.93; 192.82; 212.03; 272.96; 292.97; 415.26, 438.95). Some other peaks at lower intensity were present in the AG spectrum ( $m/z$  114.52; 121.82; 239.01; 312.14; 374.83; 487.20; 598.91; 663.17) but no mass differences of the corresponding monosaccharides (132.04; 162.05 Da etc.) were identified. Obtained results showed how the main difficulties in analyzing plant gums by mass spectrometry can be related to both two different issues: mass spectrometric detection and sample handling. Therefore, in a first step, sugar derivatization and clean-up procedures were tested in order to improve the MS detection. In a later stage variations on enzymatic digestion protocols were evaluated for each standard polysaccharide.

#### 3.1.4. Oligosaccharides derivatization by 3-aminoquinoline

Oligo- and polysaccharides cannot be considered the best analytes for MS since their polarity, thermal instability, low volatility and frequent branched structure imply a low efficient ionization. For this reason derivatization techniques, of both intact polymers or oligosaccharides obtained by different depolymerization methods, represent a fundamental step in sugar analysis by mass spectrometry. In this way simplest spectra are obtained and response increases. Multiple carbohydrate derivatization procedures for analysis by chromatography, electrophoresis and mass spectrometry have been reviewed recently by D.J. Harvey [12]. Classical carbohydrate derivatization procedures, such as permethylation, are usually time consuming and might lead to sample loss since the strategy consists in several steps and multiple purifications must be performed. For this reason, easier and quicker methods are preferred but these must be suitable for MALDI analysis in terms of sample handling and matrix efficiency.

Recently Rohmer *et al.* [50] have reconsidered the use of 3-aminoquinoline (3-AQ) which was first used by Metzger *et al.* for the analysis of plant inulins [51]. Even if 3-AQ showed better sensitivity

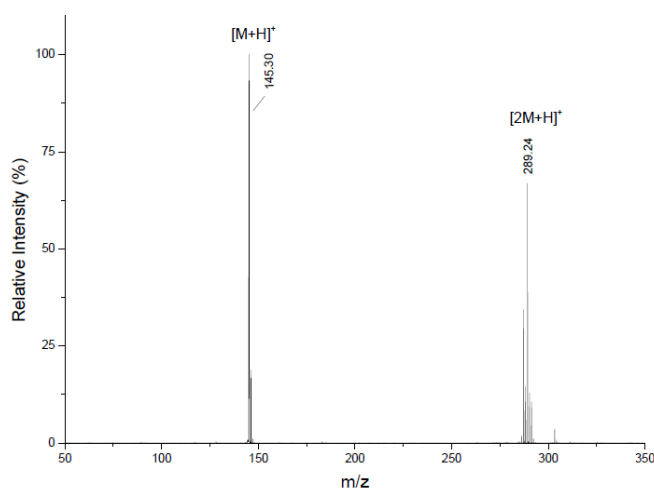
and resolution in comparison to the classical DHB matrix, it was not further used because of the formation of Schiff bases with carbohydrate reducing terminus. Rohmer *et al.* took advantage of this reaction and optimized analytical conditions in order to use 3-aminoquinoline as both a powerful matrix and derivatizing reagent. In this way oligosaccharides can be analyzed in positive mode by MALDI-TOF as their 3-AQ derivatives. Derivatization by 3-aminoquinoline consists in a reductive amination reaction. As showed in Fig. 8 (image from [50]) the carbonyl group of the sugar open chain form is converted to an amine via the matrix 3-AQ.



**Fig. 8.** Scheme of oligosaccharide derivatization by reductive amination.

According to their procedure, the sample is directly mixed with 3-aminoquinoline on the MALDI plate, so derivatization is performed directly on-target and any purification step is necessary. 3-AQ is prepared 20 mg/mL in ACN/H<sub>2</sub>O 1/2 with 0.07% of inorganic acid to reach pH = 5.

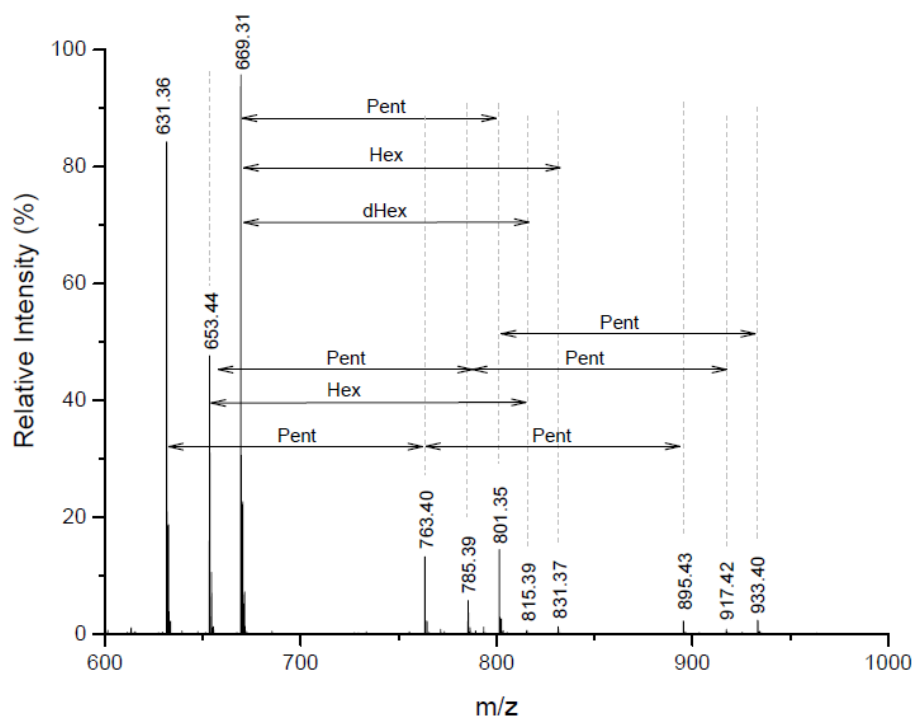
1  $\mu$ L of 3-aminoquinoline solution was spot on the MALDI plate and analyzed in reflector mode (50-1000 Da). The mass spectrum is showed in Fig. 9.



**Fig. 9.** Mass spectrum in positive reflector mode of 3-aminoquinoline in ACN/H<sub>2</sub>O 1/2 pH = 5.

The peak at 145.30 Da corresponded to the protonated 3-aminoquinoline (MW 144.1 Da) and the peak at 289.24 Da to the dimer. The trimer (435.35 Da) was visible at lower values of signal/noise.

After enzymatic digestion of the larch arabinogalactan, the sample was dried and resuspended in milliQ water. 0.5  $\mu\text{L}$  of sample were spot on the MALDI plate and 1  $\mu\text{L}$  of 3-AQ solution was added. Matrix and sample were mixed and left dried at room temperature for 1 hour. Different dilutions of the sample were prepared (range 200-10 pmol). Analysis was performed in positive reflector mode and results are shown in Fig. 10. For peaks identification see section 3.2.3.



**Fig. 10.** MALDI-TOF mass spectrum of derivatized oligosaccharides of larch arabinogalactan after enzymatic digestion (zoom 600-100Da).

The MALDI-TOF spectrum of larch arabinogalactan, after digestion with a combination of five enzymes, confirmed that the main problem encountered at the beginning of experiments was related to the detection step and, specifically, to the matrix choice and derivatization. While with DHB no spectrum was obtained, for the same sample 3-aminoquinoline allowed to detect a series of ions which differed by the specific monosaccharides residual masses. If considering the mass range over 600 Da, thus excluding the spectrum area containing the matrix ions, the most intense ions were found to be m/z 631.36; 653.44; 669.31; 763.40; 785.39; 801.35; 895.43; 933.40. It was possible to observe a mass difference of 132.04 Da between m/z 801.35 and 669.31 ion, suggesting the loss of a pentose (Pent), more likely an arabinose according to the hypothetical structure of arabinogalactans. The same mass difference could be observed between the ions 933.40 and 801.35, 917.42 and 785.39, 785.39 and 653.44, 895.43 and 763.40, 763.40 and 631.36. The mass spectrum showed also the presence of a mass difference of 162.05 Da which corresponds to an hexose (Hex), and specifically galactose in this case study. The enzymatic cleavage of a Gal was also observed between the ions 831.37 and 669.31, and 815.39 and 653.44. Furthermore the presence of a deoxyhexose (dHex), like rhamnose, in the arabinogalactan structure was proved by the mass difference of 146.08 Da between the ions 815.39 and 669.31. From these observations it could be concluded that the m/z 669.31 ion can be a combination of several oligosaccharides, thus

confirming the complexity of larch arabinogalactan structure. The mass spectrum showed how, unlike the classical enzymatic digestion approach which implies the use of only exo-acting enzymes, application of the strategy to more complex polysaccharides implies a different approach. Since both exo- and endo- acting enzymes were employed to decomplexify the polysaccharide structure, the mass spectrum was characterized by the presence of both intact oligosaccharides (e.g. side chains) released by the endo- enzymes, and possibly single monosaccharides, released by the exo- enzymes.

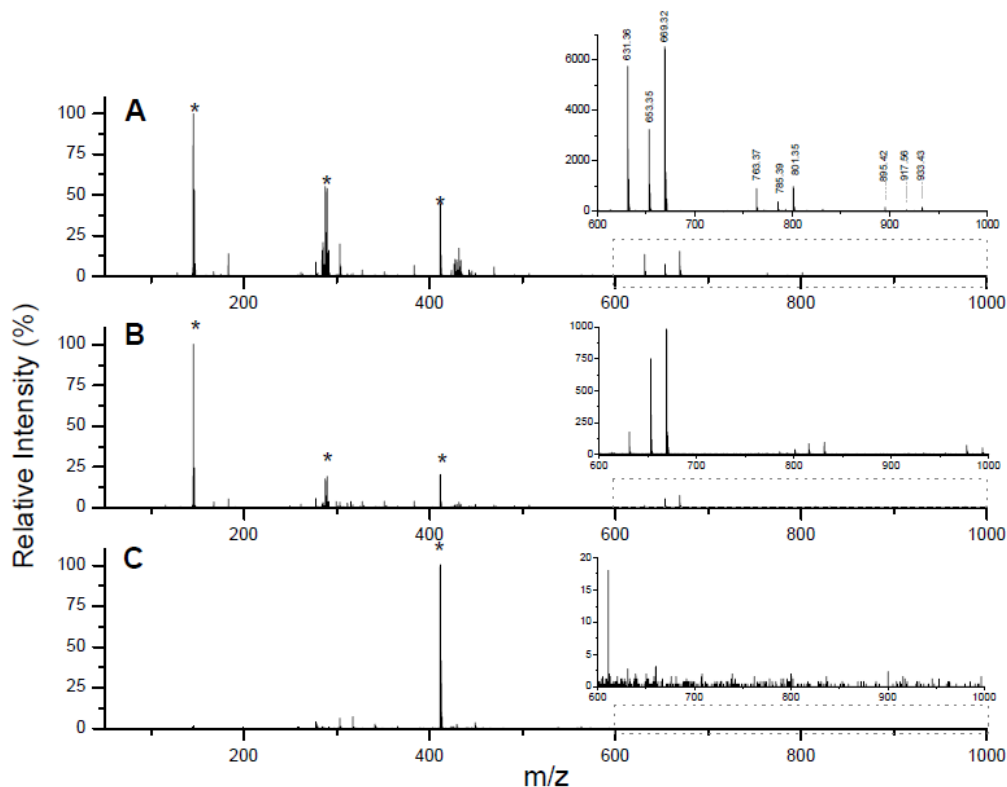
Due to these promising results, 3-aminoquinoline was selected as both matrix and derivatizing reagent. However, before proceeding to the identification of each possible oligosaccharide observed in the mass spectrum, another step in sample preparation was evaluated. After enzymatic digestion the sample can be rich in salts, enzymes and possible peptides. Therefore purification procedures were evaluated in order to improve the sample ionization.

#### 2.1.5. Oligosaccharides purification.

It is well established that the presence of salts and buffers can significantly influence the sample ionization and matrix crystallization. Among the possible procedures, dialysis cannot be used since it requires membranes with a low molecular weight cut-off. The simplest method of washing the dried target with milliQ water, usually used for proteins and peptides [52], is not applicable to carbohydrate samples because of their high solubility in water. Some conclusions can be made for the application of solid-phase extraction cartridges where a packed silica is modified in order to confer specific characteristics to the solid-phase. The classic reversed-phase (C18 or C4) shows a low retention for hydrophilic solutes so, while it is usually used for protein purification, it is actually useless for oligosaccharides. Among the different cartridges, the use of graphitized carbon was reported by Packer *et al.* as a suitable solid-phase for fractionation and desalting of mixture of oligosaccharides which are eluted with a 25% solution of ACN in water [53].

Therefore, following the enzyme digestion, the solution was purified using Hypersep Hypercarb zip tip (10-200  $\mu$ L, Thermo Scientific) according to the procedure reported in "Materials and methods". Oligosaccharides eluted with different ACN percentages were dried by a vacuum centrifuge, resuspended in water and analyzed in reflector mode (Fig. 11). Besides the main characteristic matrix ions in the mass range below 600 Da (marked in the figure by an asterisk), few peaks were observed at higher molecular weight in the mass spectra corresponding to the loading (Fig. 11a - zoom) and the washing (Fig. 11b - zoom) solutions, while no other ions were identified for elution solutions (only ACN10% elution showed). As showed in the enlarged image on the top of Fig. 11a, b and c, the observed ions corresponded to the masses obtained after enzymatic digestion of the arabinogalactan. It could be therefore concluded that, although graphitized carbon was reported as a proper material for carbohydrates purification, in this case study the released oligosaccharides were not retained by the phase and they were eluted with the washing.

Several other solid-phase cartridges were evaluated according to the physical/chemical properties of polysaccharides. Among these HILIC was tested, since it is a solid phase suitable for hydrophilic compounds, but it did not lead to any satisfactory sample clean up.



**Fig. 11.** Mass spectra (and corresponding enlarged image of the mass range 600-1000Da) of digested larch AG after fractionation by hypercarb: (A) loading solution; (B) washing solution and (C) elution with ACN 10%.

Considering the optimization of oligosaccharide detection, different parameters have been taken into account and it was concluded that:

- The use of the classical DHB matrix did not allow to obtain a mass spectrum of the type of oligosaccharides under investigation. A simple derivatization step consisting of mixing the sample directly on the MALDI plate with 3-aminoquinoline, which acts in the meantime as matrix, allowed to obtain sample spots and to observe series of ions in the mass spectrum of larch arabinogalactan. Therefore a derivatization step was showed to be necessary prior MALDI analysis.
- Even if the final oligosaccharide solution contains salts, due the buffers used, and enzymes residues, purification by solid phase cartridges (e.g. graphitized carbon, HILIC) was revealed to cause loss of oligosaccharides which are not retained by the solid phase. Therefore any purification of the sample is considered necessary before MALDI analysis.

### 3.2. OPTIMIZATION OF LARCH ARABINO GALACTAN ENZYMATIC DIGESTION

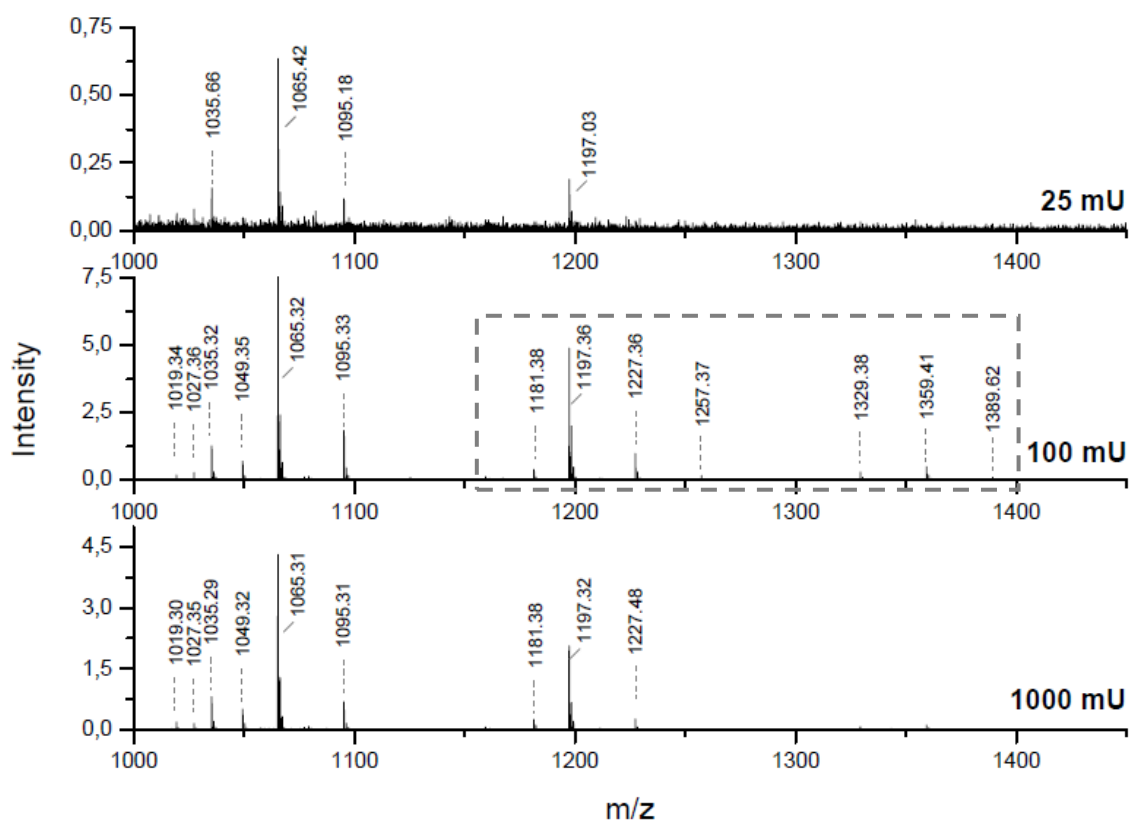
#### 3.2.1. Evaluation of enzyme amount

When working with enzymes, their unit of measure corresponds to their “activity”, which is defined as the amount of enzyme necessary to release a certain number of moles of substrate in a certain time under specific conditions. It is expressed in “Units” (U) and 1U corresponds to the



amount of enzyme required to convert 1  $\mu\text{mol}$  of substrate in the corresponding product in 1 minute under certain conditions that must be specified (e.g. T °C, pH, buffer solution etc). For example, in the case of the purchased exo- $\beta$ -1,3-galactanase, the specific activity of the enzyme is 20 U/mg where 1U is described by the supplier as: “the amount of enzyme required to release 1  $\mu\text{mol}$  galactose-reducing-sugar equivalents per minute from  $\beta$ -1,3-galactan in phosphate buffer, pH 6.0, at 60°C” ([www.nzytech.com](http://www.nzytech.com)). However, while the described substrate corresponds to a simple chain of 1,3-Gal units, the polysaccharide under investigation has a much more complex branched structure with the presence of several linkages and monosaccharides. It is reported in the literature that for example the activity of an exo- $\beta$ -1,3-galactanase obtained from *Sphingomonas* is  $23.9 \pm 0.2$  U/mg for a  $\beta$ -1,3-galactan substrate, while it decreases to  $1.8 \pm 0.1$  and  $1.2 \pm 0.1$  U/mg for respectively larch arabinogalactan and gum arabic [54]. Therefore, it was not possible to simply replicate the enzymatic digestion procedure for arabinogalactans reported in the literature because the enzymes are usually obtained from different organisms, their activity is strictly related to the substrate and the theoretical activity might not corresponds to the experimental one. Furthermore, sometimes enzyme specificity is unknown or, even if known and described, enzymes can degrade the sugar in a non predictable manner. This is mainly due to the type of substrate, the steric hindrance of branched side chains and the enzyme purity [55].

In the light of these considerations, several amounts of enzyme have been tested in order to check if the rate of hydrolysis could be related or not to the units of enzyme used. Larch arabinogalactan, prepared according to the procedure described in the previous chapters, was digested adding respectively 25, 100 and 1000 mU of each of the 5 enzymes for mg of polysaccharide. The mass spectra are showed in Fig. 12.



**Fig. 12.** MALDI-TOF enlarged spectra (1000-1500 Da) of larch arabinogalactan digested with 25, 100 and 1000 U of each enzyme.

Besides the peaks in the mass range 600 - 1000 Da, further peaks were identified at higher molecular weights. As shown in Fig. 12, just few ions and with a very low intensity appeared when the polysaccharide was digested with 25 mU of enzyme, thus revealing that the amount of enzymes was maybe not sufficient to obtain a proper digestion. 100 mU of enzyme allowed to obtain new oligosaccharides: a mass difference of 132.04 Da could be observed between the ions 1197.36 and 1329.38, 1227.36 and 1359.41, 1065.32 and 1197.36, 1049.35 and 1181.38; the  $m/z$  1389.62, 1359.41, 1257.37, 1227.36 and 1181.38 ions showed a mass loss of 162.05 Da and a mass difference of 146.06 Da was observed between the ions 1181.38 and 1035.32. The use of 1000 mU of enzyme led to similar results except for the absence of the ions 1329.38, 1359.41 and 1389.62 which may had a too low intensity to be revealed or they were further digested by the enzymes which were present in high amount.

In conclusion, an amount of 100 mU was considered to work optimally. Besides the absence of some ions at higher molecular weight ( $> 1250$  Da) when using 1000 mU, it was preferred to use a lower quantity since enzyme residues and salts, from the used buffers, will be present in the solution before analysis.

### 3.2.2. Evaluation of enzymes combination and digestion time

Currently MALDI-TOF analysis were performed at the end of the digestion protocol, so after all 5 enzymes were added. The obtained mass spectra showed the presence of specific oligosaccharides but: (i) information about the action of each enzyme is missing; (ii) the digestion procedure takes 1 week to be performed which means a significant chance to lose analytes; (iii) it is necessary to better understand the structure of the released oligosaccharide in order to get a MALDI-TOF profile of the larch arabinogalactan.

The standard polysaccharide was initially digested with  $\alpha$ -L-arabinofuranosidase for 24 hours. Reaction was quenched by boiling the solution, an aliquot was collected and analyzed by MALDI-TOF while the left dried sample was incubated for 24 hours with the second enzyme ( $\beta$ -glucuronidase). The same procedure was carried out for all the further enzymes ( $\alpha$ -L-rhamnosidase,  $\beta$ -galactosidase and  $\text{exo-}\beta$ -1,3-galactanase) so five MS spectra were finally obtained.

#### (1) $\alpha$ -L-arabinofuranosidase.

Besides the matrix ions, only a peak at  $m/z$  277.22 was observed which corresponded to a single arabinose derivatized with 3-aminoquinoline (theoretical monoisotopic mass 277.1 Da). No other oligosaccharides have been observed revealing that the enzyme had an *exo* activity and it was able to cut single Ara monosaccharides from the external chains.

#### (2) $\alpha$ -L-arabinofuranosidase + $\beta$ -glucuronidase

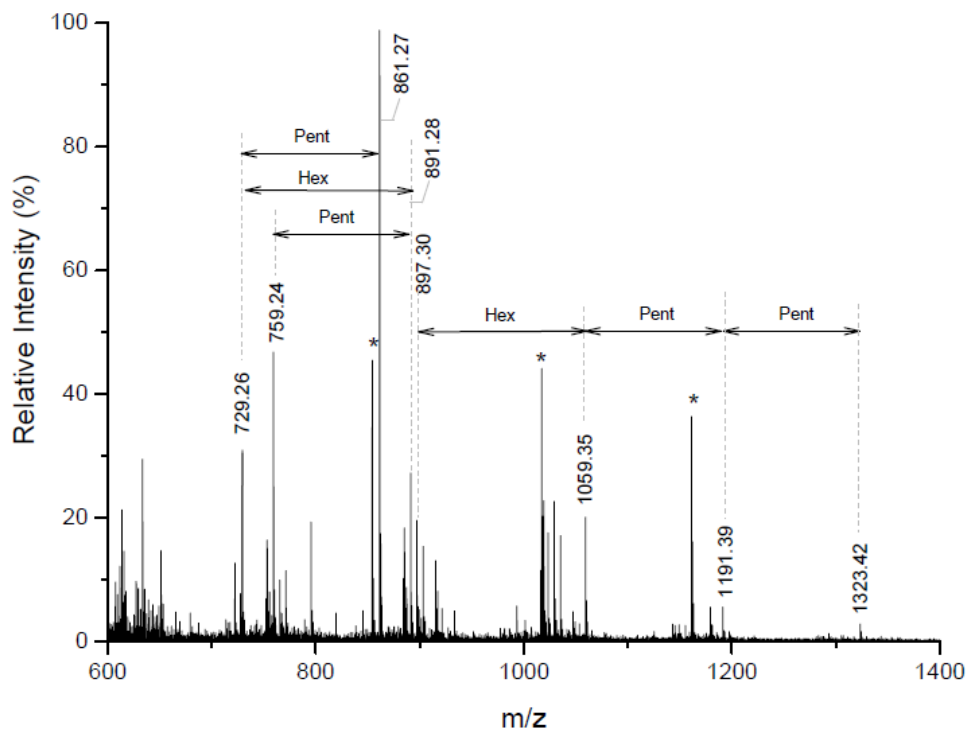
No further ions, except the matrix ones, were observed in the mass spectrum.

#### (3) $\alpha$ -L-arabinofuranosidase + $\beta$ -glucuronidase + $\alpha$ -L-rhamnosidase.

No further ions, except the matrix ones, were observed in the mass spectrum.

#### (4) $\alpha$ -L-arabinofuranosidase + $\beta$ -glucuronidase + $\alpha$ -L-rhamnosidase + $\beta$ -galactosidase.

The mass spectrum of larch arabinogalactan after adding the fourth enzyme  $\beta$ -galactosidase is reported in Fig. 13.



**Fig. 13.** MALDI-TOF enlarged spectrum (600-1400Da) of larch arabinogalactan after digestion with 4 enzymes ( $\alpha$ -L-arabinofuranosidase +  $\beta$ -glucuronidase +  $\alpha$ -L-rhamnosidase +  $\beta$ -galactosidase). (\*) described in the following section.

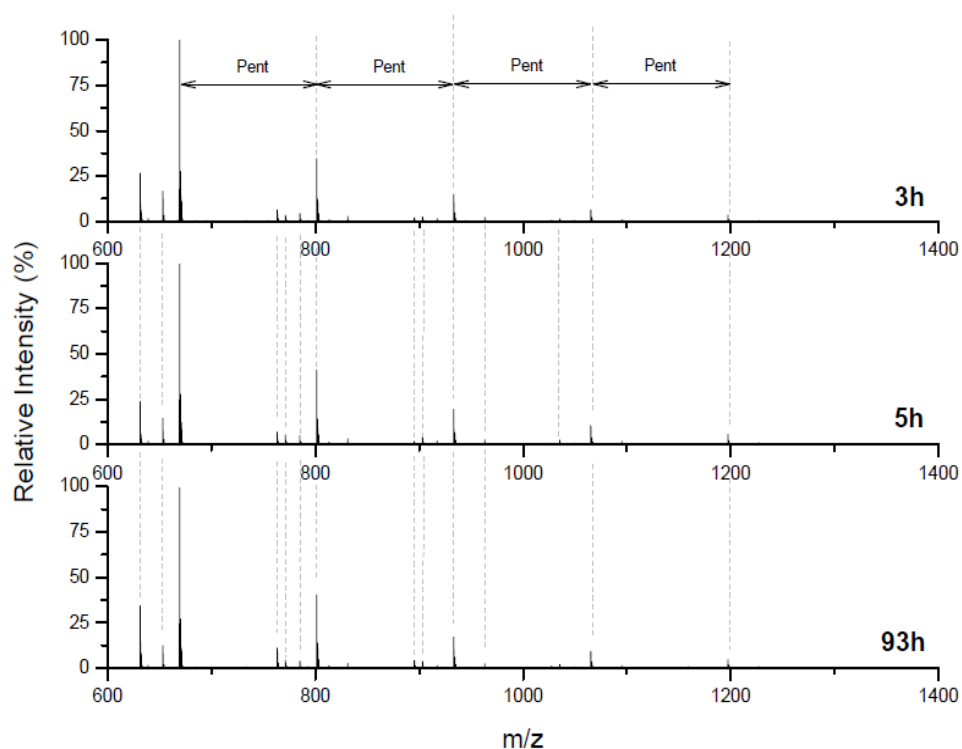
Even if  $\beta$ -galactosidase has usually a significant activity on low molecular mass substrates and it is not specific for type II AGs digestion [26], data showed how some oligosaccharides, characterized by a mass differences of 132.04 and 162.05 Da, were released from the polysaccharide. The particularity of the results was that the observed ions did not correspond to the ones obtained after digestion with all five enzymes thus showing a different activity if compared to the  $\text{exo-}\beta$ -1,3-galactanase.

#### (5) All five enzymes

Reproducibility of the digestion procedure was verified since the characteristic MALDI-MS profile of larch arabinogalactan was obtained (same ions reported in Fig. 10). Once the protocol was confirmed, aliquots of the digested sample were collected at different digestion time in order to see if the enzyme activity, and therefore the rate of hydrolysis, could be linked to the timing. In the specific the sample was analyzed after respectively 3, 5 and 93 hours. The corresponding mass spectra are reported in Fig. 14.

No differences in the mass spectra were observed when the standard polysaccharide was digested for 3, 5 or 93 hours. Therefore 3 hours digestion was considered sufficient. The enzyme activity reached the maximum in few hours and then tended to decrease with the time, so no more oligosaccharides were released. The ions found after digestion with  $\beta$ -galactosidase (Fig. 13) were not found anymore in the mass spectrum. It was therefore possible that the oligosaccharides had been further digested by the last enzyme. According to these experiments it seemed that, even if not

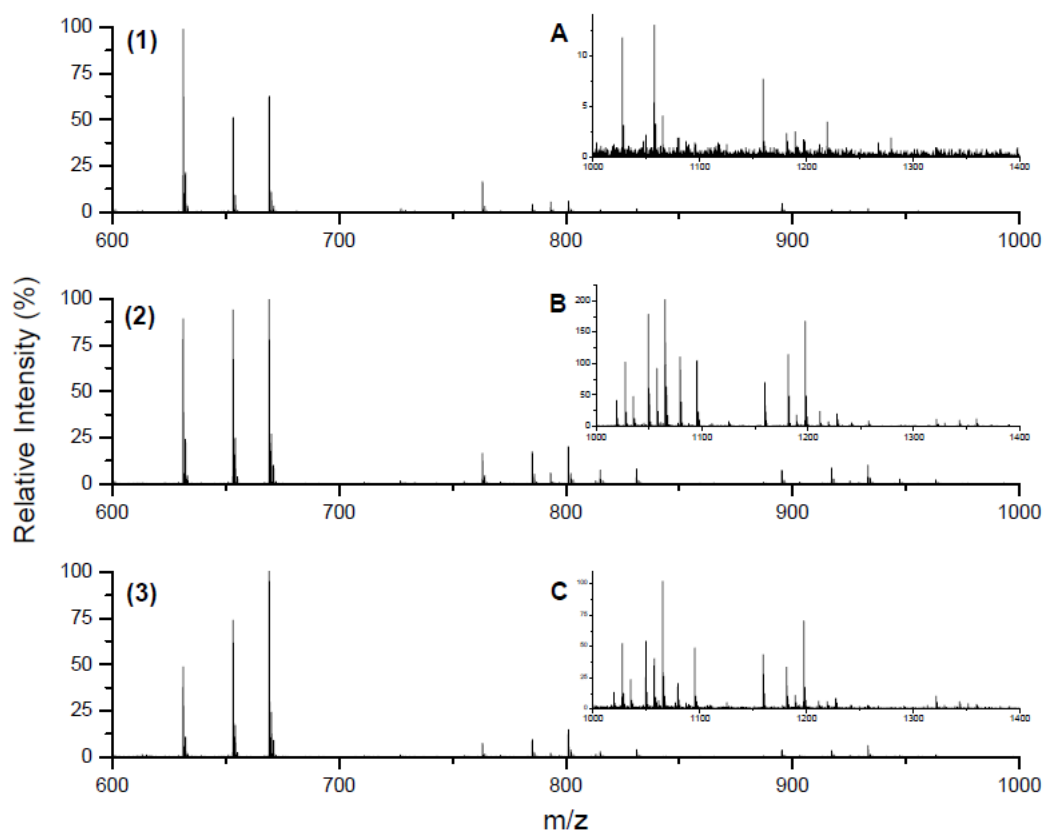
directly observable in the mass spectra, the four enzymes ( $\alpha$ -L-arabinofuranosidase,  $\beta$ -glucuronidase,  $\alpha$ -L-rhamnosidase and  $\beta$ -galactosidase) are necessary to remove the external monosaccharides units and so facilitate the access of the final exo- $\beta$ -1,3-galactanase.



**Fig. 14.** Mass spectra comparison of larch arabinogalactan digested with five enzymes for 3, 5 and 93 hours.

In order to better understand the structure of the released oligosaccharide and to shorten the digestion time, several combinations of  $\alpha$ -L-arabinofuranosidase,  $\beta$ -galactosidase and exo- $\beta$ -1,3-galactanase were tested. Among these, even if according to the literature the action of exo- $\beta$ -1,3-Gal should be prevented by the steric hindrance of the branched side chains and it should release only terminal monosaccharides, a test was performed incubating the sample only with this last enzyme for different time (5 minutes, 30 minutes, 1, 3, 5, 8, 24, 96, 120 hours). The most significant results are reported in Fig. 15.

In all the three mass spectra (Fig. 15(1), 2 and 3) it was possible to observe the characteristic ions obtained previously using five enzymes (Fig. 10). This meant that the enzyme exo- $\beta$ -1,3-galactanase was surprisingly able to partially hydrolyze the arabinogalactan without any previous removal of external side chains, thus releasing not simply monosaccharides but oligosaccharides. It is highly improbable that the enzyme cut the main chain, since it acts in a exo- way, but it is possible that it hydrolyzed some 1,3 linkages between Gal of the branched side chains. Secondly, digestion time could be extremely reduced since even after only 5 minutes the enzyme was able to release some oligosaccharides. However, if considering higher masses (images A, B and C), it could be observed how digesting the sample for more than 3 hours is preferred since other oligosaccharides at higher molecular weight were released. No significant differences were observed between a digestion of 3hours and the longer ones.



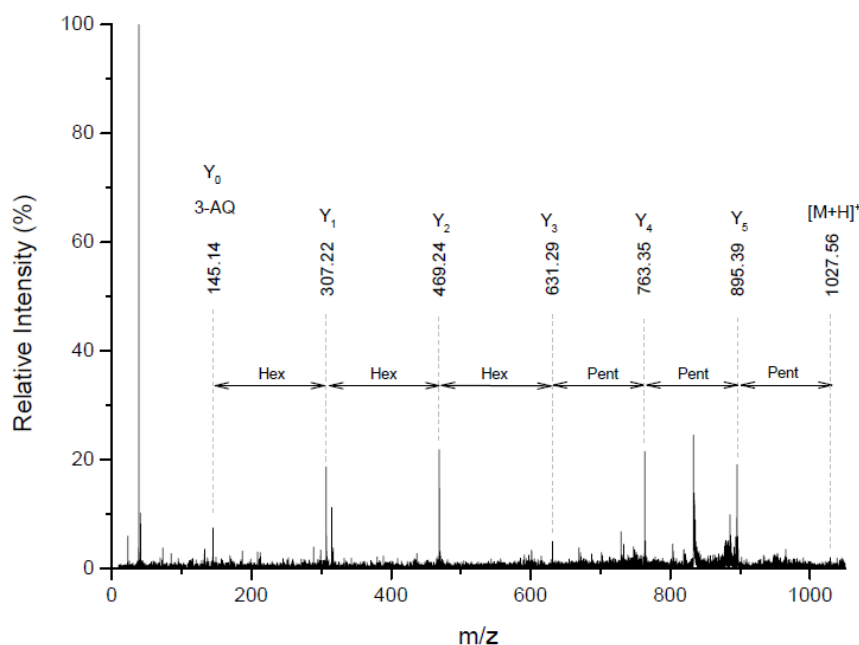
**Fig. 15.** MALDI-TOF spectra of larch arabinogalactan digested with exo- $\beta$ -1,3-galactanase for (1) 5 minutes; (2) 3 hours and (3) 120 hours. The enlargement of the mass range 1000 - 1400Da is showed respectively in the mass spectra A, B and C.

At the light of these considerations, a final digestion protocol of the standard larch arabinogalactan was established: 1 mg of polysaccharides is resuspended in a phosphate buffer 50 mM pH = 6 and incubated with 100 mU of exo- $\beta$ -1,3-galactanase for 3 hours at 50°C.

### 3.2.3. MALDI-TOF profile of the polysaccharide

In order to facilitate the interpretation of the released oligosaccharides, a basic program was built (using Fortran as program language). It calculates all the possible random combinations of the main monosaccharides of interest (Pent, Hex, dHex, HexA and GalNAc) for a maximum mass of 3000 Da and returns the mass of each combination protonated, with Na or K adducts, derivatized or non derivatized. Once the mass spectrum is obtained, the list of all peaks detected can be easily collected using the MALDI software Data Explorer (Applied Biosystem) as .txt file. This list can be subjected to the program and the final output is a .dat file with a list of all the possible oligosaccharide compositions for each m/z value. However, as it will be described later in the gum arabic section (4.2), the informatics tool represents just a first approach to highlight all the possible oligosaccharide compositions of a certain ion. Tandem mass spectrometric (MS/MS) experiments are therefore necessary to verify the correct structure of the oligosaccharide and clarify its composition. Fragmentation was performed in CID mode (Collision-Induced Dissociation) in the presence of air. Comparable results were obtained using argon. As example the MS/MS spectrum of the precursor ion m/z 1027.56 is reported in Fig. 16. As pointed out in the mass spectrum,

fragmentation analysis allowed to identify all the single monosaccharides that form the oligosaccharide under investigation, thus the oligosaccharide's sequence can be deduced. The MS/MS pattern looked simple with the presence of only Y-ions (differing in mass of 132.04 and 162.05 Da) and any ion coming from cross-ring fragmentation was observed. The formation of Y-ions is promoted by the derivatizing reagent. Quinoline has an high proton affinity and therefore the charge is retained only by the fragments that include the derivatized reducing end [50].



**Fig. 16.** MALDI-CID spectrum of 3-AQ derivatized oligosaccharide released by *exo*- $\beta$ -1,3-galactanase ( $m/z$  1027.56).

The ion with a mass of 145.14 Da corresponds to the protonated 3-aminoquinoline  $[M+H]^+$  and it indicates that the oligosaccharide is derivatized. The product ion at 895.39 Da was formed by the loss of a pentose residue (-132.04 Da). Same loss could be observed for the two product ions (763.35 and 631.29 Da) so three pentose were present in the oligosaccharide structure. The ion at 469.24 Da derived by the loss of 162.05 Da from the  $m/z$  631.29, and the same was observed for the following ion ( $m/z$  307.22) which had a mass difference of 162.07 Da with the ion representing 3-AQ. Therefore three hexoses were part of the structure. Following the fragmentation pattern, the ion 1027.56 was assigned to the 3-AQ derivatized oligosaccharide  $\text{Pent}_3\text{Hex}_3$  which, based on the general monosaccharides occurring in AGs structure, could correspond to  $\text{Ara}_3\text{Gal}_3$ . MS/MS analysis of the ion  $m/z$  895.50 showed the same fragmentation profile with a missing pentose at the end of the MS/MS spectrum. Therefore, the corresponding ion was assigned to the derivatized oligosaccharide  $\text{Pent}_2\text{Hex}_3$ .

Combination of both MS/MS results and MS data, evaluated by the developed informatics program, allowed interpretation of the MALDI-TOF profile of the arabinogalactan and identification of the released oligosaccharides. A table with the detected ion masses, the theoretical monoisotopic masses, the error and the assigned composition is reported below (Tab. 5). The considered masses range from 600 to 3000 Da. Below this range the ions were not clearly distinguishable from the matrix peaks and the background since the intensity was usually fairly low. Over 3000 Da no oligosaccharides were detected (linear mode) and, even if present in the digested solution, they may not be detected because of their difficult ionization.

**Tab. 5.** List of the assigned oligosaccharides for digested larch arabinogalactan. <sup>(a)</sup> the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline, H<sup>+</sup>, Na<sup>+</sup> or K<sup>+</sup>).

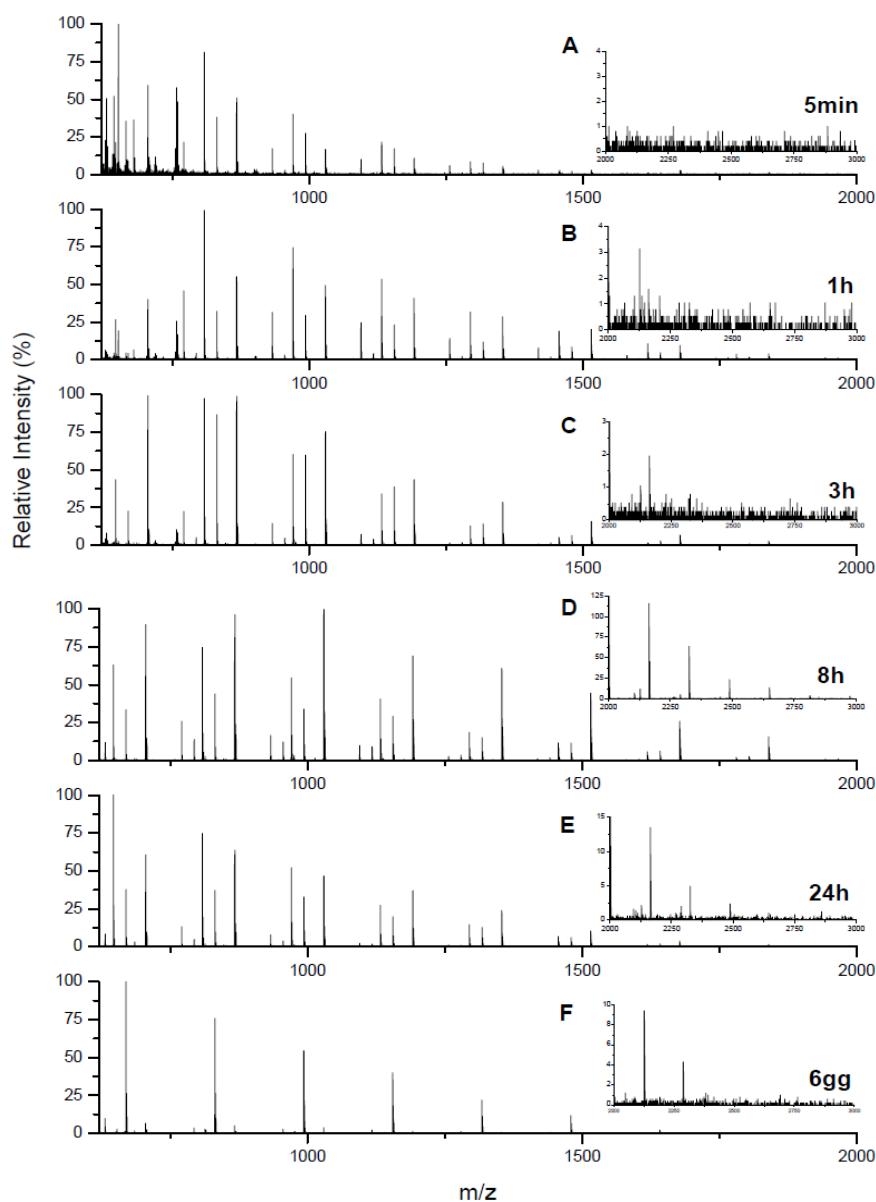
Experimental mass [Da]	Theoretical mass [Da] <sup>a</sup>	$\Delta m$ [Da]	Possible oligosaccharide
631.2867	631.2112	0.08	Hex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
653.2433	653.1932	0.05	Hex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
669.2153	669.1671	0.05	Hex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
763.3100	763.2535	0.06	PentHex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
771.2092	771.1988	0.01	Pent <sub>2</sub> Hex <sub>2</sub> [3AQ/M+K] <sup>+</sup>
785.2571	785.2354	0.02	PentHex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
793.3168	793.2641	0.05	Hex <sub>4</sub> [3AQ/M+H] <sup>+</sup>
801.2291	801.2094	0.02	PentHex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
815.2545	815.2460	0.01	Hex <sub>4</sub> [3AQ/M+Na] <sup>+</sup>
831.2254	831.2199	0.01	Hex <sub>4</sub> [3AQ/M+K] <sup>+</sup>
895.3350	895.2958	0.04	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
903.1975	903.2411	-0.04	Pent <sub>3</sub> Hex <sub>2</sub> [3AQ/M+K] <sup>+</sup>
917.2784	917.2777	0.00	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
925.3130	925.3063	0.01	PentHex <sub>4</sub> [3AQ/M+H] <sup>+</sup>
933.2505	933.2516	0.00	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
947.2711	947.2883	-0.02	PentHex <sub>4</sub> [3AQ/M+Na] <sup>+</sup>
963.2455	963.2622	-0.02	PentHex <sub>4</sub> [3AQ/M+K] <sup>+</sup>
1019.2577	1019.3090	-0.05	Pent <sub>4</sub> Hex <sub>2</sub> [3AQ/M+Na] <sup>+</sup>
1027.3125	1027.3380	-0.03	Pent <sub>3</sub> Hex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
1035.2555	1035.2830	-0.03	Pent <sub>4</sub> Hex <sub>2</sub> [3AQ/M+K] <sup>+</sup>
1049.2767	1049.3200	-0.04	Pent <sub>3</sub> Hex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
1057.3269	1057.3490	-0.02	Pent <sub>2</sub> Hex <sub>4</sub> [3AQ/M+H] <sup>+</sup>
1065.2543	1065.2940	-0.04	Pent <sub>3</sub> Hex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
1079.2898	1079.3310	-0.04	Pent <sub>2</sub> Hex <sub>4</sub> [3AQ/M+Na] <sup>+</sup>
1095.2743	1095.3040	-0.03	Pent <sub>2</sub> Hex <sub>4</sub> [3AQ/M+K] <sup>+</sup>
1159.3687	1159.3800	-0.01	Pent <sub>4</sub> Hex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
1181.3499	1181.3620	-0.01	Pent <sub>4</sub> Hex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
1197.2851	1197.3360	-0.05	Pent <sub>4</sub> Hex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
1211.3355	1211.3730	-0.04	Pent <sub>3</sub> Hex <sub>4</sub> [3AQ/M+Na] <sup>+</sup>
1227.2823	1227.3470	-0.06	Pent <sub>3</sub> Hex <sub>4</sub> [3AQ/M+K] <sup>+</sup>
1257.3164	1257.3570	-0.04	Pent <sub>2</sub> Hex <sub>5</sub> [3AQ/M+K] <sup>+</sup>
1321.4025	1321.4330	-0.03	Pent <sub>4</sub> Hex <sub>4</sub> [3AQ/M+H] <sup>+</sup>
1343.3742	1343.4150	-0.04	Pent <sub>4</sub> Hex <sub>4</sub> [3AQ/M+Na] <sup>+</sup>
1359.3415	1359.3890	-0.05	Pent <sub>4</sub> Hex <sub>4</sub> [3AQ/M+K] <sup>+</sup>

As observed from the summary, all the ions were assigned to a specific oligosaccharide which appeared to be derivatized with 3-aminoquinoline. These data are consistent with the data reported in the literature in which spotting the sample at room temperature resulted in a rate of derivatization reaching almost 100% [50]. In most cases, the protonated ions were followed by the equivalent Na and K cations adducts (e.g. m/z 631.2867 [M+H]<sup>+</sup>; 653.2433 [M+Na]<sup>+</sup> and 669.2153 [M+K]<sup>+</sup>) and the structure was mainly constituted of different combinations of pentose and hexose. From the reported arabinogalactans structure, it could be concluded that Pent corresponded to arabinose while Hex could be ascribed to galactose. The presence of rhamnose and glucuronic acid, which have been observed in type II arabinogalactan structure, could not be excluded. It is possible that the action of exo- $\beta$ -1,3-galactanase is restricted to the external side of the structure. Therefore, even if the enzyme can skip the branched point and release intact side chains, if the two monosaccharide units are linked to the more internal branches, no hydrolysis occurs.

### 3.3. OPTIMIZATION OF CAROB GALACTOMANNAN ENZYMATIC DIGESTION

The second important polysaccharide family that has been discovered in plant gums after type II arabinogalactans is the family of galactomannans. Therefore, while the mass spectrometric conditions were already established working with larch arabinogalactan, a standard carob galactomannan was used to optimize a second enzymatic digestion protocol. According to the supplier the structure of the standard reflects the classical galactomannan structure, so the two main galactomannans degrading enzymes were used: (1)  $\alpha$ -galactosidase to remove the single Gal side chains and (2) endo-1,4- $\beta$ -mannanase to cut the main chain. Digestion was performed using 100 mU of enzyme and the specific conditions of reaction are reported in Tab. 4.

Since the used enzymes are significantly specific for galactomannans, the parameter that was taken into consideration was the digestion timing. After digestion with  $\alpha$ -galactosidase for 24 hours, the sample was incubated with the second enzyme and aliquots were collected after 30 minutes, 1, 3, 8, 24 hours and 6 days. The obtained mass spectra are reported in Fig. 17.



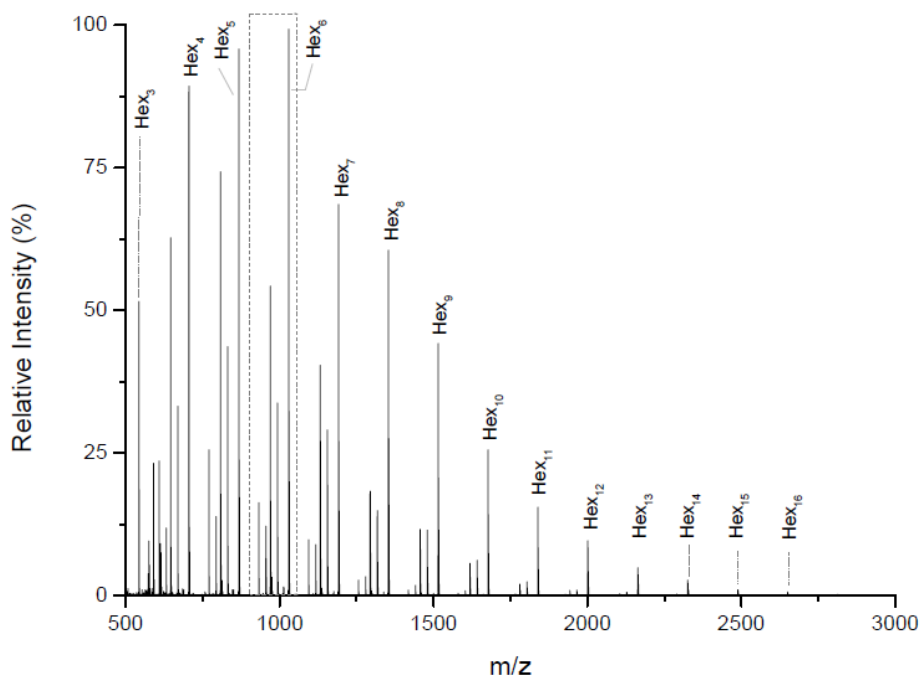
**Fig. 17.** MALDI-TOF spectra of carob galactomannan digested with  $\alpha$ -galactosidase (24 hours) and endo-1,4- $\beta$ -mannanase for: (A) 5minute; (B) 1h; (C) 3h; (D) 8h; (E) 24 hours and (F) 6 days.



Just considering the mass range between 600 and 2000 Da it was possible to observe how, unlike arabinogalactans, for galactomannans the rate of hydrolysis was linked to the digestion time. After only 5 minutes hydrolysis was already began (Fig. 17A) and some oligosaccharides with a molecular weight up to about 1500 Da were released.

Up to 3 hours of digestion, no peaks appeared over  $m/z$  2000. Increasing the digestion time, a greater number of ions were visible in the mass spectrum and mainly in the mass range 2000-3000 Da. After 8 hours digestion (Fig. 17D) oligosaccharide with a maximum molecular weight of 2487 and 2649 Da were released, while after longer time digestion (24 hours and 6 days, Fig. 17E and F), high molecular weight oligosaccharides disappeared or were barely detected because of their low intensity. It is reasonable to speculate that, if the sample is incubated for too long (over 8 hours), the enzymes begin to hydrolyze the already released oligosaccharides, leading to a decrease in the molecular weight. Therefore, even if the galactomannan profile was still recognizable after both a couple of hours and after longer reaction (e.g. 24 hours), a digestion time of 8h was preferred since it allowed to get the maximum number of oligosaccharides.

Once the enzymatic procedure was established, attention was focused on the identification of the released oligosaccharides. Since the polysaccharide is made of mannose and galactose, which are both hexose monosaccharides, all the peaks of the mass spectrum differ of 162.05 Da. Thus, once the structure of an oligosaccharide is known, it is possible to reconstruct the entire profile simply adding or removing one hexose. The mass spectrum of carob galactomannan digested with endo-1,4- $\beta$ -mannanase for 8 hours is showed in Fig. 18.



**Fig. 18.** MALDI-TOF spectrum of carob galactomannan digested with endo-1,4- $\beta$ -mannanase for 8h.

The ion at 1029.43 was attributed to the oligosaccharide Hex<sub>6</sub> [M+K]<sup>+</sup>. Therefore, since the ion at m/z 867.38 differs of 162.05 Da from the ion 1029.43, it corresponded to the oligosaccharide Hex<sub>5</sub> [M+K]<sup>+</sup>, while the further ion at m/z 1191.49 had a +162.06 Da mass so it could be assigned to the structure Hex<sub>7</sub> [M+K]<sup>+</sup>. The ion at 955.43 was attributed to the oligosaccharide Hex<sub>5</sub> [M+H]<sup>+</sup> and the ion at 993.43 to the oligosaccharide Hex<sub>5</sub> [M+Na]<sup>+</sup>. As described before, the same procedure could be used to attribute the other oligosaccharides structures. Some peaks could not be identified and they maybe corresponded to other monosaccharides different from hexoses. Simões *et al.* [56] confirmed the presence of pentose residues and acetyl groups in the main mannose backbone in locust bean gum galactomannan by ESI-MS/MS and GC-MS analysis. Since locust bean gum is obtained by the carob seeds, it is highly probable that the standard carob galactomannan structure includes this monosaccharide and group, but by know the ions have not be indentified yet. However, most of the oligosaccharides were attributed and a characteristic profile was defined.

In conclusion, the experiments run on the two standard polysaccharides allowed to:

- establish the best mass spectrometric conditions for analysis of oligosaccharides by MALDI-TOF;
- optimize the enzymatic digestion protocol for type II arabinogalactan polysaccharides in terms of type of enzymes used, amount of enzyme and timing;
- identify an oligosaccharide mass profile of larch arabinogalactan;
- establish the best digestion time for galactomannan hydrolysis;
- identify an oligosaccharide mass profile of carob galactomannan.

## 4. APPLICATION AND METHODOLOGICAL DEVELOPMENT ON PLANT GUMS SAMPLES

The focus of this section is to evaluate the best enzymatic digestion procedure that will determine the mass fingerprint for each plant gum under investigation. First the general preparation of gum samples (4.1), followed by the description of the enzymes used for each gum and the results obtained after MALDI-TOF analysis of the released oligosaccharides will be presented. The three plant gums that resemble the arabinogalactan polysaccharide structure will be described separately: gum arabic (4.2), gum tragacanth (4.3) and cherry gum (4.4). Protocol optimization for locust bean gum and guar gum analysis, which are both galactomannans, will be described in section 4.5. The oligosaccharide profile of each gum will be shown. The results suggest that the proposed analytical methodology allow to discriminate the gums by comparison of their MALDI-TOF profiles.

### 4.1. PLANT GUM SAMPLES PREPARATION

Preparation of plant gum samples is a crucial point since only gum arabic is completely soluble in water. For all the other gums, if concentration is higher than 5% w/v, the saturation point is reached and a gel-like consistence is obtained, thus preventing any further manipulation. In general all different gums were solubilized in milliQ water and the soluble part was then recovered after centrifugation at 13400 rpm for 30 minutes. The specific amount of gum and preparation procedure are reported in Tab. 6.

**Tab. 6.** Summary of amount and solubilization protocol for each plant gums.

Gum	% [w/v]	Preparation
Arabic	3%	The solution is stirred on a hot-plate magnetic stirrer at 80°C until the solution is clear (approximately 15 minutes).
Tragacanth	0.5%	Gum powder is wet with 95% EtOH (1 µL/mg of gum) before adding milliQ water. Solution is mixed o/n at room temperature.
Fruit gums	2%	Gum is solubilized in milliQ water and mixed o/n at room temperature [57].
LBG and guar	1%	Gum powder is wet with 95% MeOH (1 µL/mg of gum) before adding milliQ water. The solution is gently stirred for about 1 hour and then left over night at 4°C. The solution is then stirred on a hot-plate magnetic stirrer at 120°C for about 20 minutes [58].
Ghatti	1%	Gum is solubilized in milliQ water and mixed o/n at room temperature.
Karaya	1%	Gum is solubilized in milliQ water and mixed o/n at room temperature.

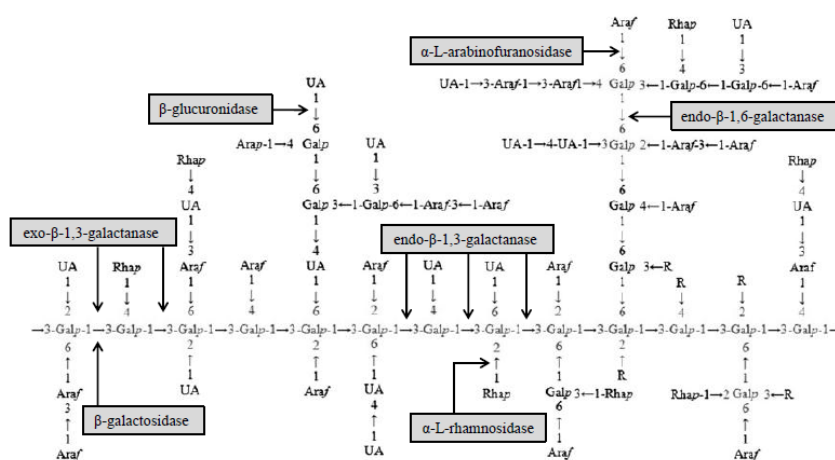
Enzymatic digestion was always performed on around 1 mg of gum so, according to the gum concentration, a proper solution aliquot was dried in a vacuum centrifuge and stocked at -20°C before enzymatic digestion.

## 4.2. GUM ARABIC

As reported previously, variations in the arabinogalactan structure in terms of monosaccharide substituents, molecular weight, side chains length and conformation are usually encountered even among the same plant species (e.g. larch [59]). It is also well known that type II AGs found in plant gums are characterized by a complex branched structure which is only presumed and not well defined. For this reason, even if the experiments on the standard larch arabinogalactan allowed optimization of the digestion procedure, (e.g. necessity of derivatization with 3-aminoquinoline), it was necessary to perform analogous tests on gum arabic samples in order to highlight its unknown arabinogalactan component.

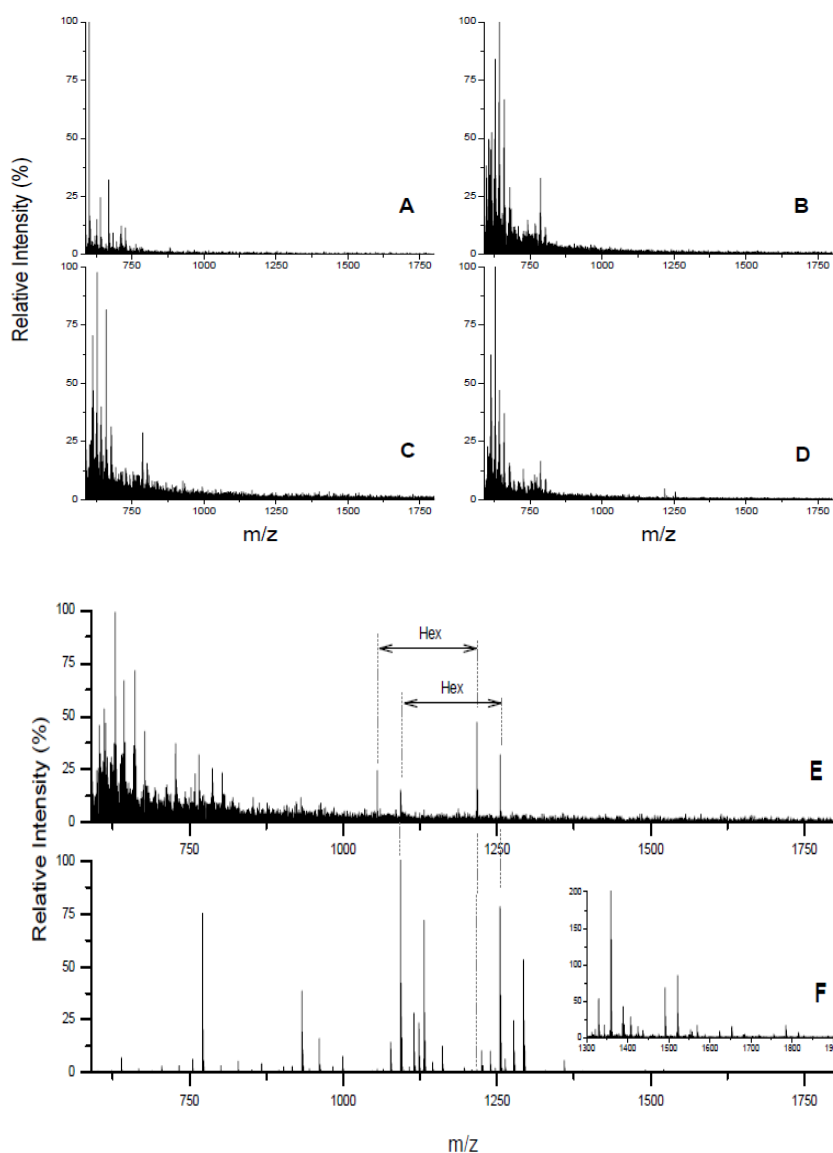
### 4.2.1. Evaluation of enzymatic digestion protocol

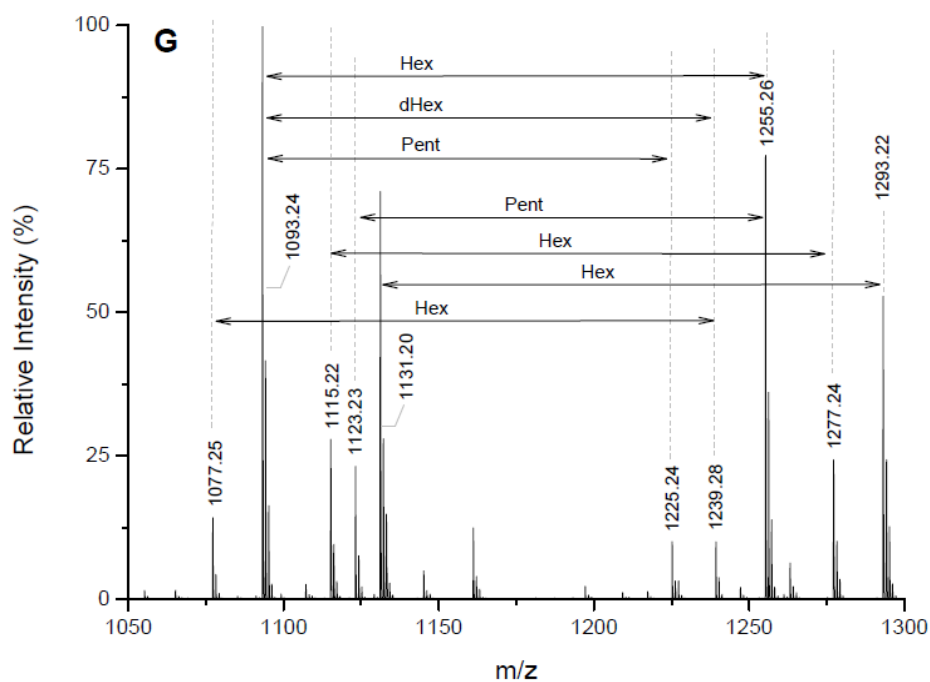
As described in chapter I, gum arabic has a complex branched structure with several and different side chains attached to the Gal residues of the main backbone [60]. For this reason the enzymatic degradation of gum arabic requires a certain number of enzymes whose combined action will hydrolyze all the different glycosidic linkages. However several problems are encountered when dealing with gum arabic degradation. First of all the structure of the polysaccharide component of gum arabic is not well known and only potential structures have been suggested in the literature. Consequently, besides a recent patent where  $\beta$ -galactosidase (EC 3.2.1.23) was successfully used to modify gum arabic structure in order to increase its emulsifying properties [61], few other information are available in the literature on enzymatic treatment of gum arabic polysaccharide component. In addition, concerning the enzymes, few information are available on the substrate specificity of the enzymes. Furthermore, the enzymes available in the market are exo-type and, even if prepared in the laboratory, only one endo-type enzyme that can degrade the 1,3-Gal main chain (endo- $\beta$ -1,3-GAL) has been characterized [62]. Taking into account all these considerations, enzymes were selected according to their theoretical action (see Tab. 3) and considering as substrate a potential gum arabic structure proposed by Nie *et al.* after analysis of the plant gum by GC-MS and 1D, 2D NMR [60]. A representation of enzymes action on the polysaccharide component of gum arabic is reported in Fig. 19.



**Fig. 19.** Illustration of various enzymes possibly involved in the degradation of gum arabic polysaccharide component. Galp galactopyranose, AraF arabinofuranose, Arap arabinopyranose, Rhap rhamnopyranose, UA uronic acids, R can be a possible terminuteal Rha, AraF, Arap or UA with different linkages.

Unfortunately the endo-type enzymes (e.g. endo- $\beta$ -1,3-GAL and endo- $\beta$ -1,6-GAL) were not commercially available so gum arabic structure could only be decomplexified from the significantly external side chains. As for the standard larch arabinogalactan, the following enzymes have been used in the listed order: (1)  $\alpha$ -L-arabinofuranosidase; (2)  $\beta$ -glucuronidase; (3)  $\alpha$ -L-rhamnosidase; (4)  $\beta$ -galactosidase; (5) exo- $\beta$ -1,3-galactanase. 1 mg of gum arabic (Zecchi, Italy) was initially digested with  $\alpha$ -L-arabinofuranosidase for 24 hours, reaction was quenched by boiling the solution and an aliquot was collected and analyzed by MALDI-TOF. The left sample was dried, re-solubilized in the proper buffer and incubated for 24 hours with the second enzyme ( $\beta$ -glucuronidase). The same procedure was carried out for all the further enzymes ( $\alpha$ -L-rhamnosidase,  $\beta$ -galactosidase) until exo- $\beta$ -1,3-galactanase that was incubated for 3, 5 and 93 hours. The corresponding mass spectra are showed in the figure below (Fig. 20).



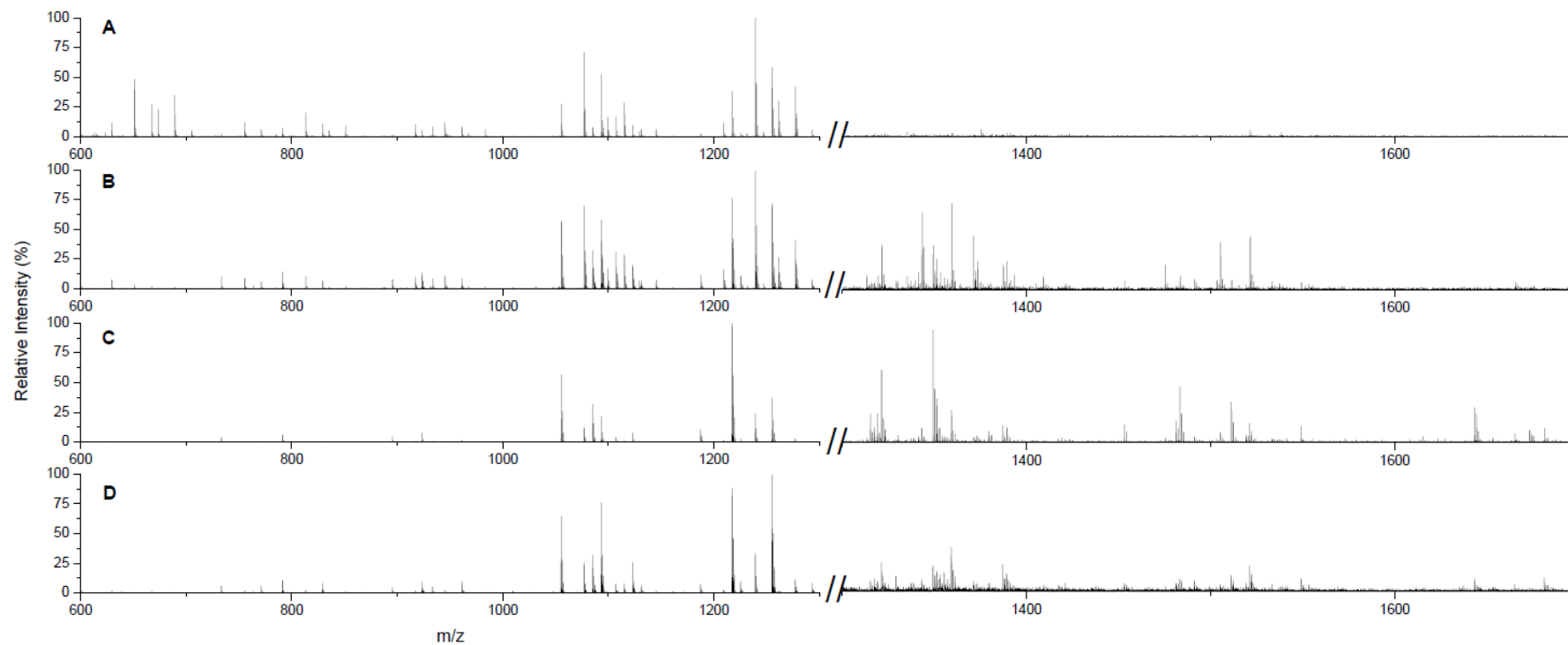


**Fig. 20.** Mass spectra (range 600-1800 Da) of gum arabic sample: (A) non digested (diluted 1/10); (B) digested with  $\alpha$ -L-arabinofuranosidase; (C) after adding  $\beta$ -glucuronidase; (D) after adding  $\alpha$ -L-rhamnosidase; (E) after adding  $\beta$ -galactosidase (93h); (F) after adding exo- $\beta$ -1,3-galactanase; (F corner) zoom 1300-1900 Da; (G) enlarged image (1050-1300Da) of Fig. 20F.

When gum arabic was spotted before enzymatic digestion, only matrix ions were observed at low molecular weights (50 - 600 Da) while at higher masses (600 - 1800 Da) only few ions at low intensity were detected (Fig. 20A). Any mass difference corresponding to monosaccharide masses were observed. Same results were obtained after analysis of the gum sample digested with the further three enzymes (Fig. 20B, C and D). After the addition of  $\beta$ -galactosidase, few oligosaccharides characterized by a mass differences of 162.05 Da appeared in the mass spectrum (Fig. 20E) indicating a certain activity of the enzyme on the gum polysaccharide component. However, the most interesting results were obtained after the addition of the last enzyme exo- $\beta$ -1,3-galactanase. As highlighted in Fig. 20F, several ions with high intensity were observed in the mass range included between 600 and 1400 Da, and other at lower intensity between 1400 and 1900 Da (image at the right corner of Fig. 20F). Some of the detected peaks can be observed in the zoom of the MALDI-TOF spectrum of gum arabic digested with five enzymes reported in Fig 20G. The mass spectrum showed a certain heterogeneity due to the presence of different carbohydrate forms. Multiple oligosaccharide were present in the sample solution and this was reflected in the mass spectrum. A mass difference of 132.02 Da between the ion at m/z 1093.24 and 1225.26 was identified, suggesting the loss of a pentose and, more likely, an arabinose according to the hypothetical structure of gum arabic polysaccharide moiety (arabinogalactan). The same mass difference could be observed between the ions at 1123.23 and 1255.26 Da. The mass spectrum showed also the presence of a mass difference of 162.05 Da, which corresponds to an hexose (Hex), and specifically galactose. The enzymatic cleavage of a Gal was observed between the ions 1077.25 and 1239.28, 1115.22 and 1277.24, and 1131.20 and 1239.28. Furthermore, the presence of a deoxyhexose (dHex) in the structure, probably rhamnose, was observed between the ions 1093.24 - 1239.22. Therefore, it could be concluded that the ion at 1093.24 m/z could be a combination of several oligosaccharides, confirming the complexity of gum arabic structure. The

peaks observed after incubation of the sample with *exo*- $\beta$ -1,3-galactanase for 93 hours (Fig. 20F), were also obtained after a shorter incubation time (3 and 5 hours). Oligosaccharides with a molecular weight of 1521.32, 1623.35, 1785.37 and 1815.39 Da were missing but, even after 93 hours digestion, their signal/noise was very low and they were barely distinguishable from the background. Results of this experiment suggested that, even if any oligosaccharide seemed to be released after using the four enzymes ( $\alpha$ -L-arabinofuranosidase,  $\beta$ -glucuronidase,  $\alpha$ -L-rhamnosidase and  $\beta$ -galactosidase), these enzymes are necessary to remove the external monosaccharides units to facilitate the access of *exo*- $\beta$ -1,3-galactanase. Even if only *exo*-type enzymes were used, it was possible to obtain some oligosaccharides that were apparently made up of mainly pentose, hexose and deoxyhexose monosaccharides. The enzymatic digestion protocol was the same applied to larch arabinogalactan digestion, but oligosaccharides with different masses were obtained, thus revealing how the arabinogalactan component of gum arabic has a different structure at least in terms of side chains composition.

Besides the interesting results obtained using 5 enzymes, the digestion time was considered too long since the complete sample preparation lasted 1 week. Furthermore, multi steps such as inactivation of the enzyme, dryness and addition of subsequent enzymes, increased the possibility of sample loss. Therefore, a less number of enzymes was tested and combined in order to shorten the digestion time. As already tested on the standard larch arabinogalactan, among the different experiments a test was performed incubating the sample only with *exo*- $\beta$ -1,3-galactanase for different time (5 minutes, 30 minutes, 1, 3, 5, 8, 24, 96, 120 hours, 1 week, 2 weeks, 3 weeks). The most significant results are reported in Fig. 21.



**Fig. 21.** MALDI-TOF spectra (600-1300 Da) of gum arabic digested with exo-β-1,3-galactanase for (A) 5 minutes; (B) 5 h; (C) 96 h and (D) 2 weeks. On the right a zoom of the mass range 1300-1700 Da is showed



In all the four mass spectra it was possible to observe, in the mass range between 1000 and 1300 Da, the same ions obtained using five enzymes. Digestion already began after 5 minutes incubation (Fig. 21A) but no oligosaccharides with a molecular weight over 1300 Da were observed. After 3, and mainly 5 hours, higher molecular weight oligosaccharides were released, as showed on the mass spectrum on the right of Fig. 21B, while the highest masses were obtained after 96 hours incubation (e.g. figure 21C;  $m/z$  1349.35; 1483.33; 1643.36). However, a longer incubation time, e.g. the mass spectrum after two weeks incubation showed in Fig 21D, did not allow to obtain any more information. As additional experiment, since the activity of the enzymes decreases with the time, more enzyme was added every 1-3 days but no differences were observed.

Further experiments of digesting the gum with  $\alpha$ -L-arabinofuranosidase and then  $\text{exo-}\beta$ -1,3-galactanase were run since, according to theory, the  $\text{exo-GAL}$  action should be impeded by the presence of side chains [26]. No differences were observed between the mass spectra of the gum digested with the two enzymes ( $\alpha$ -L-arabinofuranosidase +  $\text{exo-}\beta$ -1,3-galactanase) and only  $\text{exo-}\beta$ -1,3-galactanase. This may suggests that the detected oligosaccharides come from the side chains and that the presence of Ara units do not impede the action of  $\text{exo-}\beta$ -1,3-GAL.

In conclusion, as discovered previously with larch arabinogalactan,  $\text{exo-}\beta$ -1,3-galactanase is able to partially hydrolyze the polysaccharide without any previous removal of external side chains. This result demonstrates how the enzyme is able to digest galactose chains even on a complex substrate such as gum arabic. In terms of digestion timing, since long time incubation did not lead to a significant increase in the number of released oligosaccharides, an average digestion time of 5 hours was considered optimal.

At the light of these considerations a final digestion protocol of gum arabic was established: 1 mg of polysaccharides is resuspended in a phosphate buffer 50 mM pH 6 and incubated with 100 mU of  $\text{exo-}\beta$ -1,3-galactanase for 5 hours at 50°C.

#### 4.2.2. Identification of the released oligosaccharides

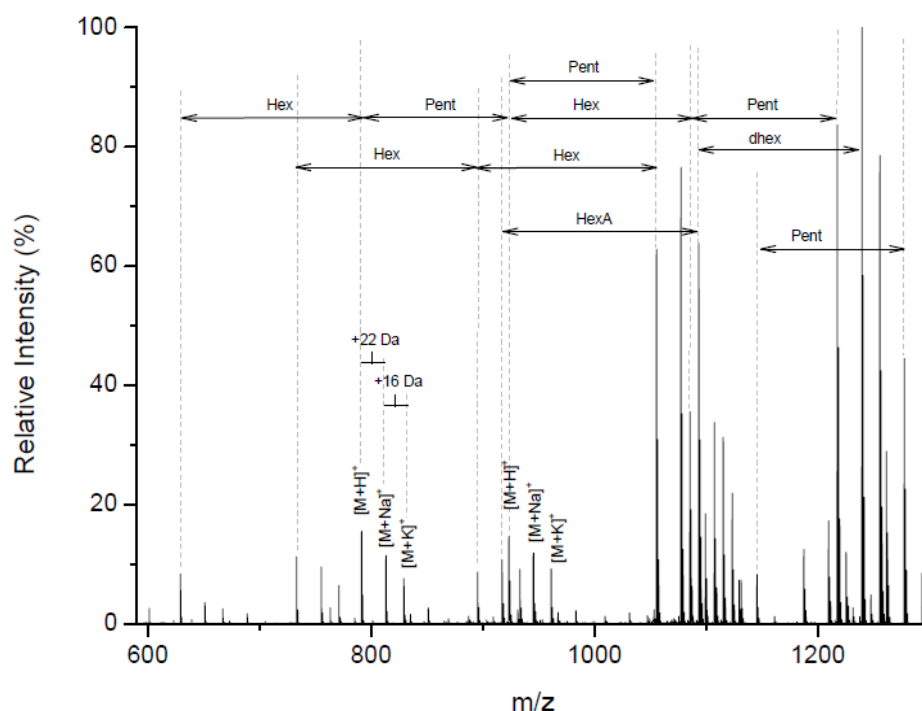
Once the digestion protocol for gum arabic was optimized, the further step was the identification of the released oligosaccharides in order to obtain a characteristic mass fingerprint of the gum polysaccharide component. To achieve this goal, different but complementary approaches were used:

1. The MS spectra were studied. This was performed by identifying mass differences between the peaks that corresponded to specific monosaccharides (e.g. 132.04 Da for pentose, 162.05 Da for hexose etc.).
2. The gum digest obtained after incubation for 5 hours with  $\text{exo-}\beta$ -1,3-galactanase was subsequently treated with other enzymes (e.g.  $\alpha$ -L-arabinofuranosidase,  $\beta$ -glucuronidase etc.). In this way the absence in the mass spectrum of some oligosaccharides after the addition of each second enzyme, would indicate the hydrolysis of the corresponding monosaccharide (e.g. *Araf* in the case of  $\alpha$ -L-arabinofuranosidase) and therefore information about the monosaccharide composition of the oligosaccharide are obtained.

- Tandem mass spectrometric experiments (MS/MS) were run in order to have a precise knowledge of the oligosaccharide composition. Furthermore, information on the monosaccharide sequence of the oligosaccharide could be obtained.

#### EVALUATION OF THE MS PROFILE

The mass spectrum of gum arabic digested for 5 hours with 100 mU of  $\alpha$ -D-galactanase (Fig. 22) resulted to be highly heterogeneous due to the presence of different oligosaccharides forms. However, the MS spectrum allowed to obtain some preliminary information about the released oligosaccharides. As already observed in the mass spectrum obtained after digestion with 5 enzymes (Fig. 20G), some of the peaks resulted to be linked. Mass differences between the ions, corresponding to the residue mass of some monosaccharides (e.g. 132.04 Da for pentose, 162.05 Da for hexose, 146.06 Da for deoxyhexose) were observed. Furthermore, the presence of ions differing in mass of 22 Da and 38 Da all along the mass range, revealed the presence in the mass spectrum of sodium  $[M+Na]^+$  and potassium  $[M+K]^+$  adduct ions. This information allowed to make the mass spectrum interpretation more simple and straightforward.

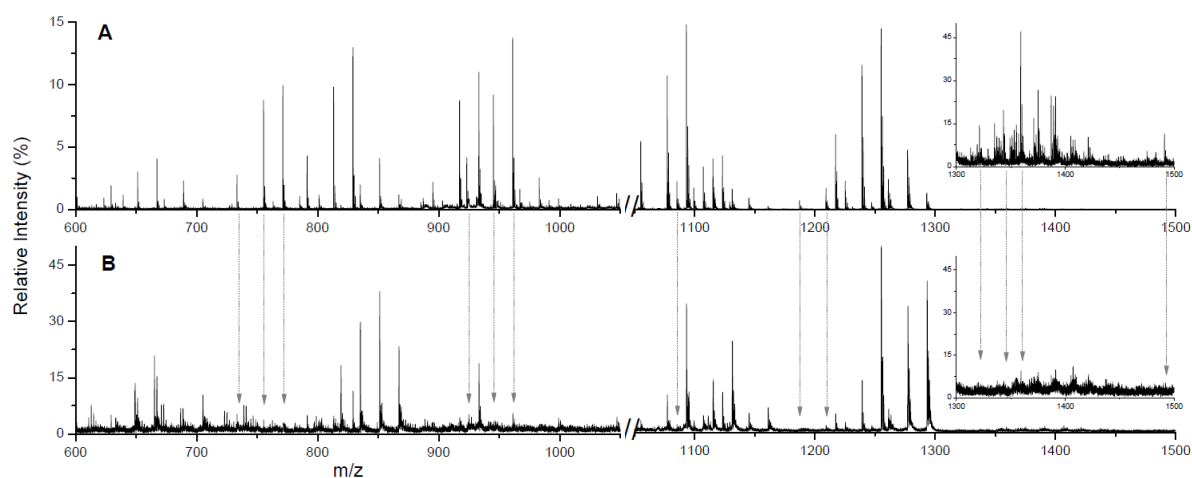


**Fig. 22.** MS spectrum (range 600-2300 Da) of gum arabic sample digested for 5 hours with  $\alpha$ -D-galactanase. Significant mass differences between peaks are reported.

#### EVALUATION OF ENZYME ADDITION

With regard to the second point, the following enzymes were added after digestion with  $\alpha$ -D-galactanase: (a)  $\alpha$ -L-arabinofuranosidase, (b)  $\beta$ -glucuronidase, (c)  $\alpha$ -L-rhamnosidase and (d)  $\beta$ -galactosidase. While with the last three enzymes no differences in the mass spectra were observed, the MALDI-TOF spectrum indicated the absence of some oligosaccharides after the addition of  $\alpha$ -L-arabinofuranosidase. The spectra of the gum digested only with  $\alpha$ -D-galactanase and after

the addition of  $\alpha$ -L-arabinofuranosidase are compared and reported respectively in the figure 23A and 23B. The MALDI-TOF spectra showed the absence of some oligosaccharides after hydrolysis with  $\alpha$ -L-arabinofuranosidase (Fig. 23B) such as  $m/z$  733.21, 755.23 and 771.21 ions. The mass difference of 22 and 16 Da among them suggested that the ions corresponded to the same oligosaccharide protonated (733.21 Da) and with adducts of different cations: Na (755.23) and K (771.21 Da). The interrogation of mass 733.21 using the informatics tool gave back  $[3AQ/Pent_2Hex_2+H]^+$  as a possible composition. The presence of Ara in the oligosaccharide structure was confirmed by the disappearance of the corresponding ion in the spectrum obtained after  $\alpha$ -L-arabinofuranosidase digestion. Similar considerations could be made for the ions at  $m/z$  1359.24 and 1491.28 which were not present anymore after adding the second enzyme. The mass difference between them consisted of 132.04 Da, so  $\alpha$ -L-arabinofuranosidase was able to hydrolyze one pentose from the chain. Regarding the possible composition of the ion at 1359.24 Da, it could correspond to the ion at  $m/z$  1321.25  $[M+H]^+$  but with a potassium adduct  $[M+K]^+$ . In regard to the mass at 1359.24, the informatics tool showed how multiple oligosaccharides could be assigned to the same ion ( $[3AQ/Hex_4HexA_3+K]^+$ ,  $[3AQ/HexdHex_6HexA+H]^+$ ,  $[3AQ/HexdHex_5HexA_2+H]^+$  and  $[3AQ/Pent_4Hex_4+K]^+$ ), thus showing how several oligosaccharides could be hidden under the same mass. However, the first three suggested structures could not be accepted since the presence of arabinose was verified by digestion experiments. Therefore, the final composition of the  $m/z$  1359.24 ion was identified as  $[3AQ/Pent_4Hex_4+K]^+$ . Another ion that was not present anymore after digestion with  $\alpha$ -L-arabinofuranosidase was at  $m/z$  1085.24. Three possible oligosaccharides were indicated by the informatics tool ( $[3AQ/Hex_4dHex_2+H]^+$ ,  $[3AQ/PentHex_3dHexHexA+H]^+$  and  $[3AQ/Pent_2Hex_2HexA_2+H]^+$ ). In accordance with the results obtained after digestion with  $\alpha$ -L-arabinofuranosidase, the attribution  $[3AQ/Hex_4dHex_2+H]^+$  could be excluded since no pentose was presumed in the sequence. However, the other two including pentose ( $[3AQ/PentHex_3dHexHexA+H]^+$  and  $[3AQ/Pent_2Hex_2HexA_2+H]^+$ ) were still probable. Similar observation could be made for the ions at  $m/z$  923.19  $[M+H]^+$ , 945.19  $[M+Na]^+$  and 961.16  $[M+K]^+$ . For example, if considering the ion at 923.19 Da, the attributed composition without any arabinose  $[3AQ/Hex_3dHex_2+H]^+$  could be excluded while the other ones ( $[3AQ/PentHex_2dHexHexA+H]^+$  and  $[3AQ/Pent_2HexHexA_2+H]^+$ ) could be considered possible.

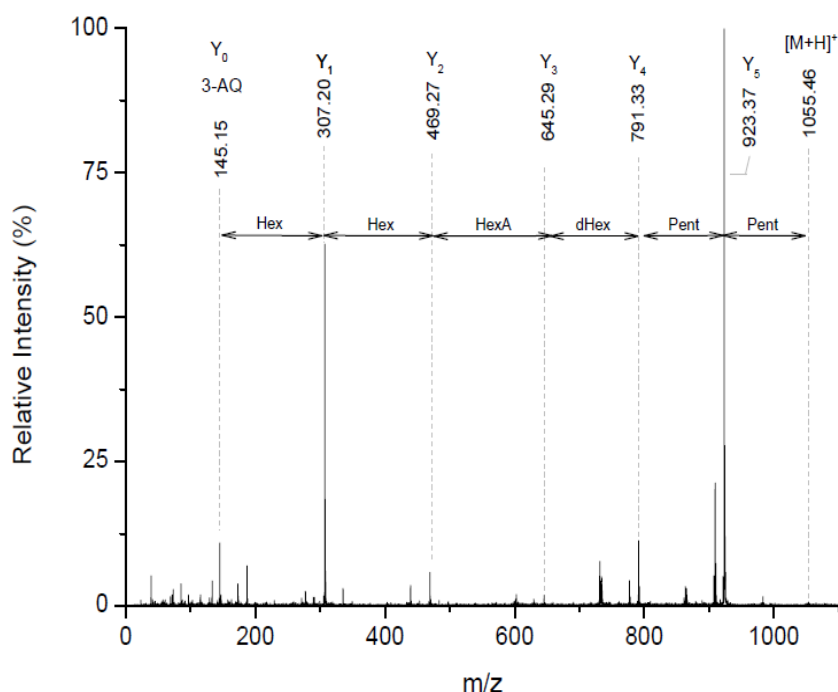


**Fig. 23.** MS spectra of the gum arabic digested with: (A) exo- $\beta$ -1,3-galactanase and (B) exo- $\beta$ -1,3-galactanase and  $\alpha$ -L-arabinofuranosidase. On the right corner the corresponding zoom of mass range 1300-1500 Da.

These results showed how the deepened investigation of MS spectra, the employment of the informatics tool and the second digestion by  $\alpha$ -L-arabinofuranosidase, provided some preliminary information about the released oligosaccharide compositions. However, as showed for the ion at  $m/z$  1085.57, several oligosaccharides could correspond to the same  $m/z$  value and the previously reported tools resulted to be not sufficient to resolve the correct oligosaccharide composition. Therefore, tandem mass spectrometry was considered the most appropriate method to further confirm the oligosaccharide attribution.

#### TANDEM MASS SPECTROMETRY

Fragmentation was performed in CID mode (Collision-Induced Dissociation) in the presence of air. As an example of the MSMS potentiality, the case of the ion at  $m/z$  1055.46 will be reported. After interrogation of the informatics tool, three possible oligosaccharides were attributed at the same mass: (1)  $[3AQ/PentHex_3dHex_2+H]^+$ , (2)  $[3AQ/Pent_2Hex_2dHexHexA+H]^+$  and (3)  $[3AQ/Pent_3HexHexA_2+H]^+$ . A second digestion with  $\alpha$ -L-arabinofuranosidase, following the optimized digestion with  $exo$ - $\beta$ -1,3-galactanase, did not allow to discriminate among the three attributions. Therefore MSMS analysis was performed and the mass spectrum is reported in Fig. 24.



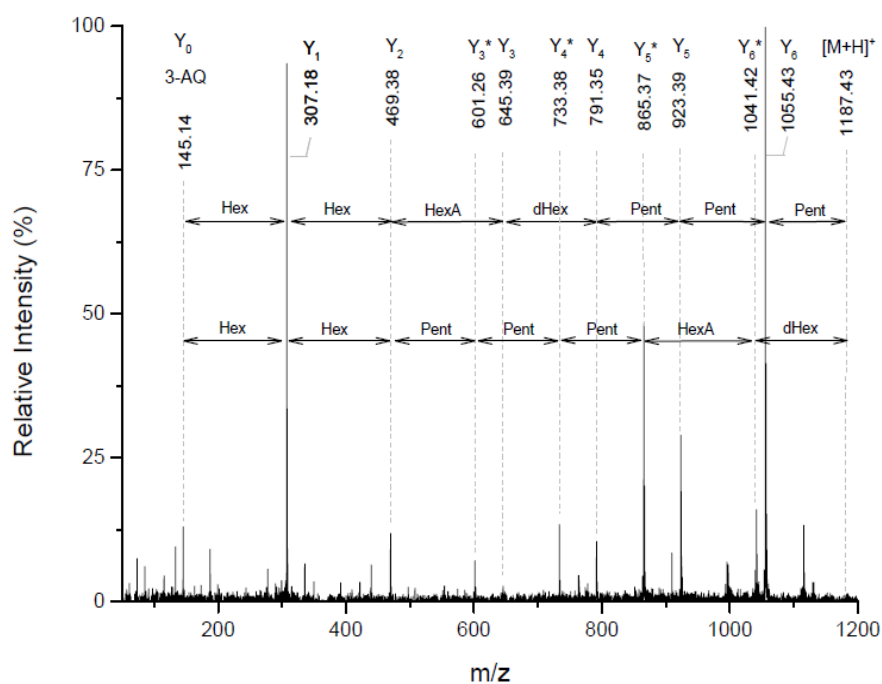
**Fig. 24.** MALDI-CID spectrum of 3-AQ derivatized oligosaccharide released by  $exo$ - $\beta$ -1,3-galactanase ( $m/z$  1055.46).

The MS/MS pattern looked simple with the presence of only Y-ions and any ion coming from cross-ring fragmentation was observed. This is due to the proton affinity of quinoline that implies the charge to be retained by the fragments including the derivatized reducing end. The MSMS spectrum showed exactly the monosaccharide sequence of the carbohydrate which corresponded to two hexoses, one hexuronic acid, one deoxyhexose and two pentoses. Therefore the analysis clearly

allowed to attribute the ion at  $m/z$  1055.46 to the second possible oligosaccharide [3AQ/Pent<sub>2</sub>Hex<sub>2</sub>dHexHexA+H]<sup>+</sup>.

Another example of the complementarity and necessity of the three different approaches for the identification of the released oligosaccharides, consisted in the ion at  $m/z$  1085.57. As discussed previously, the informatics program assigned three possible profiles to the ion: (1) [3AQ/Hex<sub>4</sub>dHex<sub>2</sub>+H]<sup>+</sup>, (2) [3AQ/PentHex<sub>3</sub>dHexHexA+H]<sup>+</sup> and (3) [3AQ/Pent<sub>2</sub>Hex<sub>2</sub>HexA<sub>2</sub>+H]<sup>+</sup>. While the first was excluded because of the results obtained after digestion of the gum by *exo*- $\beta$ -1,3-galactanase and then  $\alpha$ -L-arabinofuranosidase; the other two structures were still ascribable. In this case MS/MS analysis allowed to understand the real composition of the oligosaccharide under investigation. The presence of only one pentose, as supposed by digestion with  $\alpha$ -L-arabinofuranosidase, was demonstrated by the mass difference of 132.04 Da in the MS/MS spectrum between the precursor and the ion at  $m/z$  953.55. So the third hypothetical sequence which predicted the presence of two pentoses could not be accepted and the final assigned oligosaccharide was [3AQ/PentHex<sub>3</sub>dHexHexA+H]<sup>+</sup>.

The MSMS spectrum of the ion 1055.46 reported in Fig. 24, appeared of straightforward interpretation. However for other ions multiple peaks were observed. For example, the oligosaccharide at  $m/z$  1187.43 could be attributed, from the informatics tool, to four possible structures: (1) [3AQ/HexHexA<sub>5</sub>+H]<sup>+</sup>, (2) [3AQ/Pent<sub>2</sub>Hex<sub>3</sub>dHex<sub>2</sub>+H]<sup>+</sup>, (3) [3AQ/Pent<sub>3</sub>Hex<sub>2</sub>dHexHexA+H]<sup>+</sup> and (4) [3AQ/Pent<sub>4</sub>HexHexA<sub>2</sub>+H]<sup>+</sup>. Since the ion disappeared after digestion with  $\alpha$ -L-arabinofuranosidase, the oligosaccharide was supposed to contain at least one pentose, so the first attribution was excluded. However the other ones could not be discriminated but for MSMS analysis. The MS/MS spectrum of the precursor ion 1187.43 is reported in Fig. 25.



**Fig. 25.** MALDI-CID spectrum of 3-AQ derivatized oligosaccharide released by *exo*- $\beta$ -1,3-galactanase ( $m/z$  1187.43).

In comparison to the MSMS profile of the precursor ion at 1055.46 (Fig. 24), the fragmentation pattern of the ion at  $m/z$  1187.43 appeared more complicated. The ion with a  $m/z$  of 145.14 Da corresponded to 3-aminoquinoline  $[M+H]^+$  and confirmed the oligosaccharide to be derivatized. The product ion at 1055.43 Da was formed by the loss of a pentose residue (132.00 Da). Same loss could be observed for the following  $m/z$  923.39 and 791.35 ions (132.04 Da), therefore three pentoses were present in the oligosaccharide structure. The loss of a dHex (146 Da) was then observed between the two ions at 791.35 and 645.39 Da, and a further loss of HexA was detected between  $m/z$  645.39 and 469.38 ions. Two hexoses then completed the monosaccharide sequence. Therefore the ion 1187.43 was assigned to the 3-AQ derivatized oligosaccharide  $\text{Pent}_3\text{Hex}_2\text{dHexHexA}$  which, based on the general monosaccharides occurring in AGs structure, may corresponds to  $\text{Ara}_3\text{Gal}_3\text{RhaGluA}$ . However, two monosaccharide sequences are observed in the MS/MS spectra which lead to the same final oligosaccharide composition: (1)  $\text{HexHexHexAdHexPentPentPent}$  and (2)  $\text{HexHexPentPentPentHexAdHex}$  (indicated with an (\*) in the mass spectrum). Therefore, the MSMS analysis allowed to assign the correct oligosaccharide to the ion and it demonstrated the capability of even distinguishing the specific monosaccharide sequence.

In conclusion, combining all the information obtained by MS analysis and mass differences between the peaks, interrogation of the informatics tool, digestion with other enzymes following the optimized digestion with  $\text{exo-}\beta\text{-1,3-galactanase}$  and tandem mass spectrometry analysis, a potential MALDI-TOF profile of enzymatically digested gum arabic could be retraced. A table with the detected ion masses, the theoretical monoisotopic masses, the error and the assigned composition is reported below (Tab. 7). The mass range that is considered significant for the interpretation of the gum fingerprint goes from 600 to 3000 Da. Below this range the ions were not clearly distinguishable from the matrix peaks and the background since the intensity was usually fairly low. Over 3000 Da any oligosaccharide was detected (linear mode), and even if they were released after digestion, their ionization and therefore detection was highly difficult.

**Tab. 7.** List of the assigned oligosaccharides for digested gum arabic (Zecchi, Italy). <sup>(a)</sup> the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline,  $\text{H}^+$ ,  $\text{Na}^+$  or  $\text{K}^+$ ).

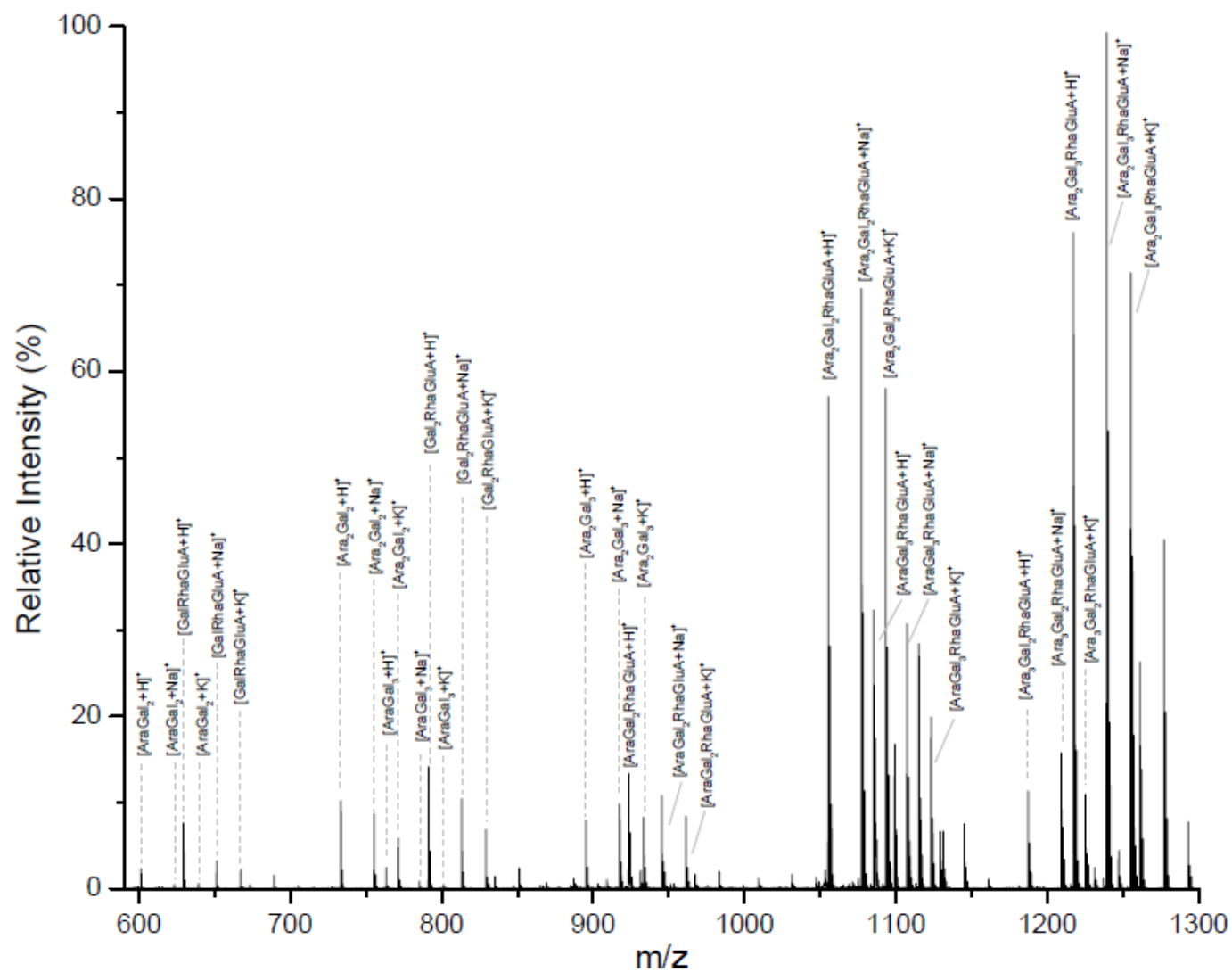
Experimental mass [Da]	Theoretical mass [Da] <sup>a</sup>	$\Delta m$ [Da]	Possible oligosaccharide
601.3003	601.2007	0.10	$\text{PentHex}_2 [3\text{AQ}/\text{M}+\text{H}]^+$
623.2650	623.1826	0.08	$\text{PentHex}_2 [3\text{AQ}/\text{M}+\text{Na}]^+$
629.2650	629.1956	0.07	$\text{HexdHexHexA} [3\text{AQ}/\text{M}+\text{H}]^+$
639.2323	639.1566	0.08	$\text{PentHex}_2 [3\text{AQ}/\text{M}+\text{K}]^+$
651.2484	651.1775	0.07	$\text{HexdHexHexA} [3\text{AQ}/\text{M}+\text{Na}]^+$
667.2224	667.1515	0.07	$\text{HexdHexHexA} [3\text{AQ}/\text{M}+\text{K}]^+$
733.3287	733.2429	0.09	$\text{Pent}_2\text{Hex}_2 [3\text{AQ}/\text{M}+\text{H}]^+$
755.2962	755.2249	0.07	$\text{Pent}_2\text{Hex}_2 [3\text{AQ}/\text{M}+\text{Na}]^+$
763.3342	763.2535	0.08	$\text{PentHex}_3 [3\text{AQ}/\text{M}+\text{H}]^+$
771.2818	771.1988	0.08	$\text{Pent}_2\text{Hex}_2 [3\text{AQ}/\text{M}+\text{K}]^+$
785.3080	785.2354	0.07	$\text{PentHex}_3 [3\text{AQ}/\text{M}+\text{Na}]^+$
791.3312	791.2484	0.08	$\text{Hex}_2\text{dHexHexA} [3\text{AQ}/\text{M}+\text{H}]^+$
801.2723	801.2094	0.06	$\text{PentHex}_3 [3\text{AQ}/\text{M}+\text{K}]^+$
813.2897	813.2304	0.06	$\text{Hex}_2\text{dHexHexA} [3\text{AQ}/\text{M}+\text{Na}]^+$
829.2632	829.2043	0.06	$\text{Hex}_2\text{dHexHexA} [3\text{AQ}/\text{M}+\text{K}]^+$
895.3676	895.2958	0.07	$\text{Pent}_2\text{Hex}_3 [3\text{AQ}/\text{M}+\text{H}]^+$
917.3582	917.2777	0.08	$\text{Pent}_2\text{Hex}_3 [3\text{AQ}/\text{M}+\text{Na}]^+$

923.3597	923.2907	0.07	PentHex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
933.3234	933.2516	0.07	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
945.3242	945.2726	0.05	PentHex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
961.3057	961.2465	0.06	PentHex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1055.3938	1055.3330	0.06	Pent <sub>2</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1077.3632	1077.3150	0.05	Pent <sub>2</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1085.4028	1085.3430	0.06	PentHex <sub>3</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1093.3425	1093.2890	0.05	Pent <sub>2</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1107.3707	1107.3250	0.05	PentHex <sub>3</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1123.3480	1123.2990	0.05	PentHex <sub>3</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1187.4246	1187.3750	0.05	Pent <sub>3</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1209.4080	1209.3570	0.05	Pent <sub>3</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1217.4409	1217.3860	0.05	Pent <sub>2</sub> Hex <sub>3</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1225.3827	1225.3310	0.05	Pent <sub>3</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1239.4120	1239.3680	0.04	Pent <sub>2</sub> Hex <sub>3</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1255.3865	1255.3420	0.04	Pent <sub>2</sub> Hex <sub>3</sub> dHexHexA [3AQ/M+K] <sup>+</sup>

Some observations can be made related to the obtained results:

- From the knowledge of the potential gum arabic structure it can be concluded that Pent corresponds to arabinose, Hex is galactose, dHex is rhamnose and HexA can be ascribed to glucuronic acid.
- Since the MALDI spots (sample mixed with 3-aminoquinoline) were dried at room temperature for at least 1 hour, the derivatization rate should be about 100% [50]. The mass spectrum demonstrated how all the oligosaccharides were in fact derivatized.
- The [M+H]<sup>+</sup> at m/z 1055 (and the corresponding [M+Na]<sup>+</sup> and [M+K]<sup>+</sup> at m/z 1077 and 1093) are composed of few pentoses but, looking at the Fig. 23B, these ions were still present in the mass spectrum after digestion with  $\alpha$ -L-arabinofuranosidase, meaning that the enzyme did not hydrolyze them. Two explanations are possible: (i) arabinose units are in the middle of the chain, so the enzyme did not have access, or (ii) the monosaccharides are *Arap* and not *Araf* so  $\alpha$ -L-arabinofuranosidase could not remove them.

At the light of all these considerations a final mass fingerprint of gum arabic digested with exo- $\beta$ -1,3-galactanase is reported in Fig. 26 with the corresponding attributed oligosaccharides. The mass spectrum shows how the combinations of the different approaches allowed to fully characterize the most intense peaks of the gum profile.



**Fig. 26.** MALDI-TOF mass spectrum of gum arabic (Zecchi - Italy) digested with exo-β-1,3-galactanase for 5h with the respective assigned oligosaccharides (all derivatized with 3-aminoquinoline).



#### 4.2.3. Analysis of different gum arabic samples from *Acacia Senegal* var. *Senegal*

The chemical composition of gum arabic, in terms of both proteinaceous and polysaccharide component, can vary according to different parameters such as plant source and species, age of the tree, area of collection and climatic conditions [63-65]. It was therefore believed necessary to verify that the MS profile obtained for the gum arabic under investigation (Zecchi - Italy) was the same of other samples from *Acacia senegal* plants. The comparison of their MS spectra is indeed of significant interest in order to check what it can be called “oligosaccharide profile reproducibility”. Only in this way it would be possible to be self-confident about gum arabic identification and then discrimination from the other plant gums.

Considering the species *Acacia senegal*, four varieties are by now recognized within this group: var. *senegal*, var. *kerensis*, var. *rostrata* and var. *leiorhachis* [66]. Due to sample availability, attentions was focused on the variety *Senegal*. The specifications of the analyzed gum arabic samples from *Acacia Senegal* var. *Senegal* (see Fig. 27), kindly supplied by Prof. P.C. Ravines (Buffalo State University – School of Arts and Humanities), are reported in Tab. 8. All the three samples were collected in Sudan in two different years.

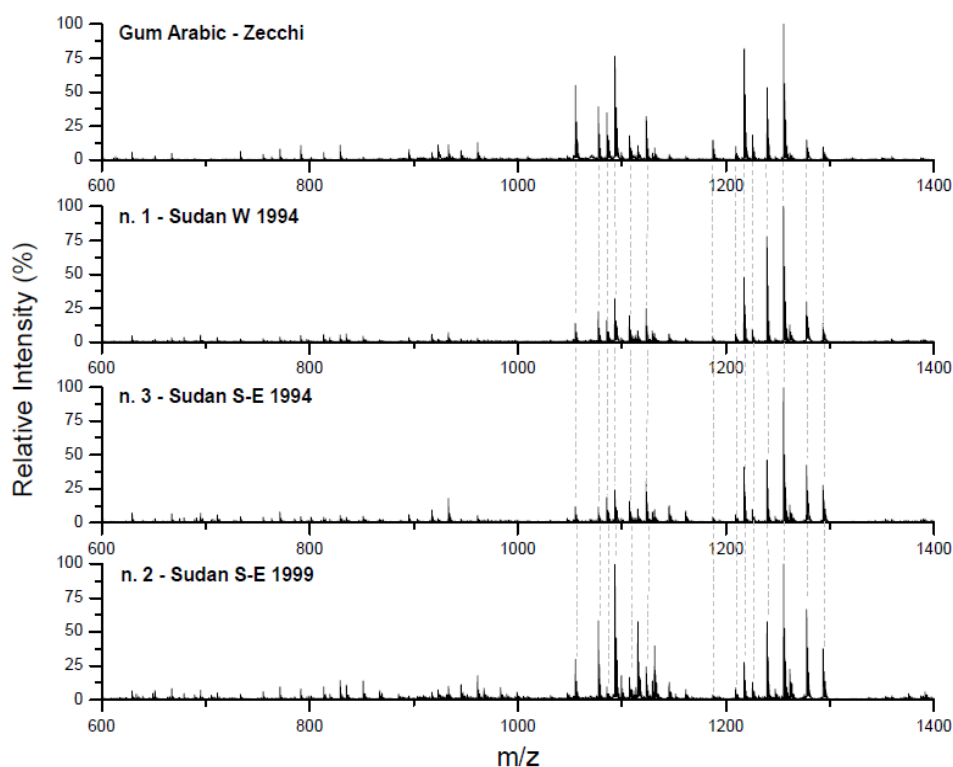
**Tab. 8.** List of the analyzed gum arabic samples from *Acacia Senegal* var. *Senegal*, area and date of collection.

N.	Plant species	Area of collection	Date of collection
1	<i>Acacia senegal</i> var. <i>senegal</i>	Kordufan (Western Sudan)	1994
2	<i>Acacia senegal</i> var. <i>senegal</i>	Damazene (Southeast Sudan)	1999
3	<i>Acacia senegal</i> var. <i>senegal</i>	Damazene (Southeast Sudan)	1994



**Fig. 27.** *Acacia senegal* var. *senegal* gum nodules from: (1) Kordufan (1994), (2) Damazene (1999) and (3) Damazene (1994).

Samples were prepared as described in the chapter “Materials and Methods” and 1 mg of gum was subjected to enzymatic digestion with exo- $\beta$ -1,3-galactanase for 5 hours. The resulted mass spectra, compared to the MS profile of the standard gum arabic (Zecchi) are reported in Fig. 28.



**Fig. 28.** Comparison of MALDI-TOF spectra of standard gum arabic (Zecchi) with gum arabic nodules (*Acacia senegal* var. *senegal*) digested for 5h with exo- $\beta$ -1,3-

The mass spectra of the three gum arabic samples (species *Acacia Senegal* var. *Senegal*) reflected the characteristic mass profile of gum arabic. The list of the observed experimental ions is reported in Tab. 9 with the corresponding mass error [Da] obtained by comparison with the oligosaccharide theoretical mass previously discussed for gum arabic in Tab. 7. Some of the ions at lower molecular weight were missing but, if considering the most significant mass range from 1000 to 1300 Da, it was possible to observe how most of the peaks could be attribute to the specific gum arabic oligosaccharides.

**Tab. 9.** Comparison of theoretical masses (gum arabic characteristic oligosaccharides, Tab.7) with the experimental masses obtained after digestion of three gum arabic samples *Acacia Senegal* var *Senegal*. The oligosaccharides which are not present in the mass spectra of the analyzed samples are indicated by ‘/’.

Theor. [Da]	Sample 1		Sample 2		Sample 3	
	Exp. [Da]	$\Delta m$ [Da]	Exp. [Da]	$\Delta m$ [Da]	Exp. [Da]	$\Delta m$ [Da]
601.2007	/	/	601.4119	0.21	601.4264	0.23
623.1826	/	/	/	/	/	/
629.1956	629.3770	0.18	629.4177	0.22	629.4227	0.23
639.1566	/	/	639.4136	0.26	639.4007	0.24
651.1775	651.3377	0.16	651.3737	0.20	651.3838	0.21
667.1515	667.3109	0.16	667.3890	0.24	667.3846	0.23
733.2429	733.4143	0.17	733.4687	0.23	733.5033	0.26
755.2249	755.4021	0.18	755.4966	0.27	755.4336	0.21
763.2535	763.2653	0.01	763.4774	0.22	763.4724	0.22
771.1988	771.3474	0.15	771.4726	0.27	771.4199	0.22
785.2354	/	/	785.4935	0.26	785.4763	0.24
791.2484	791.4005	0.15	791.4767	0.23	791.4421	0.19

801.2094	/	/	801.4722	0.26	801.4363	0.23
813.2304	813.3567	0.13	813.4533	0.22	813.3850	0.15
829.2043	829.3477	0.14	829.4393	0.24	829.3872	0.18
895.2958	895.4568	0.16	895.5126	0.22	895.4962	0.20
917.2777	917.4418	0.16	917.5651	0.29	917.5203	0.24
923.2907	923.4346	0.14	923.5294	0.24	923.4880	0.20
933.2516	933.4163	0.16	933.5371	0.29	933.4839	0.23
945.2726	945.4161	0.14	945.5103	0.24	945.4653	0.19
961.2465	961.3994	0.15	961.4934	0.25	961.4631	0.22
1055.3330	1055.4711	0.14	1055.5731	0.24	1055.5043	0.17
1077.3150	1077.4606	0.15	1077.5663	0.25	1077.5323	0.22
1085.3430	1085.4766	0.13	1085.5751	0.23	1085.5292	0.19
1093.2890	1093.4288	0.14	1093.5525	0.26	1093.4995	0.21
1107.3250	1107.4651	0.14	1107.5681	0.24	1107.5431	0.22
1123.2990	1123.4335	0.13	1123.5526	0.25	1123.5118	0.21
1187.3750	1187.4932	0.12	1187.6101	0.24	1187.5339	0.16
1209.3570	1209.4799	0.12	1209.6083	0.25	1209.5519	0.19
1217.3860	1217.5154	0.13	1217.6279	0.24	1217.5605	0.17
1225.3310	1225.4532	0.12	1225.5831	0.25	1225.5242	0.19
1239.3680	1239.5031	0.14	1239.6310	0.26	1239.5869	0.22
1255.3420	1255.4728	0.13	1255.6066	0.26	1255.5508	0.21

Any significant difference could be observed among the mass spectra of samples 1,2 and 3. Therefore, the fact that gum nodules were collected from different areas in Sudan and in different years, did not seem to affect the oligosaccharide mass profile, meaning that the gum polysaccharide structure is comparable. In addition, the gum arabic sample purchased from Zecchi (Italy) was collected in Sudan from the plant species *Acacia Senegal*, but no specifications about the variety were supplied. Therefore, due to the significant similarity of the mass spectra, it was possible to refer the sample to the *Senegal* variety.

In conclusion, the mass profile similarity among the gum arabic samples from *Acacia Senegal*, allowed to prove a profile reproducibility, thus confirming a final mass fingerprint of gum arabic (Fig. 26 and Tab. 7). In addition these results revealed how a further investigation of other *Acacia Senegal* var. *Senegal* samples collected from different regions and in different years, would be of great interest to evaluate the possibility to distinguish the origin of the polysaccharide material. Furthermore, since gum arabic could be collected from several *Acacia* species and varieties, another potential study could be related to the investigation and comparison of the MALDI mass spectra of these gums.

### 4.3. TRAGACANTH GUM

The development of the enzymatic digestion protocol for the study of gum tragacanth has shown to be challenging. To our knowledge, the structure of gum tragacanth polysaccharide moiety is still not completely established since, as described previously, multiple fractions with different properties and composition are present. In addition, sample manipulation is problematic since only the 30% of the gum is water soluble [58] and gum concentration cannot exceed the 1% otherwise a jelly like material is obtained. In order to overcome these limitations, a 0.5% w/v gum solution was prepared, as described in Tab. 6. As reported in Fig. 29, a general structure of the gum was presumed in order to select the enzymes to be used.

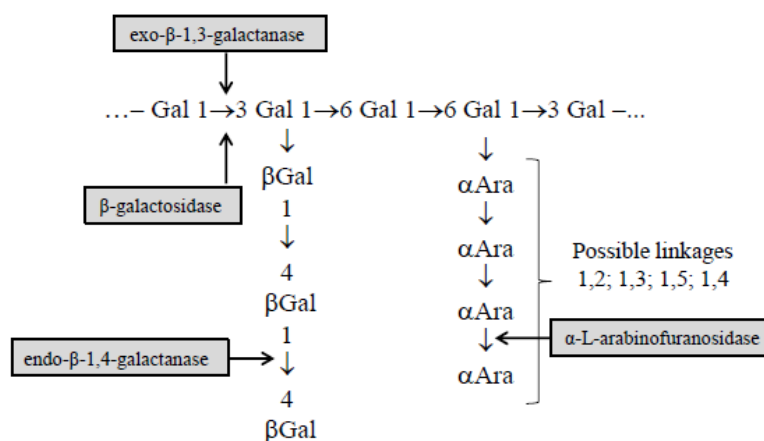
#### 4.3.1. Evaluation of enzymatic digestion protocol and oligosaccharides identification

A representation of enzymes action on the arabinogalactan fraction of tragacanth gum is reported in Fig. 29. Among the already discussed enzymes that have been used for digestion of gum arabic ( $\alpha$ -L-arabinofuranosidase;  $\beta$ -galactosidase and *exo*- $\beta$ -1,3-galactanase), another enzyme was tested: *endo*- $\beta$ -1,4-galactanase (3.2.1.89). Even if it is supposed to only hydrolyze (1 $\rightarrow$ 4)- $\beta$ -D-galactosidic linkages in type I arabinogalactans (<http://www.brenda-enzymes.info/>), it was tested anyway since the gum structure is not clear. Besides the water soluble arabinogalactan component, the presence of monosaccharides such as xylose (pentose), fucose (deoxyhexose) and mainly galacturonic acid (HexA), coming from the main backbone of the other gum fraction, have to be considered [67]. In the specific, enzymes that are involved in galacturonic acid digestion belong to the family of pectinases (e.g. endopolygalacturonase II [68] and exopolygalacturonase [69]) and they have already been successfully used for evaluating the probiotic potential of released oligosaccharides [70]. Unfortunately these enzymes were not available at the time so it was only possible to enzymatically act on the arabinogalactan component.

At the light of these consideration the following enzymes were tested in the reported order:

1.  $\alpha$ -L-arabinofuranosidase;
2. *endo*- $\beta$ -1,4-galactanase;
3. *exo*- $\beta$ -1,3-galactanase;
4.  $\beta$ -galactosidase.

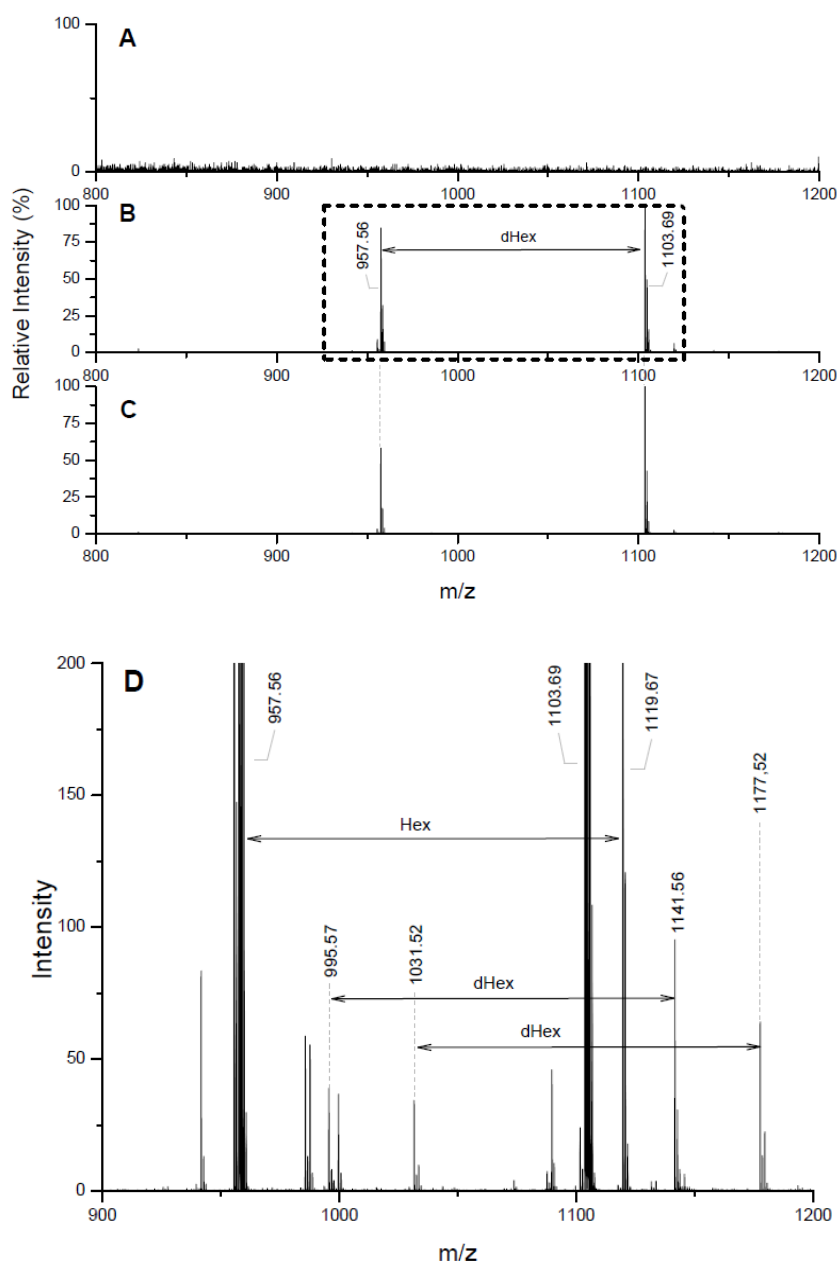
Solubilized and dried tragacanth gum samples (Bresciani, Italy) was initially digested with  $\alpha$ -L-arabinofuranosidase for 24 hours, reaction was quenched by boiling the solution and an aliquot was collected and analyzed by MALDI-TOF. The left sample was dried, re-solubilized in the proper buffer and incubated for 24 hours with the second enzyme etc. Each enzyme was incubated for 24 hours.



**Fig. 29.** Illustration of various enzymes possibly involved in the degradation of tragacanth gum arabinogalactan component.

The corresponding mass spectra are showed in the figure below (Fig. 30). When tragacanth gum was spotted and analyzed before enzymatic digestion, only matrix ions were observed at low molecular weights (50-600 Da) while at higher masses (600-1200 Da) (Fig. 30A) any ion was detected. After incubation for 24 hours with  $\alpha$ -L-arabinofuranosidase, two intense peaks, respectively at  $m/z$  957.56 and 1103.69, that differed in mass of 146.07 Da were observed (Fig.

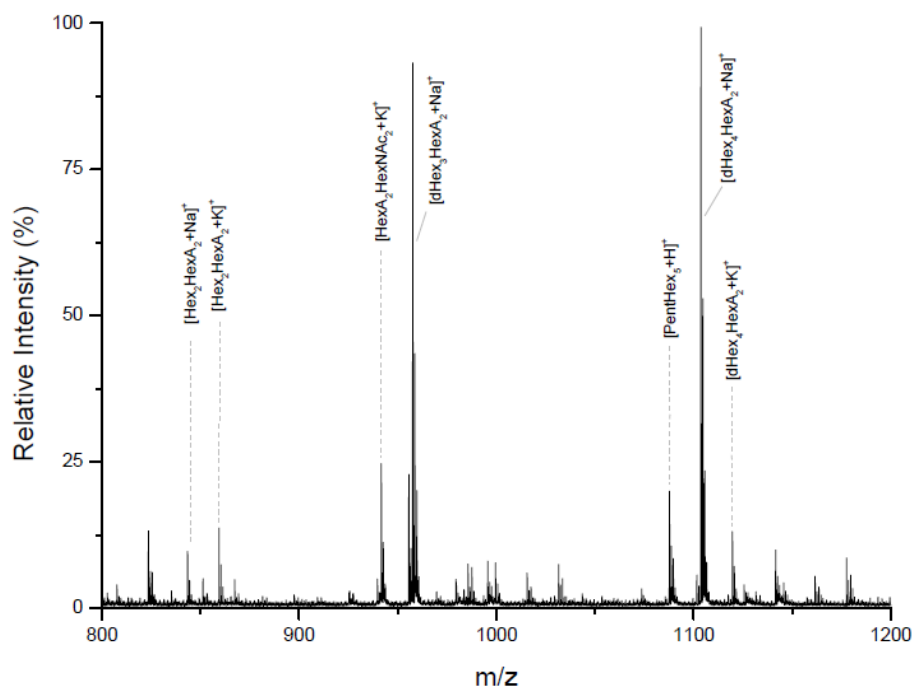
30B). The same ions were detected after addition of endo- $\beta$ -1,4-galactanase followed by exo- $\beta$ -1,3-galactanase (Fig. 30C) and  $\beta$ -galactosidase. When zooming the mass spectrum of the digested gum to check lower signal/noise peaks (Fig. 30D), other ions with mass differences corresponding to the monosaccharide masses were observed. The mass spectrum showed the presence of hexose units, for example between the peaks at  $m/z$  1119.67 and 957.56, and deoxyhexoses units (rhamnose or fucose) that were observed for some ions such as  $m/z$  1177.52 and 1031.52, and 1141.56 and 995.57. Since these ions were not detected when the gum was analyzed before digestion, it was possible to assume that the oligosaccharide were released by enzyme action of the tragacanth polysaccharide component.



**Fig. 30.** Mass spectra (range 800-1200 Da) of tragacanth gum sample: (A) non digested (diluted 1/100); (B) digested with  $\alpha$ -L-arabinofuranosidase; (C) digested with  $\alpha$ -L-arabinofuranosidase + endo- $\beta$ -1,4-galactanase + exo- $\beta$ -1,3-galactanase ; (D) zoom in signal/noise of Fig. 30B (frame).

$\alpha$ -L-arabinofuranosidase should hydrolyze terminal non-reducing residues from arabinose-containing polysaccharides so, according to the theoretical structure of gum tragacanth, chains made up of only Ara should be released. However, according to the obtained mass spectrum, more complex oligosaccharides are released and they are made up of hexoses and deoxyhexoses. This result is of difficult interpretation and two possible explanations were proposed: (i) the purchased enzyme was not completely pure, even if purity is reported by the supplier to be > 90%, so a possible residual activity on other monosaccharides is present or, (ii) the side chain structure is not made up of only arabinose but has a more complex structure.

Since the final aim is to find a common enzyme, or enzymes mixture, that is able to digest all the plant gums under investigation and to release their characteristic oligosaccharides, tragacanth gum sample was incubated only with  $\text{exo-}\beta$ -1,3-galactanase since it was already successfully used with gum arabic. During digestion aliquots were collected after 3, 5, 24 and 48 hours, and analyzed by MALDI-TOF. Since no significant differences in the digestion rate were observed at the different timing, only the mass spectrum obtained after 5 hours digestion is reported (Fig. 31).



**Fig. 31.** MALDI-TOF mass spectrum of tragacanth gum (Bresciani - Italy) digested with  $\text{exo-}\beta$ -1,3-galactanase for 5 h with the assigned oligosaccharides (all derivatized with 3-aminoquinoline).

The MALDI profile reflected the one obtained after digestion performed using four enzymes but, instead of four days preparation, the new protocol allowed to obtain the same results in only 5 hours incubation. MS spectrum was investigated by the informatics tool, which allowed to assign hypothetical oligosaccharides to each ion, and by evaluating the mass differences between the ions. A table with the detected ion masses, the theoretical monoisotopic masses, the error and the assigned composition is reported below (Tab. 10). The mass range that is considered significant for the interpretation of the gum fingerprint goes from 800 to 1200 Da. Below this range, only matrix peaks were observed and over 1200 Da any oligosaccharide was detected.

**Tab. 10.** List of the assigned oligosaccharides for digested gum tragacanth. (<sup>a</sup>) the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline, H<sup>+</sup>, Na<sup>+</sup> or K<sup>+</sup>).

Experimental mass [Da]	Theoretical mass [Da] <sup>a</sup>	$\Delta m$ [Da]	Possible oligosaccharide
843.3798	843.2045	0.18	Hex <sub>2</sub> HexA <sub>2</sub> [3AQ/M+Na] <sup>+</sup>
859.3298	859.1785	0.15	Hex <sub>2</sub> HexA <sub>2</sub> [3AQ/M+K] <sup>+</sup>
895.2809	895.2958	-0.01	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
955.6042	955.3169	0.29	Hex <sub>5</sub> [3AQ/M+H] <sup>+</sup>
957.6264	957.2726	0.35	dHex <sub>3</sub> HexA <sub>2</sub> [3AQ/M+Na] <sup>+</sup>
979.6965	979.3533	0.34	Pent <sub>3</sub> dHex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
987.5819	987.2468	0.34	dHex <sub>2</sub> HexA <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
987.5819	987.3196	0.26	Pent <sub>4</sub> dHex <sub>2</sub> [3AQ/M+Na] <sup>+</sup>
995.4845	995.2390	0.25	dHexHexA <sub>4</sub> [3AQ/M+H] <sup>+</sup>
995.4845	995.3482	0.14	Pent <sub>3</sub> HexdHex <sub>2</sub> [3AQ/M+H] <sup>+</sup>
995.4845	995.3118	0.17	Pent <sub>4</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
999.5225	999.3308	0.19	Pent <sub>2</sub> HexHexNAc <sub>2</sub> [3AQ/M+Na] <sup>+</sup>
1015.5695	1015.3510	0.22	Pent <sub>2</sub> dHex <sub>4</sub> [3AQ/M+Na] <sup>+</sup>
1031.5186	1031.3460	0.17	Pent <sub>2</sub> HexdHex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
1031.5186	1031.3090	0.21	Pent <sub>3</sub> dHex <sub>2</sub> HexA [3AQ/M+Na] <sup>+</sup>
1031.5186	1031.3250	0.19	Pent <sub>2</sub> dHex <sub>4</sub> [3AQ/M+K] <sup>+</sup>
1087.7206	1087.3590	0.36	PentHex <sub>5</sub> [3AQ/M+H] <sup>+</sup>
1103.6929	1103.3310	0.36	dHex <sub>4</sub> HexA <sub>2</sub> [3AQ/M+Na] <sup>+</sup>
1119.6757	1119.3040	0.37	dHex <sub>4</sub> HexA <sub>2</sub> [3AQ/M+K] <sup>+</sup>
1141.6387	1141.2970	0.34	dHex <sub>2</sub> HexA <sub>4</sub> [3AQ/M+H] <sup>+</sup>
1141.6387	1141.4060	0.23	Pent <sub>3</sub> HexdHex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
1141.6387	1141.3700	0.27	Pent <sub>4</sub> dHex <sub>2</sub> HexA [3AQ/M+H] <sup>+</sup>
1161.6306	1161.4090	0.22	Pent <sub>2</sub> dHex <sub>5</sub> [3AQ/M+Na] <sup>+</sup>
1177.5922	1177.3830	0.21	Pent <sub>2</sub> dHex <sub>5</sub> [3AQ/M+K] <sup>+</sup>

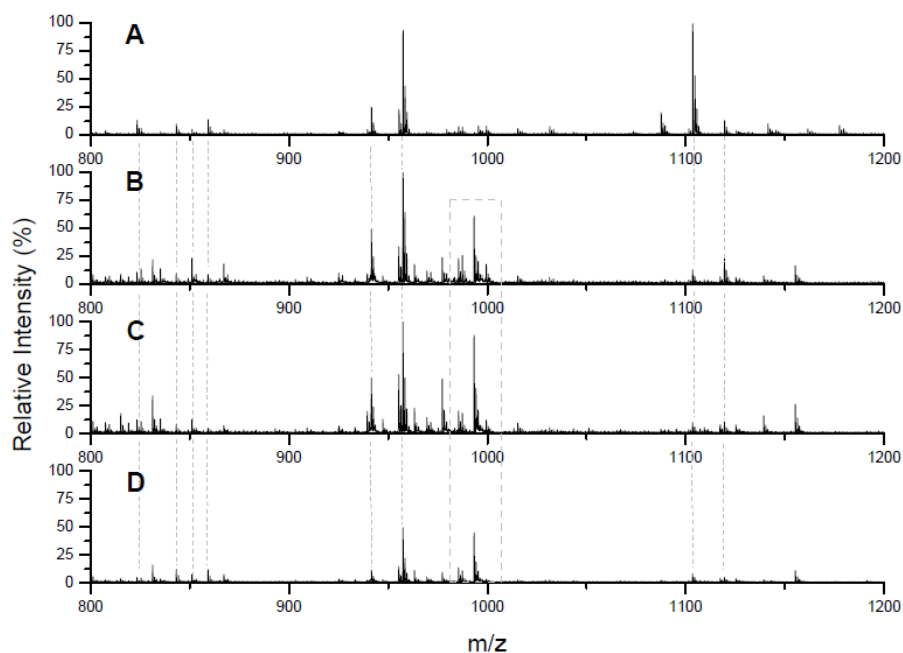
Some observations can be made related to the digestion protocol and obtained results:

- A short digestion with exo- $\beta$ -1,3-galactanase for 5 hours allows to obtain a mass spectrum of gum tragacanth.
- A MALDI-TOF pattern of gum tragacanth was obtained and it looked different from the mass fingerprint obtained for gum arabic. This first result is promising but further investigation of the released oligosaccharides has to be performed in order to have a complete profile of the gum by mass spectrometry.

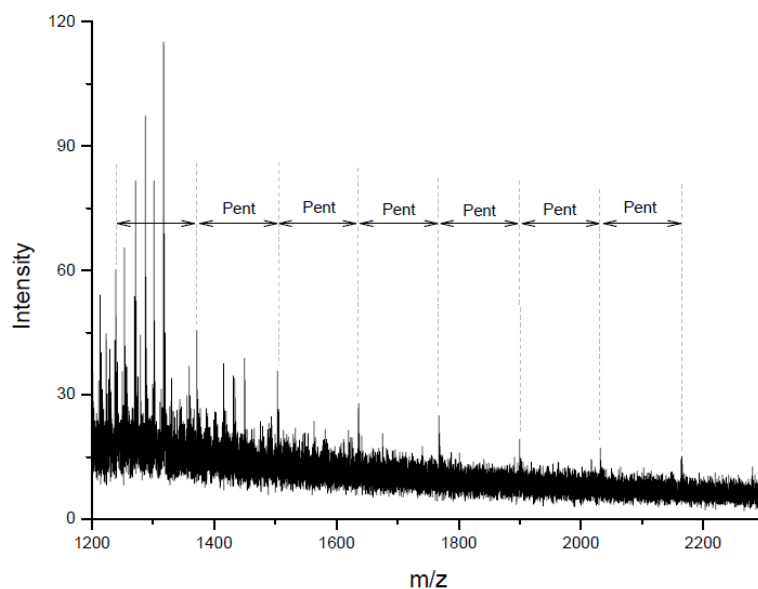
#### 4.3.2. Analysis of tragacanth samples from different brands

The monosaccharide composition of tragacanth gum samples varies among different *Astragalus* species [71]. It is therefore possible that also their structure, in terms of monosaccharide substituents, is different. For this reason other samples of tragacanth gum from different suppliers were subjected to the same digestion protocol in order to compare the MS profiles. Only in this way it would be possible to be self-confident about gum identification for a further discrimination from the other plant gums. Besides the already studied gum tragacanth from Brescaini (Italy), other gum samples from the following suppliers were investigated: Zecchi (Italy); Okhra (France); Sigma (Germany). Samples were incubated with exo- $\beta$ -1,3-galactanase for 5 and 24 hours but no

differences in the rate of digestion were observed. The corresponding mass spectra are reported in Fig. 32.



**Fig. 32.** MALDI-TOF spectra of tragacanth gum digested for 5 h with exo- $\beta$ -1,3-galactanase: (A) Bresciani; (B) Zecchi; (C) Okhra and (D) Sigma.



**Fig. 33.** Enlargement of the MS spectrum of digested tragacanth gum from Okhra (1200-2300 Da).

Gum tragacanth samples from different suppliers seemed to have a similar profile with the ion at  $m/z$  957.63 of higher intensity. If compared to the sample from Bresciani, the one used for the development of the digestion protocol, the ion at  $m/z$  1103.69 was present but its intensity was not high. On the contrary, two ions at respectively  $m/z$  of 977.40 and 993.37, were fairly intense and they had not been detected previously. Furthermore, the presence of a long chain made of pentoses enzymatically released from the Okhra gum is shown in Fig. 33. A mass difference of 132.04 Da



was in effect detected between the ions at  $m/z$  1239.39. and 1371.43, and so on up to the higher molecular weight ion at  $m/z$  2163.66. This result is consistent with the information on tragacanth gum polysaccharide structure reported in the literature.

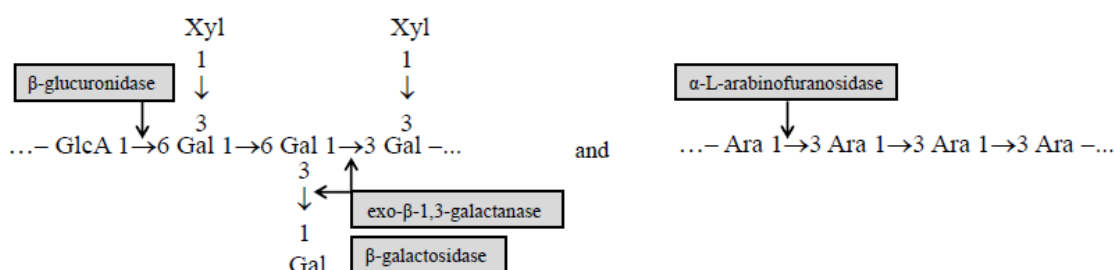
In conclusion:

- A short digestion of 5h with  $\text{exo-}\beta\text{-1,3-galactanase}$  was confirmed to be suitable to obtain a mass spectrum for tragacanth gum.
- The MS oligosaccharide profile of gum tragacanth is characterized by few intense peaks at  $m/z$  957.63 and 1103.69. These peaks allowed to distinguish tragacanth gum from gum arabic but information are not enough to reconstruct a complete mass profile of the gum. However, investigation of tragacanth gum samples purchased from different suppliers show a certain “profile reproducibility” and some additional information on the gum tragacanth structure were obtained (e.g. Okhra sample).

## 4.4 CHERRY GUM

### 3.4.1. Evaluation of enzymatic digestion protocol

In order to establish a proper digestion protocol for cherry gum, the potential structures reported in the literature were evaluated together with the available enzymes. A representation of possible enzymes action on the cherry gum polysaccharide component is reported in Fig. 34.



**Fig. 34.** Illustration of various enzymes possibly involved in the degradation of cherry gum.

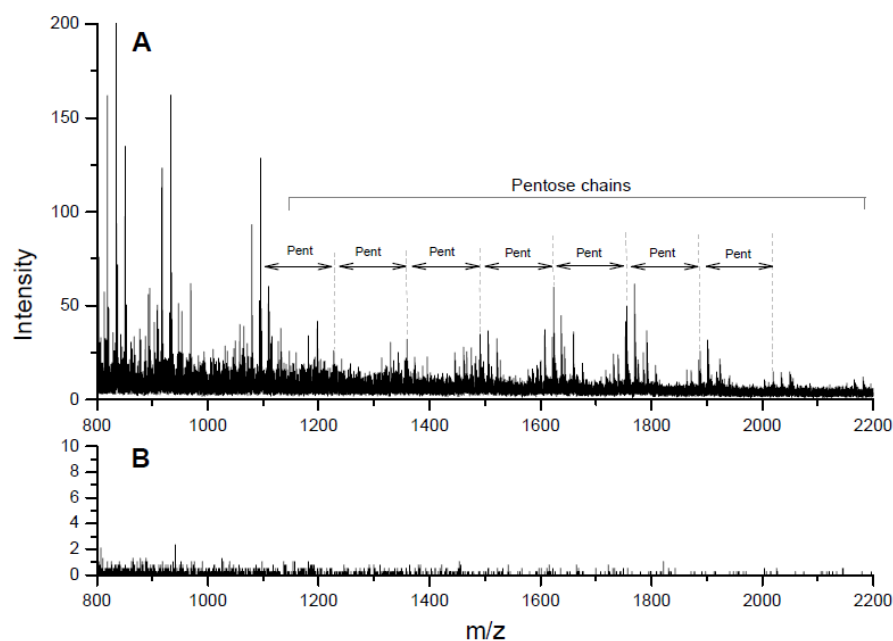
The enzymes reported in the figure above ( $\alpha\text{-L-arabinofuranosidase}$ ;  $\beta$ -glucuronidase;  $\text{exo-}\beta\text{-1,3-galactanase}$  and  $\beta$ -galactosidase) were tested on a sample of cherry gum (named cherry n.2) collected at the end of the summer 2012 in Veneto (Italy). The sample appeared to be not completely soluble in water and it was therefore feasible that a low amount of polysaccharide fraction could be present in the solution collected after centrifugation. Dried gum was initially digested with  $\alpha\text{-L-arabinofuranosidase}$  for 24 hours, reaction was quenched by boiling the solution and an aliquot was collected and analyzed by MALDI-TOF. The left sample was dried, re-solubilized in the proper buffer and incubated for 24 hours with the second enzyme etc. Each enzyme was incubated for 24 hours and different enzyme combinations were tested. Among all the different experiments only digestion performed with  $\text{exo-}\beta\text{-1,3-galactanase}$  led to some results. No differences were observed if the sample was incubated for 5, 24 or 48 hours, so the shortest time

digestion was preferred. The obtained mass spectrum is reported in Fig. 35A and, while analysis of the gum before digestion did not lead to any result, the presence of some ions along the mass range 800-2000 Da indicated that digestion worked. In the mass range of interest, over the matrix peaks, ions were characterized by low intensity, thus interpretation of the spectrum and identification of the corresponding oligosaccharides turned out to be complicated. However, the observed profile looked different from the one obtained for gum arabic and tragacanth gum, so it would still be possible to distinguish the three plant gums.

The ions at  $m/z$  835.20 and 933.30 were the most intense peaks in the mass spectrum. Both of them had also been observed in the gum arabic profile. While the oligosaccharide assigned to the ion 933 for gum arabic ( $\text{Pent}_2\text{Hex}_3$   $[\text{3AQ/M+K}]^+$ ) could be confirmed also for cherry gum since the simultaneous presence of the  $[\text{3AQ/M+Na}]^+$  (917.32 Da) and  $[\text{3AQ/M+H}]^+$  (895.32 Da), the same could not be said for the ion at  $m/z$  835.20. Composition obtained by the informatics tool referred to two possible protonated oligosaccharides:  $\text{Pent}_4\text{Hex}$  and  $\text{HexHexA}_3$   $[\text{3AQ/M+H}]^+$ , while in the cherry mass spectrum the presence of the ion at  $m/z$  819.22 (mass difference = 16Da) and  $m/z$  797.22 (mass difference = 38Da) suggested the oligosaccharide to have a potassium adduct. Another intense peak corresponded to the  $m/z$  value of 1095.33. The presence of the  $m/z$  1079.35 (mass difference = 16Da) suggested the ion to be charged with potassium ( $[\text{M+K}]^+$ ). Furthermore the mass difference between 1095.33 and 933.30 was 162.03 Da, indicating that an hexose was cleaved. Since  $m/z$  933.30 was designated as  $\text{Pent}_2\text{Gal}_3$   $[\text{3AQ/M+K}]^+$ , it was possible to assume that 1095.33  $m/z$  corresponded to the oligosaccharide  $\text{Pent}_2\text{Gal}_4$   $[\text{3AQ/M+K}]^+$ .

Beginning from this ion, a mass difference corresponding to pentose was observed for the following ions ( $m/z$  1227.37, 1359.42, 1491.45, 1623.48, 1755.52, 1887.55, 2019.58) suggesting the presence of a chain of pentose units. Consequently the above mentioned ions may respectively corresponded to the oligosaccharides  $\text{Pent}_3\text{Gal}_4$ ,  $\text{Pent}_4\text{Gal}_4$ ,  $\text{Pent}_5\text{Gal}_4$ ,  $\text{Pent}_6\text{Gal}_4$ ,  $\text{Pent}_7\text{Gal}_4$ ,  $\text{Pent}_8\text{Gal}_4$  and  $\text{Pent}_9\text{Gal}_4$   $[\text{3AQ/M+K}]^+$ . The ion at  $m/z$  value of 1131.31 could correspond to the oligosaccharide  $\text{Hex}_5\text{HexA}$   $[\text{3AQ/M+H}]^+$ . The mass difference of 162.04 Da with the ion at  $m/z$  969.27 indicated the loss of a hexose, so the oligosaccharide  $\text{Hex}_4\text{HexA}$   $[\text{3AQ/M+H}]^+$  could be assigned to this ion. However a mass difference of 22Da was observed between the  $m/z$  969.27 and 947.26, thus suggesting the former to have a sodium adduct. In this case the mass could be assigned to an oligosaccharide composed of five pentoses and two 4-O-methyl-d-glucuronic acid  $[\text{3AQ/M+Na}]^+$ . Therefore the ion at 1131.31, which differed in mass of 162.04 Dalton, could be assigned to the same oligosaccharide with the addition of a hexose.

The ion at  $m/z$  1197.36 could be assigned to both  $\text{Hex}_3\text{HexA}_3$  or  $\text{Pent}_4\text{Hex}_3$  oligosaccharides, and the presence in the mass spectrum of the peak at  $m/z$  1181.38 suggested the ion to have a potassium adduct. The ions at  $m/z$  1329.39 and 1461.43 differed in mass, from the peak at 1197.36, of respectively 132.03 and 264.07 Da, indicating that the assignable oligosaccharide corresponded to the one assigned to the ion 1197.36 added respectively of one and two pentoses. Considering the  $m/z$  values 1109.31, 1241.36, 1373.37, 1505.45, 1637.47, 1769.50, 1901.53 and 2033.78, the mass difference between neighboring peaks corresponded to pentose, suggesting the presence of another pentose chain. Same observations could be made for the  $m/z$  values 1659.44, 1791.48, 1923.52 and the series 1785.52, 1917.56, 2049.56 and 2182.62. Unfortunately no attribution could be made for the following ions.



**Fig. 35.** MALDI-MS spectrum of (A) cherry gum (n.2) digested with exo- $\beta$ -1,3-galactanase for 5h and (B) sample not digested.

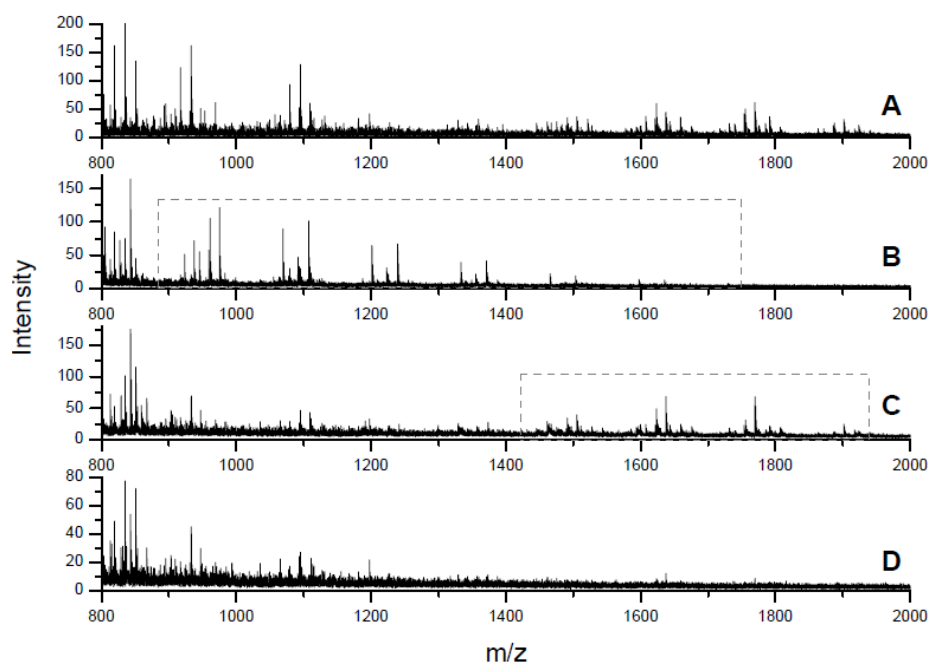
#### 4.4.2. Analysis of different cherry gum samples

Cherry gum could be simply collected by the artist directly from the cherry trees available in the territory, thus from an analytical point of view it was considered necessary to study several cherry gum samples in order to verify if a similar profile could be obtained. Furthermore, differences in the molecular weight distribution profile and average MW was measured depending on the tree species and the season of collection [72], demonstrating that differences in the chains length might be present. Gum was collected from cherry trees (the *Prunus* species is not known) of different towns around the province of Treviso (Veneto region - Italy) and were named cherry n.1 and n.3. Another sample from *Prunus Avium* var. was bought from Kremer Pigmente GmbH&Co. (gum obtained from the south-east regions of Balkans).

The samples were prepared as described in Tab. 6. In all cases, after solubilization in milliQ water, the gum appeared to be not completely soluble and, for the samples collected directly from the trees (cherry n.1 and n.3), several residues of probably wood and other materials were observed. The dried gums were digested for 5 hours with exo- $\beta$ -1,3-galactanase and the corresponding mass spectra are reported in Fig. 36.

The first mass spectrum reported in Fig. 36 reproduces the profile already described in Fig. 35 of cherry gum n.2 and it was used for comparison. It was possible to observe how the mass spectra of different cherry gum samples were characterized by few and with low intensity ions. The samples named cherry n.1 and n.3 (respectively Fig. 36C and D) showed a profile similar to the one of cherry gum n.2 (Fig. 36A). In both cases ions at m/z 819.22, 835.20 and 851.17 were observed, indicating that digestion procedure worked. In the specific, sample n.1 showed some ions in common with n.2 as highlighted with a marked area in the mass spectrum. Considering the m/z values 1505.61, 1637.62, 1769.60 and 1901.73, the mass differences between neighboring peaks

suggested the presence of a pentose chain. Same observation could be made for the  $m/z$  values 1491.66, 1623.64 and 1755.70.



**Fig. 36.** MALDI-TOF mass spectra of cherry gum samples digested with  $\text{exo-}\beta\text{-1,3-GAL}$  for 5h: (A) cherry n.2; (B) Kremer; (C) cherry n.1 and (D) cherry n.3. Marked region indicates the ions of interest.

Regarding the cherry sample from Kremer (Fig. 36B) a profile characterized by ions differing in mass of 132.04 Da was observed, as for the other samples. This mass difference was detected for the ions at  $m/z$  843.31, 975.34, 1107.39, 1239.44, 1371.48 and 1503.51, which were maybe charged with potassium since the presence of the possible  $[\text{M}+\text{Na}]^+$  peaks (mass difference = 16 Da) and  $[\text{M}+\text{H}]^+$  (mass difference = 38 Da). The ions' masses were different from the ones detected in the cherry gum n.2 spectrum and this was maybe due to the fact that the two gums were collected in different areas (Italy and Balkans). Furthermore, the peaks corresponded to the ions observed in the profile of gum tragacanth purchased by Okhra (Fig. 33) and this could impede the discrimination between the two gums. However, the two main oligosaccharides of gum tragacanth at  $m/z$  957.56 and 1103.69 were not present so these two ions might be potentially used for gums distinction.

In conclusion:

- The MS polysaccharide pattern is different from the mass pattern of gum arabic and tragacanth gum, even if all the gums were digested with the same protocol ( $\text{exo-}\beta\text{-1,3-galactanase}$  for 5 hours).
- Mass spectra of digested cherry gum samples are characterized by the presence of chains made of pentoses (more likely arabinose according to the polysaccharide structure knowledge).
- Cherry gum from Kremer showed the characteristic pentose chain. However, in order to deepen the knowledge of the gum profile, it is considered necessary to analyze more cherry gums collected from several areas of the world.

#### 4.5. LOCUST BEAN AND GUAR GUM

An enzymatic digestion protocol was developed and optimized simultaneously for locust bean and guar gum since they both resemble galactomannans. Gums were solubilized in milliQ water at a concentration of 1% w/v and prepared as described in Tab. 6. Higher amount of gum led to a jelly like material that was impossible to manipulate. In addition, a concentrated polymer solution can decrease the yield of hydrolysis due to the reduced mobility of the enzyme in the solution [73].

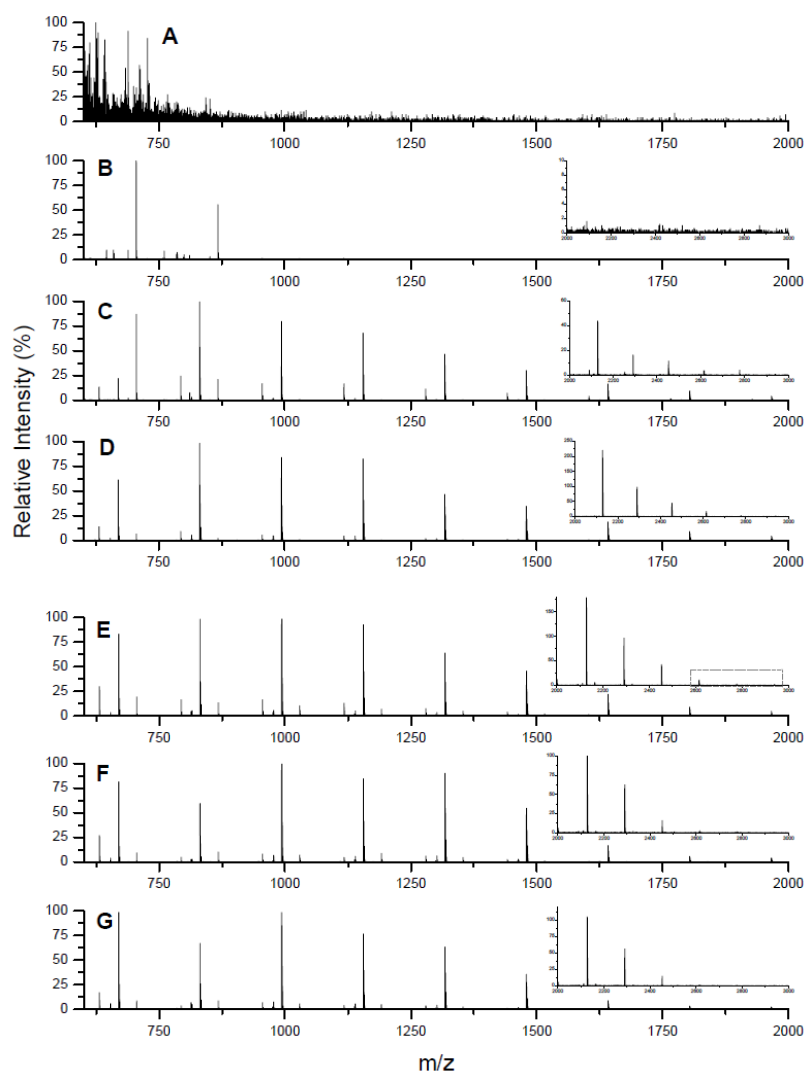
Regarding the digestion protocol, some optimizations were performed on the one developed on the standard galactomannan. As described in the previous chapters, the two enzymes that are theoretically involved in the galactomannans degradation are  $\alpha$ -galactosidase and endo-1,4- $\beta$ -mannanase. However, in most of the procedures reported in the literature, guar and LBG galactomannans are digested only with endo-1,4- $\beta$ -mannanase [74-78]. For this reason a first test on the action of  $\alpha$ -galactosidase on galactomannans substrates was performed. LBG and guar gum were digested with 100 mU of enzyme for 24 hours. Once the sample is dried after incubation, it must be then re-solubilized in few  $\mu$ L of milliQ water for being later mixed with the matrix and spotted on the MALDI plate. However it occurred that, after incubation, the digested samples were almost insoluble if a low volume of water was added, as if the enzymatic digestion did not work. For this reason any MALDI-TOF analysis could be performed. On the contrary it was also possible that polysaccharide hydrolysis occurred but, it is reported in the literature that when the percentage of galactose substituents in the galactomannan structure is  $< 12\%$ , the gum becomes almost insoluble [79].

At the light of all these considerations, digestion of LBG and guar gum was performed only with endo-1,4- $\beta$ -mannanase and digestion time was evaluated in order to evaluate connection with the rate of hydrolysis. On the contrary of the classical digestion procedure, the enzyme was denatured and so the reaction quenched by heating the sample at  $100^{\circ}\text{C}$  for 15-20 minutes since the half-life of endo-1,4- $\beta$ -mannanase was established to be of 6 minutes at  $85^{\circ}\text{C}$  [80].

##### *4.5.1. Digestion and identification of oligosaccharides released from LBG*

During digestion of locust bean gum with 100 mU of endo-1,4- $\beta$ -mannanase, aliquots were collected after 5 minutes, 30 minutes, 1, 3, 5, 8, 24 and 48 hours. The obtained mass spectra are reported in Fig. 37. Just considering the mass range between 600 and 2000 Da it was possible to observe how the rate of hydrolysis was linked to the timing. After only 5 minutes digestion (Fig. 37B) some oligosaccharides with a molecular weight up to about 1000 Da were released. Ions with  $m/z$  of 705.38, 867.44 and 1029.50 presented a mass difference of 162.06 Da indicating that a loss of hexoses (galactose and mannose) from the polysaccharide occurred. Between 5 minutes and 1 hour a significant increase in the rate of hydrolysis was observed (Fig. 37C). Considering the  $m/z$  values 669.39, 831.46, 993.52, 1155.59, 1317.65, 1479.71, 1641.77, 1803.83, the mass difference between neighboring peaks were of 162.06 Da so oligosaccharides with higher molecular weight were released. Some more oligosaccharides with a MW included between 200 and 2600 Da were also present but their intensity in the mass spectrum was fairly low (enlarged image of Fig. 37C). Similar results were obtained after 3 hours and 5 hours digestion (Fig. 37D). Incubation of locust bean gum with endo-1,4- $\beta$ -mannanase for 8 hours led the release of higher molecular weight oligosaccharides (Fig. 37E). Ions at  $m/z$  2452.09 and 2614.15 were observed in the mass spectrum

(marked area in the enlarged image) while, after longer time digestion (24 and 48 hours, Fig. 37F and G), the above-mentioned oligosaccharides were barely detectable because of their low intensity. It is reasonable that, if the sample is incubated for too long (over 8 hours), the enzymes begins to hydrolyze the already released oligosaccharides so a decrease in the molecular weight is observed. After 24 and 48 hours digestion the mass profile of the gum was still recognizable in terms of  $m/z$  values detected, with the only difference regarding the intensity of the peaks. Unfortunately any explanation could be found for this behavior. In conclusion a digestion time of 8 hours was considered a good compromise since it allowed to get the maximum number of oligosaccharides in a relative short time.

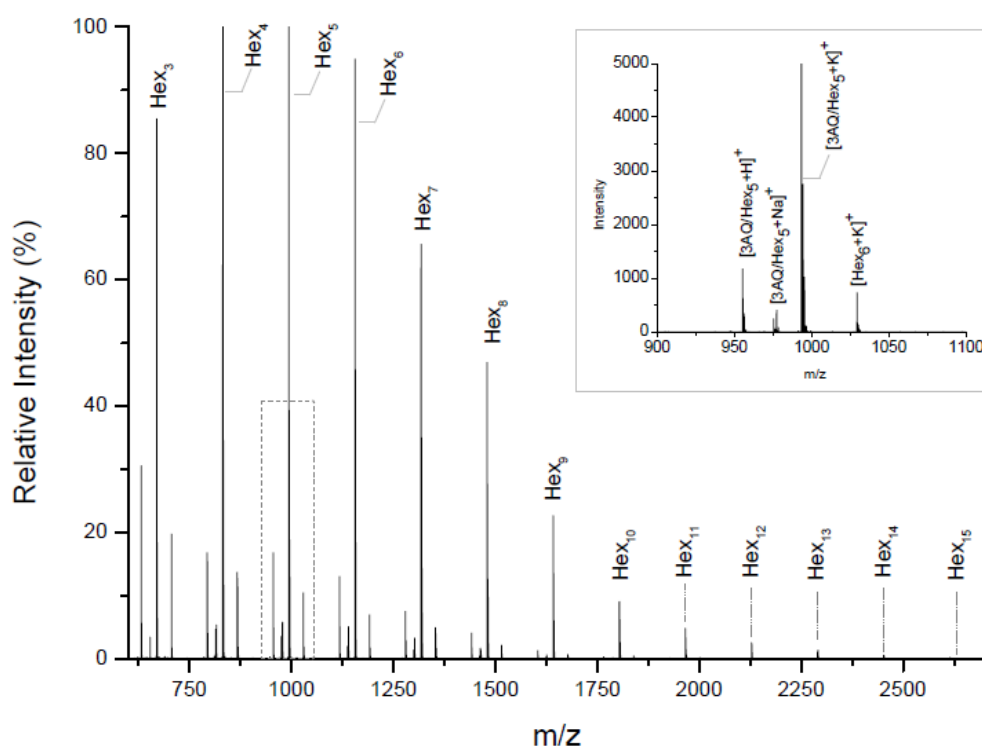


**Fig. 37.** MALDI-TOF spectra of LBG (A) not digested and digested with endo-1,4- $\beta$ -mannanase for (B) 5 minute; (C) 1 hour; (D) 5h; (E) 8h; (F) 24 hours and (G) 48h. On the right corner of each spectrum the enlarged mass range 2000-3000Da is shown.

Even if promising results were obtained digesting the gum only with endo-1,4- $\beta$ -mannanase, incubation with  $\alpha$ -galactosidase, for 24 hours, followed by endo-1,4- $\beta$ -mannanase (5 minutes, 30 minutes, 1, 3, 5, 8, 2 and 48 hours) was performed and the obtained mass spectra compared with the ones reported in Fig. 37. MALDI-TOF results showed the presence of ions at the same  $m/z$  values but just few oligosaccharides were observed in the mass range over 2000 Da. It is therefore

possible that  $\alpha$ -galactosidase significantly facilitate the access of mannanase to the backbone, therefore increasing the extent of hydrolysis. However, the objective is not to digest completely the gum, but to obtain a certain number of oligosaccharides in the mass spectrum that allow to define the gum fingerprint. For this reason, a simpler incubation using only endo-1,4- $\beta$ -mannanase was preferred.

Once the enzymatic procedure was established, attention focused on the identification of the released oligosaccharides. Since the polysaccharide is mostly made of mannose and galactose, which are both hexose monosaccharides, all the peaks of the mass spectrum differ of 162.06 Da. Thus, once an oligosaccharide is assigned to a specific ion, it is possible to reconstruct the entire profile simply adding or removing one hexose. The MS fingerprint of locust bean gum digested with endo-1,4- $\beta$ -mannanase for 8 hours is showed in Fig. 38.



**Fig. 38.** MALDI-T MALDI-TOF spectrum of LBG digested with endo-1,4- $\beta$ -mannanase for 8 hours. On the right corner: zoom of the range 900-1100 Da with the assigned oligosaccharides.

The mass spectrometric profile of digested locust bean gum is completely different from the profiles of the other arabinogalactan plant gums (gum arabic, tragacanth gum and cherry gum). The mass spectrum is characterized by neighboring ions that differ always of 162.06 Da. The presence of oligosaccharides, with 3 up to 15 hexoses, were identified and the charged with Na and K was supposed (see enlarged mass spectrum in Fig. 38). The only MALDI mass spectrum of digested LBG (and guar) was reported by Kurakake *et al.* ([78]). Gum was previously fractionated by anion-exchange chromatography and, after digestion with endo-1,4- $\beta$ -mannanase for 24 hours, the released oligosaccharides were spotted with DHB and analyzed in reflector mode. Beyond the profile similarity, the few masses reported in the paper correspond to the ones observed in the mass spectrum, even if derivatized with 3-aminoquinoline [3AQ/M+Na]<sup>+</sup> (e.g. 653.41, 815.47, 977.54,

1139.60 etc.). Therefore, if the ion at  $m/z$  977.54 was attributed to the derivatized oligosaccharide  $\text{Hex}_5$  [3AQ/M+Na]<sup>+</sup>, the one at  $m/z$  955.52 (mass difference = -22Da) was supposed to be the protonated form [M+H]<sup>+</sup> and the ion at  $m/z$  993.53 (mass difference = +16Da) the [M+K]<sup>+</sup>. The  $m/z$  values at 705.38, 867.44, 1029.50 (which continue with the other peaks at 1191.56, 1353.62, 1515.69, 1677.74, 1839.80 and 2001.88 Da), are consistent with the ions reported by Simões *et al.* [56] identified in the ESI-MS/MS spectra of LBG digested with endo-1,4- $\beta$ -mannanase for 48h. For example, the ion [Hex<sub>5</sub>+Na]<sup>+</sup> ( $m/z$  851) identified by ESI, corresponded to the non derivatized ion at  $m/z$  867.44 observed in the MALDI mass spectrum [Hex<sub>5</sub>+Na]<sup>+</sup>. According to these observations, a single oligosaccharide could be assigned to each of the ions observed in the digested LBG mass spectrum. A final list is reported in the table below (Tab. 11).

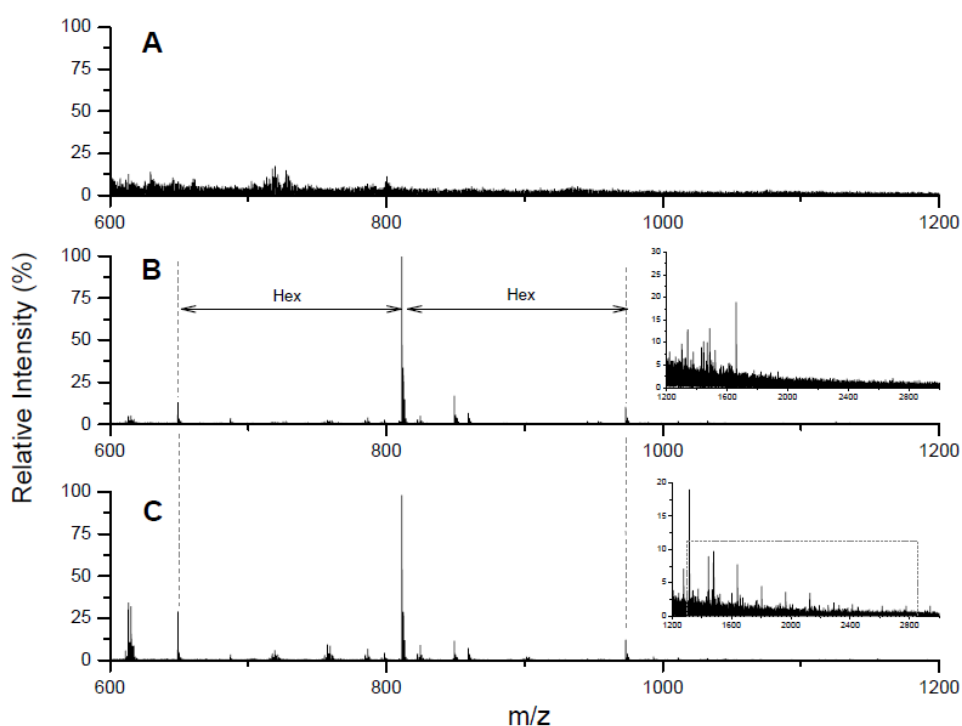
**Tab. 11.** List of the assigned oligosaccharides for digested LBG. (<sup>a</sup>) the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline, H<sup>+</sup>, Na<sup>+</sup> or K<sup>+</sup>).

Experimental mass [Da]	Theoretical mass [Da] <sup>a</sup>	$\Delta m$ [Da]	Possible oligosaccharide
631.4239	631.2112	0.21	Hex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
653.4094	653.1932	0.22	Hex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
669.3915	669.1671	0.22	Hex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
705.3797	705.1856	0.19	Hex <sub>4</sub> [M+K] <sup>+</sup>
793.4733	793.2641	0.21	Hex <sub>4</sub> [3AQ/M+H] <sup>+</sup>
815.4742	815.2460	0.23	Hex <sub>4</sub> [3AQ/M+Na] <sup>+</sup>
831.4611	831.2199	0.24	Hex <sub>4</sub> [3AQ/M+K] <sup>+</sup>
867.4442	867.2384	0.21	Hex <sub>5</sub> [M+K] <sup>+</sup>
955.5237	955.3169	0.21	Hex <sub>5</sub> [3AQ/M+H] <sup>+</sup>
977.5414	977.2988	0.24	Hex <sub>5</sub> [3AQ/M+Na] <sup>+</sup>
993.5258	993.2728	0.25	Hex <sub>5</sub> [3AQ/M+K] <sup>+</sup>
1029.5043	1029.2910	0.21	Hex <sub>6</sub> [M+K] <sup>+</sup>
1117.5739	1117.3700	0.20	Hex <sub>6</sub> [3AQ/M+H] <sup>+</sup>
1139.6024	1139.3520	0.25	Hex <sub>6</sub> [3AQ/M+Na] <sup>+</sup>
1155.5873	1155.3260	0.26	Hex <sub>6</sub> [3AQ/M+K] <sup>+</sup>
1191.5620	1191.3440	0.22	Hex <sub>7</sub> [M+K] <sup>+</sup>
1279.6315	1279.4230	0.21	Hex <sub>7</sub> [3AQ/M+H] <sup>+</sup>
1301.6532	1301.4040	0.25	Hex <sub>7</sub> [3AQ/M+Na] <sup>+</sup>
1317.6488	1317.3780	0.27	Hex <sub>7</sub> [3AQ/M+K] <sup>+</sup>
1353.6250	1353.3970	0.23	Hex <sub>8</sub> [M+K] <sup>+</sup>
1441.6869	1441.4750	0.21	Hex <sub>8</sub> [3AQ/M+H] <sup>+</sup>
1463.7179	1463.4570	0.26	Hex <sub>8</sub> [3AQ/M+Na] <sup>+</sup>
1479.7102	1479.4310	0.28	Hex <sub>8</sub> [3AQ/M+K] <sup>+</sup>
1515.6871	1515.4500	0.24	Hex <sub>9</sub> [M+K] <sup>+</sup>
1603.7467	1603.5280	0.22	Hex <sub>9</sub> [3AQ/M+H] <sup>+</sup>
1625.7755	1625.5100	0.27	Hex <sub>9</sub> [3AQ/M+Na] <sup>+</sup>
1641.7695	1641.4840	0.29	Hex <sub>9</sub> [3AQ/M+K] <sup>+</sup>
1677.7441	1677.5030	0.24	Hex <sub>10</sub> [M+K] <sup>+</sup>
1765.8083	1765.5810	0.23	Hex <sub>10</sub> [3AQ/M+H] <sup>+</sup>
1787.8324	1787.5630	0.27	Hex <sub>10</sub> [3AQ/M+Na] <sup>+</sup>
1803.8328	1803.5370	0.30	Hex <sub>10</sub> [3AQ/M+K] <sup>+</sup>
1839.8042	1839.5550	0.25	Hex <sub>11</sub> [M+K] <sup>+</sup>
1965.8973	1965.5900	0.31	Hex <sub>11</sub> [3AQ/M+K] <sup>+</sup>
2001.8806	2001.6080	0.27	Hex <sub>12</sub> [M+K] <sup>+</sup>
2127.9653	2127.6430	0.32	Hex <sub>12</sub> [3AQ/M+K] <sup>+</sup>
2290.0317	2289.6950	0.34	Hex <sub>13</sub> [3AQ/M+K] <sup>+</sup>
2452.0933	2451.7480	0.35	Hex <sub>14</sub> [3AQ/M+K] <sup>+</sup>
2614.1550	2613.8010	0.35	Hex <sub>15</sub> [3AQ/M+K] <sup>+</sup>



## 4.5.3. Digestion and identification of oligosaccharides released from guar gum

Since guar gum is supposed to have the same polysaccharide structure of locust bean gum, digestion with 100 mU of endo-1,4- $\beta$ -mannanase was performed and aliquots after 5 minutes, 30 minutes, 1, 3, 5, 8, 24 and 48 hours were collected. Besides the mass spectrum of the non-digested gum, for which only the matrix ions were observed, the most significant mass spectra of digested gum are reported in Fig. 39. Any difference was observed between 5 minutes (Fig. 39B), 1 hour and 3 hours digestion. Interesting results were instead obtained after 5 hours incubation since some ions at higher molecular weight were observed (enlarged image of Fig. 39C). Longer digestion (24 and 48 hours) did not result in any increase of the rate of hydrolysis.



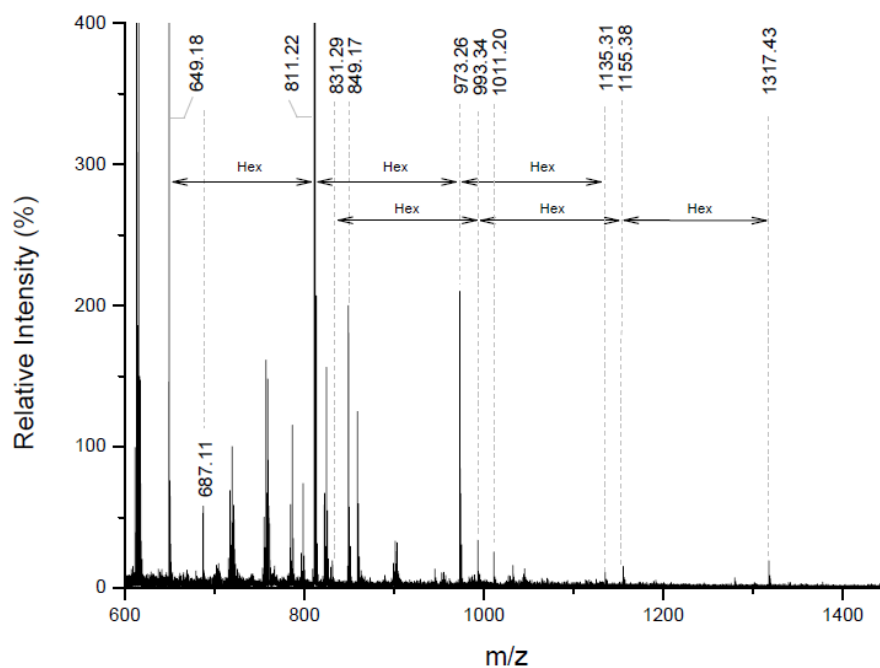
**Fig. 39.** MALDI-TOF spectra of guar gum (A) not digested and digested with endo-1,4- $\beta$ -mannanase for (B) 5 minutes; (C) 5h. On the right corner the corresponding enlarged mass range 1200-3000Da is shown.

Just considering the mass range between 600 and 1200 Da, it is possible to observe how the profile is apparently different from the LBG's one. The mass difference of the three most intense peaks, at a  $m/z$  value of 649.18, 811.23 and 973.26, demonstrates the loss of a hexose (galactose or mannose). Even if the two galactomannans gums are supposed to have the same structure, except for the number of single galactose units as side chains, the above mentioned peaks were not observed in the mass spectrum of digested locust bean gum and the other observable ions had a very low intensity. For this reason more experiments were performed.

Gum was first digested with an higher amount of enzyme (1U and 3 U/mg of gum) to see if the rate of hydrolysis could increase. In addition the sample was digested with both  $\alpha$ -galactosidase, for 24 hours, followed by endo-1,4- $\beta$ -mannanase, to check if the removal of the Gal side chains by the first enzyme could increase the digestion of the main chain by mannanase. Results showed that the amount of enzyme and the combination of  $\alpha$ -GAL and  $\beta$ -MAN did not improve the digestion of

guar gum. This result suggested that the enzyme action on this type of substrate is maybe more restricted than the one on locust bean gum.

Therefore the selected digestion procedure was: 100 mU of endo-1,4- $\beta$ -mannanase for 5 hours. A zoom of the corresponding mass spectrum is shown in Fig. 40.



**Fig. 40.** MALDI-TOF spectrum of guar gum digested with endo-1,4- $\beta$ -mannanase for 5 hours (mass range 600-1450 Da)

The mass spectrum was characterized by neighboring ions that differed always of 162.05 Da. The  $m/z$  values 831.29, 993.34, 1155.38 and 1317.43 were already observed in the mass spectrum of digested LBG and they corresponded respectively to  $\text{Hex}_4$ ,  $\text{Hex}_5$ , and  $\text{Hex}_6$   $[\text{3AQ/M+K}]^+$ . The most intense ion at  $m/z$  811.23 did not correspond to any hexose combination, thus representing a possible difference between the mass profile of LBG and guar gum.

According to the study of the MS spectrum, some of the ions could be assigned to a specific oligosaccharide. A final list is reported in the Table 12.

**Tab.12.** List of the assigned oligosaccharides for digested guar gum. (<sup>a</sup>) the theoretical mass was calculated adding the monoisotopic mass of a molecule of water to the oligosaccharide residue mass (considering also the eventual presence of 3-aminoquinoline,  $\text{H}^+$ ,  $\text{Na}^+$  or  $\text{K}^+$ ).

Experimental mass [Da]	Theoretical mass [Da] <sup>a</sup>	$\Delta m$ [Da]	Possible oligosaccharide
649.1770	649.2191	-0.04	$\text{Pent}_2\text{HexAcHex} [\text{M+H}]^+$
687.1132	687.1750	-0.06	$\text{Pent}_2\text{HexAcHex} [\text{M+K}]^+$
811.2254	811.2719	-0.05	$\text{Pent}_2\text{Hex}_2\text{AcHex} [\text{M+H}]^+$
831.2849	831.2199	0.06	$\text{Hex}_5 [\text{3AQ/M+K}]^+$
849.1661	849.2278	-0.06	$\text{Pent}_2\text{Hex}_2\text{AcHex} [\text{M+K}]^+$
973.2606	973.3248	-0.06	$\text{Pent}_2\text{Hex}_3\text{AcHex} [\text{M+H}]^+$

993.3378	993.2728	0.07	Hex <sub>5</sub> [3AQ/M+K] <sup>+</sup>
1011.203	1011.2810	-0.08	Pent <sub>2</sub> Hex <sub>3</sub> AcHex [M+K] <sup>+</sup>
1135.3062	1135.3780	-0.07	Pent <sub>2</sub> Hex <sub>4</sub> AcHex [M+H] <sup>+</sup>
1155.6704	1155.3260	0.34	Hex <sub>6</sub> [3AQ/M+K] <sup>+</sup>
1317.4275	1317.3780	0.05	Hex <sub>7</sub> [3AQ/M+K] <sup>+</sup>

On regard of the two galactomannan gums it can be concluded that:

- Digestion with endo-1,4- $\beta$ -mannanase for 8 hours resulted to be a suitable procedure for both locust bean and guar gum.
- The MS pattern of both locust bean and guar gum is different from the one of gum arabic. The results are consequently promising in the view of gums discrimination according to their mass fingerprint. Furthermore some differences were observed also between LBG and guar gum, but further investigations have to be performed in order to completely define the mass fingerprint of the two gums.

In conclusion, application of partial enzymatic digestion on the polysaccharide structure of plant gums resulted in the release of characteristic oligosaccharides that could be detected by MALDI-MS. Among the several exo- and endo- acting enzymes tested, two of them resulted in the best digestion procedure and results: exo- $\beta$ -1,3-galactanase and endo-1,4- $\beta$ -mannanase. The former worked optimally for the plant gums resembling the arabinogalactan family structure (gum arabic, tragacanth gum and cherry gum), the latter resulted to be suitable for the digestion of the two gums with a galactomannan similar structure (locust bean and guar gum). A mass profile was obtained for each gum and the attribution of several ions to a corresponding oligosaccharide was achieved by MS spectra interpretation. Data showed how it is possible to distinguish the plant gums under investigation according to their mass profile.

## 5. DEVELOPMENT OF AN ENZYME COCKTAIL

In the previous sections several enzymatic digestion protocols were optimized. MALDI-TOF analysis of the released oligosaccharides showed how each gum can be distinguished from the others according to its specific mass fingerprint. The focus of this section is to develop an enzyme cocktail that could be suitable for the digestion of all plant gums under investigation. This step is necessary in prospect of the analysis of a real art/archaeological sample, whose composition is usually unknown. The enzymes choice (5.1), followed by the analysis of digested plant gums and the corresponding results (5.2) will be discussed. The results suggest that, using the invented enzyme cocktail, digestion of gums occurs resulting in the characteristic mass fingerprints. Therefore, even if the type of polysaccharide is not known, the developed methodology seems to allow the discrimination of the different gums.

### 5.1. COCKTAIL FORMULATION AND DIGESTION PROCEDURE

Data of the previous section showed how two main enzymes are involved in plant gums digestion: exo- $\beta$ -1,3-galactanase, for digestion of arabinogalactan like polysaccharides (gum arabic, tragacanth gum and cherry gum), and endo-1,4- $\beta$ -mannanase, for hydrolysis of galactomannan polysaccharides (locust bean gum and guar gum). These enzymes were therefore selected for the formulation of an enzyme cocktail. Due to their different buffer, pH and temperature stability, specific operative conditions for sample preparation and incubation were selected (see Tab. 13).

**Tab. 13.** Description of the enzyme cocktail composition and corresponding operative conditions.

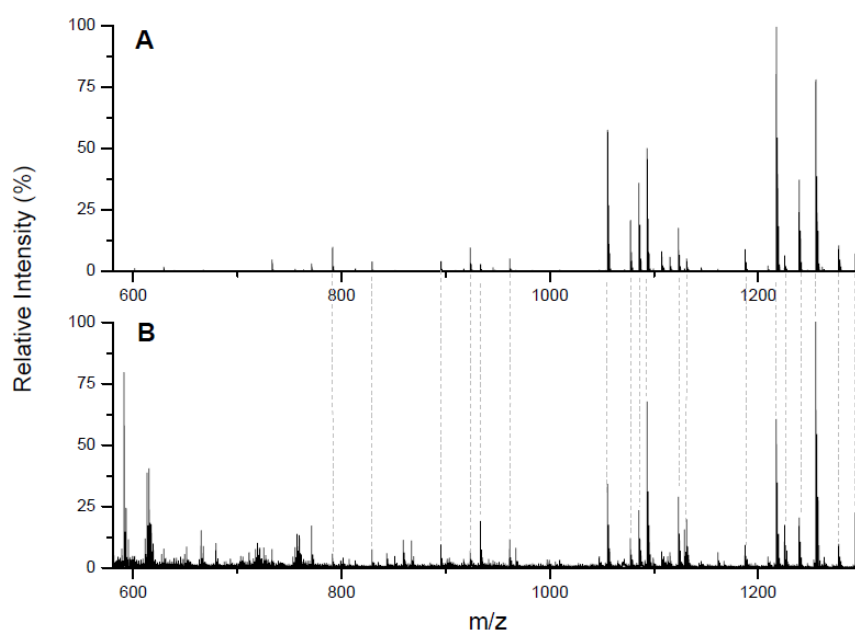
Enzyme	Temperature stability [°C]	pH stability	Operative conditions
exo- $\beta$ -1,3-galactanase	<55	3<pH<10	Phosphate buffer 50mM
endo-1,4- $\beta$ -mannanase	<50	6.5<pH<8	pH=7, incubation at 45°C

100 mU of the two enzymes were added together to the sample, previously resuspended in the selected buffer, and aliquots were collected after 5, 24 and 48 hours of incubation. The enzymes were inactivated at 100°C for 15 minutes and the sample analyzed by MALDI-TOF in reflector mode. All plant gums under investigation were digested with the enzyme cocktail and results are reported in the following section.

### 5.2. RESULTS AND DISCUSSIONS

Among the three tested digestion times, the most significant results were obtained after 24 hours digestion. The corresponding mass spectra are reported below. For each gum, the spectrum obtained after digestion with the enzyme cocktail is compared with the one obtained with the previously optimized protocol.

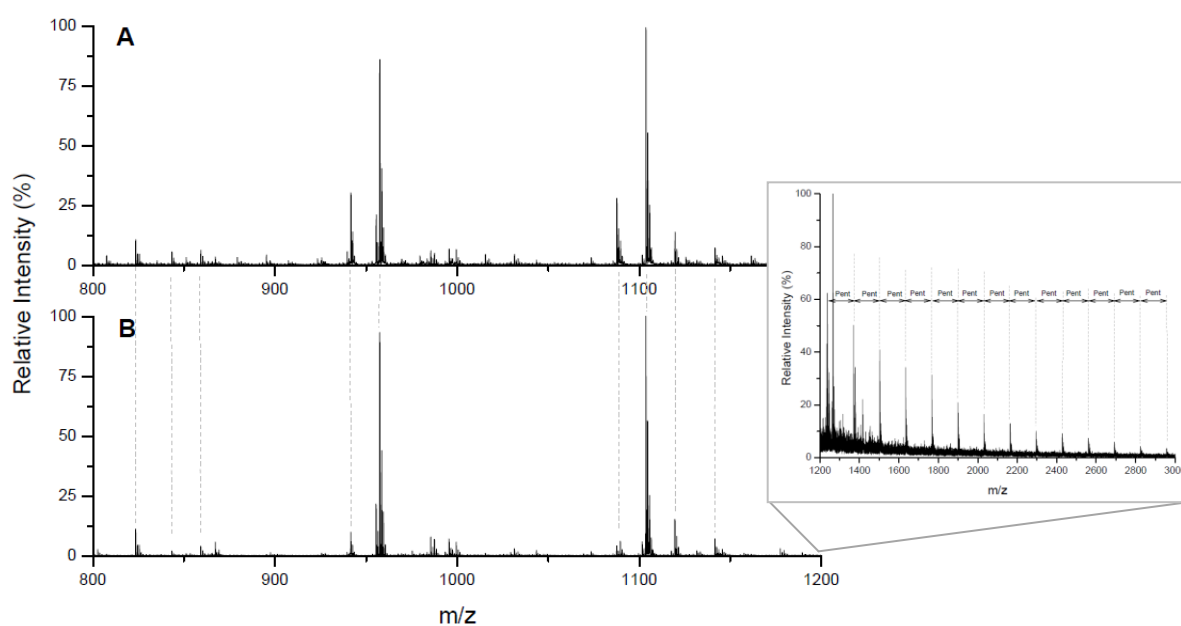
## 5.2.1. Gum arabic



**Fig. 41.** Comparison of MALDI-TOF spectra of gum arabic (Zecchi) digested for 24 hours with: (A) exo- $\beta$ -1,3-galactanase and (B) enzyme cocktail.

The mass spectrum of gum arabic, digested with the enzyme cocktail (Fig. 41B), perfectly matched with the profile obtained with the optimized digestion protocol which implied the use of only exo- $\beta$ -1,3-galactanase. Most of the released oligosaccharides which are characteristic of the gum arabic profile (see Tab. 7) were also identified in the new mass spectrum. Data showed that the new digestion conditions were optimal for the two enzymes. In addition the presence of the enzyme endo-1,4- $\beta$ -mannanase did not interfere with the exo- $\beta$ -1,3-GAL action on the arabinogalactan polysaccharide component.

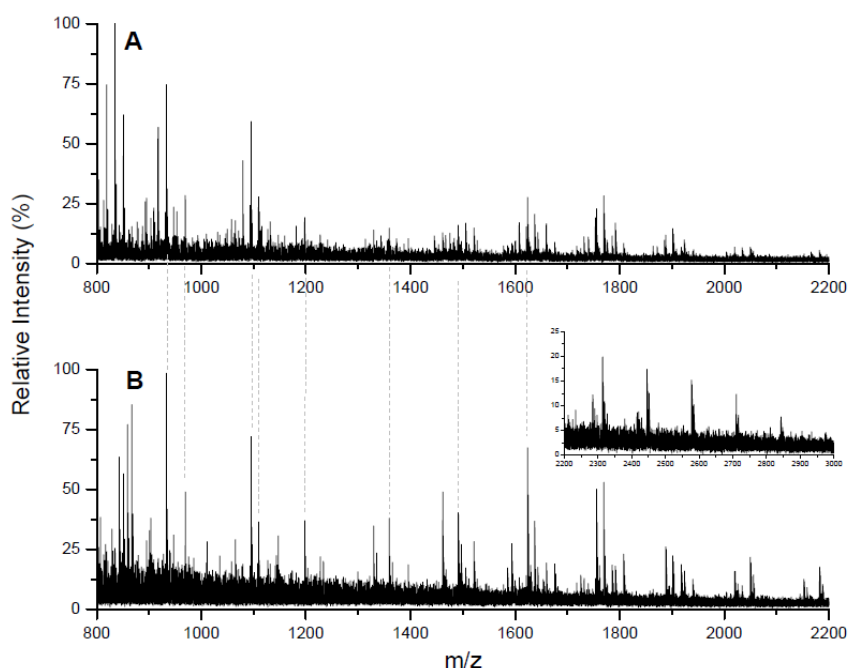
## 5.2.2. Tragacanth gum



**Fig. 42.** Comparison of MALDI-TOF spectra of tragacanth gum (Bresciani) digested for 24 hours with: (A) exo- $\beta$ -1,3-galactanase and (B) enzyme cocktail (with enlarged image of mass range 1200-3000 Da).

After digestion with the enzyme cocktail, the two characteristic intense peaks at  $m/z$  957.57 and 1103.63 were observed (Fig. 42B). Besides the other identified ions, which corresponded to the tragacanth gum profile, the presence of neighboring ions characterized by a mass difference of 132.04 Da were detected (enlarged image). The series went from the  $m/z$  value at 1239.47 up to 2692.00. This result indicates that, with the present digestion conditions, the enzyme  $\text{exo-}\beta\text{-1,3-GAL}$  could digest a long chain made of pentose. This result is consistent with the information on gum tragacanth polysaccharide structure reported in the literature.

### 5.2.3. Cherry gum

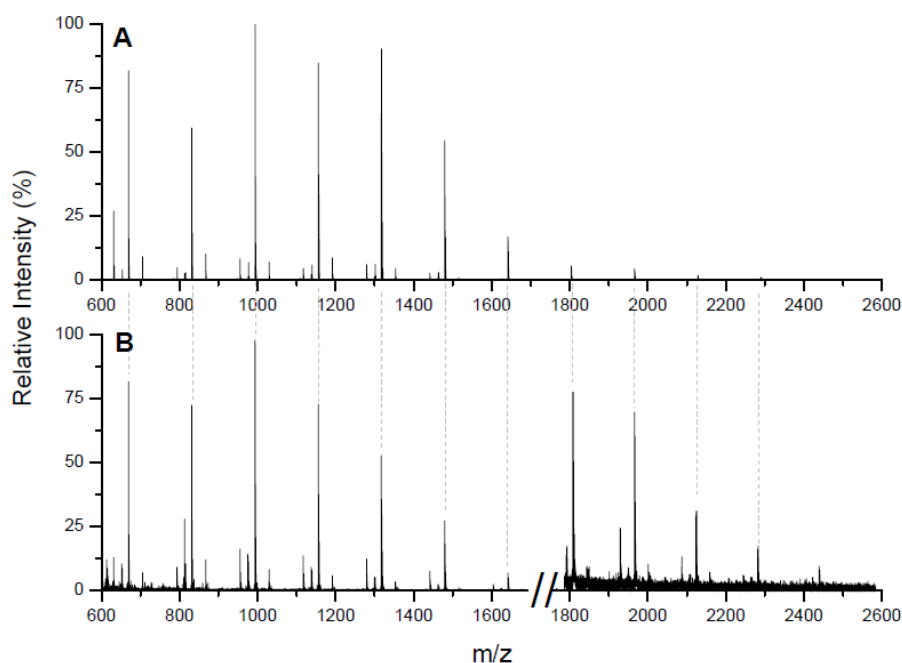


**Fig. 43.** Comparison of MALDI-TOF spectra of cherry gum n.2 digested for 24 hours with: (A)  $\text{exo-}\beta\text{-1,3-galactanase}$  and (B) enzyme cocktail (with enlarged image of mass range 2200-3000Da)

The mass spectrum of cherry gum, obtained after digestion with the cocktail (Fig. 43B), showed the same complexity observed after digestion performed using only  $\text{exo-}\beta\text{-1,3-galactanase}$ . However, the ions were the same. In addition, besides the already observed ions that differ of 132.04 Da ( $m/z$  values 1491.51, 1623.55, 1755.59, 1887.63, 2019.67 and 2151.71), ions with an higher molecular weight were observed. These peaks belonged to another pentose chain ( $m/z$  values 2049.70, 2181.75, 2313.79, 2445.82, 2577.88, 2709.94 and 2841.96), that was not observed in the mass spectrum of the gum digested with  $\text{exo-}\beta\text{-1,3-galactanase}$ . However these oligosaccharides are consistent with the possible presence of arabinose chains in the gum polysaccharide structure. The mass profile of cherry gum was therefore confirmed.

## 5.2.4. Locust bean gum

Figure 44 shows how the  $m/z$  values observed in the two MALDI-MS profiles of digested LBG are the same. This data confirmed that the experimental digestion conditions are optimal not only for  $\text{exo-}\beta\text{-1,3-galactanase}$ , as already verified for the three previous gums, but also for the  $\text{endo-1,4-}\beta\text{-mannanase}$ . Neighboring ions that differed for a value of 162.05 Da were observed up to a  $m/z$  value of 2451.64, which corresponded to an oligosaccharide with 14 hexose units. Therefore, the new digestion cocktail proved to work optimally and the oligosaccharide mass profile of locust bean gum was confirmed.



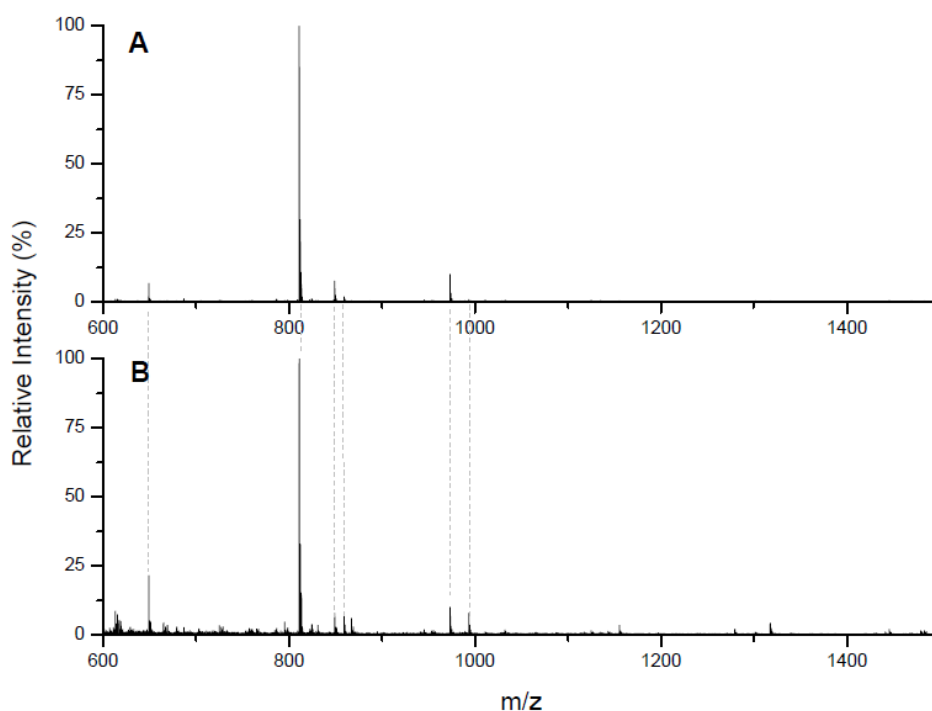
**Fig. 44.** Comparison of MALDI-TOF spectra of LBG (Sigma) digested for 24 hours with: (A)  $\text{endo-}\beta\text{-1,4-mannanase}$  and (B) enzyme cocktail.

A further experiment was performed in order to check if the enzyme  $\text{exo-}\beta\text{-1,3-galactanase}$ , which is present in the cocktail, was involved or not in the LBG digestion. The gum was incubated only with this enzyme and two ions of the profile could be observed:  $m/z$  993.34 and 1155.37 (mass difference = 162.03 Da). All the other ions were, instead, not observed. Since the presence of highly branched side chains, containing more units of galactose, was described for some galactomannans [81], it was reasonable to speculate that the observed ions corresponded to galactose side chains released by  $\text{exo-}\beta\text{-1,3-galactanase}$ . However different analytical techniques should be used to verify this hypothesis.

## 5.2.5. Guar gum

Similar results were obtained for guar gum (Fig. 45). The oligosaccharides released after digestion with the enzyme cocktail were the same that the ones obtained after digestion with  $\text{endo-1,4-}\beta\text{-}$

mannanase. Therefore the protocol resulted to be efficient also for the identification of the mass profile of guar gum.



**Fig. 45.** Comparison of MALDI-TOF spectra of guar gum (Sigma) digested for 24 hours with: (A) endo- $\beta$ -1,4-mannanase and (B) enzyme cocktail.

In conclusion, the developed enzyme cocktail resulted to work optimally for both arabinogalactan like gums (gum arabic, tragacanth and cherry gums), and for galactomannans polysaccharides (LBG and guar gum). The digestion procedure allowed to obtain the already established fingerprints for each of the analyzed gums, mainly revealing the difference between gum arabic and all the other gums. This result is of significant remark since it is the proof that, with this method, it would be possible to analyze unknown samples and being able to discriminate the different plant gums.



## 6. APPLICATION ON REAL SAMPLES

In the previous section, a new enzymatic digestion protocol was optimized and it consisted on incubation of the sample with a mixture of exo- $\beta$ -1,3-galactanase and endo-1,4- $\beta$ -mannanase for 24 hours. MALDI-TOF analysis of the released oligosaccharides showed how the mass saccharide fingerprint of each gum was preserved. It is therefore possible to discriminate gum arabic from tragacanth gum, cherry, LBG and guar gum. Since the protocols have been so far optimized on pure gum samples, the focus of this section is to apply the developed procedure to samples characterized by an increasing complexity. Analysis of a 50 years old gum arabic sample (6.1), followed by the application of the technique to the analysis of fresh watercolors (6.2) and finally, to an old watercolor sample (6.3) will be discussed. The mentioned samples were selected in order to evaluate if the presence of other organic and/or inorganic materials could interfere with the analysis and, therefore, alter the mass profile. Furthermore aged materials (pure gum and watercolors) were analyzed to verify if the developed method is applicable to ancient samples. Results suggest that, even if no clean up procedure is used, the gum arabic profile is not modified by the presence of other materials and it appears preserved even in aged samples. Furthermore, the gum arabic characteristic profile was successfully obtained for a watercolor sample dating 1870.

### 6.1. A 50 YEARS OLD GUM ARABIC SAMPLE

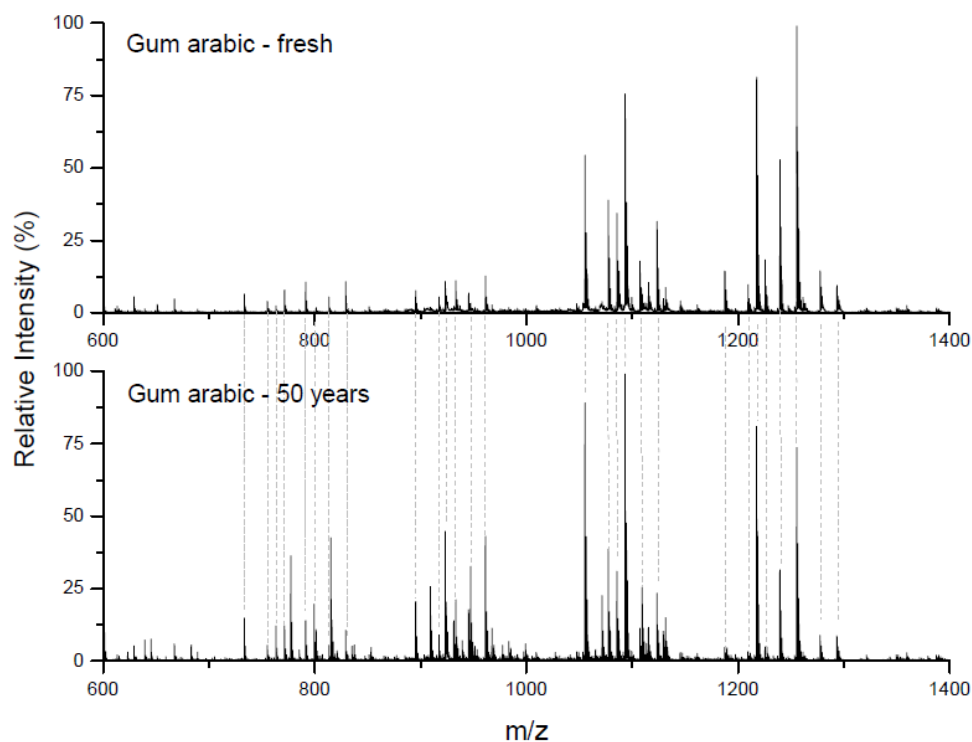
#### *6.1.1. Composition and sample preparation*

A 50 years old gum arabic sample was supplied by the Paper Conservation Department of the Metropolitan Museum of Art (New York). Gum was solubilized in water over night at room temperature, the solution was centrifuged at 13400 rpm for 20 minutes and an aliquot corresponding to 1 mg of gum was collected. Solution was dried in a vacuum-centrifuge at 30°C and subjected to enzymatic digestion with exo- $\beta$ -1,3-galactanase for 24 hours. The corresponding mass spectrum is reported below.

#### *6.1.2. Results and discussion*

The MALDI-MS spectra of respectively gum arabic (Zecchi) and the 50 years old gum arabic sample, digested with 100 mU of exo- $\beta$ -1,3-galactanase, are reported in Fig. 46.

The mass spectrum of the old gum arabic sample revealed to be consistent with the profile of standard gum arabic (Zecchi), even in terms of peak intensity. All the characteristic masses, corresponding to the oligosaccharides released specifically from the gum by enzymatic digestion, were obtained (ions listed in Tab. 7) with a  $m/z$  error  $\leq 0.2\text{Da}$  (see Tab. 14). Results demonstrated how the developed strategy will allow to univocally identify the gum arabic even in the presence of aged materials. Analysis revealed how the oligosaccharide mass spectrum of the gum resulted to be unchanged, thus confirming the possibility to employ the mass spectrum as a gum arabic fingerprint.



**Fig. 46.** Mass spectra comparison of gum arabic (Zecchi) and a 50 years old gum arabic sample (Metropolitan Museum of Art, NY) digested with  $\text{exo-}\beta\text{-1,3-galactanase}$  for 24 hours.

**Tab. 14.** List of the gum arabic oligosaccharides recognized in the digested 50 years old gum arabic sample. The oligosaccharides which are not present in the mass spectrum are indicated by ‘/’.

Experimental mass [Da]	Theoretical mass [Da]	$\Delta m$ [Da]	Possible oligosaccharide
601.3745	601.2007	0.17	PentHex <sub>2</sub> [3AQ/M+H] <sup>+</sup>
623.3514	623.1826	0.17	PentHex <sub>2</sub> [3AQ/M+Na] <sup>+</sup>
629.2803	629.1956	0.08	HexdHexHexA [3AQ/M+H] <sup>+</sup>
639.2266	639.1566	0.07	PentHex <sub>2</sub> [3AQ/M+K] <sup>+</sup>
651.3426	651.1775	0.17	HexdHexHexA [3AQ/M+Na] <sup>+</sup>
667.3340	667.1515	0.18	HexdHexHexA [3AQ/M+K] <sup>+</sup>
733.4217	733.2429	0.18	Pent <sub>2</sub> Hex <sub>2</sub> [3AQ/M+H] <sup>+</sup>
755.3169	755.2249	0.09	Pent <sub>2</sub> Hex <sub>2</sub> [3AQ/M+Na] <sup>+</sup>
763.3727	763.2535	0.12	PentHex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
771.3931	771.1988	0.19	Pent <sub>2</sub> Hex <sub>2</sub> [3AQ/M+K] <sup>+</sup>
/	785.2354	/	PentHex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
791.3426	791.2484	0.09	Hex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
801.3474	801.2094	0.14	PentHex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
813.3859	813.2304	0.16	Hex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
829.3505	829.2043	0.15	Hex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
895.4801	895.2958	0.18	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
917.4789	917.2777	0.20	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
923.4576	923.2907	0.17	PentHex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
933.4576	933.2516	0.21	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
945.4016	945.2726	0.13	PentHex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
961.4457	961.2465	0.20	PentHex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1055.5233	1055.3330	0.19	Pent <sub>2</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1077.5167	1077.3150	0.20	Pent <sub>2</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1085.5376	1085.3430	0.19	PentHex <sub>3</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1093.4911	1093.2890	0.20	Pent <sub>2</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>

1107.5076	1107.3250	0.18	PentHex <sub>3</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1123.5014	1123.2990	0.20	PentHex <sub>3</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1187.5662	1187.3750	0.19	Pent <sub>3</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1209.5230	1209.3570	0.17	Pent <sub>3</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1217.5852	1217.3860	0.20	Pent <sub>2</sub> Hex <sub>3</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1225.5333	1225.3310	0.20	Pent <sub>3</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1239.5634	1239.3680	0.20	Pent <sub>2</sub> Hex <sub>3</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1255.5447	1255.3420	0.20	Pent <sub>2</sub> Hex <sub>3</sub> dHexHexA [3AQ/M+K] <sup>+</sup>

## 6.2. FRESH WATERCOLOR SAMPLES

### 6.2.1. Composition and sample preparation

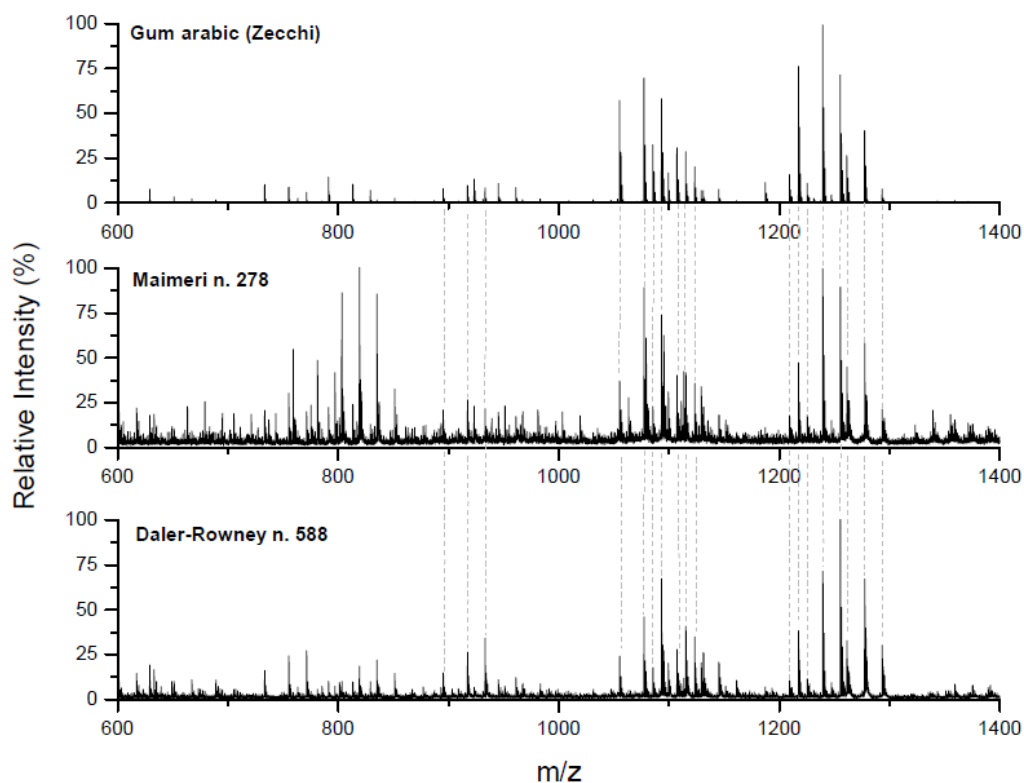
Fresh watercolors were purchased from different companies and their specific composition is reported in the table below (Tab. 15). Samples were prepared in milliQ water at a concentration of 10 mg/mL (w/v). The solution was mixed for 3 hours until the sample was completely solubilized and the solution was then centrifuged for 4-5 times (13400 rpm, 20 minutes) in order to remove the pigment. The supernatant was collected, dried in a vacuum-centrifuge at 30°C and subjected to enzymatic digestion with both exo-β-1,3-galactanase and the enzyme cocktail for 24 hours. The resulted mass spectra are reported in the following section.

**Tab. 15.** Purchased watercolors, reference number and composition as reported by the company.

Company	Number	Composition
Maimeri (Italy)	n. 278 Burnt Sienna	Gum arabic (Kurdufan, Sudan), glycerin and pigment (PBr7 calcined natural Sienna).
Daler-Rowney (England)	n. 588 Vermilion	Gum Arabic, glycerin, extenders, water, wetting agents, preservatives and pigment (diarylamide orange P034 and Naphtol red AS-OL PR9).

### 6.2.2. Results and discussion

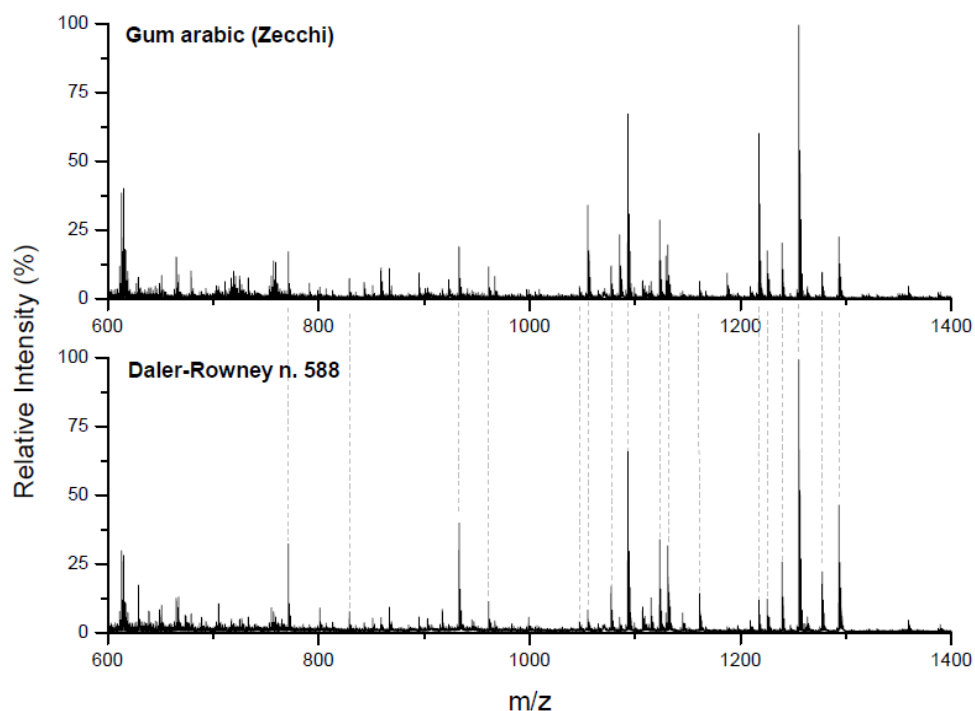
The presence of gum arabic in the watercolors was confirmed by the seller therefore, as first experiment, the samples were digested only with exo-β-1,3-galactanase for 24 hours. One of the problems encountered during digestion concerned the presence of the pigment. Multiple centrifugation steps did not allow a complete removal of the pigment particles, which were still visible after the sample was dried. Filter units in polyvinylidene difluoride (PVDF), 0.22 μm diameter (Millipore), were used to remove residual particles but, unfortunately, they resulted in a oligosaccharides loss. Therefore the extracted solution was analyzed without any clean-up step. The MALDI-MS spectra of respectively gum arabic (Zecchi) and the two watercolors samples digested with 100 mU of exo-β-1,3-galactanase are reported in Fig. 47.



**Fig. 47.** Mass spectra comparison of gum arabic (Zecchi), watercolor n. 278 (Maimeri) and watercolor n. 588 (Daler-Rowney) digested with exo- $\beta$ -1,3-galactanase for 24 hours.

The mass spectra of the two watercolor samples revealed the presence of gum arabic. If considering the most significant mass range (1000 - 1400 Da), the ions observed perfectly matched the profile of the gum obtained with the optimized digestion protocol. Same observations could not be drawn for the ions observed in the mass range that goes from 600 up to 1000 Da. Their intensity was significantly low and it was therefore difficult to identify and compare them to the gum arabic profile. Even if there is not a direct correlation between the analyte concentration and the peak intensity in MALDI-TOF spectrum [82], it was reasonable to speculate that the amount of gum arabic in the watercolor samples was close to the limit of detection. However, it has to be considered that these ions showed a low intensity even in the mass spectrum of pure gum. Therefore, the sole correspondence of the ions in the 1000-1400 Da range was considered to be significant for identifying gum arabic in the watercolor sample.

Same results were obtained when the fresh watercolor was digested with the enzyme cocktail. Its mass spectrum is showed in Fig. 48 and compared to the mass spectrum of gum arabic digested with the same protocol. The developed protocol showed to work optimally also with more complex samples since the typical fingerprint of gum arabic was revealed in the mass spectrum of the analyzed fresh watercolor.



**Fig. 48.** Mass spectra comparison of gum arabic (Zecchi) and watercolor n. 588 (Daler-Rowney) digested with enzyme cocktail (exo- $\beta$ -1,3-galactanase and endo-1,4- $\beta$ -mannanase) for 24 hours.

### 6.3. OLD WATERCOLOR SAMPLE

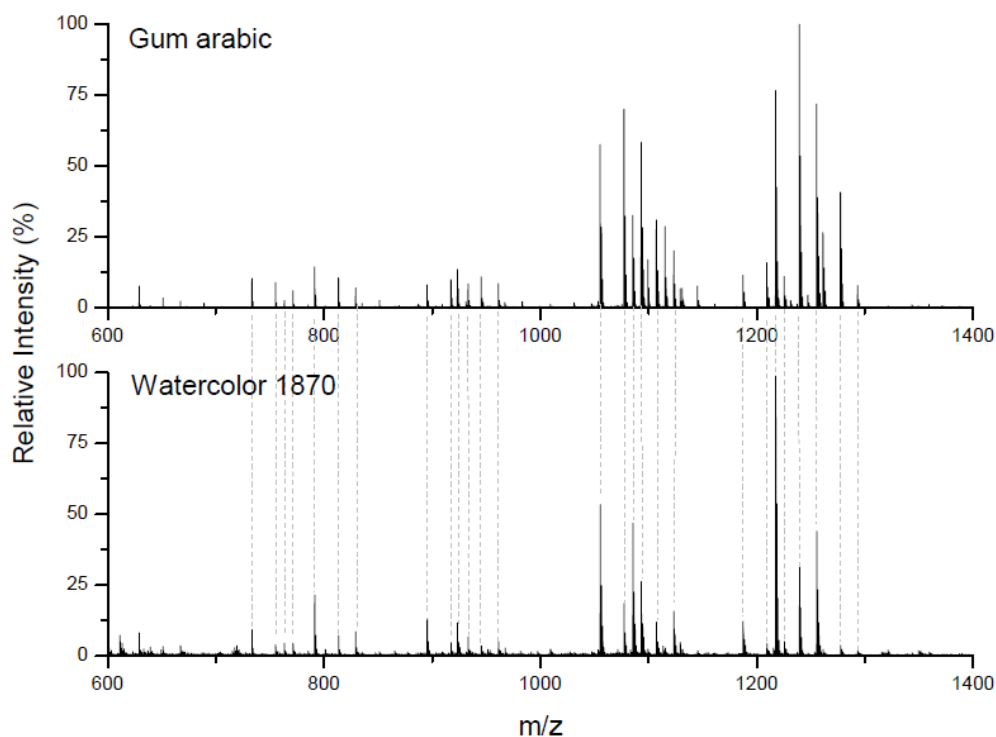
#### 6.3.1. Composition and sample preparation

An old watercolor sample, dating around 1870, was supplied by the Paper Conservation Department of the Metropolitan Museum of Art (New York). Beside the blue pigment (possibly Prussian blue), the presence of gum arabic was supposed. 5 mg of sample were solubilized in milliQ water by continuous mixing over night. Once the sample was completely solubilized, the solution was centrifuged for 2 times (13400 rpm, 20 minutes) in order to remove the pigment. The supernatant was collected, dried in a vacuum-centrifuge at 30°C and subjected to enzymatic digestion with exo- $\beta$ -1,3-galactanase for 24 hours. The resulted mass spectra are reported in the following section.

#### 6.3.2. Results and discussion

In regard to the sample preparation, pigment particles could be easily isolated from the solution by simple centrifugation. A clear solution was obtained indicating that, in comparison to the non-aged samples, the dimension of the pigment particles was higher.

The MALDI-MS spectra of respectively gum arabic (Zecchi) and the old watercolor sample, digested with 100 mU of exo- $\beta$ -1,3-galactanase, are reported in Fig. 49.



**Fig. 49.** Mass spectra comparison of gum arabic (Zecchi) and watercolor sample from 1870, digested with exo- $\beta$ -1,3-galactanase for 24 hours.

The mass spectrum of the old watercolor revealed the presence of gum arabic. The characteristic masses, corresponding to the oligosaccharides released from gum arabic, were observed (ions listed in Tab. 16). Some of the ions at lower molecular weight were missing but, as it occurred also for the fresh watercolor, it is possible to assume that they could not be detected due to the low amount of gum arabic in the sample. The only difference between the two mass spectra is related to the ion intensity. If considering the mass range 1200-1300 Da, it is possible to observe how for the watercolor the most intense peak was the ion at  $m/z$  1217.62. The ion was attributed to the oligosaccharide  $\text{Pent}_2\text{Hex}_3\text{dHexHexA}$   $[\text{3AQ/M+H}]^+$  and the corresponding  $\text{Na}^+$  and  $\text{K}^+$  charged oligosaccharides (1239.61 and 1255.59 Da) showed a lower intensity. On the contrary, the most intense peak in the mass spectrum of gum arabic corresponded to the  $[\text{3AQ/M+Na}]^+$  oligosaccharide, while the protonated molecule and the one with K adduct had lower intensity.

**Tab. 16.** List of the gum arabic oligosaccharides recognized in the digested old watercolor sample. The oligosaccharides which are not present in the mass spectrum are indicated by ‘/’.

Experimental mass [Da]	Theoretical mass [Da]	$\Delta m$ [Da]	Possible oligosaccharide
/	601.2007	/	$\text{PentHex}_2$ $[\text{3AQ/M+H}]^+$
/	623.1826	/	$\text{PentHex}_2$ $[\text{3AQ/M+Na}]^+$
629.3702	629.1956	0.17	$\text{HexdHexHexA}$ $[\text{3AQ/M+H}]^+$
/	639.1566	/	$\text{PentHex}_2$ $[\text{3AQ/M+K}]^+$
/	651.1775	/	$\text{HexdHexHexA}$ $[\text{3AQ/M+Na}]^+$
667.3432	667.1515	0.19	$\text{HexdHexHexA}$ $[\text{3AQ/M+K}]^+$
733.4186	733.2429	0.18	$\text{Pent}_2\text{Hex}_2$ $[\text{3AQ/M+H}]^+$
755.4199	755.2249	0.19	$\text{Pent}_2\text{Hex}_2$ $[\text{3AQ/M+Na}]^+$
763.4461	763.2535	0.19	$\text{PentHex}_3$ $[\text{3AQ/M+H}]^+$
771.389	771.1988	0.19	$\text{Pent}_2\text{Hex}_2$ $[\text{3AQ/M+K}]^+$

/	785.2354	/	PentHex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
791.4423	791.2484	0.19	Hex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
801.3994	801.2094	0.19	PentHex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
813.3167	813.2304	0.09	Hex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
829.399	829.2043	0.19	Hex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
895.4993	895.2958	0.20	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+H] <sup>+</sup>
917.4336	917.2777	0.16	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+Na] <sup>+</sup>
923.4984	923.2907	0.21	PentHex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
933.4885	933.2516	0.24	Pent <sub>2</sub> Hex <sub>3</sub> [3AQ/M+K] <sup>+</sup>
945.474	945.2726	0.20	PentHex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
961.4592	961.2465	0.21	PentHex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1055.5537	1055.3330	0.22	Pent <sub>2</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1077.5375	1077.3150	0.22	Pent <sub>2</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1085.5682	1085.3430	0.23	PentHex <sub>3</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1093.5198	1093.2890	0.23	Pent <sub>2</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1107.4944	1107.3250	0.17	PentHex <sub>3</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1123.5326	1123.2990	0.23	PentHex <sub>3</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1187.6079	1187.3750	0.23	Pent <sub>3</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1209.4715	1209.3570	0.11	Pent <sub>3</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1217.6238	1217.3860	0.24	Pent <sub>2</sub> Hex <sub>3</sub> dHexHexA [3AQ/M+H] <sup>+</sup>
1225.5776	1225.3310	0.25	Pent <sub>3</sub> Hex <sub>2</sub> dHexHexA [3AQ/M+K] <sup>+</sup>
1239.6096	1239.3680	0.24	Pent <sub>2</sub> Hex <sub>3</sub> dHexHexA [3AQ/M+Na] <sup>+</sup>
1255.5933	1255.3420	0.25	Pent <sub>2</sub> Hex <sub>3</sub> dHexHexA [3AQ/M+K] <sup>+</sup>

In conclusion, the developed enzymatic digestion protocol showed to work optimally also for more complex samples. On the contrary of the classical analytical methodologies used for gum analysis, the presence of pigment particles and other additives (such as glycerin, preservatives etc.) does not seem to interfere with the analysis. This aspect is of significant importance in prospect of the analysis of a real art/archaeological sample. Due to the preciousness of the objects under investigation, the amount of sample available for analysis is usually fairly limited (< 1%). In addition the amount of organic material represents a minimal portion of the sample (< 10%). Therefore, since clean-up procedures can cause the loss of the analytes, avoiding this step allows to preserve the higher amount of sample and, therefore, to maximize the recovery and analytical results. Furthermore, the analysis of aged gum arabic samples (50 and ~150 years old) did not show any significant modification in the mass profile of gum. This result is promising in the view of applying the developed methodology to more ancient samples.

## 7. CONCLUSIONS OF MALDI-TOF ANALYSIS

Partial enzymatic digestion, followed by analysis of the released oligosaccharides by MALDI-MS, was explored as a potential strategy for plant gums identification in samples from works of art. The digestion protocol was developed and optimized for individual plant gums according to their theoretical polysaccharide structure. Among the several enzymes tested, exo- $\beta$ -1,3-galactanase and endo-1,4- $\beta$ -mannanase resulted in the best digestion procedure and mass spectrometric profile. Data obtained for gum arabic, tragacanth, cherry, locust bean and guar gums suggested that the different plant gums show some differences in their polysaccharide mass profile. In the specific case of gum arabic, a complete mass fingerprint was reconstructed and the obtained ions were attributed to a corresponding oligosaccharide by means of MS spectra interpretation, employment of an informatics tool, enzymes combination and tandem mass spectrometry (MS/MS) analysis. Gum arabic resulted in a mixture of oligosaccharides ranging from 600 to 3000 Da and characterized by the presence of arabinose, galactose, rhamnose and glucuronic acid. The other gums showed some different mass profiles. Gum tragacanth pattern is defined by two intense ions at  $m/z$  957 and 1103, while the mass spectra of digested cherry gum samples are characterized by the presence of chains made of pentoses (more likely arabinose according to the polysaccharide structure knowledge). Locust bean and guar gum showed a mass pattern defined mainly by the presence of hexose units (galactose and mannose), thus resulting in a profile which is discernible from the other plant gums.

Since in the case of an art or archaeological sample the selection of a single enzyme is not feasible, because of its unknown composition, an enzyme cocktail suitable for the digestion of all of the most common plant gums, regardless of their polysaccharide structure, was developed. Exo- $\beta$ -1,3-galactanase and endo-1,4- $\beta$ -mannanase were combined together and new digestion parameters were established. Results using the enzyme cocktail indicated that effective digestion of gums occurred, and the characteristic mass fingerprints could be obtained.

After the analysis of pure gums and the characterization of their MALDI-MS profile, the proposed strategy was applied to fresh and old watercolor samples. The presence of gum arabic was successfully confirmed by comparison of the released oligosaccharides with the characteristic ions of the pure gum. Furthermore, the presence of pigment particles and other additives (such as glycerin and preservatives), as well as the ageing of the sample, did not seem to alter the gum mass profile.



## REFERENCES

1. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, et al. (1988) Protein and polymer analyses up to  $m/z$  100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* 2: 151-153.
2. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry* 60: 2299-2301.
3. Dass C (2007) *Fundamentals of contemporary mass spectrometry*; N.M. DDMN, editor: Jhon Wiley & Sons, Inc.
4. Karas M, Bachmann D, Bahr U, Hillenkamp F (1987) Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes* 78: 53-68.
5. Mock KK, Davey M, Cottrell JS (1991) The analysis of underivatized oligosaccharides by Matrix-Assisted Laser Desorption Mass Spectrometry. *Biochemical and Biophysical Research Communications* 177: 644-651.
6. Harvey DJ (2003) Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates and glycoconjugates. *International Journal of Mass Spectrometry* 226: 1-35.
7. Hanton SD (2001) *Mass Spectrometry of Polymers and Polymer Surfaces*. *Chemical Reviews* 101: 527-570.
8. Stahl B, Steup M, Karas M, Hillenkamp F (1991) Analysis of neutral oligosaccharides by matrix-assisted laser desorption ionization mass spectrometry. *Analytical Chemistry* 63: 1463-1466.
9. Harvey DJ (1999) Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates. *Mass Spectrom Rev* 18: 349-450.
10. Karas M, Ehring H, Nordhoff E, Stahl B, Strupat K, et al. (1993) Matrix-assisted laser desorption/ionization mass spectrometry with additives to 2,5-dihydroxybenzoic acid. *Organic Mass Spectrometry* 28: 1476-1481.
11. Lamari FN, Kuhn R, Karamanos NK (2003) Derivatization of carbohydrates for chromatographic, electrophoretic and mass spectrometric structure analysis. *Journal of Chromatography B* 793: 15-36.
12. Harvey DJ (2011) Derivatization of carbohydrates for analysis by chromatography; electrophoresis and mass spectrometry. *Journal of Chromatography B* 879: 1196-1225.
13. Mischnick P (2012) *Mass Spectrometric Characterization of Oligo- and Polysaccharides and Their Derivatives*. In: Hakkarainen M, editor. *Mass Spectrometry of Polymers – New Techniques*: Springer Berlin Heidelberg. pp. 105-174.
14. Ciucanu I, Kerek F (1984) A simple and rapid method for the permethylation of carbohydrates. *Carbohydrate Research* 131: 209-217.
15. Domon B, Costello C (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal* 5: 397-409.
16. Harvey DJ (2000) Collision-induced fragmentation of underivatized N-linked carbohydrates ionized by electrospray. *Journal of Mass Spectrometry* 35: 1178-1190.
17. Buchholz K, Kasche V, Bornscheuer UT (2012) *Biocatalysts & Enzyme Technology*; Stephan L, editor: Wiley-Blackwell.
18. Henrissat B, Davies G (1997) Structural and sequence-based classification of glycoside hydrolases. *Current Opinion in Structural Biology* 7: 637-644.
19. Lee KB, Loganathan D, Merchant ZM, Linhardt RJ (1990) Carbohydrate analysis of glycoproteins A review. *Applied Biochemistry and Biotechnology* 23: 53-80.
20. Davies G, Henrissat B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* 3: 853-859.
21. Jacob GS, Scudder P (1994) [17] Glycosidases in structural analysis. In: William J. Lennarz GWH, editor. *Methods in Enzymology*: Academic Press. pp. 280-299.
22. Morelle W, Michalski J-C (2007) Analysis of protein glycosylation by mass spectrometry. *Nature Protocols* 2: 1585-1602.

23. Lazar IM, Lazar AC, Cortes DF, Kabulski JL (2011) Recent advances in the MS analysis of glycoproteins: Theoretical considerations. *Electrophoresis* 32: 3-13.
24. Geyer H, Schmitt S, Wuhrer M, Geyer R (1998) Structural Analysis of Glycoconjugates by On-Target Enzymatic Digestion and MALDI-TOF-MS. *Analytical Chemistry* 71: 476-482.
25. van den Brink J, de Vries R (2011) Fungal enzyme sets for plant polysaccharide degradation. *Applied Microbiology and Biotechnology* 91: 1477-1492.
26. Sakamoto T, Ishimaru M (2013) Peculiarities and applications of galactanalytic enzymes that act on type I and II arabinogalactans. *Applied Microbiology and Biotechnology* 97: 5201-5213.
27. Mirhosseini H, Amid BT (2012) A review study on chemical composition and molecular structure of newly plant gum exudates and seed gums. *Food Research International* 46: 387-398.
28. Dumitriou S (2005) *Polysaccharides: structural diversity and functional versatility*; Dumitriou S, editor. USA.
29. Schröder R, Atkinson RG, Redgwell RJ (2009) Re-interpreting the role of endo- $\beta$ -mannanases as mannan endotransglycosylase/hydrolases in the plant cell wall. *Annals of Botany* 104: 197-204.
30. Ademark P, Varga A, Medve J, Harjunpää V, Torbjörn D, et al. (1998) Softwood hemicellulose-degrading enzymes from *Aspergillus niger*: Purification and properties of a  $\beta$ -mannanase. *Journal of Biotechnology* 63: 199-210.
31. McCleary BV, Matheson NK (1983) Action patterns and substrate-binding requirements of  $\beta$ -d-mannanase with mannosaccharides and mannan-type polysaccharides. *Carbohydrate Research* 119: 191-219.
32. McCleary BV, Nurthen E, Taravel FR, Joseleau J-P (1983) Characterisation of the oligosaccharides produced on hydrolysis of galactomannan with  $\beta$ -d-mannase. *Carbohydrate Research* 118: 91-109.
33. Dumville JC, Fry SC (2000) Uronic acid-containing oligosaccharins: Their biosynthesis, degradation and signalling roles in non-diseased plant tissues. *Plant Physiology and Biochemistry* 38: 125-140.
34. Limberg G, Körner R, Buchholt HC, Christensen TMIE, Roepstorff P, et al. (2000) Quantification of the amount of galacturonic acid residues in blocksequences in pectin homogalacturonan by enzymatic fingerprinting with exo- and endo-polygalacturonase II from *Aspergillus niger*. *Carbohydrate Research* 327: 321-332.
35. Chen EMW, Mort AJ (1996) Nature of sites hydrolyzable by endopolygalacturonase in partially-esterified homogalacturonans. *Carbohydrate Polymers* 29: 129-136.
36. Prade RA, Zhan D, Ayoubi P, Mort AJ (1999) Pectins, Pectinases and Plant-Microbe Interactions. *Biotechnology and Genetic Engineering Reviews* 16: 361-392.
37. Mutter M, Renard CMGC, Beldman G, Schols HA, Voragen AGJ (1998) Mode of action of RG-hydrolase and RG-lyase toward rhamnogalacturonan oligomers. Characterization of degradation products using RG-rhamnohydrolase and RG-galacturonohydrolase. *Carbohydrate Research* 311: 155-164.
38. Azadi P, O'Neill MA, Bergmann C, Darvill AG, Albersheim P (1995) The backbone of the pectic polysaccharide rhamnogalacturonan I is cleaved by an endohydrolase and an endolyase. *Glycobiology* 5: 783-789.
39. Øbro J, Harholt J, Scheller HV, Orfila C (2004) Rhamnogalacturonan I in *Solanum tuberosum* tubers contains complex arabinogalactan structures. *Phytochemistry* 65: 1429-1438.
40. Minic Z, Jouanin L (2006) Plant glycoside hydrolases involved in cell wall polysaccharide degradation. *Plant Physiology and Biochemistry* 44: 435-449.
41. Tryfona T, Liang H-C, Kotake T, Kaneko S, Marsh J, et al. (2010) Carbohydrate structural analysis of wheat flour arabinogalactan protein. *Carbohydrate Research* 345: 2648-2656.
42. Kotake T, Kitazawa K, Takata R, Okabe K, Ichinose H, et al. (2009) Molecular Cloning and Expression in *Pichia pastoris* of a *Irpex lacteus* Exo- $\beta$ -(1 $\rightarrow$ 3)-galactanase Gene. *Bioscience, Biotechnology, and Biochemistry* 73: 2303-2309.
43. Tsumuraya Y, Mochizuki N, Hashimoto Y, Kovac P (1990) Purification of an exo-beta-(1 $\rightarrow$ 3)-D-galactanase of *Irpex lacteus* (*Polyporus tulipiferae*) and its action on arabinogalactan-proteins. *J Biol Chem* 265: 7207-7215.

44. Konishi T, Kotake T, Soraya D, Matsuoka K, Koyama T, et al. (2008) Properties of family 79  $\beta$ -glucuronidases that hydrolyze  $\beta$ -glucuronosyl and 4-O-methyl- $\beta$ -glucuronosyl residues of arabinogalactan-protein. *Carbohydrate Research* 343: 1191-1201.
45. Tryfona T, Liang H-C, Kotake T, Tsumuraya Y, Stephens E, et al. (2012) Structural characterization of Arabidopsis leaf arabinogalactan polysaccharides. *Plant Physiol* 160: 653-666.
46. Huisman MMH, Brüll LP, Thomas-Oates JE, Haverkamp J, Schols HA, et al. (2001) The occurrence of internal (1 $\rightarrow$ 5)-linked arabinofuranose and arabinopyranose residues in arabinogalactan side chains from soybean pectic substances. *Carbohydrate Research* 330: 103-114.
47. Huisman MMH, Schols HA, Voragen AGJ (1999) Enzymatic degradation of cell wall polysaccharides from soybean meal. *Carbohydrate Polymers* 38: 299-307.
48. Mohr MD, OlafBörnsen K, Widmer HM (1995) Matrix-assisted laser desorption/ionization mass spectrometry: Improved matrix for oligosaccharides. *Rapid Communications in Mass Spectrometry* 9: 809-814.
49. Hsu N-Y, Yang W-B, Wong C-H, Lee Y-C, Lee RT, et al. (2007) Matrix-assisted laser desorption/ionization mass spectrometry of polysaccharides with 2',4',6'-trihydroxy-acetophenone as matrix. *Rapid Commun Mass Spectrom* 21: 2137-2146.
50. Rohmer M, Meyer B, Mank M, Stahl B, Bahr U, et al. (2010) 3-Aminoquinoline Acting as Matrix and Derivatizing Agent for MALDI MS Analysis of Oligosaccharides. *Analytical Chemistry* 82: 3719-3726.
51. Metzger J, Woisch R, Tuszynski W, Angermann R (1994) New type of matrix for matrix-assisted laser desorption mass spectrometry of polysaccharides and proteins. *Fresenius' Journal of Analytical Chemistry* 349: 473-474.
52. Beavis RC, Chait BT (1990) High-accuracy molecular mass determination of proteins using matrix-assisted laser desorption mass spectrometry. *Analytical Chemistry* 62: 1836-1840.
53. Packer N, Lawson M, Jardine D, Redmond J (1998) A general approach to desalting oligosaccharides released from glycoproteins. *Glycoconjugate Journal* 15: 737-747.
54. Sakamoto T, Tanaka H, Nishimura Y, Ishimaru M, Kasai N (2011) Characterization of an exo- $\beta$ -1,3-D-galactanase from *Sphingomonas* sp. 24T and its application to structural analysis of larch wood arabinogalactan. *Appl Microbiol Biotechnol* 90: 1701-1710.
55. Geyer H, Geyer R (2006) Strategies for analysis of glycoprotein glycosylation. *Biochimica Et Biophysica Acta-Proteins and Proteomics* 1764: 1853-1869.
56. Simoes J, Nunes FM, Rosario Domingues M, Coimbra MA (2011) Demonstration of the presence of acetylation and arabinose branching as structural features of locust bean gum galactomannans. *Carbohydrate Polymers* 86: 1476-1483.
57. Simas-Tosin FF, Barraza RR, Petkowicz CLO, Silveira JLM, Sasaki GL, et al. (2010) Rheological and structural characteristics of peach tree gum exudate. *Food Hydrocolloids* 24: 486-493.
58. Mohammadifar MA, Musavi SM, Kiumarsi A, Williams PA (2006) Solution properties of targacanthin (water-soluble part of gum tragacanth exudate from *Astragalus gossypinus*). *International Journal of Biological Macromolecules* 38: 31-39.
59. Goellner EM, Utermoehlen J, Kramer R, Classen B (2011) Structure of arabinogalactan from *Larix laricina* and its reactivity with antibodies directed against type-II-arabinogalactans. *Carbohydrate Polymers* 86: 1739-1744.
60. Nie S-P, Wang C, Cui SW, Wang Q, Xie M-Y, et al. (2013) A further amendment to the classical core structure of gum arabic (*Acacia senegal*). *Food Hydrocolloids* 31: 42-48.
61. Heidebach T, Sass M, De WA (2013) Enzymatic treatment of gum arabic. Google Patents.
62. Kotake T, Hirata N, Degi Y, Ishiguro M, Kitazawa K, et al. (2011) Endo- $\beta$ -1,3-galactanase from Winter Mushroom *Flammulina velutipes*. *Journal of Biological Chemistry* 286: 27848-27854.
63. Al-Assaf S, Phillips GO, Williams PA (2005) Studies on Acacia exudate gums: part II. Molecular weight comparison of the Vulgares and Gummiferae series of Acacia gums. *Food Hydrocolloids* 19: 661-667.

64. Idris OHM, Williams PA, Phillips GO (1998) Characterisation of gum from *Acacia senegal* trees of different age and location using multidetection gel permeation chromatography. *Food Hydrocolloids* 12: 379-388.
65. Motlagh S, Ravines P, Karamallah KA, Ma QF (2006) The analysis of *Acacia* gums using electrophoresis. *Food Hydrocolloids* 20: 848-854.
66. Brenan JPM (1983) Manual of taxonomy of *Acacia* species: Present taxonomy of four species of *Acacia* (*A. albida*, *A. senegal*, *A. nilotica*, *A. tortilis*); FAO, editor. Rome.
67. Balaghi S, Mohammadifar MA, Zargaraan A, Gavlighi HA, Mohammadi M (2011) Compositional analysis and rheological characterization of gum tragacanth exudates from six species of Iranian *Astragalus*. *Food Hydrocolloids* 25: 1775-1784.
68. Daas PJH, Meyer-Hansen K, Schols HA, De Ruiter GA, Voragen AGJ (1999) Investigation of the non-esterified galacturonic acid distribution in pectin with endopolygalacturonase. *Carbohydrate Research* 318: 135-145.
69. Limberg G, Körner R, Buchholt HC, Christensen TMIE, Roepstorff P, et al. (2000) Analysis of different de-esterification mechanisms for pectin by enzymatic fingerprinting using endopectin lyase and endopolygalacturonase II from *A. Niger*. *Carbohydrate Research* 327: 293-307.
70. Gavlighi HA, Michalak M, Meyer AS, Mikkelsen JD (2013) Enzymatic Depolymerization of Gum Tragacanth: Bifidogenic Potential of Low Molecular Weight Oligosaccharides. *Journal of Agricultural and Food Chemistry* 61: 1272-1278.
71. Gavlighi AH, Meyer AS, Zaidel DNA, Mohammadifar MA, Mikkelsen JD (2013) Stabilization of emulsions by gum tragacanth (*Astragalus* spp.) correlates to the galacturonic acid content and methoxylation degree of the gum. *Food Hydrocolloids* 31: 5-14.
72. Amarioarei G, Lungu M, Sorin C (2012) Molar mass characterization of cherry tree exudate gums of different seasons. *Cellulose Chem Technol* 46: 583-588.
73. Cheng Y, Prud'homme RK (2000) Enzymatic Degradation of Guar and Substituted Guar Galactomannans. *Biomacromolecules* 1: 782-788.
74. Daas PJH, Schols HA, de Jongh HHJ (2000) On the galactosyl distribution of commercial galactomannans. *Carbohydrate Research* 329: 609-619.
75. Jian H-L, Zhu L-W, Zhang W-M, Sun D-F, Jiang J-X (2013) Enzymatic production and characterization of manno-oligosaccharides from *Gleditsia sinensis* galactomannan gum. *International Journal of Biological Macromolecules* 55: 282-288.
76. Cheng Y, Brown KM, Prud'homme RK (2002) Preparation and characterization of molecular weight fractions of guar galactomannans using acid and enzymatic hydrolysis. *International Journal of Biological Macromolecules* 31: 29-35.
77. McCleary BV, Clark AH, Dea ICM, Rees DA (1985) The fine structures of carob and guar galactomannans. *Carbohydrate Research* 139: 237-260.
78. Kurakake M, Sumida T, Masuda D, Oonishi S, Komaki T (2006) Production of Galactomanno-oligosaccharides from Guar Gum by  $\beta$ -Mannanase from *Penicillium oxalicum* SO. *Journal of Agricultural and Food Chemistry* 54: 7885-7889.
79. Wielinga WC (2009) Galactomannans. In: Williams GOPaPA, editor. *Handbook of hydrocolloids*: CRC Press. pp. 228-297.
80. Tayal A, Kelly RM, Khan SA (1999) Rheology and Molecular Weight Changes during Enzymatic Degradation of a Water-Soluble Polymer. *Macromolecules* 32: 294-300.
81. Srivastava M, Kapoor VP (2005) Seed Galactomannans: An Overview. *Chemistry & Biodiversity* 2: 295-317.
82. Duncan MW, Roder H, Hunsucker SW (2008) Quantitative matrix-assisted laser desorption/ionization mass spectrometry. *Briefings in Functional Genomics & Proteomics* 7: 355-370.

## Final conclusions

This thesis aimed to introduce alternative methods for the identification of plant gums in samples from works of art and to better understand the double proteinaceous-polysaccharide nature of plant gums. The classical methods employed for gum analysis in art and archaeological samples suffer from some limitations due to the presence of inorganic pigments, gum mixtures and other sample components, as well as the ageing of the sample. Furthermore, analysis of these materials proved to be particularly challenging due to the fact that their structure has not so far been completely elucidated.

The present research was thus intended to compensate for some of the lacks related to both plant gums features and identification. Size exclusion chromatography and polyacrylamide gel electrophoresis were applied to investigate the chromatographic and electrophoretic profile of the protein component of several plant gums and to better understand the systems under investigation. In addition, a method based on enzymatic digestion of gums' polysaccharide moiety and MALDI-MS/MS was developed for identification of different plant gums.

SDS-PAGE and SEC proved to be two complementary techniques and allowed to get the widest view and knowledge of plant gums glycoproteins. The two methods allowed to obtain information on the molecular weight of the proteinaceous fractions constituting the structure of plant gums, from 15 kDa to millions of Dalton. Polyacrylamide gel electrophoresis confirmed the presence of highly glycosylated proteins while size exclusion chromatography have been able to clearly demonstrate the polydisperse nature of the gums. Furthermore, each gum was showed to have a different SEC and SDS-PAGE profile, thus revealing that differences might not only result in terms of amino acid composition, as already showed in the literature, but also in terms of glycosylation degree and/or structure.

The identification of high molecular weight fractions demonstrated the necessity to decomplexify the gum structure before analysis by mass spectrometry. The developed method, involving partial enzymatic digestion of the plant gums, followed by analysis of the released oligosaccharides by MALDI-MS, represented a possible novel strategy for plant gums identification. The method, named "saccharide mass fingerprint", since its analogy to the peptide mass fingerprinting approach widely used for identification of proteins, was demonstrated to be applicable for discriminating several pure plant gums such as gum arabic, tragacanth, cherry, locust bean and guar gums. Furthermore, an enzyme cocktail suitable for the digestion of all of the most common plant gums, regardless of their polysaccharide structure, was developed. Results confirmed the characteristic mass profiles of the gums, thus revealing a great potentiality of applying the strategy to unknown samples. A list of specific ions with corresponding carbohydrate attributions has been compiled for each gum and this represents a prototype database for the analysis of plant gums by enzyme digestion and MALDI-MS. In this way an unknown polysaccharide material can thus be identified by comparison with the tabulated list of ions derived from the analysis of reference materials and identification of the oligosaccharides detected.

In addition, the applicability of the developed method was tested by analyzing a watercolor sample dating 1870. The presence of gum arabic was successfully verified by comparison of the obtained mass spectrum with the one of pure gum arabic. This application, together with the analysis of fresh watercolor samples, suggested how the presence of inorganic materials (e.g. pigments) or other additives, as well as the ageing, did not alter the gum mass fingerprint. These data represent a first confirmation of the possible applicability of the method to real and much more complex samples from works of art and cultural artefacts.

# **Materials and Methods**





## **TABLE OF CONTENTS**

1. PROTEIN QUANTITATION AND SEPARATION .....	187
1.1. Protein quantitation by bicinchoninic acid (BCA) .....	187
1.2. Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis.....	187
1.3. SDS agarose-acrylamide composite gel electrophoresis .....	188
1.4. Size Exclusion Chromatography (SEC).....	188
2. PURIFICATION PROCEDURE .....	188
3. ENZYMATIC DIGESTION .....	189
4. MALDI-MS AND MS/MS ANALYSIS .....	189



## 1. PROTEIN QUANTITATION AND SEPARATION

### 1.1. PROTEIN QUANTITATION BY BICINCHONIC ACID (BCA)

Standard Bovine Serum Albumin BSA (Sigma Aldrich) is prepared 1 mg/mL in milliQ water and serial dilutions are performed to set a working range from 1000 to 20  $\mu\text{g/mL}$ . Working reagent is prepared by mixing reagent A with reagent B (both containing BCA) in a proportion 50/1. Samples are mixed with the working reagent in a ratio 1/20 (v/v) and incubated at 37°C for 30 minutes. Samples are cooled down and absorbance is measured at  $\lambda = 562 \text{ nm}$ . A standard curve is built by plotting the blank-corrected absorbance of each BSA standard dilution vs. its concentration. The curve is therefore used to determine the protein concentration of the unknown sample.

### 1.2. SODIUM DODECYLSULFATE POLYACRYLAMIDE GEL ELECTROPHORESIS.

Plant gum samples are dissolved directly in Laemmli 4x (denaturing) and mixed over night at room temperature. The soluble component is recovered by centrifugation at 13400rpm for 30 minutes. 35  $\mu\text{L}$  of gum sample are heated at 90°C for 5 minutes and separated in denaturing conditions on a mini polyacrylamide gel. The solutions used for gel preparation are reported in Tab. 1. 2% v/v of APS 30% and 0.2% v/v of tetramethylethylenediamin (TEMED) are added to respectively initiate and accelerate the gel polymerization. The maximum volume of sample is loaded to each well in the stacking gel (~35  $\mu\text{L}$ ), while in a different well standard proteins with a MW in the range of 14.4 - 220 kDa are loaded (see Tab. 1). Acrylamide concentration varies from 15 % to 4.7% according to the protein molecular weight of interest. Electrophoresis is run at 200 V and stopped when the bromophenol blue dye front reaches the end of the gel (around 1.5 hour). Gels are fixed with a water solution containing 7% acetic acid (v/v) and 10% ethanol (v/v). Fixing is performed with a total of 250 mL solution for around 1 hour. Proteins are stained over night, with continuous agitation, with 100 mL of SYPRO<sup>®</sup> Ruby (20 mg/mL) prepared in a water solution containing 14% acetic acid and 25% ethanol. Sypro solution is removed and gels are rinsed three times with the fixing solution (for around 1 hour) and 20 minutes with milliQ water. Gels are scanned at 200 dpi resolution by a Typhoon 9000 scanner (GE Healthcare). Image is acquired with a blue laser ( $\lambda_{\text{exc}} = 488\text{nm}$ ;  $\lambda_{\text{em}} = 610 \text{ nm}$ ).

**Tab. 1.** List of materials, solutions for acrylamide gel preparation and corresponding composition.

Solutions	Composition
Separation buffer	: Tris-HCl buffer 1.5 M (pH 8.8), SDS 0.7% w/v.
Concentration buffer	: Tris-base buffer 0.5 M (pH 6.8), SDS 0.4% w/v.
Running buffer	: Tris-base buffer 20 mM, SDS 0.1% w/v, Glycine 0.2% w/v.
Laemmli 4x (Denaturing)	: 24% separation buffer v/v, SDS 0.3 M, 40% glycerol v/v, DTT 0.3M, 0.01% Bromophenol Blue w/v, 36 % milliQ water.
Acrylamide/Bisacrylamide 30%	Acrylamide 29.2% w/v, Bisacrylamide (N,N'-Methylenebisacrylamide) 0.8% w/v, Dextran 500kDa 0.66% w/v. Solution prepared in milliQ water.
Ammonium persulfate	30% w/v in milliQ water
Standard proteins	Myosin (220 kDa), $\alpha_2$ -Macroglobulin (170 kDa), $\beta$ -

---

galactosidase (116 kDa), Phosphorylase b (97 kDa), Transferrin (76 kDa), Albumin (66 kDa), Glutamic dehydrogenase (53 kDa), Ovalbumin (45 kDa), Carbonic anhydrase (30 kDa), Trypsin inhibitor (20.1 kDa),  $\alpha$ -Lactalbumin (14.4 kDa) (Amersham Low Molecular Weight and High Molecular Weight Calibration Kit for SDS Electrophoresis, GE Healthcare).

---

### 1.3. SDS AGAROSE-ACRYLAMIDE COMPOSITE GEL ELECTROPHORESIS

Agarose (1.5% w/v) is stirred in deionized water and completely dissolved by heating the solution at 100°C on a hot plate. When the solution appears clear, it is let cooling down and kept in a warm bath at 45°C. Acrylamide solution is prepared as reported in section 1.2 but without the addition of TEMED. The solution is kept at 45°C for 20 minutes to equilibrate and it is then mixed to the agarose solution with gentle stirring ( $T = 45^{\circ}\text{C}$ ). Around 0.1 % v/v of TEMED is added and the gel solution is rapidly casted in the electrophoresis system. The system is transferred at 4°C for 10 minutes, in order to prevent acrylamide from polymerizing until the end of agarose gelation, and then brought to room temperature for 1 hour until acrylamide is completely polymerization. All the solution volumes are dependant of the final acrylamide and agarose percentage desired.

### 1.4. SIZE EXCLUSION CHROMATOGRAPHY (SEC)

Plant gum samples are dissolved at a concentration of 1% w/v in Tris-HCl 50 mM with the addition of NaCl 50 mM ( $\text{pH} = 7$ ), which correspond to the mobile base used for chromatographic separation. The solution is mixed over night at room temperature for complete gum solubilization. The soluble part is then recovered after centrifugation at 13400 rpm for 30 minutes and 100  $\mu\text{L}$  of solution are injected in the SEC column. An isocratic elution is performed for a total run time of 40 minutes. The flow is fixed at 0.5 mL/min for the column BioSuite HR SEC (higher MWs), and 0.2 mL/min for the column Biosep S2000 (lower MWs). The eluted proteins were monitored using a Diode Array Detector (HPLC-DAD system, Agilent 1200 Series) operating at  $\lambda = 280$  and 214 nm (BW 4 nm).

## 2. PURIFICATION PROCEDURE

### PURIFICATION BY ZIP TIP HYPERCARB

The sample is resuspended in milliQ water containing 0.1% of trifluoroacetic acid (TFA). The tip (maximum volume 200  $\mu\text{L}$ ) is equilibrate with 2 mL of ACN 100% containing 0.1% TFA and then 2 mL of milliQ water with 0.1% of TFA. The analytes are bound to the tip phase by aspirating and dispensing the sample 30 - 40 times. The phase is then washed with 2 mL of milliQ water acidified with 0.1% FA and the oligosaccharides eluted with increasing percentage of ACN in milliQ water

acidified with 0.1% of TFA (0.1; 0.5; 1; 5; 10; 30; 50; 70; 100 %). The samples are then dried in a vacuum centrifuge and resuspended in milliQ water.

### 3. ENZYMATIC DIGESTION

#### POLYSACCHARIDE SEQUENTIAL ENZYMATIC DIGESTION

Gum powder, or nodule, is dissolved in milliQ water at a concentration that goes from 20 mg/mL to 0.8 mg/mL. Gum solution is mixed over night at room temperature and the supernatant is recovered after centrifugation at 13400 rpm for 30minute. For specific preparation procedures see Tab 6 chapter III. A 100  $\mu$ L aliquot is collected and dried in a vacuum centrifuge at 30°C (concentrator 5301, Eppendorf, Hamburg, Germany). The sample is resuspended in 100  $\mu$ L of the buffer which is specific for the selected enzyme. 100 mU of enzyme is added to the sample, mixed and incubated for the selected time. Addition of subsequent enzyme requires inactivation of the previous one by heating at 100°C for 15 minutes. The solution is then dried, the sample dissolved in the proper new buffer and the following enzyme is added. Buffer composition and digestion temperature are reported in Tab. 2. If samples were not subjected immediately to enzymatic digestion, they were stocked, dried and kept at -20°C.

**Tab. 2.** Experimental conditions for enzymatic digestion.

Enzyme	Buffer	pH	Temperature [°C]
Exo- $\beta$ -1,3-galactanase	Phosphate buffer 50 mM	6	50
$\alpha$ -L-Arabinofuranosidase	Phosphate buffer 50 mM	7	65
Endo- $\beta$ -1,4-mannanase	Phosphate buffer 50 mM	7.5	45
$\alpha$ -Galactosidase	Tris-HCl buffer 25 mM	8.5	37
$\beta$ -Galactosidase	Acetate buffer 50 mM	4.5	60
$\beta$ -Glucuronidase	Phosphate buffer 100 mM	6.8	37
$\alpha$ -L-rhamnosidase	Phosphate buffer 20 mM	6.5	50
Endo- $\beta$ -1,4-galactanase	Phosphate buffer 50 mM	7	45
Enzyme cocktail	Phosphate buffer 50 mM	7	45

### 4. MALDI-MS AND MS/MS ANALYSIS

Analysis were carried out on a MALDI-TOF-TOF ABI 4800 (Applied Biosystems) equipped with a Nd-YAG laser of 355 nm wavelength. Spectra were obtained with a delayed extraction technology, in reflector positive mode with a grid voltage of 16 kV and the extraction delay at 420 ns. A total of 3000 laser shots were accumulated for each spectrum. Identification of MS and MSMS spectra was performed using the instrument software Data Explorer (Applied Biosystems). MS/MS analysis were performed in CID mode (Collision-Induced Dissociation) in the presence of air. Grid voltage was fixed at 7.3 kV and collision cell at 7 kV.

### **OLIGOSACCHARIDE MS ANALYSIS**

Oligosaccharides, released after enzymatic digestion, are directly mixed with the matrix 3-aminoquinoline on the MALDI plate. 3-AQ is prepared 20 mg/mL in ACN/water 1/2 (v/v) with 0.07% of inorganic acid to reach pH = 5. 0.5  $\mu$ L of sample were spotted on the MALDI plate and 1  $\mu$ L of 3-AQ solution was added. Matrix and sample were mixed and left dried at room temperature for 1 hour. Different dilutions of the sample were performed (200-10 pmol).

# Abstract

Plant gums are naturally occurring compounds coming from several plant species and their application as binding media and adhesives in the field of cultural heritage traces back to the third millennium B.C. On the contrary of other organic materials, information on gums are fairly limited since their complex nature which consists of a mixture of polysaccharide and polypeptide chains. The analytical study of these materials is important to shed light on artists' techniques, on the methods of manufacture of ancient artefacts, and to inform appropriate strategies for their preservation. In the cultural heritage field, characterization of plant gums is usually based on the monosaccharide composition as determined by gas chromatography mass spectrometry (GCMS) after complete acid hydrolysis of the gum. However, inorganic and/or organic compounds simultaneously present in the sample, were demonstrated to alter the gum chromatographic profile, thus preventing the correct identification.

The present research aims to evaluate a proper analytical approach for plant gums identification in samples from works of art based on their chromatographic, electrophoretic and mass spectrometric profile. Analytical parameters have been optimized for the analysis of arabic, tragacanth, karaya, guar, ghatti, locust bean and fruit tree gums raw samples. In the first part of the thesis, size exclusion chromatography (SEC) was used to highlight the proteinaceous component of plant gums thus revealing the presence of glycoproteins with a molecular weight of 1-2 million Da. Complementary information were obtained by higher resolution techniques such as polyacrylamide gel electrophoresis (PAGE) that revealed to be a promising method for discriminating plant gums according to their protein electrophoretic profile. In the second part of the research, an optimized method involving partial enzymatic digestion of the plant gums polysaccharide component, followed by analysis of the released oligosaccharides by mass spectrometry was developed. Due to the different polysaccharide structure of the gums, the obtained MALDI mass spectra represented the characteristic profile of the gums, thus allowing their discrimination. The efficacy of the strategy was also successfully tested on an aged gum arabic sample and on fresh and old watercolors.