Universidade de Aveiro Departamento de Química
Ano 2016

Ana Sofia Pereira Moreira

Estudo de modificações induzidas por tratamento térmico e oxidativo em oligo e polissacarídeos do café por espectrometria de massa

Study of modifications induced by thermal and oxidative treatment in oligo and polysaccharides of coffee by mass spectrometry

Universidade de Aveiro Departamento de Química
Ano 2016

Ana Sofia Pereira Moreira

Estudo de modificações induzidas por tratamento térmico e oxidativo em oligo e polissacarídeos do café por espectrometria de massa

Study of modifications induced by thermal and oxidative treatment in oligo and polysaccharides of coffee by mass spectrometry

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Bioquímica, realizada sob a orientação científica da Doutora Maria do Rosário Gonçalves dos Reis Marques Domingues, Professora Associada com Agregação do Departamento de Química da Universidade de Aveiro, e do Doutor Manuel António Coimbra Rodrigues da Silva, Professor Associado com Agregação do Departamento de Química da Universidade de Aveiro

Apoio financeiro do POPH-QREN no âmbito do III Quadro Comunitário de Apoio.

Apoio financeiro da FCT e do FSE no âmbito do III Quadro Comunitário de Apoio.

Dedico este trabalho aos meus pais, aos meus irmãos e ao Vasco.

o júri

presidente

Prof. Doutor Carlos Alberto Diogo Soares Borrego
Professor Catedrático da Universidade de Aveiro

vogais

Prof. Doutora Amélia Pilar Grases Santos Silva Rauter
Professora Associada com Agregação da Faculdade de Ciências da Universidade de Lisboa

Prof. Doutora Maria do Rosário Gonçalves dos Reis Marques Domingues
Professora Associada com Agregação da Universidade de Aveiro (orientadora)

Prof. Doutor Carlos Manuel Ferreira de Sousa Borges
Professor Auxiliar da Faculdade de Ciências da Universidade de Lisboa

Prof. Doutora Susana Isabel Pereira Casal Vicente
Professora Auxiliar da Faculdade de Farmácia da Universidade do Porto

Prof. Doutor Fernando Hermínio Ferreira Milheiro Nunes
Professor Auxiliar da Escola de Ciências da Vida e do Ambiente da Universidade de Trás-os-Montes e Alto Douro

agradecimentos

Agradeço à Professora Rosário Domingues e ao Professor Manuel António Coimbra pela orientação deste trabalho, pelos conhecimentos transmitidos, pelas entusiastas discussões de resultados, pela disponibilidade e pelo incentivo constante.

Aos restantes coautores dos trabalhos apresentados nesta tese agradeço pelo seu contributo. Agradeço ao Professor Fernando Nunes da Universidade de Trás-os-Montes e Alto Douro. Agradeço à Professora Maria Rangel e ao André Silva da Universidade do Porto. Agradeço aos Professores Armando Silvestre, Dmitry Evtuguin e Pedro Domingues e às colegas Cláudia Passos, Elisabete Costa, Elisabete Maciel, Joana Simões e Sónia Santos do Departamento de Química da Universidade de Aveiro.

Agradeço à Doutora Laura Bravo por me ter recebido no seu laboratório, localizado no *Instituto de Ciencia y Tecnología de Alimentos y Nutrición* (ICTAN-CSIC) em Madrid. À Doutora Beatriz Sarriá e à Shenli Wang agradeço pela disponibilidade, pelos conhecimentos partilhados e pelo apoio no desenvolvimento do trabalho nesse laboratório.

De uma forma geral, agradeço aos colegas do grupo de Espectrometria de Massa e do grupo de Bioquímica e Química Alimentar pelo auxílio e pelo agradável ambiente de trabalho. À Dra. Cristina Barros agradeço o apoio nos espectrómetros e no laboratório.

Agradeço aos meus pais, aos meus irmãos e ao Vasco pelo amor e pelo apoio incondicional.

Agradeço à unidade de investigação de Química Orgânica, Produtos Naturais e Agroalimentares (QOPNA) da Universidade de Aveiro.

Agradeço à Fundação para a Ciência e a Tecnologia (FCT) pelo financiamento através de uma bolsa de doutoramento (SFRH/BD/80553/2011).

palavras-chave

Café, polissacarídeos, galactomananas, arabinogalactanas, torra, alterações estruturais, melanoidinas, espectrometria de massa

resumo

Os polissacarídeos são os componentes maioritários dos grãos de café verde e torrado e da bebida de café. Os mais abundantes são as galactomananas, seguindo-se as arabinogalactanas. Durante o processo de torra, as galactomananas e arabinogalactanas sofrem modificações estruturais, as quais estão longe de estar completamente elucidadas devido à sua diversidade e à complexidade estrutural dos compostos formados. Durante o processo de torra, as galactomananas e arabinogalactanas reagem com proteínas, ácidos clorogénicos e sacarose, originando compostos castanhos de alto peso molecular contendo nitrogénio, designados de melanoidinas. As melanoidinas do café apresentam diversas atividades biológicas e efeitos benéficos para a saúde. No entanto, a sua estrutura exata e os mecanismos envolvidos na sua formação permanecem desconhecidos, bem como a relação estrutura-atividade biológica.

A utilização de sistemas modelo e a análise por espectrometria de massa permitem obter uma visão global e, simultaneamente, detalhada das modificações estruturais nos polissacarídeos do café promovidas pela torra, contribuindo para a elucidação das estruturas e mecanismos de formação das melanoidinas. Com base nesta tese, oligossacarídeos estruturalmente relacionados com a cadeia principal das galactomananas, (β1→4)-D-manotriose (Man_3), e as cadeias laterais das arabinogalactanas, (α1→5)-L-arabinotriose (Ara_3), isoladamente ou em misturas com ácido 5-*O*-cafeoilquínico (5-CQA), o ácido clorogénico mais abundante nos grãos de café verde, e péptidos compostos por tirosina e leucina, usados como modelos das proteínas, foram sujeitos a tratamento térmico a seco, mimetizando o processo de torra. A oxidação induzida por radicais hidroxilo ($HO^{\bullet}$) foi também estudada, uma vez que estes radicais parecem estar envolvidos na modificação dos polissacarídeos durante a torra. A identificação das modificações estruturais induzidas por tratamento térmico e oxidativo dos compostos modelo foi feita por estratégias analíticas baseadas principalmente em espectrometria de massa, mas também em cromatografia líquida. A cromatografia de gás foi usada na análise de açúcares neutros e ligações glicosídicas. Para validar as conclusões obtidas com os compostos modelo, foram também analisadas amostras de polissacarídeos do café obtidas a partir de resíduo de café e café instantâneo.

Os resultados obtidos a partir dos oligossacarídeos modelo quando submetidos a tratamento térmico (seco), assim como à oxidação induzida por $HO^{\bullet}$ (em solução), indicam a ocorrência de despolimerização, o que está de acordo com estudos anteriores que reportam a despolimerização das galactomananas e arabinogalactanas do café durante a torra. Foram ainda identificados outros compostos resultantes da quebra do anel de açúcares formados durante o tratamento térmico e oxidativo da Ara_3.

Por outro lado, o tratamento térmico a seco dos oligossacarídeos modelo (individualmente ou quando misturados) promoveu a formação de oligossacarídeos com um maior grau de polimerização, e também polissacarídeos com novos tipos de ligações glicosídicas, evidenciando a ocorrência de polimerização através reações de transglicosilação não enzimática induzidas por tratamento térmico a seco.

As reações de transglicosilação induzidas por tratamento térmico a seco podem ocorrer entre resíduos de açúcares provenientes da mesma origem, mas também de origens diferentes com formação de estruturas híbridas, contendo arabinose e manose como observado nos casos dos compostos modelo usados. Os resultados obtidos a partir de amostras do resíduo de café e de café instantâneo sugerem a presença de polissacarídeos híbridos nestas amostras de café processado, corroborando a ocorrência de transglicosilação durante o processo de torra. Além disso, o estudo de misturas contendo diferentes proporções de cada oligossacarídeo modelo, mimetizando regiões do grão de café com composição distinta em polissacarídeos, sujeitos a diferentes períodos de tratamento térmico, permitiu inferir que diferentes estruturas híbridas e não híbridas podem ser formadas a partir das arabinogalactanas e galactomananas, dependendo da sua distribuição nas paredes celulares do grão e das condições de torra. Estes resultados podem explicar a heterogeneidade de estruturas de melanoidinas formadas durante a torra do café.

Os resultados obtidos a partir de misturas modelo contendo um oligossacarídeo (Ara_3 ou Man_3) e 5-CQA sujeitas a tratamento térmico a seco, assim como de amostras provenientes do resíduo de café, mostraram a formação de compostos híbridos compostos por moléculas de CQA ligadas covalentemente a um número variável de resíduos de açúcar. Além disso, os resultados obtidos a partir da mistura contendo Man_3 e 5-CQA mostraram que o CQA atua como catalisador das reações de transglicosilação. Por outro lado, nas misturas modelo contendo um péptido, mesmo contendo também 5-CQA e sujeitas ao mesmo tratamento, observou-se uma diminuição na extensão das reações transglicosilação. Este resultado pode explicar a baixa extensão das reações de transglicosilação não enzimáticas durante a torra nas regiões do grão de café mais ricas em proteínas, apesar dos polissacarídeos serem os componentes maioritários dos grãos de café. A diminuição das reações de transglicosilação na presença de péptidos/proteínas pode dever-se ao facto de os resíduos de açúcares redutores reagirem preferencialmente com os grupos amina de péptidos/proteínas por reação de Maillard, diminuindo o número de resíduos de açúcares redutores disponíveis para as reações de transglicosilação. Além dos compostos já descritos, uma diversidade de outros compostos foram formados a partir dos sistemas modelo, nomeadamente derivados de desidratação formados durante o tratamento térmico a seco.

Em conclusão, a tipificação das modificações estruturais promovidas pela torra nos polissacarídeos do café abre o caminho para a compreensão dos mecanismos de formação das melanoidinas e da relação estrutura-atividade destes compostos.

keywords

Coffee, polysaccharides, galactomannans, arabinogalactans, roasting, structural changes, melanoidins, mass spectrometry

abstract

Polysaccharides are the major components of green and roasted coffee beans, and coffee brew. The most abundant ones are galactomannans, followed by arabinogalactans. During the roasting process, galactomannans and arabinogalactans undergo structural modifications that are far to be completely elucidated due to their diversity and complexity of the compounds formed. During the roasting process, galactomannans and arabinogalactans react with proteins, chlorogenic acids, and sucrose, originating high molecular weight brown compounds containing nitrogen, known as melanoidins. Several biological activities and beneficial health effects have been attributed to coffee melanoidins. However, their exact structures and the mechanisms involved in their formation remain unknown, as well as the structure-biological activity relationship.

The use of model systems and mass spectrometry analysis allow to obtain an overall view and, simultaneously, detailed, of the structural modifications in coffee polysaccharides promoted by roasting, contributing to the elucidation of the structures and formation mechanisms of melanoidins. Based on this thesis, oligosaccharides structurally related to the backbone of galactomannans, (β1→4)-D-mannotriose, and the side chains of arabinogalactans, (α1→5)-L-arabinotriose, alone or in mixtures with 5-*O*-caffeoylquinic acid, the most abundant chlorogenic acid in green coffee beans, and dipeptides composed by tyrosine and leucine, used as models of proteins, were submitted to dry thermal treatments, mimicking the coffee roasting process. The oxidation induced by hydroxyl radicals (HO$^{\bullet}$) was also studied, since these radicals seem to be involved in the modification of the polysaccharides during roasting. The identification of the structural modifications induced by thermal and oxidative treatment of the model compounds was performed mostly by mass spectrometry-based analytical strategies, but also using liquid chromatography. Gas chromatography was used in the analysis of neutral sugars and glycosidic linkages. To validate the conclusions achieved with the model compounds, coffee polysaccharide samples obtained from spent coffee grounds and instant coffee were also analysed.

The results obtained from the model oligosaccharides when submitted to thermal treatment (dry) or oxidation induced by HO$^{\bullet}$ (in solution) indicate the occurrence of depolymerization, which is in line with previous studies reporting the depolymerization of coffee galactomannans and arabinogalactans during roasting. Compounds resulting from sugar ring cleavage were also formed during thermal treatment and oxidative treatment of Ara_3.

On the other hand, the dry thermal treatment of the model oligosaccharides (alone or when mixed) promoted the formation of oligosaccharides with a higher degree of polymerization, and also polysaccharides with new type of glycosidic linkages, evidencing the occurrence of polymerization via non-enzymatic transglycosylation reactions induced by dry thermal treatment.

The transglycosylation reactions induced by dry thermal treatment can occur between sugar residues from the same origin, but also of different origins, with formation of hybrid structures, containing arabinose and mannose in the case of the model compounds used. The results obtained from spent coffee grounds and instant coffee samples suggest the presence of hybrid polysaccharides in these processed coffee samples, corroborating the occurrence of transglycosylation during the roasting process. Furthermore, the study of mixtures containing different proportions of each model oligosaccharide, mimicking coffee bean regions with distinct polysaccharide composition, subjected to different periods of thermal treatment, allowed to infer that different hybrid and non-hybrid structures may be formed from arabinogalactans and galactomannans, depending on their distribution in the bean cell walls and on roasting conditions. These results may explain the heterogeneity of melanoidins structures formed during coffee roasting.

The results obtained from model mixtures containing an oligosaccharide (Ara_3 or Man_3) and 5-CQA and subjected to dry thermal treatment, as well as samples derived from spent coffee grounds, showed the formation of hybrid compounds composed by CQA molecules covalently linked to a variable number of sugar residues. Moreover, the results obtained from the mixture containing Man_3 and 5-CQA showed that CQA acts as catalyst of transglycosylation reactions. On the other hand, in the model mixtures containing a peptide, even if containing 5-CQA and subjected to the same treatment, it was observed a decrease in the extent of transglycosylation reactions. This outcome can explain the low extent of non-enzymatic transglycosylation reactions during roasting in coffee bean regions enriched in proteins, although polysaccharides are the major components of the coffee beans. The decrease of transglycosylation reactions in the presence of peptides/proteins can be related with the preferential reactivity of reducing residues with the amino groups of peptides/proteins by Maillard reaction, decreasing the number of reducing residues available to be directly involved in the transglycosylation reactions. In addition to the compounds already described, a diversity of other compounds were formed from model systems, namely dehydrated derivatives formed during dry thermal treatment.

In conclusion, the identification of the structural modifications in coffee polysaccharides promoted by roasting pave the way to the understanding of the mechanisms of formation of melanoidins and structure-activity relationship of these compounds.

Publications and communications

The results presented in this thesis originated several publications in international scientific journals with Referee, as well as oral and poster communications (as first author and presenter) in national and international meetings.

Publications in international scientific journals with Referee

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. (2013). Roasting-induced changes in arabinotriose, a model of coffee arabinogalactan side chains. *Food Chemistry*, *138*, 2291-2299.

Moreira, A. S. P., da Costa, E. V., Evtuguin, D. V., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. (2014). Neutral and acidic products derived from hydroxyl radical-induced oxidation of arabinotriose assessed by electrospray ionisation mass spectrometry. *Journal of Mass Spectrometry*, *49*, 280-290.

Moreira, A. S. P., Simões, J., Pereira, A. T., Passos, C. P., Nunes, F. M., Domingues, M. R. M., & Coimbra, M. A. (2014). Transglycosylation reactions between galactomannans and arabinogalactans during dry thermal treatment. *Carbohydrate Polymers*, *112*, 48-55.

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., Passos, C. P., Santos, S. A. O., Silvestre, A. J. D., Silva, A. M. N., Rangel, M., & Domingues, M. R. M. (2015). Chlorogenic acid-arabinose hybrid domains in coffee melanoidins: evidences from a model system. *Food Chemistry*, *185*, 135-144.

Moreira, A. S. P., Simões, J., Nunes, F. M., Evtuguin, D. V., Domingues, P., Coimbra, M. A., & Domingues, M. R. M. (2016). Nonenzymatic transglycosylation reactions induced by roasting: new insights from models mimicking coffee bean regions with distinct polysaccharide composition. *Journal of Agricultural and Food Chemistry*, *64*, 1831-1840.

Moreira, A. S. P., Maciel, E., Domingues, P., Nunes, F. M., Domingues, M. R. M., & Coimbra, M. A. Acid catalyses, and Maillard reaction prevents, non-enzymatic transglycosylation reactions during coffee roasting (in preparation).

Oral communications

Moreira, A. S. P., Coimbra, M. A., & Domingues, M. R. M. MS and structural changes promoted by roasting in model coffee arabinogalactans. I Workshop of MS and carbohydrates. Aveiro, Portugal, 18 April 2012.

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. Transglycosylation reactions between coffee polysaccharides during roasting: a model study. Glupor 10 - 10th International Meeting of the Portuguese Carbohydrate Chemistry Group & 1º Workshop do Grupo da Química dos Glúcidos. Covilhã, Portugal, 1 December 2013.

Moreira, A. S. P., da Costa, E. V., Evtuguin, D. V., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. Hydroxyl radical-induced oxidation of arabinotriose, a model of coffee arabinogalactan side chains. 3rd Workshop – Mass spectrometry and carbohydrates. Aveiro, Portugal, 23 April 2014.

Moreira, A. S. P., Domingues, M. R. M., & Coimbra, M. A. Coffee polysaccharides and changes induced by roasting: an overview. International workshop and mini-conference - Development of mitigation strategies for acrylamide in diabetic bakery products (Collaboration between Portugal and Slovakia). Aveiro, Portugal, 3 December 2014.

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., Passos, C. P., Santos, S. A. O., Silvestre, A. J. D., Silva, A. M. N., Rangel, M., & Domingues, M. R. M. Chlorogenic acid-arabinose hybrids as part of coffee melanoidin structures: evidences from a model system. 3rd International Congress on Cocoa Coffee and Tea. Aveiro, Portugal, 22-24 June 2015.

Poster communications

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. Roasting-induced changes in side chains of coffee arabinogalactans: a model study to evaluate changes in arabinose residues. 26th International Carbohydrate Symposium (ICS). Madrid, Spain, 22-27 July 2012.

Moreira, A. S. P., Costa, E. V., Evtuguin, D. V., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. Hydroxyl radical-induced oxidation of arabinotriose, a model of coffee arabinogalactan side chains: neutral and acidic products assessed by electrospray mass spectrometry. XXIII Encontro Nacional da Sociedade Portuguesa de Química. Aveiro, Portugal, 12-14 June 2013.

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. Chlorogenic acid-arabinose derivatives as part of coffee melanoidin structures. XX Encontro Luso-Galego de Química. Complexo FFUP/ICBAS, Porto, Portugal, 26-28 November 2014.

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., Passos, C. P., Santos, S. A. O., Silvestre, A. J. D., Silva, A. M. N., Rangel, M., & Domingues, M. R. M. Chlorogenic acid-arabinose hybrids as part of coffee melanoidin structures: evidences from a model system. 3rd International Congress on Cocoa Coffee and Tea. Aveiro, Portugal, 22-24 June 2015.

Contents

List of Figures

CHAPTER I

CHAPTER II

CHAPTER III

List of Tables

CHAPTER III

Abbreviations

$[M+H]^+$	Protonated molecule
$[M+K]^+$	Potassium adduct ion
$[M+Na]^+$	Sodium adduct ion
$[M-H]^-$	Deprotonated molecule
A25M75	25% of Ara_3 and 75% of Man_3
A50M50	50% of Ara_3 and 50% of Man_3
A75M25	75% of Ara_3 and 25% of Man_3
AGP	Arabinogalactan-protein
Ara	Arabinose
Ara_3	(α1→5)-L-arabinotriose
CA	Caffeic acid
CAD	Collision-activated dissociation
CA_{res}	Caffeic acid residue
CEL	N^{ε}-(carboxyethyl)lysine
CGA	Chlorogenic acid
CID	Collision-induced dissociation
CML	N^{ε}-(carboxymethyl)lysine
CQA	Caffeoylquinic acid
DB	Degree of branching
DHB	2,5-dihydroxybenzoic acid
DMSO	Dimethylsulfoxide
DP	Degree of polymerization
DTG	Derivative thermogravimetric (curve)
EDTA	Ethylenediaminetetraacetic acid
EI	Electron ionization (also referred as electron impact)
ESI	Electrospray ionization
-f (Suffix)	Furanose
FID	Flame ionization detector
FL	N^{ε}-(fructosyl)lysine
Gal	Galactose
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
Glc	Glucose
GlcA	Glucuronic acid

Hex	Hexose
Hex_{res}	Hexose residue
HMWM	High molecular weight material
HPLC	High-performance liquid chromatography
LC	Liquid chromatography
LEX/SEC	Ligand exchange/size-exclusion chromatography
LIT	Linear ion trap
LMW	Low molecular weight
LY	Leucine-tyrosine peptide
Lyx	Lyxose
m/z	Mass-to-charge ratio
MALDI	Matrix assisted laser desorption/ionization
Man	Mannose
Man_3	(β1→4)-D-mannotriose
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS^n	Multistage tandem mass spectrometry
NMR	Nuclear magnetic resonance
-p (Suffix)	Pyranose
PDA	Photodiode array detection
Pent	Pentose
$Pent_{res}$	Pentose residue
PMAA	Partially *O*-methylated alditol acetate
Q	Quadrupole
QA	Quinic acid
QA_{res}	Quinic acid residue
Rha	Rhamnose
Rib	Ribose
SCG	Spent coffee grounds
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
TFA	Trifluoroacetic acid
TG	Thermogravimetric (curve)
TOF	Time-of-flight
Xyl	Xylose
YL	Tyrosine-leucine peptide

CHAPTER I. INTRODUCTION

Parts of the text of this chapter as well as Figures I.2 and I.3 were published in the following publications:

Moreira, A. S. P., Nunes, F. M., Domingues, M. R., & Coimbra, M. A. (2012). Coffee melanoidins: structures, mechanisms of formation and potential health impacts. *Food & Function, 3*, 903-915.

Moreira, A. S. P., Nunes, F. M., Domingues, M. R. M., & Coimbra, M. A. (2015). Galactomannans in coffee. In V. R. Preedy (Ed.), *Coffee in Health and Disease Prevention* (pp. 173-182). San Diego: Academic Press.

Moreira, A. S. P., Simões, J., Passos, C. P., Nunes, F. M., Domingues, M. R. M., & Coimbra, M. A. (2016). Melanoidins. In A. Farah (Ed.), *Coffee: Chemistry, Quality and Health Implications*: Royal Society of Chemistry, in press.

Simões, J., Moreira, A. S. P., Passos, C. P., Nunes, F. M., Domingues, M. R. M., & Coimbra, M. A. (2016). Polysaccharides and other carbohydrates. In A. Farah (Ed.), *Coffee: Chemistry, Quality and Health Implications*: Royal Society of Chemistry, in press.

In food industry, thermal processing persists as the most widely used industrial procedure (Bhattacharya, 2015; Sun, 2012). However, high temperatures can promote reactions that change the structure of food constituents, often with formation of unknown compounds and, consequently, without knowing their properties and effects on human health. Thus, the study of structural modifications occurring in food constituents as consequence of thermal processing is of great interest.

The roasting of green coffee beans - the seeds from the fruit of coffee trees - is an important step in the coffee processing. The characteristic aroma, taste, and colour of the coffee brews, prepared by hot water extraction from roasted and ground coffee beans, are to a great extent determined by the roasting process (Illy et al., 1995; Kumazawa & Masuda, 2003; Moon & Shibamoto, 2009; Schenker et al., 2002). During the roasting process, green coffee beans undergo several physical changes, including the colour change to brown. Also, several chemical reactions take place inside of the beans changing the structure of their components (Figure I.1), namely the polysaccharides (Illy et al., 1995). Despite the studies already performed, these structural modifications are far to be completely elucidated.

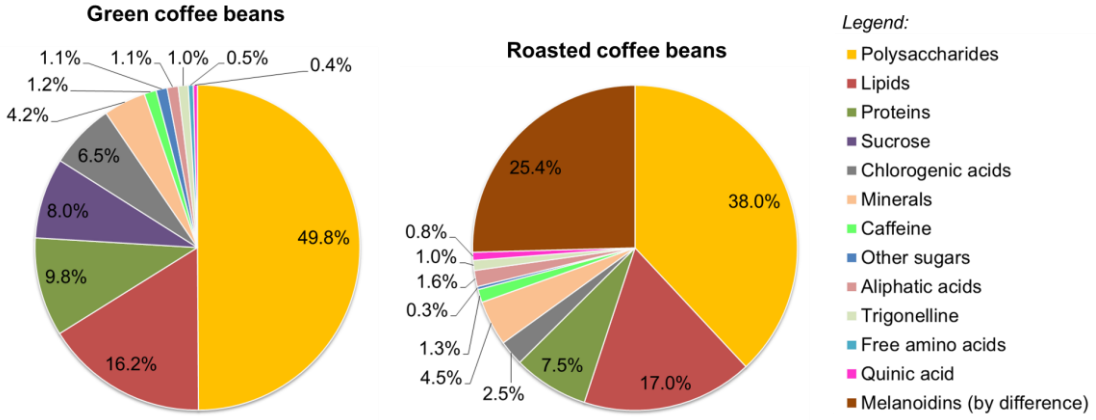

Figure I.1. Chemical composition of green and roasted Arabica beans (% dry matter). Adapted from Illy et al. (1995).

I.1. Green coffee polysaccharides

Polysaccharides are the most abundant constituents of green coffee beans, accounting for about 50% of their dry weight. The whole polysaccharide fraction of green coffee beans is mainly composed by mannose (Man), followed by galactose (Gal), glucose (Glc), and arabinose (Ara) (Bradbury & Halliday, 1990; Fischer, Reimann,

Trovato & Redgwell, 2001; Redgwell, Trovato, Curti & Fischer, 2002b). These monosaccharide residues are the building blocks of three main polysaccharides present in green coffee beans: galactomannans, type II arabinogalactans, and cellulose, which were estimated to account respectively for 22, 14-17, and 8% of the dry weight of green beans of the most commercialized species, *Coffea arabica* (Arabica) and *Coffea canephora* (Robusta) (Bradbury & Halliday, 1990).

Despite several studies since, at least, the 1960s (Wolfrom, Laver & Patin, 1961; Wolfrom & Patin, 1965), a complete structural characterization of the green coffee bean polysaccharides, namely of the galactomannans and arabinogalactans, is far to be achieved. This task is hampered by the high proportion of insoluble polysaccharides in hot water, or even in concentrated alkaline solutions. For example, only 7% of the polysaccharides were extracted from green coffee beans by sequential extraction with hot water, EDTA, and 0.05-4 M NaOH solutions (Oosterveld, Harmsen, Voragen & Schols, 2003a). It is, therefore, important to have in mind that most of the published studies on structural characterization of green coffee polysaccharides are focused on a minor fraction of the total polysaccharides extracted in specific conditions and this is not necessarily representative of the whole bean.

Although higher yields could be obtained using more drastic extraction conditions, it is important to note that the study of hot water soluble coffee polysaccharides is of special interest due the high consumption of coffee brews, prepared by hot water extraction from roasted and ground coffee beans. In this context, the study of hot water soluble coffee polysaccharides extracted from green beans is often used as a control, allowing distinguishing structural features of green coffee polysaccharides from changes induced by roasting. Typically, the polysaccharides are isolated from the hot water soluble material taking advantage of their high molecular weight, namely by dialysis. In the hot water extractable high molecular weight material (HMWM) recovered from green coffee beans, galactomannans and arabinogalactans were shown to be present in mixture with other compounds, namely, proteins (Nunes & Coimbra, 2001, 2002b). Thus, to obtain a detailed structural characterization of green coffee polysaccharides present in the hot water soluble HMWM, additional purification procedures are required. Despite all the difficulties, some structural features of green coffee polysaccharides, particularly of galactomannans and arabinogalactans, have been disclosed so far. The following sections summarize what is known to date about the structure of these polysaccharides.

I.1.1. Galactomannans

Generically, the structure of galactomannans (from green coffee beans or other seeds) consists of a linear backbone of (β1→4)-linked D-Man residues, some of them substituted at *O*-6 by single residues of α-D-Gal (Dey, 1978; Prajapati et al., 2013).

In one of the studies on structural characterization of green coffee polysaccharides, in the 90s, Bradbury & Halliday (1990) used a sequential fractionation procedure, based on a delignification treatment, an acid wash, and subsequent alkali extraction followed by neutralization with HCl, to isolate a fraction with high content of galactomannans from Robusta beans. This fraction represented 9.4% of the beans' dry weight, and had 94.0% of Man, 3.3% of Gal, and 1.7% of Glc. As found in other studies, these authors preferred to use the term "mannans" instead of "galactomannans" due the low content of Gal in this fraction. According to Aspinall (1959), the term "galactomannans" should be used for mannans containing more than 5% of D-Gal.

Regarding the mannan fraction isolated by Bradbury & Halliday (1990), glycosidic linkage analysis allowed to estimate a ratio of total Man/terminally linked Man of 13 and a ratio of (1→4,6)-linked/Total Man of 0.007, which are, respectively, diagnostic of the degree of polymerization (DP) and degree of branching (DB), considering the usual structure of galactomannans. In other words, these ratios are diagnostic of the number of Man residues constituent of the main backbone and the proportion of the substituted Man residues, respectively. However, the glycosidic linkage analysis of whole beans showed that the galactomannans were more polymerized (DP=36) and branched (DB=0.062) than the isolated ones. For green Arabica beans, the values of DP 24 and DB 0.049 were estimated for the galactomannans, also based on the whole-bean glycosidic linkage analysis (Oosterveld, et al., 2003a), and the glycosidic linkage analysis of a galactomannan-enriched fraction containing 50.8% of Man residues showed a DP of 69 and a DB of 0.066 (Redgwell et al., 2002b). The different values of DP and DB are not necessarily due to the coffee species, as they may be due to the different methodologies used. It is possible that some polysaccharide degradation has occurred under more drastic extraction conditions. On the other hand, even assuming that there was no degradation, it is possible that the different methodologies have allowed extracting fractions of galactomannans having distinct structures. Indeed, the structural heterogeneity of coffee galactomannans was corroborated by labelling green coffee bean sections with a mannan-specific monoclonal antibody, suggesting the existence of galactomannans with different

DBs (Sutherland et al., 2004). Also, the different DBs reported in the literature for green coffee galactomannans may be due to differences in the exact developmental stage of the samples used, since the Gal/Man ratio of coffee galactomannans changes during the bean development (Marraccini et al., 2005; Redgwell et al., 2003).

Detailed studies on hot water soluble green coffee galactomannans allowed assigning the structural features shown in Figure I.2 using the symbol notation proposed by the Consortium for Functional Glycomics (Berger, McBride, Razi & Paulson, 2008).

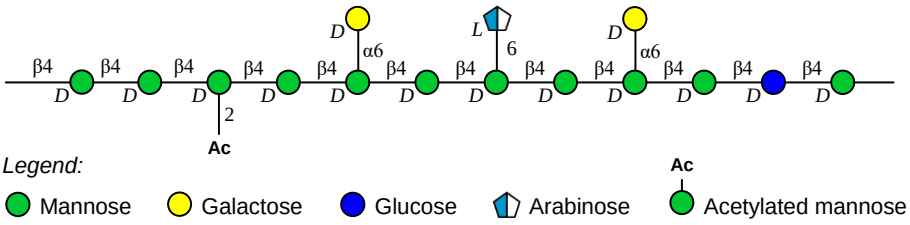

Figure I.2. Illustration of the main structural features of hot water soluble green coffee galactomannans.

In a study of Oosterveld et al. (2004), the galactomannans present in the high molecular weight material (HMWM) extracted with hot water from green coffee beans were separated into two neutral fractions by anion-exchange chromatography. These fractions had 45 and 48% of sugars, of which Man was the most abundant (70 and 73 mol%). The further characterization of these fractions by nuclear magnetic resonance (NMR) spectroscopy and by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) after enzymatic hydrolysis with an *endo*-mannanase, which cleaves the (*β*1→4)-D-mannan backbone between adjacent unsubstituted (β1→4)-linked Man residues, confirmed the presence of acetyl groups linked to the galactomannans, as suggested by using a commercial acetyl kit. In the fraction of higher average molecular weight, 9% of the Man residues were acetylated, whereas in the other fraction there were 4%.

In order to obtain more purified hot water soluble galactomannans while keeping intact their structural features, a soft fractionation procedure was developed by Nunes et al. (2005). Briefly, galactomannans were purified from the hot water extractable HMWM by precipitation in 50% ethanol solutions, followed by the recovery of the non-retained fraction from the anion-exchange chromatography on a Q-Sepharose FF stationary phase. This non-retained fraction was further fractionated by phenylboronic acid immobilized chromatography. Due to the interaction of the phenylboronic acid with free *cis*-diol

groups present in terminally linked Gal residues and in (1→4)-linked and terminally linked Man residues, galactomannans are retained and selectively eluted with mannitol.

For green coffee beans, the percentage of mannan residues in the fraction eluted with mannitol corresponded only to 0.76% of the total Man residues present in the green coffee beans. Based on the results of glycosidic linkage analysis, side chains of arabinose (Ara) residues (2%) were suggested as structural features of hot water soluble green coffee galactomannans. The analysis by electrospray ionization tandem mass spectrometry (ESI-MS/MS) of pentose substituted oligosaccharides produced by enzymatic hydrolysis with an *endo*-β-mannanase allowed to unambiguously concluding that terminally linked Ara residues are linked at *O*-6 to the (β1→4)-linked Man residues in the backbone. These results are in agreement with earlier findings from hot water extracts of roasted coffee, suggesting that Ara residues are constituents of coffee bean galactomannans (Navarini et al., 1999). As determined by gas chromatography (GC) and enzymatic analysis after alkaline hydrolysis, there were 11 mol % of acetylated Man residues. The ESI-MS/MS analysis of acetylated oligosaccharides produced by enzymatic hydrolysis indicated that the acetylation occurs at *O*-2 and/or *O*-3 of Man residues. Single acetylated residues, doubly acetylated residues, and consecutively acetylated residues were found. Also, glycosidic linkage analysis, hydrolysis with specific glycosidases, and analysis of the reduced and methylated oligosaccharides by GC with electron impact mass spectrometry (EI-MS), allowed concluding that (β1→4)-linked Glc residues (6 mol%) are also components of the mannan backbone (Nunes et al., 2005). This opens the possibility that the 1.7% of Glc quantified by Bradbury & Halliday (1990), although not considered as component of the mannan backbone by these authors, could in fact be a structural feature of this coffee bean polysaccharide.

I.1.2. Type II arabinogalactans

Generically, the structure of type II arabinogalactans (from green coffee beans and other plant sources) consist of a main backbone of (β1→3)-linked D-Gal residues, some of them substituted at *O*-6 by (β1→6)-linked D-Gal residues, which are in turn substituted by Ara and other less abundant sugar residues. These polysaccharides are usually found covalently linked to proteins, and therefore referred as the carbohydrate moiety of arabinogalactan-proteins (AGPs) (Showalter, 2001).

In the study of Bradbury & Halliday (1990), an arabinogalactan fraction was also isolated from green coffee beans. This had 48.2% Gal, 19.8% Ara, 1.1% rhamnose (Rha), and 0.8% Man. Glycosidic linkage analysis of whole beans showed that the Gal residues were mainly (1→3)- and (1→3,6)-linked, whereas the Ara residues (all identified in the furanosidic form) were mainly terminally and (1→5)-linked. The possibility of covalent linkages to proteins was suggested by the increase of soluble arabinogalactans after the treatment with a protease.

Redgwell et al. (2002a) were able to solubilize over 90% of the arabinogalactan content of green coffee beans by a combination of chemical and enzymatic treatments. All the recovered fractions gave a positive result for the AGP-specific β-glucosyl Yariv reagent, demonstrating that almost all of the coffee arabinogalactans are covalently linked to proteins. For three fractions recovered from Arabica beans, it was shown that the protein moieties accounted for 0.4-1.9% of the AGPs and contained 7-12% of hydroxyproline, one of the most abundant amino acids found in AGPs of other sources. In agreement with earlier findings (Fischer et al., 2001), the arabinogalactan moieties were shown to be heterogeneous with regard to their degree of branching (DB), as well as the monosaccharide composition and extent of their side chains. The heterogeneity of the arabinogalactan moieties, particularly related to the extent of their Ara side chains, was corroborated in a later study by labelling of green coffee bean sections with two monoclonal antibodies: one recognizing contiguous (α1→5)-linked Ara residues and other one recognizing a carbohydrate epitope containing β-linked glucuronic acid (GlcA) residues (Sutherland et al., 2004). Single GlcA residues have been shown to occur in the terminal position of the (β1→6)-linked D-Gal side chains (Redgwell et al., 2002a).

Nunes et al. (2008) also developed a soft fractionation procedure, based on graded ethanol fractionation, anion exchange chromatography on Q-Sepharose FF, and size exclusion chromatography (SEC) on Sephacryl S-300 HR, obtaining a pure fraction of hot water soluble arabinogalactans. This fraction had 54% Gal, 37% Ara, 9% Rha, and traces of Man (0.4%), as determined by sugar analysis. Linkage analysis showed that the Rha residues were exclusively terminally linked. Methanolysis analysis showed the presence of GlcA in an amount of nearly 2%, a very low percentage when compared to the 7-10% that has been previously reported (Redgwell et al., 2002a). In contrast, the amount of protein (4%) was higher than that previously reported (0.4-1.9%). The further characterization of the purified arabinogalactans by NMR and, after partial acid hydrolysis, by electrospray ionization tandem mass spectrometry (ESI-MS/MS), showed

the occurrence of single α-L-Ara*f* residues and [L-Ara*f*-(α1→5)-L-Ara*f*-(α1→] disaccharides at *O*-3 position of the (β1→6)-linked D-Gal residues. Also, [L-Rha*p*-(α1→5)-L-Ara*f*-(α1→] disaccharides and [L-Rha*p*-(α1→5)-L-Ara*f*-(α1→5)-L-Ara*f*-(α1→] trisaccharides were shown to be present at *O*-3 position of the (β1→6)-linked D-Gal residues (Figure I.3).

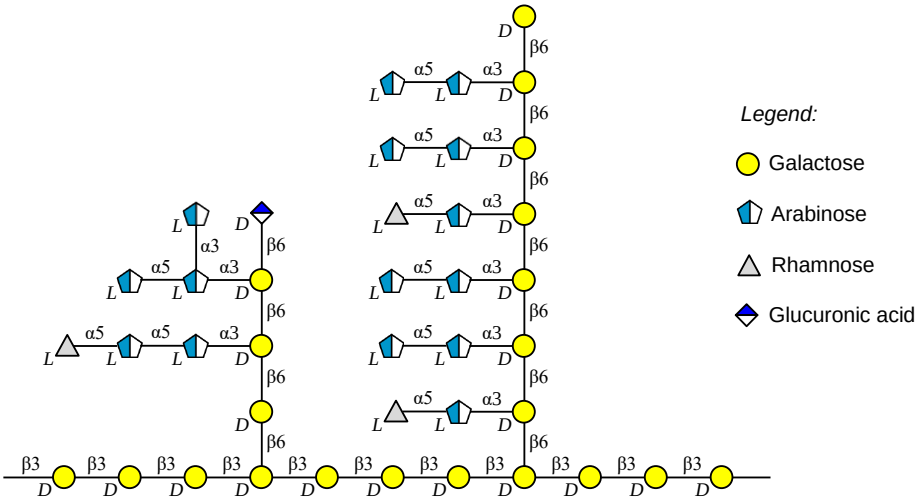

Figure I.3. Illustration of the main structural features of hot water soluble green coffee arabinogalactans.

I.2. Structural modifications induced by roasting

During roasting, the green coffee beans are submitted to high temperatures for a certain period of time. In most of the conventional roasters, hot air is used to roast. The temperature inside the beans rises to about 200 °C. Beyond other physical changes, the beans undergo weight loss and acquired a brown coloration. These physical changes are often used as indicators of the degree of roast, which is commonly categorized as light, medium, or dark roast (Illy et al., 1995; Wei & Tanokura, 2015).

The roasting process changes the amount of extractable polysaccharides, but also the relative amount of extracted polysaccharides. For instance, Oosterveld et al. (2003a) were able to extract 29% of the polysaccharides from roasted Arabica coffee beans by sequential extraction with hot water, EDTA, and 0.05-4M NaOH solutions, whereas only 7% were extracted from the corresponding green beans using the same procedure. Nunes & Coimbra (2001, 2002a, 2002b) analysed the high molecular weight material (HMWM) extracted with hot water from Arabica and Robusta coffee beans. From data of glycosidic linkage analysis, the amount of galactomannans was estimated considering that all Man

residues are components of the galactomannans and that an amount of Gal residues equal to the amount of (1→4,6)-linked Man residues is also a component of galactomannans. Similarly, the amount of arabinogalactans was estimated considering that their components are all Ara residues and all Gal residues except those components of galactomannans. According to these criteria, galactomannans represented only 24% of the polysaccharides of the HMWM extracted from green coffee beans, whereas arabinogalactans represented 57-69%.

In contrast to what occurs with green coffee beans, the hot water extractable polysaccharides from roasted coffee beans are mainly galactomannans (Nunes & Coimbra, 2001, 2002a, 2002b; Oosterveld, Voragen & Schols, 2003b). However, as estimated based on the results of glycosidic linkage analysis, different amounts of galactomannans can be extracted with hot water, depending of the coffee species as well as the degree of roast. For three degrees of roast, Arabica coffee infusions showed higher amount of galactomannans (60-71%) than Robusta (45-60%). On other hand, the amount of water extractable galactomannans from Robusta coffees increased with increasing degree of roast, whereas the amount of water extractable galactomannans from Arabica coffees was about the same for different degrees of roast. Contrarily, the amount of arabinogalactans extracted from Robusta and Arabica roasted coffees was similar. Also, for both Arabica (Brazil) and Robusta (Uganda), the amount of arabinogalactans extracted for the roasted coffees decreased with increasing degree of roast (Nunes & Coimbra, 2001, 2002a, 2002b).

The roasting of coffee beans promotes an increase of the bean volume and the appearance of larger micropores in the cell walls (Schenker et al., 2000). The differences in the extraction yield of galactomannans between Arabica and Robusta roasted coffees are possibly due to the different susceptibility of the cell walls during the roasting process. Using the same roasting conditions, the cell walls of Arabica coffee beans were shown to be more degraded and contain more pores than the cell walls of Robusta (Gutiérrez, Ortolá, Chiralt & Fito, 1993).

Beyond the modification of the structure of coffee beans during the roasting process, the changes in the hot water extractability of galactomannans and arabinogalactans is also possibly due to the changes on their structures. The polysaccharide content of Arabica coffee beans was estimated to decrease by 12-40%, depending on the degree of roast (Redgwell et al., 2002b). However, the thermal stability of the galactomannans and arabinogalactans is markedly different. Arabinogalactans are

the most susceptible of the coffee polysaccharides to degradation during roasting (Simões et al., 2014). Particularly, the Ara in their side chains is the sugar unit most labile (Oosterveld et al., 2003b; Redgwell et al., 2002b). Although the higher thermal stability of the galactomannans compared to that of the arabinogalactans, both polysaccharides undergo structural modifications during roasting, as described in the following sections.

I.2.1. Depolymerization, debranching and transglycosylation reactions

In the studies of Nunes & Coimbra (2001, 2002a, 2002b), independently of the coffee species (Arabica or Robusta) and the degree of roast, the relative amount of terminally linked mannopyranose (Man*p*) residues in the galactomannan-enriched fraction recovered from the hot water extractable high molecular weight material (HMWM) of roasted coffees by precipitation in 50% ethanol solutions was higher than that obtained for the correspondent green coffees. This showed that lower molecular weight galactomannans are extracted upon roasting and corroborate the occurrence of depolymerization during roasting. The relative amount of (1→4,6)-linked Man*p* residues were also lower than those found for the correspondent green coffees, showing that less branched galactomannans were also extracted upon roasting. For both Arabica and Robusta, the linkage analysis of the HMWM extracted from roasted beans, when compared with that of the corresponding green beans, showed a higher ratio of (1→3)-linked galactopyranose (Gal*p*)/(1→3,6)-linked Gal*p*, suggesting that less branched arabinogalactans were also extracted upon roasting. Also, an increase of terminally linked arabinofuranose (Ara*f*) in relation to the (1→5)-linked Ara*f* with the increasing degree of roast was observed, suggesting the decrease of the size of the Ara side chains.

Although the evidences for depolymerization and debranching of the galactomannans and arabinogalactans, higher molecular weight derivatives can also be formed during coffee roasting via transglycosylation reactions. The roasting of mannosyl and galactomannosyl oligosaccharides, structurally related to coffee galactomannans, promoted the formation of oligosaccharides with a higher degree of polymerization (DP) and new types of glycosidic linkages (Moreira et al., 2011). More recently, the roasting of spent coffee grounds (SCG) galactomannans also promoted the formation of new types of glycosidic linkages (Simões et al., 2014). Together, these studies support the hypothesis of the occurrence of non-enzymatic transglycosylation reactions during coffee roasting. In particular, the occurrence of transglycosylation reactions between

galactomannans and arabinogalactans, especially involving Ara residues of arabinogalactan side chains, was hypothesized and investigated in this work (Section III.2). The relative amount of Ara residues found as side chains of hot water soluble roasted coffee galactomannans (<0.9 mol%) was lower than that of green coffee (2 mol%) (Nunes et al., 2005). This finding, however, does not exclude the presented hypothesis, because galactomannans containing a higher amount of Ara residues can occur in the unextractable polysaccharide fraction, as supported by sugar and glycosidic linkage analyses of the galactomannans recovered from SCG (Simões, Nunes, Domingues & Coimbra, 2010). Also, the possible products formed during coffee roasting by transglycosylation reactions between galactomannans and arabinogalactans can be incorporated in melanoidin structures, since these both polysaccharides are known to be involved in their formation (Section I.2.3).

I.2.2. Modifications occurring at the reducing end of galactomannans

Using the same procedure used for green coffee galactomannans (Section I.1.1), roasted coffee galactomannans were purified from the HMWM. Their structural characterization also corroborated that the roasting process is responsible for a change in the structural features of coffee galactomannans (Nunes et al., 2005; Nunes, Reis, Domingues & Coimbra, 2006) (Figure I.4).

Whereas Glc residues were identified in the green coffee as a constituent of the mannan backbone, in the roasted coffee they were detected only at the reducing end of the mannan backbone. The (1→4)-linked Glc residues at the reducing end of the mannan can formed during roasting from reducing Man residues via isomerization at C2. This hypothesis was supported by the increase of (1→4)-linked Glc residues observed when mannosyl and galactomannosyl oligosaccharides were submitted to dry thermal treatments at 200 °C (Moreira et al., 2011), and a coffee galactomannan-rich fraction was submitted to dry thermal treatments at different temperatures (180-240 ºC) (Simões et al., 2014). In addition, the modification of coffee galactomannans by isomerization reactions during roasting was also corroborated with the identification of fructose, which can be formed by ring opening of the reducing end Man residues through a 1,2-keto-enol tautomery. On other hand, the identification by electrospray ionization tandem mass spectrometry (ESI-MS/MS) of 1,6-β-anhydromannose, 2-deoxyarabinose (2-deoxy-*erythro*-pentose), acetylformoin, mannonic acid, arabinonic acid, and Amadori

compounds at the reducing end of oligosaccharides produced by enzymatic hydrolysis with an *endo*-β-mannanase showed that the coffee galactomannans are also modified at the reducing end by the occurrence of caramelization, oxidation/decarboxylation, and Maillard reactions (Nunes et al., 2006).

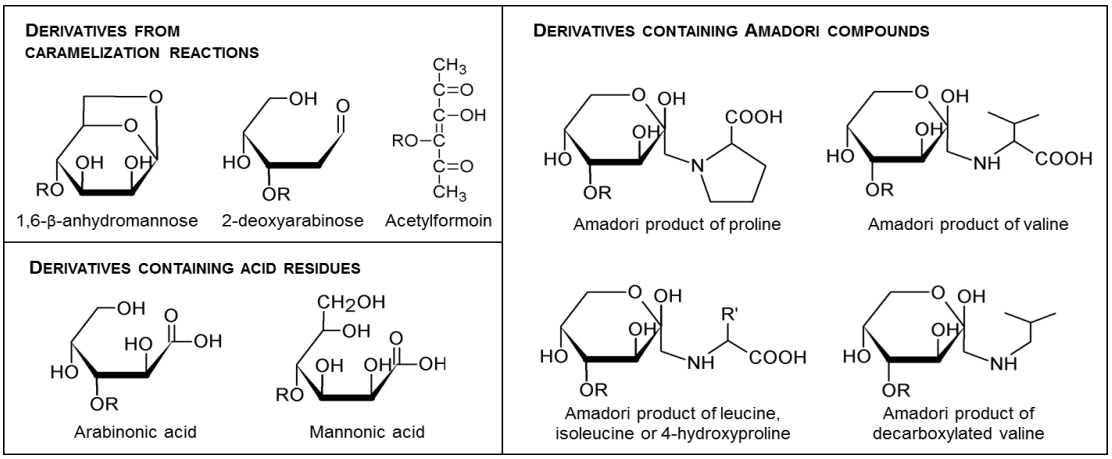

Figure I.4. Structural modifications identified at the reducing end of galactomannans isolated by hot water extraction of roasted coffee beans. Adapted from Nunes et al. (2006).

I.2.2.1. Hydroxyl radical-induced oxidation

Despite the oxidative modifications identified at the reducing end of galactomannans, the exact mechanisms of oxidation are still unknown. However, studies by electron paramagnetic resonance spectroscopy have demonstrated the formation of free radicals during coffee roasting (Goodman, Pascual & Yeretzian, 2011; Levěque, Godechal & Gallez, 2008; Nebesny & Budryn, 2006), as well as the presence of free radicals in soluble coffee (Pascual, Goodman & Yeretzian, 2002). These studies suggest that free radical reactions may occur extensively during coffee roasting. Particularly, hydroxyl radicals ($HO^{\bullet}$) were proposed to be involved in the oxidation of coffee galactomannans occurring during roasting, as in the thermal degradation of cellulose (Scheirs, Camino & Tumiatti, 2001; Shen & Gu, 2009). Also, the formation of the $HO^{\bullet}$ during coffee roasting was inferred by the detection of 8-oxocaffeine in roasted coffee powders (Stadler & Fay, 1995). In accordance, the oxidative modifications identified at the reducing end of roasted coffee galactomannans, namely mannonic and 2-ketogluconic acids (Nunes et al., 2006), were identified in the $HO^{\bullet}$-induced oxidation of mannosyl and galactomannosyl oligosaccharides, structurally related to coffee galactomannans (Tudella et al., 2011). As for galactomannans, it can be postulated that these reactive oxygen

species are also involved in the oxidation of coffee arabinogalactans during roasting. $HO^{\bullet}$ have also been implicated in the structural modification of polysaccharides in other industrial processes, such as the bleaching in the papermaking industry (Guay et al., 2001), and in the scission of polysaccharides in biological processes (Duan & Kasper, 2011; Fry, 1998; Schweikert, Liszkay & Schopfer, 2000).

The presence of hydrogen peroxide in roasted coffee extracts, and its absence in green extracts, have also been reported (Hegele, Münch & Pischetsrieder, 2009; Mueller et al., 2011). The occurrence of hydrogen peroxide during coffee roasting is of particular interest, since $HO^{\bullet}$ can be formed via the decomposition of hydrogen peroxide in the presence of transition metal ions such as Fe^{2+} (Eq. (1) by Fenton reaction). $HO^{\bullet}$ are able to abstract hydrogen atoms from C–H bonds of carbohydrates (R–H), generating carbohydrate radicals ($R^{\bullet}$) (Eq. (2)). They are specifically carbon-centered radicals as α-hydroxyalkyl radicals (HO–C$^{\bullet}$<) (Arts, Mombarg, van Bekkum & Sheldon, 1997; Isbell & Frush, 1987).

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^- + HO^{\bullet} \quad (1)$$

$$R–H + HO^{\bullet} \rightarrow R^{\bullet} + H_2O \quad (2)$$

The hydrogen abstraction from C–H bonds of carbohydrates is rather non-specific though stereo dependent. For example, evidences were obtained for random hydrogen abstraction at all ring C–H bonds of the hyaluronan polysaccharide, except at C2 of N-acetylglucosamine units (Šoltés et al., 2006). Despite the stereo effects limiting the formation of carbohydrate radicals, a vast number of oxidation products can be formed during carbohydrate oxidation induced by $HO^{\bullet}$, as previously observed for the mannosyl and galactomannosyl oligosaccharides submitted to conditions of Fenton reaction (Fe^{2+}/H_2O_2) (Tudella et al., 2011). However, previous studies have reported the carbohydrate oxidation with $HO^{\bullet}$ generated simply by the presence of Fe^{2+} (Faure, Andersen & Nyström, 2012; Faure, Werder & Nyström, 2013; Gutteridge, 1987; Uchiyama, Dobashi, Ohkouchi & Nagasawa, 1990). The most reported mechanism by which Fe^{2+} is involved in formation of $HO^{\bullet}$ includes three reactions (Ryan & Aust, 1992; Welch, Davis & Aust, 2002): first, superoxide radical ($O_2^{\bullet-}$) is formed through the oxidation of Fe^{2+} (Eq. (3)); then, the dismutation of superoxide leads to the formation of H_2O_2 (Eq. (4)); in the presence of H_2O_2 and Fe^{2+}, the formation of $HO^{\bullet}$ occurs according to the Fenton reaction (Eq. (1)).

$$Fe^{2+} + O_2 \rightarrow Fe^{3+} + O_2^{\bullet-} \quad (3)$$

$$2O_2^{\bullet-} + 2H^+ \rightarrow H_2O_2 + O_2 \quad (4)$$

I.2.2.2. Caramelization and Maillard reactions

The Maillard reaction and caramelization are known as non-enzymatic browning reactions which produce brown coloured compounds during thermal processing of a large range of food products, including coffee beans. Despite several studies, these non-enzymatic browning reactions are far to be completely understood. They are extremely complex (Eskin, Ho & Shahidi, 2013; Purlis, 2010).

The Maillard reaction encompasses not one reaction but a network of various reactions between carbonyl compounds, especially reducing sugars, and compounds with free amino groups, such as amino acids, peptides, and proteins (Eskin et al., 2013). Hodge (1953) published the first comprehensive scheme of this network. Accordingly, the first step in the Maillard reaction involves a condensation between a carbonyl compound (e.g. an aldose) and a compound having a free amino group (e.g. an amino acid), with formation of an imine (*Schiff* base), which is subsequently rearranged to form a more stable product. In the case of an aldose, it is formed an aminoketose called "Amadori compound". This compound is subsequently transformed. In the final stage of the MR, the intermediate compounds polymerize, leading to the formation of brown nitrogenous polymers, named as melanoidins. The Amadori compounds of different amino acid residues identified at the reducing end of roasted coffee galactomannans support the occurrence of the Maillard reaction during coffee roasting (Nunes et al., 2006). Melanoidins are also known to be formed during coffee roasting (Section I.2.3).

Caramelization is a term for describing a complex group of reactions that occur when polyhydroxycarbonyl compounds, namely reducing sugars, are exposed to high temperatures. Distinctly of the Maillard reaction, amino compounds are not involved (Eskin et al., 2013; Hodge, 1953). As described by Kroh (1994), the caramelization products consist of volatile and non-volatile compounds of low and high molecular weight. Also, Kroh (1994) described a sequence of sugar (aldose) degradation reactions characterized by the initial enolization, followed by dehydration, dicarbonyl cleavage, retro-aldol reaction, aldol condensation and, finally, a radical reaction. From these reactions, the osuloses (α-dicarbonyl compounds) were identified as key intermediates of the caramelization, which give rise to volatile products responsible for the characteristic caramel flavor. It is worth to note that some products, such as the acetylformoin identified at the reducing end of roasted coffee galactomannans (Nunes et al., 2006), can be formed under caramelization and Maillard reaction conditions. Also, according to model studies,

caramelization characteristically requires more energy to get started than the Maillard reaction (Hodge, 1953).

I.2.3. Formation of melanoidins

Melanoidins are generically defined as high molecular weight nitrogenous brown-coloured compounds generated in the final stage of the Maillard reaction. Scarce information is available about their chemical structures, although they are formed during the heat processing of a large range of food products beyond coffee, such as bread, malt and meat. Due the uncertainties about their structures, they are usually quantified by difference, subtracting the total percentage of known compounds from 100 percent. Using this criterion, they were estimated to account for up to around 25% (w/w) of the dry weight of roasted coffee beans (Illy et al., 1995).

Indeed, coffee brews, prepared by hot water extraction from ground and roasted coffee beans, are considered one of the main sources of melanoidins in human diet (Fogliano & Morales, 2011). Consequently, their health implications are of great interest. In fact, several biological activities, such as antioxidant, antimicrobial, anticariogenic, anti-inflammatory, antihypertensive and antiglycative activities (Rufián-Henares & Morales, 2007; Stauder et al., 2010; Verzelloni et al., 2011; Vitaglione et al., 2010), have been attributed to coffee melanoidins. However, to understand their biological activities, it is essential to know their chemical structures.

Since at least the 1960s (Maier, Diemair & Ganssmann, 1968), several attempts have been made to elucidate the structure of coffee brew melanoidins. In recent years, their structural diversity has been evidenced with the purification of different melanoidin populations by applying chromatographic separation techniques and other specific isolation procedures. On the other hand, the chemical characterization of these purified melanoidin populations has given increasing evidences that polysaccharides (galactomannans and arabinogalactans), proteins, and chlorogenic acids are involved in the formation of coffee melanoidins (Bekedam et al., 2007; Bekedam et al., 2008a; Bekedam, Schols, Van Boekel & Smit, 2008b; Gniechwitz et al., 2008b; Nunes & Coimbra, 2007; Nunes & Coimbra, 2010). More recently, evidences for the contribution of sucrose to coffee melanoidin formation were obtained by using of modified “in bean” green coffee models (Nunes, Cruz & Coimbra, 2012). However, it is still unclear how these different constituents (or their derivatives) are linked. Figure I.5 is a simplistic

illustration of coffee melanoidin formation, since the exact mechanisms involved also remain unclear.

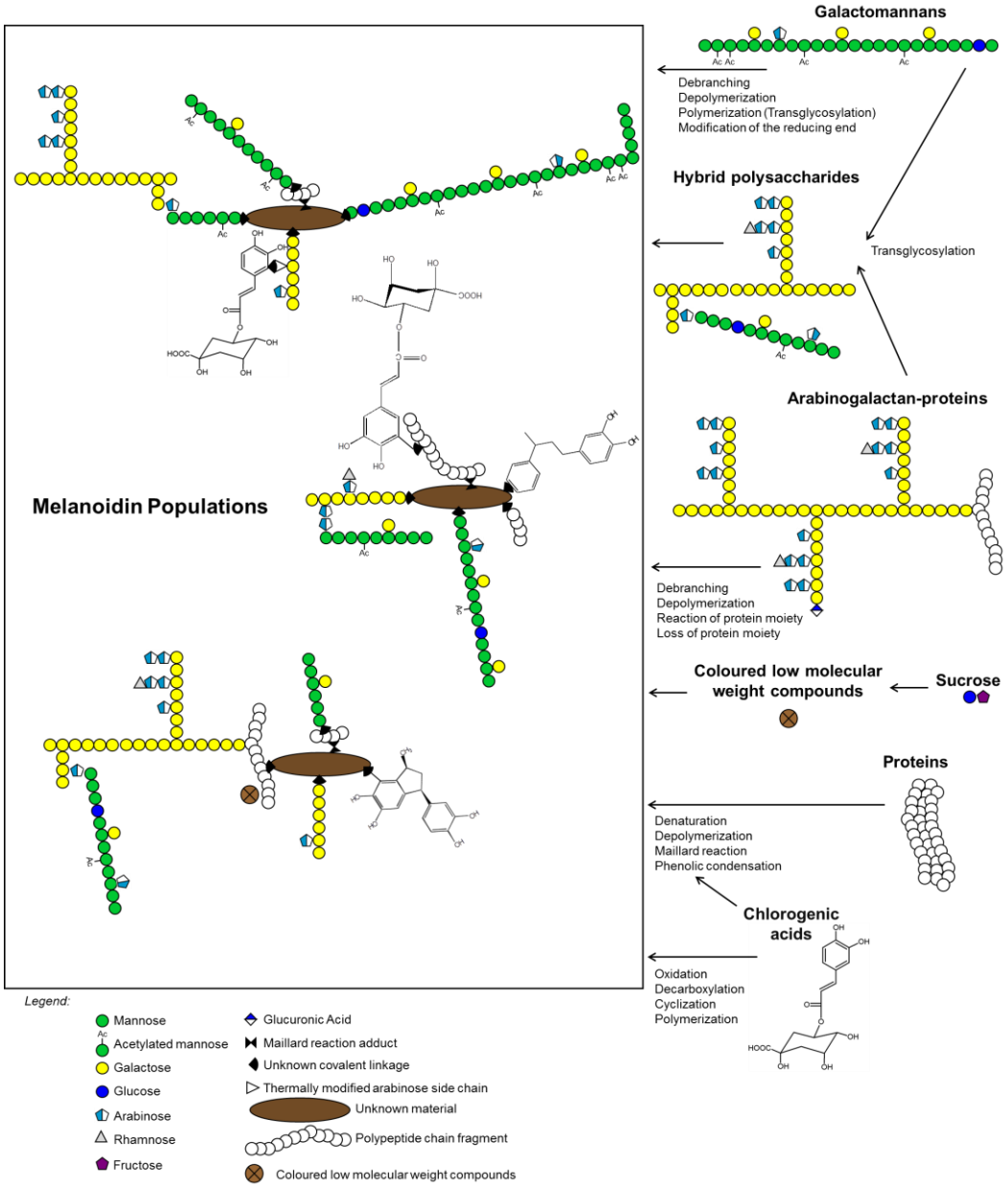

Figure I.5. Simplified illustration of coffee melanoidins formation. Adapted from Nunes et al. (2010) and Moreira et al. (2012).

I.2.3.1. Contribution of polysaccharides

Melanoidin populations with different polysaccharide composition have been isolated from the HMWM of coffee brews (Bekedam et al., 2007; Bekedam et al., 2008a; Gniechwitz et al., 2008a; Nunes & Coimbra, 2007). The existence of polysaccharides

covalently linked in melanoidins was first proven by Nunes et al. (2007) by the observation of galactomannans and arabinogalactans in anionic fractions isolated from the ethanol fractions of roasted coffee brews by anion exchange chromatography, which contrasted with the negligible retention observed for the green coffee polysaccharides under the same chromatographic conditions. As demonstrated by the addition of the Yariv reagent to the HMWM of a roasted coffee brew, arabinogalactans can be incorporated in coffee melanoidin structures in the form of intact arabinogalactan-proteins (AGPs). Also, it was shown that the AGP fraction accounts for approximately half of the melanoidins present in the HMWM of the coffee brew (Bekedam et al., 2007).

I.2.3.2. Contribution of proteins

Proteins of green coffee beans account for about 10% of their dry weight. The roasting process leads to protein degradation, as inferred by the decrease of total amount of amino acids identified in coffee beans upon roasting (Illy et al., 1995). The changes in green coffee proteins promoted by roasting have also been followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Montavon, Mauron & Duruz, 2003; Nunes & Coimbra, 2001, 2002a, 2002b). The SDS-PAGE patterns obtained under non-reducing conditions of the HMWM isolated from green and roasted coffee brews are clearly distinct. While green coffees presented a major protein band at 58 kDa and a second one at 38 kDa, roasted coffees presented a defined band with ≤14 kDa and a diffuse band with >200 kDa (Nunes & Coimbra, 2001, 2002a, 2002b), the later possibly due to the involvement of proteins in melanoidins formation.

Different melanoidin fractions have been isolated from the HMWM of coffee brews containing variable amounts of protein-like material, quantified by amino acid analysis after acid hydrolysis (Bekedam, Schols, van Boekel & Smit, 2006; Gniechwitz et al., 2008b; Nunes & Coimbra, 2007). Among other amino acids, various melanoidin fractions contain hydroxyproline (Nunes & Coimbra, 2007), an amino acid found in high amounts in arabinogalactan-proteins (AGPs) isolated from green coffee beans (Redgwell et al., 2002a) and green coffee brews (Nunes et al., 2008). The presence of hydroxyproline in melanoidin fractions containing also Ara and Gal residues is another evidence for the existence of AGPs in coffee melanoidin structures, as discussed earlier. Apart from the intact ones, amino acids modified during coffee roasting are also incorporated in melanoidin structures, possibly contributing to the non-amino acid nitrogen identified in

different melanoidin fractions (Bekedam et al., 2006). Although at very low amounts, Maillard reaction products derived from lysine, including N^{ε}-(fructosyl)lysine (FL, detected as furosine after acid hydrolysis), N^{ε}-(carboxymethyl)lysine (CML), and N^{ε}-(carboxyethyl)lysine (CEL), were identified in the HMWM and melanoidin fractions (Nunes & Coimbra, 2007).

I.2.3.3. Contribution of chlorogenic acids

Chlorogenic acids (CGAs) consist of a quinic acid (QA) moiety esterified to one or more *trans*-cinnamic acids, such as caffeic, *p*-coumaric, and ferulic acids (Clifford, 2000). In green coffee beans, the most abundant is a mono-ester of caffeic acid (CA) named as 5-O-caffeoylquinic acid (5-CQA) (Moon & Shibamoto, 2009; Perrone et al., 2008), according to the IUPAC nomenclature for numbering of the cyclohexane ring of the QA (IUPAC, 1976) (Figure I.6A).

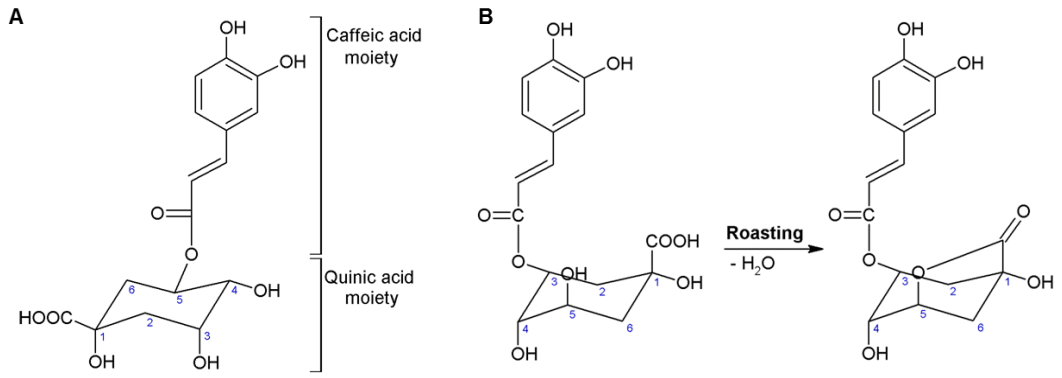

Figure I.6. (A) Structure of 5-*O*-caffeoylquinic acid (5-CQA). (B) Formation of a 1,5-γ-quinolactone from 3-*O*-caffeoylquinic acid (3-CQA) occurring during roasting, as proposed by Farah et al. (2006).

The total CGA content found in green coffee beans accounts for up to 10% of their dry weight (Illy et al., 1995). Upon roasting, the amount of hot water extractable CGAs decreases by 50% or more, depending on roasting intensity (Moon, Yoo & Shibamoto, 2009). In respect to the products derived from the thermal degradation of CGAs during coffee roasting, it is already known that they are diverse and range from simple phenols to condensation products of high structural complexity, as reviewed by Nunes et al. (2010). Briefly, the coffee roasting process promotes the isomerization of CGAs, as well as their hydrolysis yielding QA (non-phenolic moiety) and various cinnamic acids (phenolic moieties) (Clifford, 2000; Farah, de Paulis, Trugo & Martin, 2005; Leloup,

Louvrier & Liardon, 1995). Model studies on dry thermal treatment of 5-CQA, CA, and ferulic acid suggested the occurrence of decarboxylation of QA and cinnamic acids during coffee roasting, yielding a range of simple phenols (Fiddler, Parker, Wasserman & Doerr, 1967; Sharma, Fisher & Hajaligol, 2002; Stadler, Welti, Stämpfli & Fay, 1996). For CA, several condensation products were also identified (Stadler et al., 1996). Accordingly, bitter compounds present in coffee brews seem to be generated by oligomerization of a simple phenol (4-vinylcatechol) released from CA moieties upon roasting (Frank et al., 2007). On the other hand, QA isomers and γ-quinides (lactones) seem to be generated from QA moieties (Scholz & Maier, 1990; Wei et al., 2012).

Lactones are formed by intramolecular esterification of the corresponding hydroxycarboxylic acids with the release of a water molecule. In addition to the QA lactones, CGA lactones are formed during coffee roasting (Farah et al., 2006; Farah et al., 2005; Perrone et al., 2008) (Figure I.6B). However, it was shown that the loss of water from the QA moiety of CGAs also occurs by other pathway, yielding a shikimic acid moiety (Jaiswal et al., 2012; Jaiswal, Matei, Subedi & Kuhnert, 2014). Beyond dehydration, coffee CGAs seem to be modified during roasting by acyl migration, epimerization, and transesterification with acetic, quinic and shikimic acid (Jaiswal et al., 2012).

The presence of covalently-linked CGA derivatives in coffee melanoidin fractions was first demonstrated by using the alkaline fusion method, known as an efficient method to release condensed phenolic structures (Nunes & Coimbra, 2007; Takenaka et al., 2005). The incorporation of covalently-linked CGA derivatives, namely ester-linked phenolic and QA moieties, as well as the presence of intact CGAs incorporated via CA moiety through mainly non-ester linkages, was corroborated by subsequent studies using saponification and enzymatic treatments (Bekedam et al., 2008a; Bekedam et al., 2008b; Perrone, Farah & Donangelo, 2012), and other study using a modified "in bean" coffee model enriched with a typical green coffee CGA mixture (Coelho et al., 2014).

Although the strong evidences that CGA derivatives are components of coffee melanoidins, it is not yet known how they are linked within the melanoidin structure. Based on studies demonstrating that the roasting-induced changes in coffee protein profiles are similar to those observed when oxidized chlorogenic acids react with green coffee proteins in model systems (Montavon et al., 2003; Rawel, Rohn & Kroll, 2005), proteins were considered as a possible binding site the CGA derivatives (Nunes & Coimbra, 2010). Also, Ara from arabinogalactan side chains was hypothesized as a

possible binding site for the CGA derivatives (Bekedam et al., 2008b). This hypothesis was proposed based on previous studies demonstrating that the Ara is quite susceptible to the degradation induced during coffee roasting (Oosterveld et al., 2003b; Redgwell et al., 2002b), and the reaction of a polyphenolic compound (epicatechin) with sugar fragments under simulated roasting conditions (Totlani & Peterson, 2007).

I.2.3.4. Contribution of sucrose

Sucrose is the most abundant low molecular weight (LMW) carbohydrate found in green coffee beans, accounting for up to 9% of the beans' dry weight (Illy et al., 1995; Wei & Tanokura, 2015). Using a temperature of 220 °C and roasting time up to 9 min, sucrose is destroyed quickly by roasting for 2-4 min (Wei et al., 2012). The degradation is complete after roasting for 5 min. It is known that sucrose acts as an aroma precursor during coffee roasting, generating several classes of compounds, such as carboxylic acids, furans, and aldehydes (Wei & Tanokura, 2015). Furthermore, the use of modified "in bean" green coffee models, produced by the incorporation of sucrose into the green beans, allowed to obtain evidences for the contribution of sucrose to coffee melanoidin formation, which is probably related to the formation of coloured LMW structures attached to the polysaccharides and proteins, or to protein fragments (Nunes et al., 2012).

In conclusion, more work is needed to understand the roasting-induced modifications in the structure of the different coffee components, especially of polysaccharides, and their reactivity with proteins and chlorogenic acids, which will allow to have new insights into the structural details of coffee melanoidins. Mass spectrometry (MS)-based approaches are required to fulfil these goals.

I.3. Mass spectrometry analysis of polysaccharides and their derivatives

Mass spectrometry (MS) has become an indispensable technique for the structural analysis of carbohydrates (Bauer, 2012; Mischnick, 2012; Rodrigues et al., 2007). This is due to important instrumental developments, especially the introduction of two soft ionization techniques: electrospray ionization (ESI) (Yamashita & Fenn, 1984) and matrix-assisted laser desorption/ionization (MALDI) (Karas & Hillenkamp, 1988; Tanaka et al., 1988). In particular, structural features of green coffee galactomannans and

arabinogalactans (Section I.1.1), as well as modifications at the reducing end of roasted coffee galactomannans (Section I.2.2) were disclosed by ESI-MS and tandem MS (ESI-MS/MS) analyses.

I.3.1. Basics of mass spectrometry

The mass spectrometer is an instrument designed to separate gas phase ions (positively or negatively charged) according to their mass-to-charge ratios (*m/z*). Therefore, the basic components of a mass spectrometer, as illustrated in Figure I.7, include: an ionization source to produce gas phase ions from the sample; one or several mass analysers to separate the various ions according to their *m/z* values; and a detector to measure and amplify the ion current of the mass-resolved ions. The sample can be introduced into the ionization source through a direct inlet system or a direct coupling of a separation technique, such as gas chromatography (GC) or high-performance liquid chromatography (HPLC). As shown in Figure I.7, a data system records the signal registered by the detector and converts it into a mass spectrum. This can be presented in graphic or table format. In either format, the intensity of the ions is usually represented in percent relative to the most intense ion, called the base peak. Accordingly, in graphic format, the abscissa represents the *m/z* values of all ions, while the ordinate represents their relative abundances (Dass, 2007; Ekman et al., 2009).

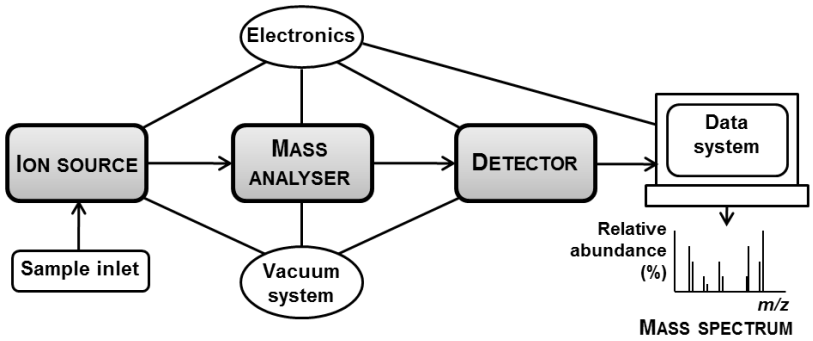

Figure I.7. Basic components of a mass spectrometer. Adapted from Dass (2007).

Mass spectrometers are currently available in numerous configurations with various combinations of ionization techniques and mass analysers. The most used ionization techniques for the analysis of carbohydrates are electrospray ionization (ESI), matrix-assisted laser desorption/ionization (MALDI), and electron ionization (EI, formerly

called electron impact) (Pettolino, Walsh, Fincher & Bacic, 2012; Sassaki & Souza, 2013). EI is the technique of choice for mass spectrometers coupled to gas chromatographs. Also, there are several types of mass analysers used in the mass spectrometers, such as quadrupole (Q), time-of-flight (TOF), linear ion trap (LIT), and Orbitrap. They can be used alone in a mass spectrometer, or can be combined. When different types of analysers are combined in a mass spectrometer, it is named as a hybrid instrument. A common type of hybrid instrument, also used in this work, is the Q-TOF (Dass, 2007; Ekman et al., 2009).

I.3.2. Ionization of oligosaccharides

Electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) are soft ionization techniques. In this context, the term soft indicates that very little excess energy is imparted to the analytes during the ionization process, producing almost no fragmentation. Consequently, the ions observed in the ESI-MS and MALDI-MS spectra correspond to intact analytes. For this reason, ESI-MS and MALDI-MS are suitable for analysis of mixtures. Nevertheless, the coupling of a separation technique, especially high-performance liquid chromatography (HPLC), with MS, either online or offline, can be advantageous in the analysis of complex mixtures, especially when it provides the separation of isomers, i.e. compounds having the same elemental composition, and thus not differentiable by direct MS analysis.

Although both ESI and MALDI represent soft ionization techniques, there are differences between these techniques. One of the main differences between ESI and MALDI is the state in which the sample is introduced into the ion source. In ESI (Figure I.8A), the sample solution in a suitable solvent mixture is continuously pumped into the source through a capillary. A high voltage is applied to the capillary tip, resulting in the formation of a spray of highly charged droplets. The size of these droplets is continuously diminished by solvent evaporation, assisted by a flow of hot gas (usually nitrogen). When the charge density at the droplet surface reaches a critical value (the Rayleigh limit), a so called Coulomb explosion occurs and several even smaller droplets are formed. The exact mechanism of ion formation from these droplets is still not fully elucidated and there are different theories proposed. In terms of sample preparation, ESI-MS analysis requires only dissolution of the sample to an appropriate concentration in a suitable solvent mixture. In MALDI-MS, ions are generated from the solid phase. A mixture of the sample

and an appropriate matrix, both dissolved in a suitable solvent, is first spotted on a MALDI plate. In carbohydrate analysis, the most frequently used matrix is 2,5-dihydroxybenzoic acid (DHB). After solvents evaporation, the sample/matrix mixture is irradiated with a laser beam, often a nitrogen laser emitting in the ultraviolet range at 337 nm. The matrix absorbs and transfers some of the energy to the analyte which ionizes (Figure I.8B). The mechanisms by which ions are formed in MALDI are not yet clear.

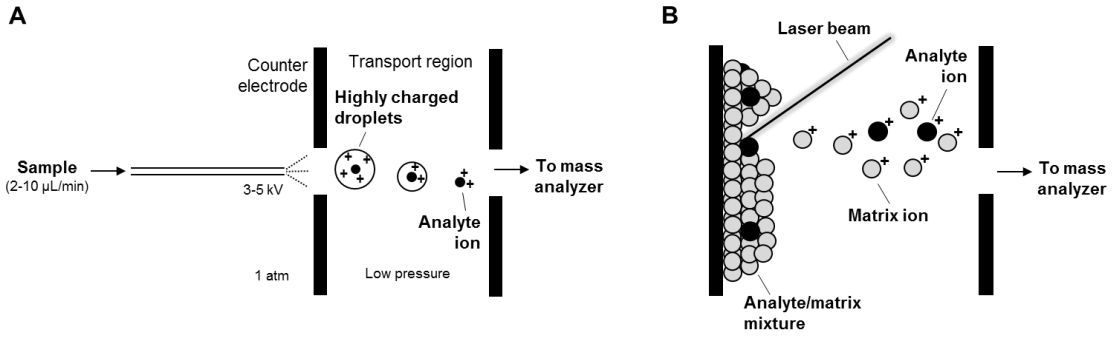

Figure I.8. Schematic representation of the (A) electrospray ionization (ESI) and (B) matrix-assisted laser desorption/ionization (MALDI) process.

Either with ESI or MALDI, entire polysaccharides are not suitable to be analysed. Unlike proteins, which have a defined molecular mass, the polysaccharides are polydisperse. Thus, in order to obtain lower molecular weight oligosaccharides, polysaccharides are usually subjected to partial acid or enzymatic hydrolysis prior to MS analysis. Compared the acid hydrolysis, the enzymatic hydrolysis has the advantage of being selective, since the enzymes cleave specific glycosidic linkages (Sassaki & Souza, 2013). Typically, the hydrolysate obtained, either by acid or enzymatic hydrolysis, is further fractionated by size exclusion chromatography (SEC) and the selected fractions are then analysed by ESI-MS and/or MALDI-MS (Nunes et al., 2006; Nunes et al., 2008; Simões et al., 2012). The ESI-MS is suitable for the analysis of oligosaccharides with molecular weight lower than 2,000 Da. Higher molecular weight oligosaccharides can be analysed by MALDI-MS, but it has the disadvantage of not allowing the analysis of low molecular weight oligosaccharides due the presence of background matrix ions. For this reason, in this work, the MALDI-MS spectra were acquired from *m/z* 450, whereas the ESI-MS spectra were acquired from *m/z* 50.

Under ESI-MS and MALDI-MS conditions, ions can be generated either in positive or negative ion mode, depending on the nature of the sample. Neutral oligosaccharides

ionize preferentially in positive ion mode as sodium adduct ions ($[M+Na]^+$). Potassium adduct ions ($[M+K]^+$) are also observed in the positive MS spectra of neutral oligosaccharides, but having a lower relative abundance than the corresponding $[M+Na]^+$ ions. As an exception, $[M+K]^+$ ions with a higher relative abundance than the corresponding $[M+Na]^+$ ions were identified in the MALDI-MS spectra of fractions recovered from spent coffee grounds by microwave assisted extraction, but it was consequence of the high abundance of potassium ions in coffee beans (Passos et al., 2014b). Protonated molecules ($[M+H]^+$) are usually absent, or have a very low relative abundance in the MS spectra of neutral oligosaccharides (Moreira et al., 2011; Reis et al., 2004a). Contrarily, oligosaccharides (or their derivatives) containing nitrogen, as the case of the oligosaccharide Amadori compounds obtained from roasted coffee galactomannans (Nunes et al., 2006), produce abundant $[M+H]^+$ ions by the protonation of the amino group. The negative ion mode is used in the analysis of oligosaccharides containing acidic groups, as in the case of oligosaccharides containing sialic acids. In this case, abundant $[M-H]^-$ ions are produced by the deprotonation of the carboxyl group (Casal et al., 2010; Zaia, 2004).

As ESI and MALDI allow the ionization without fragmentation, the molecular mass of an analyte can readily be inferred with the known of the type of ions formed during the ionization process. For example, neutral oligosaccharides, as previously mentioned, ionize preferentially as $[M+Na]^+$. In this case, the molecular mass is obtained by subtracting the mass of a sodium (23 Da) from the *m/z* value of the $[M+Na]^+$ ion of the oligosaccharide in study. The information of the molecular mass, when combined with the monosaccharide composition, usually obtained by GC-based methods (Section II.5), allows to propose a number of possible structures for the oligosaccharide in study. Of note, isomeric hexoses with the molecular formula $C_6H_{12}O_6$, like Gal, Glc, and Man, and isomeric pentoses with the molecular formula $C_5H_{10}O_5$, like xylose (Xyl) and Ara, are not differentiable by direct MS analysis. Thus, the interpretation of the MS data is usually made taking into account the monosaccharide composition obtained by GC-based methods.

In the analysis of oligosaccharides bearing modified sugar residues, as those obtained from roasted coffee galactomannans (Nunes et al., 2006), the identification of possible structures can be more complicated. According to the nitrogen rule, if a compound has an odd number of nitrogen atoms, its molecular mass will be an odd number; and if it contains no nitrogen atoms or an even number of nitrogen atoms, its

molecular mass will be an even number (Dass, 2007). The consequence of this rule is that the *m/z* value of oligosaccharide ions, including $[M+Na]^+$, $[M+K]^+$, and $[M+H]^+$, or in negative ion mode as $[M-H]^-$, with an odd number of nitrogen atoms is even, whereas the *m/z* value of oligosaccharide ions with no nitrogen atoms or an even number of nitrogen atoms is odd. In the analysis of unknown/modified compounds, as those formed from coffee polysaccharides during roasting, it can provide some idea of the presence or absence of nitrogen atoms in a compound.

Finally, it is important have in mind that certain mass spectrometers, such as Q-TOF and Orbitrap-based mass spectrometers used in this work, can provide much more accurate spectra than mass spectrometers equipped with a linear ion trap (LIT), also used in this work. Accurate mass measurements made the assignment of elemental compositions possible. However, the confirmation of the proposed assignments for the ions identified by MS, as well as more information about their structures, namely the monosaccharide sequence, branching, and modified residues, is obtainable by tandem mass spectrometry (MS/MS). The MS/MS experiments can also help in identifying which are the ions of interest and the ions due to impurities.

I.3.3. Fragmentation of oligosaccharides under tandem mass spectrometry

Product ion scanning is the most common MS/MS experiment. It is based on the isolation of a selected ion (*precursor ion*, formerly called *parent ion*), that is then induced to dissociate, and the resulting ions (*product ions*, formerly called *daughter ions*) are further analysed. The terms *parent ions* and *daughter ions* are discouraged by International Union of Pure and Applied Chemistry (IUPAC) (Murray et al., 2013). Specific mass spectrometers, as those equipped with a linear ion trap (LIT), allow performing a sequential selection and fragmentation of ions, which is called multistage tandem mass spectrometry (MS^n).

The dissociation of the precursor ion can be promoted using different techniques, but collision-induced dissociation (CID) is by far the most commonly used. In this technique, also known as collision-activated dissociation (CAD), the precursor ion is fragmented by collisions with atoms of an inert gas, such as helium, nitrogen, or argon (Dass, 2007). The CID fragmentation of oligosaccharide ions generates product ions from the cleavage of glycosidic linkages, cross-ring cleavages (cleavage of two bonds within the sugar ring), and loss of water. The glycosidic cleavages, resulting from the cleavage

of a bond between two adjacent sugar residues, provide information about the monosaccharide composition, sequence and branching. The cross-ring cleavages provide additional information that is important to reveal the type of linkage between monosaccharides. The cross-ring cleavages are usually referred as neutral losses of $C_2H_4O_2$ (-60 Da), $C_3H_6O_3$ (-90 Da), and $C_4H_8O_4$ (-120 Da) (Simões et al., 2007). However, the neutral loss of CH_2O (-30 Da) has also been observed (da Costa et al., 2012). Furthermore, studies performed on the fragmentation of ^{18}O-labelled oligosaccharides have been shown that there is a strong preference for glycosidic bond cleavage on the non-reducing side of the glycosidic oxygen with elimination of the residues located at the non-reducing end, and for cross-ring cleavages on the reducing end residue (Asam & Glish, 1997; da Costa et al., 2012).

The glycosidic and cross-ring ions resulting from oligosaccharide fragmentation are usually named according to the nomenclature introduced by Domon & Costello (1988). As shown in Figure I.9, product ions that contain the non-reducing end sugar unit are labelled with uppercase letters from the beginning of the alphabet (A, B, C), and those that contain the reducing end sugar unit (or the aglycone moiety in the case of glycoconjugates) are labelled with letters from the end of the alphabet (X, Y, Z). The A and X ions result from cross-ring cleavages, and are labelled with two superscript numbers showing the sugar ring bonds cleaved. The ring bonds are conventionally denoted in the clockwise direction from 0 to 5 in pyranose rings, or 0 to 4 in furanose rings. The other ions (B, C, Y, Z) result from glycosidic cleavages. All product ions are labelled with a subscript number identifying the number of sugar residues retained in the product ion. In the case of branched oligosaccharides, products ions from side chains are supplemented by a Greek letter, which varies according to the molecular weight of the side chain ($\alpha \geq \beta \geq \gamma$).

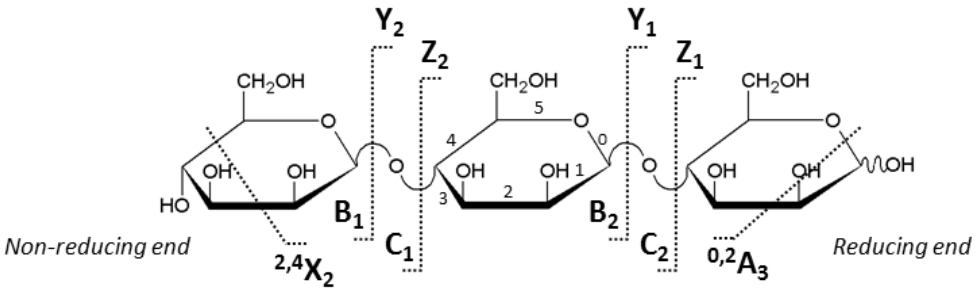

Figure I.9. Nomenclature for the assignment of glycosidic and cross-ring product ions according to Domon & Costello (1988).

I.4. Aim of the work

Although it is known that coffee polysaccharides are modified during the roasting process, namely by reaction with proteins, chlorogenic acids, and sucrose, leading to the formation of melanoidins, the structures of the roasting-induced compounds derived from coffee polysaccharides are far to be completely elucidated. On the basis of the current knowledge, it can be stated that the complete identification of the roasting-induced modifications in coffee polysaccharides is hampered by their diversity, and the high structural complexity of the compounds formed. The use of model systems, comprising few populations of molecules able to provide information concerning the specific reactions that can take place inside the coffee beans, and that can change the structure of their polysaccharides, when using mass spectrometry, allow to obtain an overall view and, simultaneously, detailed, of the structural modifications in coffee polysaccharides promoted by roasting. Based on this thesis, selected compounds were submitted to dry thermal and oxidative treatments, mimicking the coffee roasting process.

As stated in the literature review, mass spectrometry (MS), either with electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI), combined with tandem mass spectrometry (MS/MS or MS^n) has been successfully used to identify structural features of green coffee polysaccharides (Section I.1), and modifications induced by roasting (Section I.2). Thus, in order to obtain an overall view of the structural modifications in coffee polysaccharides promoted by roasting, mostly MS-based approaches, complemented with other analytical techniques, were employed to identify the structural modifications induced by dry thermal and oxidative treatments in model compounds. Also, roasted coffee polysaccharides were analysed to screen for the structural modifications identified from the model compounds, using the same analytical strategies. All the experimental details are presented in **Chapter II** of this thesis. The results obtained and their discussion are presented in **Chapter III**, grouped into three subchapters, according to the specific aims of the different studies performed, as follow:

III.1. Understanding roasting-induced modifications of coffee arabinogalactan side chains

III.2. Understanding non-enzymatic transglycosylation reactions

III.3. Understanding coffee melanoidin formation

In **Chapter VI** of this thesis are presented the general concluding remarks. The references cited in the text are presented in **Chapter V**, and the supplementary material in **Appendices A-E**.

CHAPTER II. MATERIAL AND METHODS

II.1. Samples

II.1.1. Models of green coffee bean components

II.1.1.1. Oligosaccharides

(α1→5)-L-arabinotriose (Ara_3, syrup) and (β1→4)-D-mannotriose (Man_3, powder), having a purity ≥ 95%, were purchased from Megazyme (County Wicklow, Ireland). These commercial oligosaccharides were used as models of coffee polysaccharides. Ara_3 is an oligosaccharide structurally related to the side chains of arabinogalactans, and it was used in the studies presented in Sections III.1.1-2, III.2.1-2 and III.3.1. Man_3 is an oligosaccharide structurally related to the backbone of galactomannans, and it was used in the studies presented in Section III.2.1-2 and III.3.2. The syrup sample commercially provided was diluted with ultrapure water (Millipore Synergy system, Billerica, MA) and freeze-dried. The resulting material was ground with an agate mortar and pestle and then stored at room temperature in a desiccator under a phosphorous pentoxide environment until to be used.

II.1.1.2. 5-*O*-Caffeoylquinic acid

Because 5-*O*-caffeoylquinic acid (5-CQA) is the most abundant chlorogenic acid present in green coffee beans, a commercial standard of 5-CQA was used in the studies presented in Sections III.3.1 and III.3.2. It had a purity ≥ 95%, and was obtained from Sigma (St. Louis, MO, USA).

II.1.1.3. Peptides

Dipeptides, tyrosine-leucine (YL) and D-leucine-tyrosine (LY), were purchased from Sigma (St. Louis, MO, USA). Attending that tyrosine and leucine are amino acids present in green coffee beans, YL and LY were used as model of green coffee proteins (Section III.3.3).

II.1.1.4. Aliphatic acids

DL-malic acid (MalA, purity > 99%) was purchased from Fluka Chemie AG (Buchs, Switzerland), and citric acid (CitA, 100%) from Fisher Scientific (Fair Lawn, NJ, USA).

These commercial standards were used as models because MalA and CitA are the most abundant aliphatic acids present in green coffee beans (Section III.3.2).

II.1.2. Preparation of the model mixtures

Solutions of Ara_3 (53.3 mg, 0.129 mmol) and Man_3 (64.9 mg, 0.129 mmol) in 200 mL of ultrapure water (Millipore Synergy system, Billerica, MA) were prepared. Using these solutions, three mixtures were prepared containing different molar proportions of Ara_3 and Man_3 as follows: 25% of Ara_3 and 75% of Man_3, 50% of Ara_3 and 50% of Man_3, and 75% of Ara_3 and 25% of Man_3. These mixtures are designated as A25M75, A50M50, and A75M25, respectively. Also, solutions containing equimolar amounts of Ara_3-CQA, Man_3-YL, Man_3-LY, Man_3-CQA, Man_3-MalA, Man_3-CitA, YL-CQA, and Man_3-YL-CQA were prepared using ultrapure water. After freeze-drying, the solid mixtures were powdered with an agate mortar and pestle, and then stored at room temperature in a desiccator until use.

II.1.3. Spent coffee grounds and instant coffee samples

In order to screen for the structural modifications identified by using model samples submitted to dry thermal treatment or oxidative conditions (Fenton reaction) in real coffee samples, polysaccharide-rich fractions recovered from spent coffee grounds (SCG) and instant coffee samples were analysed.

II.1.3.1. SCG polysaccharides recovered by alkali extraction and neutralization, solubilization by roasting treatments, and enzymatic hydrolysis

SCG, obtained after a commercial espresso coffee preparation, were submitted to a roasting pre-treatment at 160 or 220 °C, and then sequentially extracted with water at 90 °C and 4 M NaOH solutions at 20, 60 and 120 °C. A galactomannan-rich fraction, containing 89% of mannose, was recovered from SCG roasted at 160 °C upon extraction with 4 M NaOH at 120 °C, and became water insoluble upon neutralization (Simões, Nunes, Domingues & Coimbra, 2013).

To convert the insoluble material recovered from SCG into cold water soluble material, it was submitted to a roasting treatment at 160 ºC in a pre-heated oven (Binder) with an internal volume of 115 L during 1 h, as described by Simões et al. (2014). After the roasting procedure, the material was suspended in 100 mL of water at room temperature (20 ºC) with stirring during 1 h. The suspension was then centrifuged and the supernatant was separated from the insoluble residue and freeze-dried. These processes of roasting at 160 ºC plus solubilization in water were repetitively done in a total of 7 times. A sample of the solubilized coffee polysaccharides (15 mg) was hydrolysed with 1 U of a α-galactosidase from Guar seed (Megazyme, EC 3.2.1.22) during 24 h at 40 ºC with continuous stirring in a 100 mM sodium acetate buffer, pH 4.5, containing 0.02% sodium azide. The coffee polysaccharide sample (15 mg) was also hydrolysed with 1 U of a β-galactosidase from *Escherichia coli* (Sigma, EC 3.2.1.23) during 24 h at 37 ºC with continuous stirring in a 100 mM sodium acetate buffer, pH 7.3, containing 0.02% sodium azide. The hydrolysed materials were dialysed for 3 days using a membrane with a molecular weight (MW) cut-off of 12-14 kDa, with several changes of distilled water. The high molecular weight materials were collected, frozen, and freeze-dried for further linkage analysis, as described in Section II.5.2. The results obtained by linkage analysis are presented and discussed in Section III.2.1.

To obtain cold water soluble material, the insoluble material recovered from SCG was also submitted to sequential roasting treatments of 1 h at 200 ºC. After each roasting procedure, the material was suspended in water at room temperature with stirring during 1 h. The suspension was then centrifuged, and the cold water-soluble material was recovered and freeze-dried (Simões et al., 2014). A sample (2.4 mg) of the cold water-soluble material recovered after the second roasting procedure, referred to as R2W20sn (Simões et al., 2014), was hydrolysed with 0.3 U of a pure *endo*-(β1→4)-D-mannanase preparation (*Aspergillus niger*, EC 3.2.1.78) (Megazyme, County Wicklow, Ireland) during 60 h at 37 ºC with continuous stirring in 600 μL of 100 mM sodium acetate buffer, pH 5.5, containing 0.02% sodium azide. The hydrolysed sample was filtrated using a Cronus nylon syringe filter (0.2 μm pore size and 13 mm diameter) and then fractioned by ligand exchange/size-exclusion chromatography (LEX/SEC), as described in Section II.6.1. To check its LEX-SEC elution profile, the enzyme (0.3 U in 600 μL of 100 mM sodium acetate buffer) was injected and eluted using the same conditions used for the hydrolysed SCG sample. The results obtained are presented and discussed in Section III.2.2

II.1.3.2. Instant coffee sample, enzymatic hydrolysis, and fractionation by ethanol precipitation

The instant coffee powder was obtained in a local supermarket from a commercial batch (Néscafe, Nestlé). A sample (400 g) was dissolved in 1 L of distilled water at 80 ºC, and stirred during 1 h. The solution was cooled down at room temperature, and then was placed at 4 °C for a period of 48 h, prior to the decantation. The cold water insoluble material recovered was discarded. The supernatant was then centrifuged at 15,000 rpm for 15 min at 4 °C and the precipitate and supernatant recovered were frozen and freeze-dried.

The precipitate (5 g) was partially hydrolysed with 11.25 U of a pure *endo*-(β1→4)-D-mannanase preparation (Megazyme, *A. niger*, EC 3.2.1.78) during 1 h at 37 °C with continuous stirring in 25 mL of 10 mM phosphate buffer, pH 5.5. The precipitate and supernatant recovered after hydrolysis were named respectively as Mannanase_Ppt and Mannanase_Sn. The Mannanase_Sn (40 mg in 1 mL of distilled water) was loaded on a XK26-100 column (Pharmacia) containing Bio-Gel P-30 (Bio-Rad, 2.5-40 kDa fractionation range) using 0.1 M NaCl as eluent. The collected fractions were assayed for sugars according to the phenol-sulphuric acid method (Dubois et al., 1956). The Mannanase_Sn (150 mg in 15 mL of distilled water) was also fractionated by graded ethanol precipitation. First, 15 mL of absolute ethanol were added to reach an aqueous solution containing 50% ethanol (v/v). The solution was then centrifuged at 15,000 rpm for 10 min at 4 ºC. The precipitate recovered was named as Et50. To the supernatant was added absolute ethanol to reach an aqueous solution containing 75% ethanol (v/v). After centrifugation, the precipitate (Et75) and supernatant (Sn) were recovered. As presented in Section III.2.1, the sugar composition of the initial instant coffee sample, as well as of all the precipitates and supernatants was further determined according to the procedure of neutral sugar analysis described in Section II.5.1.

II.1.3.3. SCG fractions recovered by microwave assisted extraction

SCG samples, previously dried at 105 ºC for 8 h, were suspended in water in 1 g portion to 10 mL of water in a total volume of 70 mL in each one of 6 individual containers. Microwave irradiation was performed with a MicroSYNTH Labstation (Milestone Inc.), using operating conditions similar to the ones previously described (Passos & Coimbra, 2013; Passos et al., 2014b). The fractions recovered after a third

microwave assisted extraction (MAE3) at 170 °C for 5 min and a fourth microwave assisted extraction (MAE4) at 200 °C for 2 min were centrifuged at 15,000 rpm for 20 min at 4 °C. The supernatant solution was filtered using MN GF-3 glass fibre filter, frozen, freeze-dried, and stored under an anhydrous atmosphere.

Each fraction was dissolved in the minimum amount of water, stirring during 10 min at room temperature, and then absolute ethanol was added to reach an aqueous solution containing 75% ethanol (v/v). The solution was then centrifuged at 15,000 rpm for 10 min at 4 °C and the precipitated material recovered (PptEt). To check the presence of chlorogenic acid-arabinose structures, in the study presented in Section III.3.1, samples (2 mg) of MAE3_PptEt and MAE4_PptEt were dissolved in 1 mL of ultrapure water and then kept frozen at -20 °C. Before MS analysis, they were filtered using 0.2 μm syringe filters. Negative ion ESI-MS and tandem MS (ESI-MS^n) spectra were acquired on a LIT mass spectrometer (Section II.7.1).

II.2. Thermal treatments

As in a previous study with mannosyl and galactomannosyl oligosaccharides (Moreira et al., 2011), thermal treatments were performed on a TGA-50 thermogravimetric analyser (Shimadzu, Kyoto, Japan), operating with a controlled air flow of 20 mL/min and a heating rate of 10 ºC/min.

The thermal stability of the Ara_3 (Section III.1.1), as well as of the 5-CQA and the mixture Ara_3-CQA (Section III.3.1) was studied by submitting a sample of 2-5 mg to a temperature program from ambient temperature to 600 °C.

To study the roasting-induced compounds, in a different set of experiments, three samples (5-10 mg) of Ara_3, as well as of each oligosaccharide mixture (A25M75, A50M50, and A75M25) were heated from room temperature up to 200 °C and maintained at 200 °C for different periods: 0, 30, and 60 min. These three different thermal treatments are respectively designated as 200T1, 200T2, and 200T3, whereas the respective untreated sample is designated as T0. Similarly, a sample of the 5-CQA (8 mg) and the mixture Ara_3-CQA (7 mg) was also submitted to the thermal treatment 200T1 (Section III.3.1). Aiming to obtain a lower degradation extent and, if possible, to identify intermediate degradation products, two additional samples (5-8 mg) of the 5-CQA and the mixture Ara_3-CQA were heated to 175 °C and maintained at this temperature for two

different periods: 0 (175T1) and 30 min (175T2) (Section III.3.1). In the study presented in Section III.3.2, a sample (3-11 mg) of the Man_3 and each mixture Man_3-YL, Man_3-LY, Man_3-CQA, Man_3-MalA, Man_3-CitA, YL-CQA, and Man_3-YL-CQA was submitted to the thermal treatment 175T1.

The samples submitted to the thermal treatments at 175 °C and 200 °C were recuperated, weighed, and suspended in ultrapure water (5 mg sample/mL solvent). To facilitate their dissolution, they were stirred at 37 ºC for 3 h. The samples were then centrifuged to separate the water-soluble and water-insoluble fractions, designated as WSF and WIF, respectively. Both WSF and WIF were kept frozen at -20 °C until analysis.

A solution of each untreated sample (1 mg/mL in ultrapure water) was also prepared. As used for the thermally treated samples, each solution was stirred at 37 ºC for 3 h, and then kept frozen at -20 °C until analysis.

II.2.1. Fractionation of water-insoluble fractions

A portion (approx. 2 mg) of each WIF resulting from T2 and T3 treatments of Ara_3 (Section III.1.1) was completely dissolved in 0.5 mL of anhydrous dimethyl sulfoxide (DMSO). The most hydrophobic material was precipitated by addition of distilled water (5 mL). The supernatant was recovered after centrifugation and the precipitate was washed two times with 2 mL of distilled water. To remove the DMSO, the combined supernatants were dialysed using a membrane with a MW cut-off of 1 kDa during 8 h against 300 mL of distilled water, with two water renewals. The retentate obtained was divided in two parts, one used for sugar analysis and the other for linkage analysis.

II.3. Oxidation by Fenton reaction

A stock solution of Ara_3 (4 mg/mL in ultrapure water; approx. 10 mM) was prepared and stored at -20 °C. For the Fenton system, aqueous solutions of 0.5 M H_2O_2 and 5 mM $FeCl_2 \cdot 4H_2O$ were also freshly prepared. The reaction mixture was prepared by adding to 300 µL of the trisaccharide stock solution, 60 µL of the H_2O_2 solution, 12 µL of the $FeCl_2 \cdot 4H_2O$ solution, and ultrapure water to reach a total volume of 600 µL. The mixture was vortex mixed, sonicated for 10 min, vortex mixed again, and left to react at 50 °C in the dark for 6h under stirring. After that, to destroy the remaining H_2O_2, catalase was added to the mixture. This was then vortex mixed and kept frozen at -20 °C until

further analysis. A control mixture was prepared by replacing H_2O_2 by an equal volume of ultrapure water, submitted to the same conditions. $FeCl_2 \cdot 4H_2O$ (99.99%), H_2O_2 (30 wt. % in H_2O), and catalase from bovine liver (lyophilized powder, ≥10,000 units/mg protein) were purchased from Sigma-Aldrich (St. Louis, MO).

The extent of the oxidation was monitored by ESI-MS analysis. The oxidation products were subsequently separated by LEX-SEC (Section II.6.1), identified by ESI-MS, and characterized by ESI-MS/MS. Both ESI-MS and ESI-MS/MS experiments were carried out on a Q-TOF mass spectrometer, using the operating conditions described in Section II.7.1.

II.4. Labelling with oxygen-18

To label with oxygen-18 (^{18}O) the anomeric oxygen of non-modified reducing sugar residues, or new carbonyl groups formed by dry thermal processing, 50 µL of the Ara_3 solution (1 mg/mL) (Section III.1.1), as well as of untreated and thermally treated oligosaccharide mixtures (A25M75, A50M50 and A75M25) (Section III.2.2), were dried and redissolved in 50 µL of ^{18}O-enriched water ($H_2{}^{18}O$, 97 %, Sigma-Aldrich, St. Louis, MO). Each solution was then kept under stirring at 500 rpm for 4 h at 37 °C in a sealed vial, after which time it was frozen at -20 °C until analysis by ESI-MS (4 days). To check the ^{18}O-labelling of a keto group (Section III.2.2), 2 µL of 3-octanone (99%, Aldrich-Chemie, Steinheim, Germany) were diluted in 50 µL of $H_2{}^{18}O$, and then submitted to the same conditions of stirring and freezing before ESI-MS analysis on a LIT mass spectrometer (Section II.7.1).

II.5. Gas chromatography-based methods

In this work, the monosaccharide composition of oligo- and polysaccharide samples and the glycosidic linkage positions between the monosaccharide residues were determined by gas chromatography (GC)-based methods, named as sugar and linkage analyses, respectively. The latter is also commonly referred as methylation analysis.

The basic operating principle of GC involves volatilization of the sample in a heated inlet or injector of a gas chromatograph, followed by separation of the components of the sample within a heated hollow tube (the column) that contains an appropriate stationary phase. The separation is determined by the distribution of each component between the

stationary phase and a carrier gas (the mobile phase). After elution from the column, each component still in the carrier gas flows into a detector, or in the case of a GC-MS system, a mass spectrometer (Sparkman, Penton & Kitson, 2011).

Due to the low volatility of carbohydrates, they have to be converted into volatile derivatives, before GC and GC-MS analyses. Thus, even with derivatization, GC and GC-MS analyses are limited to carbohydrates of low molecular weight, mainly mono-, di- and trisaccharides (Ruiz-Matute et al., 2011).

For neutral sugar analysis, monosaccharides are released by acid hydrolysis, and converted to alditol acetates by reduction and acetylation, as illustrated in Figure II.1. The reduction step, yielding alditol derivatives, is used to simplify the GC chromatograms, avoiding the mutarotation and, consequently, the α, β, pyranose and furanose isomers to each monosaccharide. The alditol derivatives are then converted into volatile derivatives by acetylation of the free hydroxyl groups. The alditol derivatives are usually separated by GC and identified by comparing the retention times with those of standards (Pettolino et al., 2012; Sassaki & Souza, 2013). For quantification purpose, in our laboratory, 2-deoxy-*arabino*-hexose (commonly named as 2-deoxyglucose) is used as internal standard.

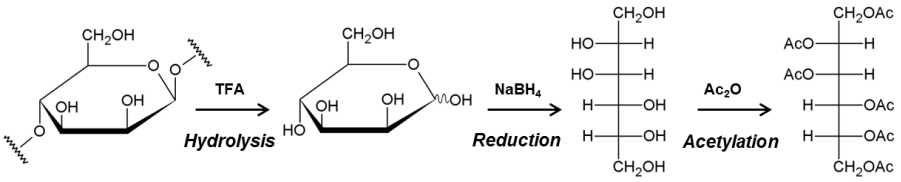

Figure II.1. Schematic representation of the derivatization procedure for sugar analysis on a polysaccharide composed by (β1→4)-linked D-mannose residues.

For linkage analysis, firstly, carbohydrates are per-*O*-methylated in dimethylsulfoxide (DMSO), using a fine suspension of sodium hydroxide (NaOH) and methyl iodide (CH_3I), according to the method introduced by Ciucanu & Kerek (1984). The per-*O*-methylated carbohydrates are then converted to partially *O*-methylated alditol acetates (PMAAs) by successive hydrolysis, reduction, and acetylation (Figure II.2). The position of the *O*-acetyl and *O*-methyl groups on the PMAA derivatives reflects the linkage positions, but also the ring forms of the monosaccharides. The PMAA derivatives are separated by GC-MS and identified according to their GC retention times and mass spectra. Due the similarity of mass spectra from different PMAA derivatives, it also

important to compare the GC retention times with those of standards (Pettolino et al., 2012; Sassaki & Souza, 2013).

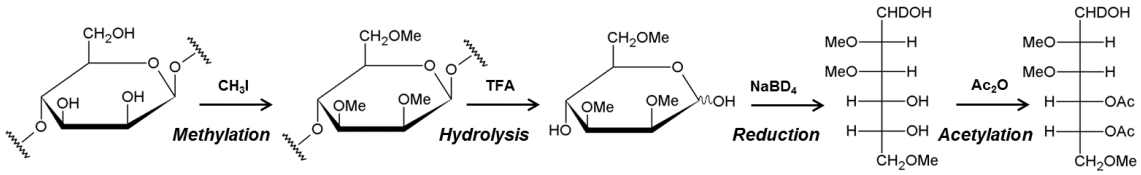

Figure II.2. Schematic representation of the derivatization procedure for linkage analysis on a polysaccharide composed by (β1→4)-linked D-mannose residues.

The procedures of neutral sugar and glycosidic linkage analyses used in this work, described briefly above, are presented in detail in the following sections.

II.5.1. Neutral sugar analysis

For neutral sugar analysis, each sample was hydrolysed with 1 mL of trifluoroacetic acid (TFA) 2 M during 1 h at 120 °C, and the acid was then evaporated to dryness. After addition of the internal standard solution (1 mg/mL of 2-deoxyglucose in distilled water), monosaccharides were reduced with 200 µL of $NaBH_4$ (15% in 3 M NH_3) during 1 h at 30 ºC. After cooling, the excess of borohydride was destroyed by the addition of glacial acetic acid (2 x 50 µL). The alditol derivatives were acetylated with acetic anhydride (3 mL) in the presence of 1-methylimidazole (450 µL) during 30 min at 30 ºC. To decompose the excess of acetic anhydride, distilled water (3 mL) was added. Dichloromethane (2.5 mL) was then added and, upon vigorous manual shaking and centrifugation for phase separation, the dichloromethane phase was recovered. The addition of water (3 mL) and dichloromethane (2.5 mL), and the recovery of the organic phase were performed once more. The dichloromethane phase containing the alditol acetate derivatives was washed two times by addition of distilled water (3 mL) and evaporated to dryness. The dried material was dissolved with anhydrous acetone (2 x 1 mL) followed by the evaporation of the acetone to dryness.

The alditol acetate derivatives, previously redissolved with anhydrous acetone, were analysed on a PerkinElmer Clarus 400 gas chromatograph (Waltham, MA) equipped with a flame ionization detector (FID) and a DB-225 column with 30 m of length, 0.25 mm of internal diameter, and 0.15 µm of film thickness (J&W Scientific, Folsom, CA).

The oven temperature program used was: 200 ºC for 1 min, a linear increase of 40 ºC/min up to 220 ºC, standing at this temperature for 7 min, followed by linear increase of 20 ºC/min up to 230 ºC, standing at this temperature for 1 min. The injector and detector temperatures were 220 and 230 °C, respectively. Hydrogen was used as carrier gas at a flow rate of 1.7 mL/min.

To check the presence of other sugars not identified by GC-FID, alditol acetate derivatives obtained from thermally treated mixtures of Ara_3 and Man_3 (Section III.2.1) were also analysed by gas chromatography-mass spectrometry (GC-MS) on an Agilent Technologies 6890N Network (Santa Clara, CA) equipped with a DB-1 column with 30 m of length, 0.25 mm of internal diameter, and 0.1 µm of film thickness (J&W Scientific, Folsom, CA). The GC was connected to an Agilent 5973 Network Mass Selective Detector operating with an electron impact mode at 70 eV and scanning the range *m/z* 40-500 in a 1 s cycle in a full scan mode acquisition. The oven temperature program used was: initial temperature 90 ºC, a linear increase of 10 ºC/min up to 127 ºC, followed by linear increase of 0.5 ºC/min up to 150 ºC, followed by linear increase of 40 ºC/min up to 250 ºC. The injector and detector temperatures were 220 and 230 °C, respectively. Helium was used as carrier gas at a flow rate of 1.7 mL/min.

II.5.2. Glycosidic linkage analysis

Each sample was dissolved with DMSO (1 mL), and then powdered NaOH (40 mg) was added to the solution. After 30 min at room temperature with continuous agitation, samples were methylated by adding of CH_3I (80 µL), allowed to react 20 min under vigorous stirring. Distilled water (2 mL) was then added, and the solution was neutralized using 1 M HCl. Dichloromethane (3 mL) was then added and, upon vigorous manual shaking and centrifugation, the dichloromethane phase was recovered. The dichloromethane phase was washed two times by addition of distilled water (2-3 mL). The organic phase was evaporated to dryness, and the resulting material was then remethylated using the same procedure.

The remethylated material was hydrolysed with 500 µL of 2 M TFA at 121 °C for 1 h, and the acid was then evaporated to dryness. For carbonyl-reduction, the resulting material was then suspended in 300 µL of 2 M NH_3 and 20 mg of $NaBD_4$ were added. The reaction mixture was incubated at 30 °C for 1 h. After cooling, the excess of

borodeuteride was destroyed by the addition of glacial acetic acid (2 x 50 µL). The acetylation was then performed as previously described for sugar analysis.

The PMAAs were redissolved with anhydrous acetone and identified by GC-MS, using the DB-1 column and the operating conditions (except the temperature program) previous described for sugar analysis (Section II.5.1). The oven temperature program used was: initial temperature 50 ºC, a linear increase of 8 ºC/min up to 140 ºC, standing at this temperature for 5 min, followed by linear increase of 0.5 ºC/min up to 150 ºC, followed by linear increase of 40 ºC/min up to 250 ºC, standing at this temperature for 1 min. Because they were not separated using the DB-1 column, the PMAAs derivatives obtained from D-xylose and D-lyxose standards (see Section III.2.1) were also analysed by GC-MS using the same temperature program previous described, but with a DB-FFAP column with 30 m of length, 0.32 mm of internal diameter, and 0.25 µm of film thickness (J&W Scientific, Folsom, CA).

II.6. High-performance liquid chromatography

Additional separation techniques are often used in conjunction with mass spectrometers. In particular, liquid chromatography (LC), either online or offline, can be performed prior to the analysis by mass spectrometry (MS). This strategy can be advantageous in the analysis of complex mixtures, especially when it provides the separation of isomeric compounds, which are not differentiable by direct MS analysis.

In LC the mobile phase is a liquid, as opposed to GC where a gas is utilized as a mobile phase. As in the GC, the separation is determined by the distribution of each component between the two phases. High-performance liquid chromatography (HPLC) is the term used to describe LC in which the liquid mobile phase is mechanically pumped through a column that contains the stationary phase. Therefore, the basic components of an HPLC instrument include: an injector, a pump, a column, and a detector (Sewell, 2000).

In this PhD work, selected samples were fractionated by offline ligand exchange/size-exclusion chromatography (LEX/SEC) before subsequent MS analysis.

II.6.1. Ligand exchange/size-exclusion chromatography

The oxidation products of Ara_3 formed under conditions of Fenton reaction (Section II.3) were separated by ligand exchange/size-exclusion chromatography (LEX/SEC) on a high-performance liquid chromatograph equipped with a Shodex sugar KS 2002 column (300 mm of length and 20 mm of internal diameter) from Showa Denko K. K. (Tokyo, Japan). The column was maintained at 30 °C, the injected sample volume was 500 µL, and ultrapure water was used as eluent at a flow rate of 2.80 mL/min. A refractive index (RI) detector (Knauer K-2401, Berlin, Germany) was used. All collected fractions were dried, redissolved in 100 µL of ultrapure water, and kept frozen at -20 °C until analysis by MS. Before MS analysis, the acidic fraction (F1) was incubated with cation exchange resin (AG 50W-X8, Bio-Rad Laboratories, Hercules, CA) for 20 min at room temperature. To obtain the retention time of Ara_3, the stock solution (4 mg/mL) was injected using the chromatographic conditions used for separation of the oxidation products.

As described in Section II.1.2.3, the galactomannan-rich fraction isolated from spent coffee grounds (SCG) and submitted to two additional roasting treatments at 200 °C was treated with an *endo*-(β1→4)-D-mannanase. This hydrolysed sample was fractionated by LEX-SEC, using the conditions previously described for LEX-SEC analysis of Ara_3 and their oxidation products, and all collected fractions were also dried, redissolved in 100 µL of ultrapure water, and kept frozen at -20 °C until analysis by MS. To check its LEX-SEC elution profile, the enzyme (0.3 U in 600 µL of 100 mM sodium acetate buffer) was also injected and eluted using the same conditions used for the hydrolysed SCG sample.

II.7. Soft ionization mass spectrometry

Mass spectrometry (MS) with electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI), combined with tandem MS (MS/MS or MS^n), in which the fragmentation of the precursor ion is usually achieved by collision-induced dissociation (CID), is a key technique for the structural analysis of carbohydrates, especially oligosaccharides and their derivatives (Section I.3). Also, it is of note that the liquid chromatography (LC), either online or offline, can be coupled to the soft ionization

MS analysis. In this work, different mass spectrometers equipped with a soft ionization source, ESI or MALDI, were used. As detailed below, the ESI-MS data were obtained either by direct infusion of the sample into the mass spectrometer or online coupling to LC.

II.7.1. Electrospray ionization mass spectrometry

Depending on the sample under study, as detailed below, one, or two, of four mass spectrometers equipped a ESI source, but different mass analysers were used for direct infusion ESI-MS analysis: LXQ linear ion trap (LIT) mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA); Q-TOF 2 hybrid quadrupole-time-of-flight mass spectrometer (Micromass, Manchester, UK); LTQ-Orbitrap XL hybrid ion trap-Orbitrap mass spectrometer (Thermo Fisher Scientific, Germany); and Q Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Germany). Note that the LIT spectrometer has as advantage the capability to perform multistage tandem mass spectrometry (MS^n) analysis. However, the Q-TOF spectrometer provided positive ion MS spectra with a better signal/noise ratio than those acquired using the LIT spectrometer. The Orbitrap-based mass spectrometers were used due to their analytical performance in terms of resolution and mass accuracy.

Immediately prior to direct infusion ESI-MS analysis, each sample, previously dissolved in water, was diluted in methanol/water (1:1, v/v) containing formic acid (0.1%, v/v).

II.7.1.1. Linear ion trap (LIT) mass spectrometry

Both untreated and thermally treated samples of the Ara_3 and the oligosaccharide mixtures (A25M75, A50M50, and A75M25), as well as the LEX-SEC fraction that eluted between 16-17 min from the enzymatically treated SCG galactomannan-rich sample were analysed by ESI-MS on the LIT mass spectrometer in the positive ion mode, using the following operating conditions: spray voltage was 5kV; capillary temperature was 275 °C; capillary voltage was 1V; and tube lens voltage was 40V. Samples were introduced at a flow rate of 8 µL/min into the ESI source. Nitrogen was used as nebulizing and drying gas. ESI-MS spectra were acquired over the *m/z* range 100-1500. ESI-CID-MS^n spectra were acquired with the energy collision set between 19 and 37 (arbitrary units). Data were acquired and analysed using Xcalibur software (version 2.0).

Both untreated and thermally treated samples of the mixture Ara_3-CQA were also analysed by ESI-MS analysis on the LIT mass spectrometer, but in negative ion mode. The operating conditions were as follow: spray voltage, 4.7 kV; capillary temperature, 275 ºC; capillary voltage, -22 V; tube lens voltage, -45 V. In the ESI-CID-MS^n experiments, the collision energy was set between 14 and 21 (arbitrary units).

II.7.1.2. Quadrupole-time-of-flight (Q-TOF) mass spectrometry

Positive ion ESI-MS and ESI-MS/MS spectra were acquired on the Q-TOF mass spectrometer from the control and reaction mixtures used in the study of the Ara_3 oxidation (Section II.3), as well as the respective fractions obtained by LEX/SEC (Section II.6.1). Samples were introduced into the ESI source on a flow rate of 10 µL min^{-1}. The cone voltage was set at 30 V and the capillary voltage was maintained at 3 kV. The source temperature was 80 °C and the desolvation temperature was 150 °C. ESI-MS spectra were acquired scanning the mass range from *m/z* 100 to 1500. MS/MS spectra were acquired by CID using argon as the collision gas. The collision energy used was set between 25 and 40 eV. All spectra were accumulated for 1 min. The data were processed using MassLynx software (version 4.0).

The identification of the new ions resulting from oxidation of Ara_3 was confirmed by exact mass measurement and elemental composition determination using ESI-MS spectra acquired for the LEX/SEC fractions, processed using MassLynx software. For fractions F2-F4, the lock mass in each spectrum was the calculated monoisotopic mass of the most abundant ion corresponding to the sodium adducts ($[M+Na]^+$) of a non-modified pentose oligosaccharide. For F1 and F5, the lock mass was the calculated monoisotopic mass of $[M+Na]^+$ ions of (β1→4)-D-mannotriose, an internal standard added before data acquisition.

Beyond the ESI-MS spectrum acquired on the LIT instrument, positive ion ESI-MS spectrum of the Ara_3 heated to 200 °C (T1) was also acquired on the Q-TOF mass spectrometer, using the operating conditions previously described for control and reaction mixtures used in the study of the Ara_3 oxidation. This spectrum was also processed using MassLynx software (version 4.0) to determine exact mass and elemental composition of the ions assigned as modified pentose oligosaccharides identified after the thermal treatment. For the exact mass measurements, the lock mass used for each ion was the

calculated monoisotopic mass of the corresponding non-modified pentose oligosaccharide in sodiated form ($[M+Na]^+$).

For untreated and thermally treated samples of the mixture Ara_3-CQA, beyond negative ion ESI-MS and ESI-MS^n spectra acquired on the LIT instrument, positive ion ESI-MS and ESI-MS/MS spectra were acquired on the Q-TOF mass spectrometer, also using the operating conditions previously described for control and reaction mixtures used in the study of the Ara_3 oxidation. In CID-ESI-MS/MS experiments, the energy collision was set between 30 and 43 eV.

II.7.1.3. Ion trap-Orbitrap mass spectrometry

For accurate mass measurements, negative ion ESI-MS spectra of the mixtures Ara_3-CQA, Man_3-CQA and Man_3-YL heated to 175 ºC (175T1) were also acquired using the LTQ-Orbitrap XL mass spectrometer, controlled by the LTQ Tune Plus 2.5.5 software and the Xcalibur 2.1.0 for data processing. The operating conditions were as follows: sheath gas flow, 5 (arbitrary units); spray voltage, 2.8 kV; capillary temperature, 275 ºC; capillary voltage, -35 V; and tube lens voltage, -200 V.

II.7.1.4. Quadrupole-Orbitrap mass spectrometry

The Q Exactive hybrid quadrupole-Orbitrap mass spectrometer, interfaced with H-ESI II ion source, was employed for accurate mass measurements of the oligosaccharide mixtures (A25M75, A50M50, and A75M25) submitted to the T3 treatment and the A50M50 mixture submitted to the T1 treatment, as well as the LEX-SEC fraction that eluted between 16-17 min from the enzymatically treated SCG galactomannan-rich sample. The acquisition method was set with a full scan and 140,000 resolution (relative to *m/z* 200) in positive mode. The method parameters were: AGC 3e6, IT 100 ms, scan range: 100-1000, spray voltage 3.0 kV, sheath gas 5, aux gas 1, capillary temperature 250 °C, S-lens RF level 50, probe heater temperature 50 °C and flow rate of 5 µL/min. The Q Exactive system was tuned and calibrated in positive mode using peaks of known mass from a calibration solution (Thermo Scientific) to achieve a mass accuracy of < 0.5 ppm RMS. The data were processed with Xcalibur 3.0.63 software.

II.7.2. Matrix assisted laser desorption/ionization mass spectrometry

MALDI-MS spectra were acquired using a MALDI-TOF/TOF Applied Biosystems 4800 Proteomics Analyser (Applied Biosystems, Framingham, MA) instrument equipped with a nitrogen laser emitting at 337 nm and operating in a reflectron mode. Full-scan mass spectra ranging from *m/z* 450 (or 500) to 4000 were acquired in the positive mode. Sample preparation for MALDI-MS was performed using 2,5-dihydroxybenzoic (DHB) acid as matrix, but using two different procedures.

The untreated and thermally treated samples of the Ara_3 were prepared for MALDI-MS analysis by deposition of 0.5 μL of sample, previously dissolved in water, on top of a layer of crystals formed by deposition of 0.5 μL of DHB solution on the MALDI plate and letting it dry at ambient conditions. The matrix solution was prepared by dissolving 5 mg of DHB in a 1 mL mixture of acetonitrile/methanol/aqueous trifluoroacetic acid (TFA) (1%, v/v) (1:1:1, v/v/v).

The untreated and thermally treated samples of the oligosaccharide mixtures (A25M75, A50M50, and A75M25), previously dissolved in water, were prepared for MALDI-MS analysis by mixing 5 μL of each sample to 5 μL of DHB solution (10 mg/mL in methanol). From this mixture, 0.5 μL were deposited on the MALDI plate, allowed to dry at ambient conditions.

II.7.3. High-performance liquid chromatography online coupled with electrospray ionization mass spectrometry

Because the electrospray ionization (ESI) source produces gas-phase ions directly from a liquid solution (Section I.3.2), it can be easily online coupled to a high-performance liquid chromatography (HPLC) system, allowing the separation of complex mixtures. In this PhD work, as detailed below, some of the model samples were analysed by HPLC-ESI-MS. For this analysis, C18 and C5 columns were used.

II.7.3.1. HPLC-ESI-MS analysis with a C18 column

The separation of compounds in the Ara_3-CQA mixture heated to 175 °C (175T1) was carried out on a Thermo Scientific Hypersil Gold RP C18 column (100 mm × 2.1 mm, 1.9 μm particle size) at controlled temperature (15 °C) using a HPLC system

equipped with an Accela autosampler (set at 16 °C), an Accela 600 LC pump, and an Accela 80 Hz photodiode array (PDA) detector. The roasted mixture was dissolved in water:methanol (50:50; v/v) and a volume of 10 µL was introduced into the column, using a flow rate of 400 µL/min. The mobile phases consisted of water:acetonitrile (99:1, v/v) (A) and acetonitrile (B), both with 0.1% of formic acid. The percentage of B was kept at 3% from 0 to 5 min, then reached 12% from 5 to 14 min, 12.8% from 14 to 14.50 min, and 100% from 14.5 to 16 min. Before the following run, the column was re-equilibrated by decreasing the percentage of B from 100% to 3% during 4.5 min, and held it at 3% for 3.5 min. Double online detection was carried out in the PDA detector, at 280 and 325 nm, and UV spectra were also recorded from 210 to 600 nm. The HPLC was coupled to a LCQ Fleet ion trap mass spectrometer (ThermoFinnigan, San Jose, CA, USA), equipped with an ESI source and operating in negative mode. The flow rate of nitrogen sheath and auxiliary gas were 40 and 5 (arbitrary units), respectively. The spray voltage was 5 kV and the capillary temperature, 300 °C. The capillary and tune lens voltages were set at -28 V and -115 V, respectively. The CID-MS^n experiments were performed on mass-selected precursor ions in the range of *m/z* 100-2000. The isolation width of precursor ions was 1.0 mass units. The scan time was equal to 100 ms and the collision energy was optimized between 15 and 45 (arbitrary units), using helium as collision gas. The data acquisition and processing were carried out using Xcalibur 2.1.0.

II.7.3.2. HPLC-ESI-MS analysis with a C5 column

The compounds formed from the Man_3 and mixtures Man_3-YL, Man_3-LY, Man_3-CQA, YL-CQA and Man_3-YL-CQA roasted by reaching a temperature of 175 °C (175T1) were identified by HPLC-ESI-MS and HPLC-ESI-MS^n. These experiments were performed on a Waters Alliance 2690 HPLC system (Milford, MA, USA) coupled to the LXQ linear ion trap (LIT) mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Each unroasted and roasted sample (previously dissolved in water) was diluted in mobile phase A, and then filtered using a PVDF Millex-GV syringe filter with a 0.22 µm pore size (Millipore, Billerica, MA, USA). For analysis, 5 µL of each diluted and filtered sample were introduced into a Discovery Bio Wide Pore C5 column (15 cm x 0.5 mm, 5 µm particle size; Supelco, Bellefonte, PA, USA), using a flow rate of 8 µL/min. The mobile phases consisted of water:acetonitrile (95:5, v/v) (A) and acetonitrile (B), both with 0.1% of formic acid. Prior to use, the mobile phases were filtered using

Supelco Nylon 66 membranes (0.45 μm pore size and 47 mm diameter), and degassed by ultrasonication for 15 min. The mobile phase gradient was programmed as follows: 0-5 min, 0% B; 5-8.75 min, a linear increase to 15% B; 8.75-30 min, isocratic mode with 15% B; 30-35 min, a linear increase to 50% B; 35-45 min, isocratic mode with 50% B. The mobile phase was brought back to the initial conditions in 10 min and allowed to equilibrate for 15 min until the next injection. Methanol was introduced post-column by a syringe pump (flow rate 5 μL/min) through a T union into the HPLC flow before entering the mass spectrometer.

The LIT mass spectrometer was operated in positive mode. Typical ESI conditions were as follows: source voltage, 5 kV; capillary temperature, 275 °C, and sheath gas (nitrogen) flow rate, 8 (arbitrary units). To obtain the product ion spectra of the major components during LC experiments, cycles consisting of one full scan mass spectrum (*m/z* 150-1500) and three data dependent MS^n (*n*=2-4) scans were repeated continuously throughout the experiments with the following dynamic exclusion settings: repeat count 1; repeat duration 30 s; exclusion duration 45 s. An isolation width of 0.9 Da was used with a 30 ms activation time for MS^n experiments, using helium as collision gas. Normalized collision energy was varied between 24 and 35 (arbitrary units). Data acquisition and processing were carried out using Xcalibur software.

CHAPTER III. RESULTS AND DISCUSSION

III.1. UNDERSTANDING ROASTING-INDUCED MODIFICATIONS OF COFFEE ARABINOGALACTAN SIDE CHAINS

III.1.1. ROASTING-INDUCED CHANGES IN ARABINOTRIOSE, A MODEL OF COFFEE ARABINOGALACTAN SIDE CHAINS

III.1.2. NEUTRAL AND ACIDIC PRODUCTS DERIVED FROM HYDROXYL RADICAL-INDUCED OXIDATION OF ARABINOTRIOSE ASSESSED BY ELECTROSPRAY IONIZATION MASS SPECTROMETRY

III.2. UNDERSTANDING NON-ENZYMATIC TRANSGLYCOSYLATION REACTIONS

III.2.1. TRANSGLYCOSYLATION REACTIONS BETWEEN GALACTOMANNANS AND ARABINOGALACTANS DURING DRY HEAT TREATMENT

III.2.2. NON-ENZYMATIC TRANSGLYCOSYLATION REACTIONS INDUCED BY ROASTING: NEW INSIGHTS FROM MODELS MIMICKING COFFEE BEAN REGIONS WITH DISTINCT POLYSACCHARIDE COMPOSITION

III. 3. UNDERSTANDING COFFEE MELANOIDIN FORMATION

III.3.1. CHLOROGENIC ACID–ARABINOSE HYBRID DOMAINS IN COFFEE MELANOIDINS: EVIDENCES FROM A MODEL SYSTEM

III.3.2. ACID CATALYSES, AND MAILLARD REACTION PREVENTS, NON-ENZYMATIC TRANSGLYCOSYLATION REACTIONS DURING COFFEE ROASTING

CHAPTER III. RESULTS AND DISCUSSION

III.1. UNDERSTANDING ROASTING-INDUCED MODIFICATIONS OF COFFEE ARABINOGALACTAN SIDE CHAINS

Polysaccharides comprise about 50% of the dry weight of green coffee beans, in which 22% comprise galactomannans and 14-17% type II arabinogalactans (Section I.1). During the roasting process, coffee polysaccharides undergo structural modifications. In particular, arabinose (Ara), which occurs mainly in the side chains of arabinogalactans, is the polysaccharide sugar unit most degraded (Section I.2). However, the structural modifications promoted by roasting in Ara residues of coffee arabinogalactan side chains are not completely elucidated.

In order to better understand the roasting-induced modifications in Ara side chains of arabinogalactans, (α1→5)-L-arabinotriose (Ara_3), an oligosaccharide structurally related to the Ara side chains of coffee arabinogalactans, was submitted to dry thermal treatment (Section III.1.1) and oxidation induced by hydroxyl radicals (Section III.1.2).

CHAPTER III. RESULTS AND DISCUSSION

III.1. UNDERSTANDING ROASTING-INDUCED MODIFICATIONS OF COFFEE ARABINOGALACTAN SIDE CHAINS

III.1.1. ROASTING-INDUCED CHANGES IN ARABINOTRIOSE, A MODEL OF COFFEE ARABINOGALACTAN SIDE CHAINS

The results and discussion presented in this section were integrally published as follow:

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. (2013). Roasting-induced changes in arabinotriose, a model of coffee arabinogalactan side chains. *Food Chemistry*, *138*, 2291-2299.

III.1.1.1. Background and aim of the study

During coffee roasting, arabinose (Ara), which occurs mainly in the side chains of arabinogalactans (Section I.1.2), is the most sensitive sugar residue to degradation (Section I.2). In spite of the data already available, the structural modifications promoted by roasting in Ara residues of coffee arabinogalactan side chains are not well elucidated. Aiming to identify these structural modifications, (α1→5)-L-arabinotriose (Ara_3) was subjected to dry thermal treatments at 200 ºC during different periods. In order to obtain information concerning their sugar and glycosidic linkage compositions, sugar and linkage analyses were performed in untreated and thermally treated samples. Electrospray ionization mass spectrometry (ESI-MS), matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS), and electrospray ionization tandem mass spectrometry (ESI-MS^n) were also used for detailed analysis of structural modifications promoted by dry thermal processing of Ara_3.

III.1.1.2. Results and discussion

Thermal stability of (α1→5)-L-arabinotriose

Thermogravimetric (TG) and derivative thermogravimetric (DTG) curves of (α1→5)-L-arabinotriose (Ara_3) obtained from room temperature to 600 °C showed that it starts to degrade before 200 °C (Figure S1 in Appendix A). These results allowed concluding that Ara_3 has a lower thermal stability than that observed for (β1→4)-D-mannotriose (Man_3) under the same conditions (Moreira et al., 2011). The total relative mass loss observed when Ara_3 was heated to 200 °C (T1) and maintained at this temperature during 30 min (T2) and 60 min (T3) was 6.0, 19.4, and 21.7%, respectively. The longer the thermal treatment, a more intense brown coloration was observed, darker than that observed for Man_3 subjected to the same treatment (Moreira et al., 2011). In relation to the water-solubility, the material resulting from T1 treatment of Ara_3 was completely soluble in water. Nevertheless, and contrarily to that observed for Man_3 (Moreira et al., 2011), the materials resulting from T2 and T3 treatments were only partially soluble in water (35.3% for T2 and 24.6% for T3). To further characterize the structural changes associated with each thermal treatment, sugar and linkage analyses were performed.

Changes in sugar and glycosidic linkage compositions

The sugar and glycosidic linkage compositions of Ara_3 before and after thermal processing are presented in Table III.1. For T0, methylation analysis revealed the presence of terminally linked Ara*f* (T-Ara*f*, 29.3%) and (1→5)-linked Ara*f* (5-Ara*f*, 69.1%) approximately in the proportion of 1 to 2 and small amounts of 3,5- and 2,5-Ara*f* (0.1%), T-Xyl*p* (0.3%), 4-Xyl*p* (0.4%), and 4-Glc*p* (0.5%). These results and those obtained by sugar analysis allowed concluding that the unroasted sample was mainly composed by (α1→5)-L-arabinotriose oligosaccharides (>98%).

Table III.1. Sugar and glycosidic-linkage composition of Ara_3 before (T0) and after thermal processing (T1, T2 and T3).

Linkage	T0	T1	T2 - WSF[a]	T3 - WSF	T2 - WIF[b]	T3 - WIF
T-Ara*f*	29.3	12.9	16.6	24.7	29.6	29.8
2-Ara*f*	-	1.2	4.3	8.6	2.7	3.4
3-Ara*f*	-	2.3	8.4	11.6	6.8	6.4
5-Ara*f*	69.1	68.6	37.5	32.7	44.4	41.8
3,5-Ara*f*	0.1	2.8	12.7	7.3	8.6	8.7
2,5-Ara*f*	0.1	3.0	13.7	8.7	5.4	6.1
Total	98.6[c] (99.5)[d]	90.8 (99.4)	93.2 (98.8)	93.6 (99.0)	97.5 (99.9)	96.2 (99.0)
T-Xyl*p*	0.3	0.1	1.0	1.7	2.4	3.7
4-Xyl*p*	0.4	2.3	1.2	1.3	-	-
Total	0.7 (0.2)	2.4 (0.3)	2.2 (0.4)	3.0 (0.4)	2.4 (0.1)	3.7 (0.1)
4-Glc*p*	0.5 (0.3)	6.7 (0.2)	4.6 (0.8)	3.6 (0.5)	-	-
%Total sugars[d]	99.4	60.5	nd[e]	nd	47.2	46.8

[a]WSF, water-soluble fraction. [b]WIF, water-insoluble fraction. [c]Molar percentage obtained by linkage analysis. [d]Molar percentage (values in brackets) and total sugar percentage obtained by sugar analysis. [e]Not determined.

The percentage of T- and 5-Ara*f* (12.9% and 68.6%, respectively) decreased from T0 to T1, whereas 3,5- and 2,5-Ara*f* (2.8% and 3.0%, respectively) increased and 2-Ara*f* (1.2%) and 3-Ara*f* (2.3%) were formed. The percentage of total sugars decreased from 99.4% to 60.5%. The material resulting from T1 treatment was completely water soluble. However, for T2 and T3, a water insoluble fraction (WIF) was recovered and separated from the water soluble fraction (WSF). As observed for T1, the WSF from T2 and T3 treatments when compared with T0 showed a decrease of T- and 5-Ara*f*, a increase of 3,5-

and 2,5-Ara*f*, and the appearance of 2- and 3-Ara*f*. The same was observed for the WIF from T2 and T3 treatments with the exception of the content of T-Ara*f*, which was similar to T0. Also, the percentage of total sugars in WIF was much lower than in T0 (47.2% and 46.8%, respectively). Both WIF were soluble in dimethyl sulfoxide and precipitate upon addition of water without changing their glycosidic-linkage composition (data not shown).

Overall, the decrease of total sugars quantified after thermal processing allowed inferring their degradation. Also, the identification of new types of glycosidic linkages allowed inferring the occurrence of transglycosylation reactions. A more detailed analysis of structural modifications was made by mass spectrometry (MS).

MS analysis of (α1→5)-L-arabinotriose

In order to identify the products formed by dry thermal processing of Ara_3, unroasted sample was firstly analysed by ESI-MS in the positive mode, without any salt addition. Under these conditions neutral oligosaccharides ionize preferentially as sodium adducts ($[M+Na]^+$) (Moreira et al., 2011; Reis et al., 2004a; Tudella et al., 2011). Accordingly, the ESI-MS spectrum of unroasted sample (Figure S2A in Appendix A) showed a predominant ion at *m/z* 437, attributed to $[Ara_3+Na]^+$, and minor ions at *m/z* 453 and 851, attributed to $[Ara_3+K]^+$ and $[2Ara_3+Na]^+$, respectively.

The typical product ions observed in ESI-MS^n spectra of sodiated oligosaccharides ($[M+Na]^+$) result from the cleavage of glycosidic linkages, cross-ring cleavages (cleavage of two bonds within the sugar ring), and loss of water. Their relative abundance depends on structural details as monosaccharide composition, type of glycosidic linkages, and anomeric configurations (Asam & Glish, 1997; Simões et al., 2007). Thus, ESI-MS^n experiments were initiated studying the fragmentation pattern of unroasted sample. In order to mass-discriminate the reducing end product ions from those derived from non-reducing end of the oligosaccharide, ^{18}O-labelling at the carbonyl group of the reducing sugar residue was performed (Asam & Glish, 1997; Fang & Bendiak, 2007; Fang, Zirrolli & Bendiak, 2007; Hofmeister, Zhou & Leary, 1991; Konda, Bendiak & Xia, 2012). The ESI-MS^2 spectra of $[M+Na]^+$ ions of unlabelled (*m/z* 437) and ^{18}O-labelled (*m/z* 439) Ara_3 are shown in Figure III.1.

The MS^2 spectrum of unlabelled Ara_3 (Figure III.1A) showed two predominant product ions, at *m/z* 377 and 347, resulting from cross-ring cleavages with the loss of

$C_2H_4O_2$ (-60 Da) and $C_3H_6O_3$ (-90 Da) from the precursor ion, respectively. Also, the product ion observed at *m/z* 407 can result from a cross-ring cleavage between two adjacent bonds with the loss of CH_2O (-30 Da). Glycosidic cleavage product ions were also observed at *m/z* 305, formed by loss of one pentose residue ($Pent_{res}$, -132 Da); at *m/z* 287, formed by loss of a pentose sugar (Pent, -150 Da); at *m/z* 173, formed by loss of two $Pent_{res}$; and at *m/z* 155, formed by combined loss of Pent and $Pent_{res}$. The product ion observed at *m/z* 419 was formed by loss of H_2O (-18 Da). Minor abundant product ions were also found at *m/z* 245 (-192 Da) and 215 (-222 Da), formed by combined loss of $Pent_{res}$ and $C_2H_4O_2$ or $C_3H_6O_3$, respectively.

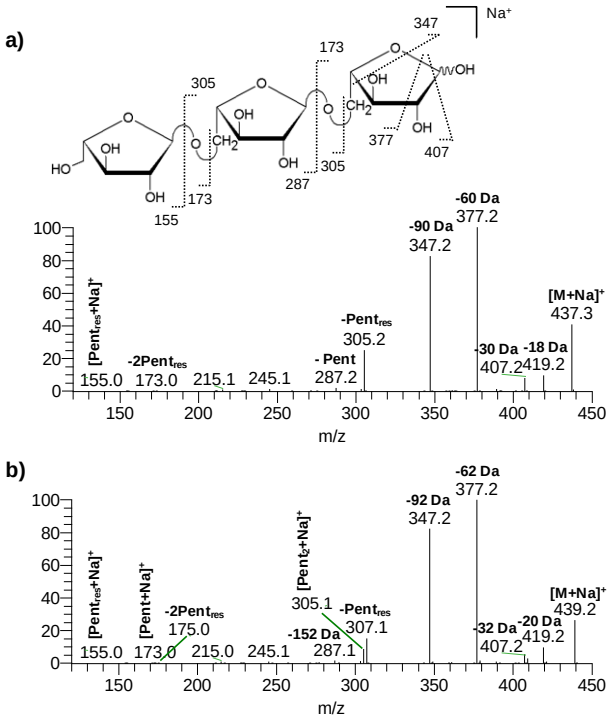

Figure III.1. ESI-MS^2 spectra of $[M+Na]^+$ ions of (A) unlabelled and (B) ^{18}O-labelled Ara_3.

The MS^2 spectrum of ^{18}O-labelled Ara_3 (Figure III.1B) showed two predominant product ions at *m/z* 377 and 347, formed by loss of $C_2H_4O^{18}O$ (-62 Da) and $C_3H_6O_2{}^{18}O$ (-92 Da), respectively. Also, the product ion at *m/z* 407, formed by loss of $CH_2{}^{18}O$ (-32 Da), showed a higher relative abundance compared to that at *m/z* 409, formed by loss of CH_2O (-30 Da). These observations allowed concluding that cross-ring cleavages occur predominantly at the reducing terminal residue with loss of the anomeric oxygen. In other words, and using the nomenclature proposed by Domon & Costello (1988), the product ions at *m/z* 407, 377, and 347 in the MS^2 spectrum of unlabelled Ara_3 (Figure III.1A) are

mainly attributable to $^{0,1}A_3^+$, $^{0,2}A_3^+$, and $^{0,3}A_3^+$ ions, respectively. In the MS^2 spectrum of ^{18}O-labelled Ara_3 (Figure III.1B), the higher relative abundance of the product ion at *m/z* 307 (-132 Da) compared to that at *m/z* 305 (-134 Da), and the occurrence of the product ion at *m/z* 287 (-152 Da) instead of the *m/z* 289 (-150 Da) allowed concluding that glycosidic cleavages occur preferentially (but not exclusively) between the anomeric carbon and the glycosidic oxygen. Also, the higher abundance of the product ion at *m/z* 419, formed by loss of $H_2{}^{18}O$ (-20 Da), compared to that at *m/z* 421 (-18 Da) allowed concluding that the water loss occurs preferentially at the reducing terminal residue with loss of the anomeric oxygen. All these conclusions are in agreement with previous studies on the fragmentation pattern of ^{18}O-labelled glucopyranose disaccharides (Asam & Glish, 1997; Hofmeister et al., 1991).

MS analysis of roasted samples

After heating Ara_3 to 200 °C (T1), the ESI-MS spectrum (Figure S2B in Appendix A) showed that the ion at *m/z* 437 remained the most abundant, but several new ions were also showed. All these ions are summarized in Table III.2, presenting the *m/z* values and relative abundances of all ions identified.

For longer thermal treatments (T2 and T3), the same ions were observed, although with different relative abundances (Figure S2 and Table S1 in Appendix A). To get additional information about their structures and to confirm the proposed structural assignments, ESI-MS^n (n=2-3) experiments were also performed for all ions identified after thermal processing of Ara_3 (T1, T2, and T3 treatments).

All ion series identified after thermal processing of Ara_3 (Table III.2) and their fragmentation patterns are described below, organized in the following sections grouping the series according to the different types of products formed: a) Products of depolymerization and polymerization; b) Dehydration products; c) Oxidation products; and d) Products of carbon-carbon bond cleavage. For each series, the MS^2 spectrum of the ion corresponding to the sodium adduct of the triose derivative ($Pent_n$ with n=3 in Table III.2) acquired from T1 sample is shown as an example.

a) Products of depolymerization and polymerization

The ions at *m/z* 173, 305, 437, 569, 701, 833, 965, 1097, and 1229 (Table III.2) were attributed to $[Pent_{1\text{-}9}+Na]^+$. The identification of pentose and pentose

oligosaccharides with a lower and higher degree of polymerization (DP) than 3 is an indication of the occurrence of depolymerization and polymerization reactions. These reactions were previously reported to occur during T2 and T3 treatments of the mannosyl and galactomannosyl oligosaccharides, while T1 treatment almost did not promote structural modifications (Moreira et al., 2011). This can be explained by the lower thermal stability of Ara_3 when compared with Man_3, which was also supported by the thermogravimetric data obtained. Depolymerization of coffee polysaccharides is a well-known reaction occurring during roasting (Nunes & Coimbra, 2002a; Oosterveld et al., 2003a). The occurrence of polymerization has also been supported using model compounds (Moreira et al., 2011).

Table III.2. Summary of the ions observed in the ESI-MS spectrum of Ara_3 heated to 200 °C (T1).

n	1	2	3	4	5	6	7	8	9
Products of depolymerization and polymerization									
$[Pent_n+Na]^+$	**173**[a] 6.3 ± 3.5[b]	**305** 28.5 ± 2.7	**437** 100 %	**569** 42.8 ± 1.5	**701** 30.9 ± 5.2	**833** 23.0 ± 3.9	**965** 11.5 ± 1.4	**1097** 5.3 ± 0.6	**1229** 2.2 ± 0.6
Dehydration products									
$[Pent_n-18Da+Na]^+$		**287** 8.3 ± 0.8	**419** 28.5 ± 1.1	**551** 19.5 ± 0.3	**683** 12.2 ± 1.1	**815** 9.9 ± 0.4	**947** 5.6 ± 1.2	**1079** 1.3 ± 1.0	**1211** 1.0 ± 0.8
$[Pent_n-36Da+Na]^+$		**269** 1.4 ± 0.2	**401** 5.5 ± 0.7	**533** 5.7 ± 0.2	**665** 4.6 ± 1.2	**797** 4.0 ± 0.6	**929** 2.6 ± 0.4	**1061** 1.3 ± 0.8	
$[Pent_n-54Da+Na]^+$			**383** 2.0 ± 0.5	**515** 3.6 ± 0.6	**647** 4.2 ± 0.6	**779** 3.7 ± 1.3	**911** 2.6 ± 0.8	**1043** 1.3 ± 1.1	
Oxidation products									
$[Pent_n+16Da+Na]^+$ (and $[Pent_n+K]^+$)		**321** 4.7 ± 1.1	**453** 14.4 ± 5.8	**585** 6.9 ± 1.8	**717** 6.4 ± 1.3	**849** 5.8 ± 1.9	**981** 3.7 ± 1.5	**1113** 1.6 ± 1.0	
$[Pent_n-2Da+Na]^+$			**435** 8.3 ± 1.5	**567** 5.8 ± 1.7	**699** 5.9 ± 1.7	**831** 6.0 ± 2.9	**963** 3.5 ± 2.7		
Products of carbon-carbon bond cleavage									
$[Pent_n-28Da+Na]^+$		**277** 1.2 ± 0.5	**409** 3.3 ± 0.5	**541** 2.8 ± 0.8	**673** 2.6 ± 0.8	**805** 2.6 ± 1.5	**937** 1.6 ± 0.9		
$[Pent_n-30Da+Na]^+$			**407** 3.7 ± 0.5	**539** 3.0 ± 0.9	**671** 3.4 ± 1.5	**803** 6.4 ± 5.9	**935** 2.8 ± 1.6		
$[Pent_n-44Da+Na]^+$			**393** 4.0 ± 1.9	**525** 3.4 ± 1.2	**657** 3.4 ± 1.3	**789** 3.0 ± 1.0	**921** 2.5 ± 1.1	**1053** 1.2 ± 0.8	
$[Pent_n-46Da+Na]^+$		**259** 2.4 ± 1.5	**391** 8.1 ± 6.2	**523** 4.3 ± 2.0	**655** 3.7 ± 1.2	**787** 3.7 ± 1.6	**919** 2.5 ± 1.5		
$[Pent_n-58Da+Na]^+$		**247** 4.5 ± 1.3	**379** 8.3 ± 0.6	**511** 5.2 ± 0.9	**643** 3.8 ± 0.8	**775** 3.8 ± 1.1	**907** 2.7 ± 0.8		
$[Pent_n-60Da+Na]^+$		**245** 2.3 ± 0.6	**377** 7.0 ± 1.5	**509** 9.0 ± 0.3	**641** 6.0 ± 4.0	**773** 6.3 ± 4.8	**905** 5.6 ± 2.4	**1037** 3.4 ± 2.6	
$[Pent_n-72Da+Na]^+$			**365** 1.8 ± 0.4	**497** 3.4 ± 0.8	**629** 3.8 ± 1.1	**761** 4.0 ± 1.1	**893** 3.4 ± 1.1	**1025** 1.8 ± 1.1	
$[Pent_n-74Da+Na]^+$			**363** 3.9 ± 1.0	**495** 4.4 ± 0.5	**627** 4.1 ± 1.1	**759** 4.3 ± 1.3	**891** 3.4 ± 1.5	**1023** 1.9 ± 1.4	
$[Pent_n-76Da+Na]^+$		**229** 2.0 ± 0.5	**361** 3.5 ± 0.7	**493** 6.0 ± 2.6	**625** 4.7 ± 1.0	**757** 4.1 ± 1.0	**889** 3.0 ± 1.3	**1021** 1.6 ± 1.0	
$[Pent_n-78Da+Na]^+$			**359** 3.6 ± 1.2	**491** 6.5 ± 1.0	**623** 5.1 ± 1.5	**755** 5.2 ± 2.5	**887** 3.8 ± 1.8	**1019** 2.0 ± 1.4	
$[Pent_n-88Da+Na]^+$		**217** 1.6 ± 0.6	**349** 1.7 ± 0.6	**481** 4.0 ± 1.5	**613** 3.7 ± 1.1	**745** 3.3 ± 0.8	**877** 2.8 ± 1.2	**1009** 1.8 ± 1.4	
$[Pent_n-90Da+Na]^+$		**215** 2.6 ± 0.7	**347** 6.2 ± 1.4	**479** 10.5 ± 1.3	**611** 7.0 ± 1.1	**743** 6.9 ± 3.1	**875** 4.9 ± 2.3	**1007** 2.8 ± 1.8	
$[Pent_n-104Da+Na]^+$			**333** 2.9 ± 0.4	**465** 5.6 ± 0.1	**597** 4.6 ± 0.1	**729** 4.4 ± 0.8	**861** 4.0 ± 1.3	**993** 2.2 ± 1.2	**1125** 1.3 ± 1.5
$[Pent_n-112Da+Na]^+$		**193** 1.1 ± 0.6	**325** 3.1 ± 0.6	**457** 11.3 ± 1.6	**589** 3.8 ± 1.8	**721** 4.6 ± 1.6	**853** 3.6 ± 1.2	**985** 2.3 ± 1.3	**1117** 1.3 ± 1.5
$[Pent_n-118Da+Na]^+$			**319** 6.4 ± 2.4	**451** 8.5 ± 2.0	**583** 4.6 ± 1.4	**715** 4.2 ± 1.6	**847** 3.8 ± 1.8	**979** 2.3 ± 1.4	

[a] *m/z* values of the ions are depicted in bold; [b]Mean ± standard deviation of three replicate spectra acquisitions made in different days.

To evaluate the presence of oligosaccharides with higher DP than those observed by ESI-MS, MALDI-MS analysis was performed. The ions at *m/z* 569, 701, 833, 965, 1097, 1229, 1361, 1493, 1625, 1757, 1889, 2021, and 2153, attributed to $[Pent_{4\text{-}16}+Na]^+$, were observed in the MALDI-MS spectrum acquired from T1 sample (Figure S3 in Appendix A), as well as in those acquired from T2 and T3 (data not shown).

The ESI-MS^2 spectra of all ions of the $[Pent_n+Na]^+$ series with $n \geq 2$ showed the product ions resulting from glycosidic cleavages, cross-ring cleavages, and water loss. The product ions observed for the ion at *m/z* 437 ($[Pent_3+Na]^+$) (Figure S4 in Appendix A) were the same observed for the untreated sample (Figure III.1A). However, the relative abundance of the product ions at *m/z* 377 (-60 Da) and 347 (-90 Da) was, respectively, lower and higher in the MS^2 spectrum of the roasted sample. Studies with hexose disaccharides have demonstrated that the relative abundance of cross-ring cleavage ions is linkage-dependent (Asam & Glish, 1997; Simões et al., 2007). Thus, the change in the relative abundance of these product ions suggests, in accordance with the results obtained by methylation analysis, that new types of glycosidic linkages are formed during thermal processing.

Other series of ions identified by ESI-MS after thermal processing of Ara_3 were identified as resulting from modification of the $[Pent_n+Na]^+$ series (Table III.2). The elemental composition of all these ions was confirmed by exact mass measurement and elemental composition determination from the ESI-MS spectrum of T1 sample acquired using a Q-TOF instrument (Table S2 in Appendix A). For all these ions, the differences (in absolute values) between the observed and calculated masses ranging from 0.0 to 20.4 mDa, which correspond to relative errors between 0.0 and 47.7 ppm, corroborate with high confidence the proposed modifications that will be described below. However, the coexistence of different isomers cannot be excluded.

b) Dehydration products

The series of ions with less 18, 36, and 54 Da when compared with the corresponding ion of the $[Pent_n+Na]^+$ series can be assigned as sodium adducts of pentose oligosaccharides modified by dehydration, due to the loss of one, two, or three water molecules, respectively. Dehydration products resulting from loss of one or three water molecules, but not those resulting from loss of two water molecules, were also formed during dry thermal processing of mannosyl and galactomannosyl oligosaccharides (Moreira et al., 2011). More recently, dehydration products resulting from loss of two

water molecules were reported based on the MS analysis of sugar samples heated for 2h at 140 °C in the case of fructose and at 180 °C for glucose and sucrose (Golon & Kuhnert, 2012).

The most abundant product ion observed in the ESI-MS^2 spectra (Figure S5 in Appendix A) of the ions at *m/z* 419 ($[Pent_3-H_2O+Na]^+$), 401 ($[Pent_3-2H_2O+Na]^+$), and 383 ($[Pent_3-3H_2O+Na]^+$) was formed by loss of a $Pent_{res}$. As glycosidic cleavages were shown through the ^{18}O-labelling experiment to occur preferentially between the anomeric carbon and the glycosidic oxygen, the occurrence of this product ion suggests that the dehydration products have a non-modified $Pent_{res}$ located at the non-reducing end. For the ions at *m/z* 419 and 401, the MS^3 spectrum of the $[M-Pent_{res}+Na]^+$ ion (Figure S5) also showed as the most abundant product ion that resulting from loss of a $Pent_{res}$, indicating that the loss of one and two water molecules occurred at the reducing end sugar residue of the corresponding non-modified oligosaccharide. The MS^2 spectrum of the dehydration product resulting from the loss of three water molecules ($[Pent_3-3H_2O+Na]^+$) showed the product ions at *m/z* 287 and 269 that can be formed by loss of ($Pent_{res}-2H_2O$) and ($Pent-2H_2O$), respectively. Because a non-modified $Pent_{res}$ was shown to be located at the non-reducing end, the presence of these ions suggests the loss of two water molecules at the reducing end and other one at the middle $Pent_{res}$ of the oligosaccharide.

c) Oxidation products

The series of ions with more 16 Da when compared with the corresponding ion of the $[Pent_n+Na]^+$ series can be assigned as sodium adducts of pentose oligosaccharides modified by oxidation containing an extra oxygen atom. These products can result from the formation of a pentonic acid moiety at the reducing end sugar residue, similar to that occurring during coffee roasting at the reducing end of galactomannans with formation of mannonic acid (Nunes et al., 2006). Oxidation products containing an additional oxygen atom and possessing an acid character were also identified by ESI-MS after oxidation of mannosyl and galactomannosyl oligosaccharides induced by the hydroxyl radicals generated by Fenton reaction (Tudella et al., 2011). Nevertheless, according to the results of elemental composition determination (Table S2 in Appendix A), $[Pent_n+O+Na]^+$ is the second most probable assignment for some ions of this series, the first corresponding to the potassium adducts of pentose oligosaccharides ($[Pent_n+K]^+$). Also, the increase in the relative abundance of the ion at *m/z* 453 ($[Pent_3+O+Na]^+$ and $[Pent_3+K]^+$) in the ESI-MS spectrum acquired from T1 sample compared to that in the

ESI-MS spectrum of untreated sample (Figure S2 in Appendix A) reinforces that these two ion assignments are possible.

The ESI-MS^2 spectrum of the ion at *m/z* 453 acquired from T1 sample (Figure S6B in Appendix A) showed the product ions at *m/z* 321 and 189, formed by loss of one and two $Pent_{res}$, respectively. These products ions were also observed in the ESI-MS^2 spectrum of the $[M+K]^+$ ion (*m/z* 453) acquired from untreated sample (Figure S6A). The presence of the oxidation product at *m/z* 453 ($[Pent_3+O+Na]^+$) in the thermally treated sample was supported by the product ions at *m/z* 409, 407, 305, and 287. The ions at *m/z* 305 and 287 were formed by loss of $(Pent+O)_{res}$ and (Pent+O), confirming the presence of an additional oxygen. The ions at *m/z* 409 and 407 were formed by loss of HCOOH and CO_2, confirming the presence of a carboxylic acid group. The relative abundance of the ions at *m/z* 407 and 409 was higher in the MS^2 spectra acquired from T2 and T3 samples (Figure S6C-D), suggesting an increase in the proportion of $[Pent_3+O+Na]^+$ ions in respect to $[Pent_3+K]^+$ ions for longer times of treatment.

As shown in Table III.2, a series of ions with less 2 Da when compared with the corresponding ion of the $[Pent_n+Na]^+$ series was also identified. Oxidation products with less 2 Da than the corresponding non-modified oligosaccharide due the formation of a keto group, also identified after oxidation of mannosyl and galactomannosyl oligosaccharides with hydroxyl radicals (Tudella et al., 2011), are well-known products of sugar oxidation (de Lederkremer & Marino, 2003). The ESI-MS^2 spectrum (Figure S7 in Appendix A) of the ion at *m/z* 435 ($[Pent_3-2H+Na]^+$) showed the product ions at *m/z* 303 and 171, formed by loss of one and two $Pent_{res}$, at *m/z* 287, formed by loss of (Pent-2H), and at *m/z* 155, attributed to $[Pent_{res}+Na]^+$, supporting the modification of the sugar residue located at the reducing end.

d) Products of carbon-carbon bond cleavage

Several series of ions designated as products of carbon-carbon bond cleavage were also observed after thermal processing of Ara_3 (Table III.2). These compounds are summarized in Figure III.2 and described below, grouped according to the proposed C-C bond undergoing cleavage.

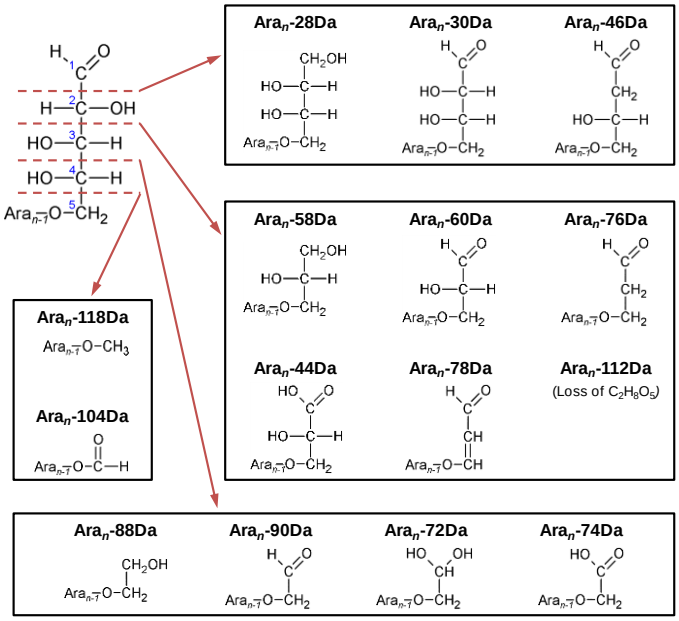

Figure III.2. Summary of the products resulting from the cleavage of a carbon-carbon bond.

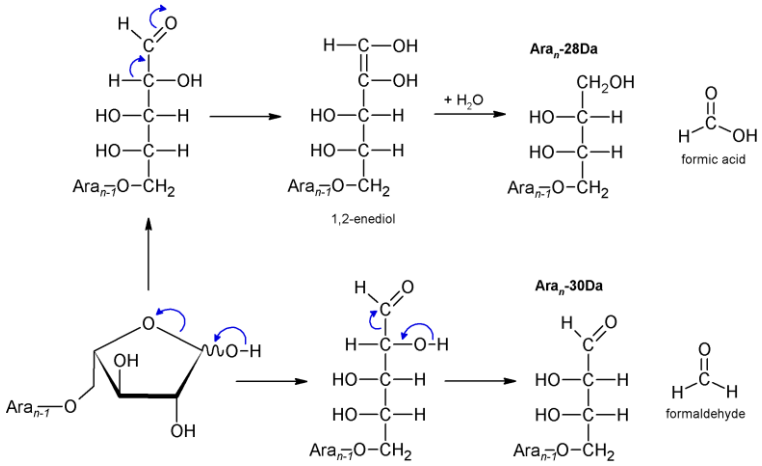

Figure III.3. Proposed mechanism for the formation of the products with less 28 Da and 30 Da than the corresponding non-modified oligosaccharide.

C1-C2 cleavage: The ions with less 28, 30, and 46 Da when compared with the corresponding ion of the $[Pent_n+Na]^+$ series were attributed to $[Pent_n-CO+Na]^+$, $[Pent_n-CH_2O+Na]^+$, and $[Pent_n-CH_2O_2+Na]^+$, respectively. The formation of the $[Pent_n-28Da+Na]^+$ series can be explained by hydrolytic cleavage of the C1-C2 bond of an 1,2-enediol intermediate (Figure III.3), as proposed to occur during dry thermal treatment (240 °C for 15 min) of glucose (Ginz, Balzer, Bradbury & Maier, 2000). Also, the formation of the $[Pent_n-46Da+Na]^+$ series can be explained by the hydrolytic cleavage of the C1-C2 bond of a 1,2-dicarbonyl intermediate (3-deoxypentosone) formed by dehydration of the 1,2-enediol intermediate (Ginz et al., 2000). Both proposed

mechanisms include the formation of formic acid (CH_2O_2), suggesting that the degradation of arabinose residues contributes to the formation of the formic acid emitted from green coffee beans during roasting (Yeretzian, Jordan, Badoud & Lindinger, 2002; Yeretzian, Jordan & Lindinger, 2003). The formation of the $[Pent_n\text{-}30Da+Na]^+$ series can be explained by loss of formaldehyde (CH_2O) (Figure III.3), other volatile organic compound known to be released from green coffee beans during roasting (Yeretzian et al., 2002).

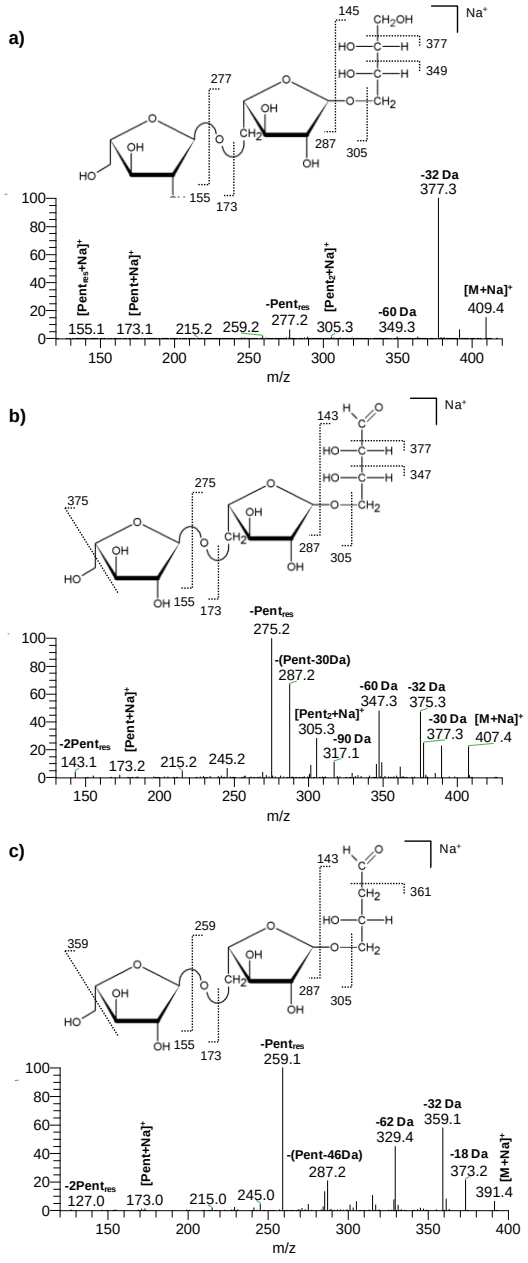

Figure III.4. ESI-MS^2 spectra of the ions at *m/z* (A) 409 ($[Pent_3\text{-}28Da+Na]^+$), (B) 407 ($[Pent_3\text{-}30Da+Na]^+$), and (B) 391 ($[Pent_3\text{-}46Da+Na]^+$) acquired from Ara_3 heated to 200 °C (T1).

The ESI-MS^2 spectrum of the ion at *m/z* 409 ($[Pent_3\text{-}28Da+Na]^+$, Figure III.4A) showed the product ion at *m/z* 277, formed by loss of a $Pent_{res}$. Also, its MS^3 spectrum (data not shown) showed the product ion at *m/z* 145, formed by loss of another $Pent_{res}$, supporting that the cleavage occurred at the sugar residue located at the reducing end. The ion at *m/z* 377, formed by loss of CH_3OH (-32 Da), due to its high abundance, should be probably resultant from the reducing end. The ESI-MS^2 spectra of the ions at *m/z* 407 ($[Pent_3\text{-}30Da+Na]^+$, Figure III.4B) and 391 ($[Pent_3\text{-}46Da+Na]^+$, Figure III.4C) showed the product ions formed by loss of one and two $Pent_{res}$, also supporting the cleavage of the sugar residue located at the reducing end. This assumption is also supported by the presence of the product ion at *m/z* 287, formed respectively by loss of (Pent-30Da) and (Pent-46Da) from the reducing end. The loss of CH_3OH (-32Da) from non-reducing end was also observed in both spectra.

C2-C3 cleavage: The ions with less 58 and 60 Da when compared with the corresponding ion of the $[Pent_n+Na]^+$ series were attributed to $[Pent_n\text{-}C_2H_2O_2+Na]^+$ and $[Pent_n\text{-}C_2H_4O_2+Na]^+$, respectively. The formation of $[Pent_n\text{-}58Da+Na]^+$ series can be explained by hydrolytic cleavage of the C2-C3 bond of an 2,3-enediol intermediate with the additional formation of glycolic acid ($C_2H_4O_3$). The products with less 58 Da than the corresponding non-modified pentose oligosaccharide can also be formed by loss of glyoxal ($C_2H_2O_2$). The formation of the $[Pent_n\text{-}60Da+Na]^+$ series can be explained by hydrolytic cleavage of the C2-C3 bond of a 2,3-dicarbonyl intermediate (1-deoxypentosone) formed by dehydration of the 2,3-enediol intermediate. According to this mechanism, acetic acid ($C_2H_4O_2$) is also released from the hydrolytic cleavage (Ginz et al., 2000). The formation of products with less 60 Da than the corresponding non-modified oligosaccharide can also be explained by loss of glycolaldehyde ($C_2H_4O_2$). The series of ions with less 76 and 112 Da when compared with the corresponding ion of the $[Pent_n+Na]^+$ series, attributed to $[Pent_n\text{-}C_2H_4O_3+Na]^+$ and $[Pent_n\text{-}C_2H_8O_5+Na]^+$, can be formed by dehydration from the products belonging to the $[Pent_n\text{-}58Da+Na]^+$ series with loss of one and three water molecules, respectively. Similarly, the ions of the $[Pent_n\text{-}78Da+Na]^+$ series can be formed by loss of one water molecule from the products belonging to the $[Pent_n\text{-}60Da+Na]^+$ series. On other hand, the ions of the $[Pent_n\text{-}44Da+Na]^+$ series, attributed to $[Pent_n\text{-}C_2H_4O+Na]^+$, can be explained by oxidation from

those belonging to the $[Pent_n\text{-}60Da+Na]^+$ series with the formation of a carboxylic acid group.

The ESI-MS^2 spectra, and casually MS^3 spectra, acquired for the ion series derived from C2-C3 cleavages are shown in the Supplementary Material (Figure S8-13, Appendix A). All spectra, except those acquired for the $[Pent\text{-}112Da+Na]^+$ series, supported the modification of only one of the three pentose residues, as well as the location of the modified residue at the reducing end. The MS^3 spectrum (Figure S11B) of the ion at *m/z* 193 ($[Pent_2\text{-}112Da+Na]^+$) showed the product ion at *m/z* 137, formed by loss of ($Pent\text{-}58Da\text{-}2H_2O$) and attributed to $[Pent_{res}\text{-}H_2O+Na]^+$, suggesting the presence of two modified residues: ($Pent\text{-}H_2O$) and ($Pent\text{-}58Da\text{-}2H_2O$).

Glycolic acid (Galli & Barbas, 2004), glyoxal (Daglia et al., 2007), and acetic acid (Yeretzian et al., 2002; Yeretzian et al., 2003) are known to be released from green coffee beans during roasting. Glyoxal and glycolaldehyde were also identified as oxidation products of laminaran under Fenton conditions (Ovalle et al., 2001). Also, glycolaldehyde was shown to be released predominantly from the C1-C2 of glucose monomers during pyrolysis of ^{13}C-labelled glucans (Ponder & Richards, 1993).

C3-C4 cleavage: The ions with less 88 and 90 Da when compared with the corresponding ion of the $[Pent_n+Na]^+$ series, attributed to $[Pent_n\text{-}C_3H_4O_3+Na]^+$ and $[Pent_n\text{-}C_3H_6O_3+Na]^+$, were identified as products resulting from the cleavage of the C3-C4. The formation of these series can be explained by loss of 2-hydroxypropanedialdehyde ($C_3H_4O_3$) and glyceraldehyde ($C_3H_6O_3$), as proposed for the oxidation of laminaran (Ovalle et al., 2001). The $[Pent_n\text{-}90Da+Na]^+$ series can also be explained by the formation of lactic acid ($C_3H_6O_3$) through the hydrolytic cleavage of a 3,4-dicarbonyl intermediate formed by rearrangement of the 1-deoxypentosone (Ginz et al., 2000). This is in accordance with the occurrence of lactic acid in coffee brews (Bähre & Maier, 1996; Galli & Barbas, 2004). Also, two series of ions with less 72 and 74 Da when compared with the $[Pent_n+Na]^+$ series, attributed to $[Pent_n\text{-}C_3H_4O_2+Na]^+$ and $[Pent_n\text{-}C_3H_6O_2+Na]^+$, were identified. These hydrate and carboxylic acid derivatives can be formed respectively by hydration and oxidation of the products belonging to the $[Pent_n\text{-}90Da+Na]^+$ series. These structures are supported by the correspondent ESI-MS^n data (Figures S14-17 in Appendix A).

C4-C5 cleavage: The ions with less 118 Da when compared with the corresponding ion of the $[Pent_n+Na]^+$ series, attributed to $[Pent_n-C_4H_6O_4+Na]^+$, were identified as products resulting from the cleavage of the C4-C5 bond. The ions of the $[Pent_n-104Da+Na]^+$ series, attributed to $[Pent_n-C_4H_8O_3+Na]^+$, can be explained by oxidation from those belonging to the $[Pent_n-118Da+Na]^+$ series with the formation of a carbonyl group (Figure III.2). These modifications are supported by the correspondent ESI-MS^2 spectra (Figures S18-19 in Appendix A).

III.1.1.3. Concluding remarks

This work showed that (α1→5)-linked arabinosyl trisaccharides are extensively modified by dry thermal processing. Together with the formation of depolymerized and polymerized products, it was also observed the formation of dehydration and oxidation products, and products of carbon-carbon bond cleavages with probably release of formaldehyde, formic acid, glycolaldehyde, glyoxal, acetic acid, glycolic acid, glyceraldehyde, 2-hydroxypropanedialdehyde, and lactic acid. The roasting of arabinotriose as model of coffee arabinogalactan side chains showed that the arabinose residues can be a source of aldehydes, dialdehydes, and acids previously reported to occur in coffee brews.

CHAPTER III. RESULTS AND DISCUSSION

III.1. UNDERSTANDING ROASTING-INDUCED MODIFICATIONS OF COFFEE ARABINOGALACTAN SIDE CHAINS

III.1.2. NEUTRAL AND ACIDIC PRODUCTS DERIVED FROM HYDROXYL RADICAL-INDUCED OXIDATION OF ARABINOTRIOSE ASSESSED BY ELECTROSPRAY IONIZATION MASS SPECTROMETRY

The results and discussion presented in this section were integrally published as follow:

Moreira, A. S. P., da Costa, E. V., Evtuguin, D. V., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. (2014). Neutral and acidic products derived from hydroxyl radical-induced oxidation of arabinotriose assessed by electrospray ionisation mass spectrometry. *Journal of Mass Spectrometry*, *49*, 280-290.

III.1.2.1. Background and aim of the study

Despite the knowledge gained so far, the modifications promoted by roasting on the structure of coffee polysaccharides are not completely elucidated, as well as the chemical reactions responsible for these modifications. However, studies by electron paramagnetic resonance spectroscopy suggest that free radical reactions may occur extensively during coffee roasting. In particular, hydroxyl radicals ($HO^{\bullet}$) were proposed to be involved in the oxidation of coffee galactomannans. Probably, these reactive oxygen species are also involved in the oxidation of coffee arabinogalactans during roasting (Section I.2.2.1).

In this work, the oxidation of (α1→5)-L-arabinotriose (Ara_3), an oligosaccharide structurally related to Ara side chains of coffee arabinogalactans, was studied in reaction with $HO^{\bullet}$ generated under conditions of Fenton reaction. To check the occurrence of oxidation induced by the presence of Fe^{2+}, a control mixture ($Ara_3/Fe^{2+}/H_2O$) was prepared. The extent of the oxidation of Ara_3 was monitored by electrospray ionization mass spectrometry (ESI-MS). The oxidation products were subsequently separated by ligand exchange/size-exclusion chromatography (LEX/SEC), identified by ESI-MS, and characterized by tandem MS (ESI-MS/MS).

III.1.2.2. Results and discussion

Oxidation of arabinotriose monitored by ESI-MS

The oxidation extent of (α1→5)-L-arabinotriose (Ara_3) was monitored by ESI-MS in positive mode. Under these MS conditions, oligosaccharides ionize preferentially as sodium adducts ($[M+Na]^+$) (Moreira et al., 2011; Zaia, 2004).

The ESI-MS spectra of the control ($Ara_3/Fe^{2+}/H_2O$) and reaction ($Ara_3/Fe^{2+}/H_2O_2$) mixtures are shown in Supplementary (Figure S1 in Appendix B). The MS spectrum of the control mixture (Figure S1A) shows a predominant ion at *m/z* 437, attributed to $[Ara_3+Na]^+$, and two minor ions at *m/z* 434 and 453, attributed to $[2Ara_3+K+H]^{2+}$ and $[Ara_3+K]^+$, respectively. As no other ions were identified, it was inferred that no oxidation occurred in the absence of H_2O_2. In contrast, the spectrum of the reaction mixture containing H_2O_2 (Figure S1B) shows several ions at *m/z* values lower and higher than 437, suggesting that Ara_3 was strongly oxidized. The ion at *m/z* 305, with less 132 Da than $[Ara_3+Na]^+$, due to the loss of a pentose residue ($Pent_{res}$), suggests the

depolymerization of Ara_3. The depolymerization of coffee polysaccharides is a well-known reaction occurring during roasting (Nunes & Coimbra, 2002a; Oosterveld et al., 2003a), that possibly occurs via a radical mechanism. The ions at *m/z* 453, 469, and 485, with more 16, 32, and 48 Da than $[Ara_3+Na]^+$, can correspond to oxidized species with plus one, two, and three oxygen atoms, respectively. The derivatives with additional oxygen atoms are well-known oxidation products of mono- and oligosaccharides formed during the reaction with $HO^{\bullet}$ generated by Fenton reaction (Moody, 1963; Tudella et al., 2011). Also, the reducing end of coffee galactomannans is modified during roasting with formation of a mannonic acid moiety, which contains an additional oxygen atom compared to the mannose unit (Nunes et al., 2006). In order to observe if both acidic and neutral products can result from the oxidative modifications, as previously observed for the oxidation of mannosyl and galactomannosyl oligosaccharides (Tudella et al., 2011), the reaction mixture of Ara_3 was fractionated by LEX/SEC.

Separation of neutral and acidic products by LEX/SEC

Figure III.5A shows the LEX/SEC chromatograms of the stock solution of Ara_3 (4 mg/mL) and the reaction mixture ($Ara_3/Fe^{2+}/H_2O_2$). According to results of previous studies (Gonçalves, Evtuguin & Domingues, 2008; Prozil et al., 2012; Reis et al., 2005; Tudella et al., 2011), the peak with lowest elution time (10-12 min; F1) in the chromatogram of the reaction mixture was assigned to acidic products of Ara_3. Also considering the elution time of the Ara_3 (maximum of the peak at 17 min), the fraction eluted at 16-18 min was mainly assigned to neutral trisaccharides. The fractions eluted at 18-19 min (F3) and 19-21 min (F4) were mainly assigned to disaccharide products, and the fraction at 21-23 min (F5) to monosaccharide products, both of neutral character. Thus, in agreement with the ESI-MS analysis, the occurrence of depolymerization was also supported by the elution profile of the reaction mixture. Also, it allowed confirming the presence of acidic and neutral products, subsequently identified by ESI-MS analysis of the fraction F1 and fractions F2-F5, respectively.

ESI-MS analysis of the acidic fraction

The ESI-MS spectrum obtained for the acidic fraction (F1) from oxidized Ara_3 is shown in Figure III.5B. The oxidation products inferred from this spectrum as $[M+Na]^+$

ions are summarized in Table III.3, in which the ions that showed a relative abundance higher than 20% are marked with a symbol.

For all ions, the proposed assignments were confirmed by exact mass measurement and elemental composition determination (Table S1 in Appendix B). The differences (in absolute values) between the exact and theoretical masses ranged from 0 to 15.6 mDa, which correspond to relative errors between 0.1 and 34.3 ppm, corroborating with high confidence the proposed assignments (Table III.3). These assignments were also corroborated with the ESI-CID-MS/MS spectra acquired for all tri- (*n*=3) and disaccharide (*n*=2) derivatives, which will be later described.

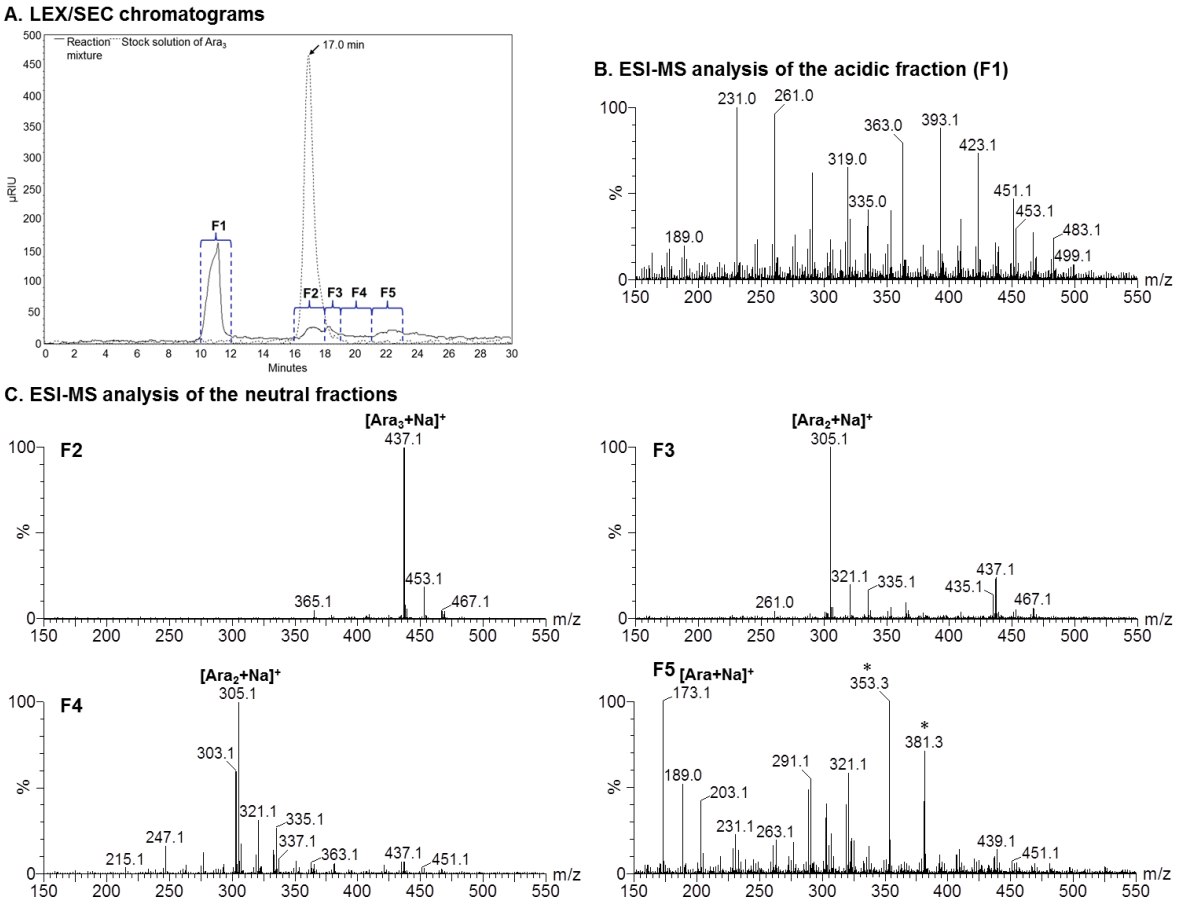

Figure III.5. (A) LEX/SEC chromatograms of the stock solution of Ara_3 (4 mg/mL) and the reaction mixture ($Ara_3/Fe^{2+}/H_2O_2$). (B) ESI-MS spectrum of the acidic fraction. (C) ESI-MS spectra of the neutral fractions. The ions belonging to impurities are marked with an asterisk (*).

Table III.3. Acidic oxidation products of Ara_3 identified by ESI-MS with respective *m/z* values of the $[M+Na]^+$ ions and proposed assignments (see Figures III.5B and III.6).

n		3	2	1
Arabinonic acid series				
$Pent_n$+O	[a](+16)	[b]♦453	♦321	♦189
$Pent_n$+O-2H	(+14)	♦451	♦319	187
$Pent_n$+O-4H	(+12)	449	♦317	
$Pent_n$+2O	(+32)	469	337	205
$Pent_n$+2O-2H	(+30)	♦467	♦335	203
$Pent_n$+2O-4H	(+28)	465	333	
$Pent_n$+3O	(+48)	485		
$Pent_n$+3O-2H	(+46)	♦483	351	
$Pent_n$+3O-4H	(+44)	481		
$Pent_n$+4O-2H	(+62)	499		
$Pent_n$+4O-4H	(+60)	497		
Erythronic acid series (from C1-C2 cleavage)				
$Pent_n$-CH_2O+O	(-14)	♦423	♦291	
$Pent_n$-CH_2O+O-2H	(-16)	421	♦289	
$Pent_n$-CH_2O+2O	(+2)	439	307	
$Pent_n$-CH_2O+2O-2H	(0)	♦437	♦305	
$Pent_n$-CH_2O+2O-4H	(-2)	435	303	
Glyceric acid series (from C2-C3 cleavage)				
$Pent_n$-$C_2H_4O_2$+O	(-44)	♦393	♦261	
$Pent_n$-$C_2H_4O_2$+O-2H	(-46)	391	♦259	
$Pent_n$-$C_2H_4O_2$+2O	(-28)	♦409	♦277	
$Pent_n$-$C_2H_4O_2$+2O-2H	(-30)	♦407	275	
Glycolic acid series (from C3-C4 cleavage)				
$Pent_n$-$C_3H_6O_3$+O	(-74)	♦363	♦231	
$Pent_n$-$C_3H_6O_3$+O-2H	(-76)	361	229	
$Pent_n$-$C_3H_6O_3$+2O	(-58)	♦379	♦247	
$Pent_n$-$C_3H_6O_3$+2O-2H	(-60)	377	♦245	

[a]Values in brackets are the *m/z* value differences compared to the $[Pent_n+Na]^+$ ions. [b]The ions marked with a symbol (♦) showed a relative abundance higher than 20%.

All the ions attributed to acidic oxidation products of Ara_3 are briefly described below, organized as in Table III.3 by grouping of different ion series according the proposed acidic residue: a) Arabinonic acid series, b) Erythronic acid series (from C1-C2 cleavage), c) Glyceric acid series (from C2-C3 cleavage), and d) Glycolic acid series (from C3-C4 cleavage). In Figure III.6 are represented the oxidation products containing an acidic residue, without further modifications.

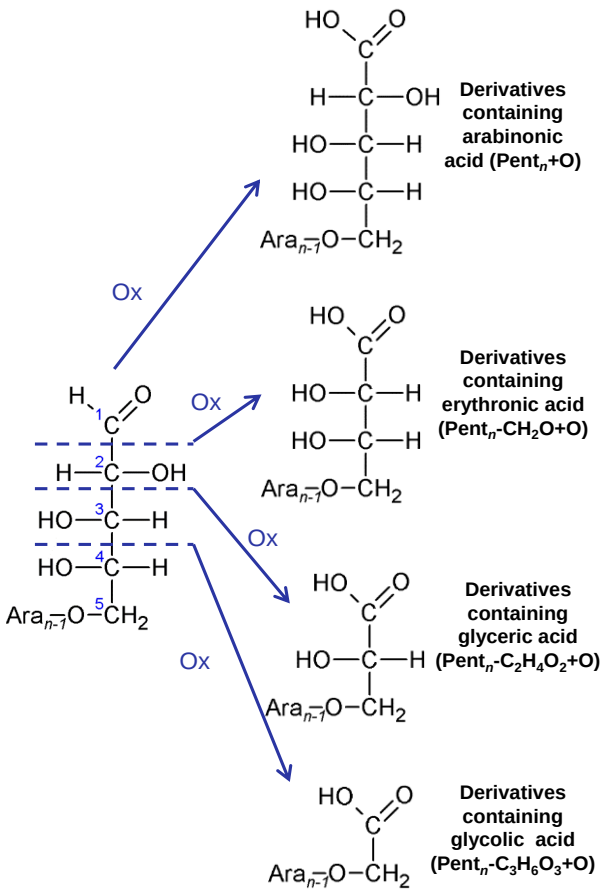

Figure III.6. Oxidation products containing an acidic residue at the reducing end of the corresponding non-modified oligosaccharide.

a) Arabinonic acid series: The ion at *m/z* 453 was assigned to $[Pent_3+O+Na]^+$. The corresponding di- and monosaccharide products (ions at *m/z* 321 and 189) were also identified, corroborating the occurrence of depolymerization. Due to their acidic character, these products can result from the formation of an arabinonic acid moiety (Figure III.6), similar to the formation of a mannonic acid moiety at the corresponding reducing end of coffee galactomannans occurring during roasting (Nunes et al., 2006), and of mannosyl and galactomannosyl oligosaccharides during oxidation with HO$^\bullet$ (Tudella et al., 2011). Also, oligosaccharides containing an arabinonic acid were proposed to be formed during roasting of Ara_3 (Section III.1.1). The formation of these oxidation products can result from the proton abstraction at C1 of the reducing end residue. The further reaction of the oligosaccharide radical with HO$^\bullet$ gives primarily the L-arabino-1,4-lactone derivative, producing the arabinonic acid derivative after hydrolysis of the lactone (Tudella et al., 2011).

Several series of ions were identified as products formed by the further oxidation of the derivatives containing an arabinonic acid moiety (Pent$_n$+O). Two series, assigned

as $[Pent_n+O-2H+Na]^+$ (n=3-1) and $[Pent_n+O-4H+Na]^+$ (n=3-2), were observed with less 2 and 4 Da than the $[Pent_n+O+Na]^+$ series, possibly due to the formation of one and two keto groups, respectively. Similarly, 2-keto-gluconic acid (Hex+O-2H) was identified after oxidation of glucose (Moody, 1963). Two series of ions with more 16 and 32 Da than the $[Pent_n+O+Na]^+$ series, assigned respectively as $[Pent_n+2O+Na]^+$ (n=3-1) and $[Pent_n+3O+Na]^+$ (n=3), were also observed. These oxidation products can be formed, respectively, by the further formation of an additional hydroxyl (–OH) group, and two –OH groups or a hydroperoxy (–OOH) group. Ion series with less 2 and 4 Da (due the formation of one and two keto groups, respectively) than the $[Pent_n+2O+Na]^+$ and $[Pent_n+3O+Na]^+$ series were also observed. Also, $[Pent_n+4O-2H+Na]^+$ (n=3) and $[Pent_n+4O-4H+Na]^+$ (n=3) series were identified, possibly corresponding to oxidation products containing additional keto, –OH, and –OOH groups (with different possible combinations for the number of –OH and –OOH groups).

b) Erythronic acid series (from C1-C2 cleavage): The ion at *m/z* 423, with less 14 Da than $[Ara_3+Na]^+$, was assigned to $[Pent_3-CH_2O+O+Na]^+$. The corresponding disaccharide product (ion at *m/z* 291) was also identified. The acidic character of these ions allowed suggesting the presence of an erythronic acid moiety resulting from oxidative cleavage of the C1-C2 bond of the sugar residue at the reducing end of the corresponding non-modified oligosaccharide (Figure III.6), as proposed for the formation of the pentonic acid moiety at the reducing end of coffee galactomannans occurring during roasting (Nunes et al., 2006), and of mannosyl and galactomannosyl oligosaccharides during oxidation with $HO^{\bullet}$ (Tudella et al., 2011).

c) Glyceric acid series (from C2-C3 cleavage): The ions at *m/z* 393 (less 44 Da than $[Ara_3+Na]^+$) and 261 were assigned to $[Pent_n-C_2H_4O_2+O+Na]^+$ (n=3 and 2, respectively). The acidic character of these ions allowed suggesting the presence of a glyceric acid moiety resulting from oxidative cleavage of the C2-C3 bond of the sugar residue at the reducing end of the corresponding non-modified oligosaccharide (Figure III.6), as proposed to the $Pent_n$-44 Da series identified after roasting of Ara_3 (Section III.1.1). Similarly, an erythronic acid derivative was identified after oxidation of mannosyl and galactomannosyl oligosaccharides (Tudella et al., 2011).

d) Glycolic acid series (from C3-C4 cleavage): The ions at *m/z* 363 (less 74 Da than $[Ara_3+Na]^+$) and 231 were assigned to $[Pent_n\text{-}C_3H_6O_3+O+Na]^+$ (*n*=3 and 2, respectively). Due to their acidic character, and as proposed to the $Pent_n$-74 Da series identified after roasting of Ara_3 (Section III.1.1), these ions can be attributed to oligosaccharides containing a glycolic acid moiety resulting from oxidative cleavage of the C3-C4 bond of the reducing end sugar residue (Figure III.6).

As described for arabinonic acid series, additional ion series were identified as products formed by the further oxidation of the derivatives containing erythronic, glyceric, or glycolic acid moieties with the formation of additional keto, –OH, and/or –OOH groups. As shown in Table III.3, they are as follows: $[Pent_n\text{-}CH_2O+O\text{-}2H+Na]^+$, $[Pent_n\text{-}CH_2O+2O+Na]^+$, $[Pent_n\text{-}CH_2O+2O\text{-}2H+Na]^+$, and $[Pent_n\text{-}CH_2O+2O\text{-}4H+Na]^+$; $Pent_n\text{-}C_2H_4O_2+O\text{-}2H+Na]^+$, $[Pent_n\text{-}C_2H_4O_2+2O+Na]^+$, and $[Pent_n\text{-}C_2H_4O_2+2O\text{-}2H+Na]^+$; $[Pent_n\text{-}C_3H_6O_3+O\text{-}2H+Na]^+$, $[Pent_n\text{-}C_3H_6O_3+2O+Na]^+$, and $[Pent_n\text{-}C_3H_6O_3+2O\text{-}2H+Na]^+$.

ESI-MS analysis of the neutral fractions

The ESI-MS spectra of the neutral fractions (F2-F5) are shown in Figure III.5C. All the ions assigned to neutral oxidation products of Ara_3 are summarized in Table III.4. They are briefly described below, organized by grouping of different ion series according the product type: a) Ara_3 and depolymerized products, b) Derivatives with additional keto, –OH, and –OOH groups, and c) Derivatives from C*m*-C*m+1* (*m* = 1-4) cleavage. The identification of the neutral products was confirmed by exact mass measurement and elemental composition determination (Tables S2-S5 in Appendix B). The differences (in absolute values) between the exact and theoretical masses ranged from 0 to 13.2 mDa, which correspond to relative errors between 0 and 37.3 ppm, corroborating with high confidence the proposed assignments (Table III.4). Also, the ESI-CID-MS/MS spectra acquired for all tri- and disaccharide products of neutral character allowed confirming the proposed assignments, as will be later described. It is also important to note that some of the neutral (Table III.4) and acidic (Table III.3) products are isobaric, some of them isomeric.

Table III.4. Neutral oxidation products of Ara_3 identified by ESI-MS with respective *m/z* values of the $[M+Na]^+$ ions and proposed assignments (see Figures III.5C and III.7).

n		3	2	1
Ara_3 and depolymerized products				
$Pent_n$		437	305	173
Derivatives with additional keto, –OH, and –OOH groups				
$Pent_n$-2H	[a](-2)	435	303	171
$Pent_n$-4H	(-4)	433		
$Pent_n$+O	(+16)	453	321	189
$Pent_n$+O-2H	(+14)	451	319	187
$Pent_n$+O-4H	(+12)	449	317	
$Pent_n$+2O	(+32)	469	337	
$Pent_n$+2O-2H	(+30)	467		
$Pent_n$+3O	(+48)	485		
$Pent_n$+3O-2H	(+46)	483		
$Pent_n$+3O-4H	(+44)	481		
Derivatives from C1-C2 cleavage				
$Pent_n$-CO	(-28)	409	277	
$Pent_n$-CO+O	(-12)	425	293	
$Pent_n$-CH_2O	(-30)	407	275	
$Pent_n$-CH_2O+O	(-14)	423	291	
$Pent_n$-CH_2O+O-2H	(-16)	421	289	
$Pent_n$-CH_2O+2O	(+2)	439	307	
Derivatives from C2-C3 cleavage				
$Pent_n$-$C_2H_2O_2$	(-58)	379	247	
$Pent_n$-$C_2H_2O_2$+O	(-42)		263	
$Pent_n$-$C_2H_4O_2$	(-60)	377	245	
$Pent_n$-$C_2H_4O_2$+O	(-44)	393	261	
$Pent_n$-$C_2H_4O_2$+O-2H	(-46)		259	
Derivatives from C3-C4 cleavage				
$Pent_n$-$C_3H_4O_3$+O	(-72)	365	233	
$Pent_n$-$C_3H_6O_3$	(-90)	347	215	
$Pent_n$-$C_3H_6O_3$+O	(-74)	363	231	
Derivatives from C4-C5 cleavage				
$Pent_n$-$C_4H_6O_4$+O	(-102)	335	203	
$Pent_n$-$C_4H_6O_4$+O-2H	(-104)	333	201	

[a]Values in brackets are the *m/z* value differences compared to the $[Pent_n+Na]^+$ ions.

a) Ara_3 and depolymerized products: As expected considering its LEX/SEC elution time, the most abundant ion observed in the ESI-MS spectrum of F2 was at *m/z* 437 ($[Pent_3+Na]^+$). It allowed inferring the presence of non-modified Ara_3. On the other hand, the most abundant ion observed in the ESI-MS spectra of F3 and F4 (*m/z* 305, $[Pent_2+Na]^+$), and that observed in the spectrum of F5 (*m/z* 173, $[Pent_1+Na]^+$) allowed

confirming the occurrence of depolymerization (Figure III.5C). As proposed for the depolymerization of cellobiose under Fenton conditions (Kane & Timpa, 1992; Varela, 2003), the formation of these depolymerized products can be due to the abstraction of hydrogen from the anomeric carbon involved in glycosidic linkage which underwent cleavage.

b) Derivatives with additional keto, –OH, and –OOH groups: Two ion series were observed with less 2 and 4 Da than the $[Pent_n+Na]^+$ series, assigned respectively as $[Pent_n-2H+Na]^+$ (n=3-1) and $[Pent_n-4H+Na]^+$ (n=3), containing one and two additional keto groups compared to the corresponding ion of the $Pent_n$ series. The $Pent_n$-2H series may correspond to derivatives containing a L-arabino-1,4-lactone moiety. Also, series of ions assigned as $[Pent_n+O+Na]^+$ (n=3-1), $[Pent_n+2O+Na]^+$ (n=3-2), and $[Pent_n+3O+Na]^+$ (n=3) were identified, and can result from the formation of new –OH and/or –OOH groups. These type of oxidation products were also observed after oxidation of mannosyl and galactomannosyl oligosaccharides (Tudella et al., 2011). Also, ion series were observed with less 2 and 4 Da than the $Pent_n+mO$ series (m=1-3): $[Pent_n+O-2H+Na]^+$ (n=3-1), $[Pent_n+O-4H+Na]^+$ (n=3-2), $[Pent_n+2O-2H+Na]^+$ (n=3), $[Pent_n+3O-2H+Na]^+$ (n=3), and $[Pent_n+3O-4H+Na]^+$ (n=3).

Regarding possible mechanisms, the formation of the neutral oxidation products containing an additional –OOH group ($Pent_n+2O$) can be due to the reaction of the carbohydrate radical ($R^{\bullet}$) with O_2 to form a peroxyl radical ($ROO^{\bullet}$). This $ROO^{\bullet}$ abstracts a hydrogen atom from another molecule to form a hydroperoxide derivative (ROOH). On the other hand, the neutral oxidation products containing an additional –OH group ($Pent_n+O$) can result from the cleavage of the ROOH catalysed by the Fe^{2+} ions, forming hydrates. Alternatively, the formation of an additional –OH group can occur through the reaction of the $R^{\bullet}$ with H_2O_2 (Eq. (5)) (Tudella et al., 2011).

$$R^{\bullet} + H_2O_2 \rightarrow ROH + HO^{\bullet} \quad (5)$$

c) Derivatives from Cm-Cm+1 (*m* = 1-4) *cleavage*: Similar to the acidic oxidation products, several neutral derivatives were identified as resulting from the cleavage of a carbon-carbon bond, grouped in Table III.4 according to the proposed bond undergoing cleavage. The proposed structures for some of these derivatives are represented in Figure III.7.

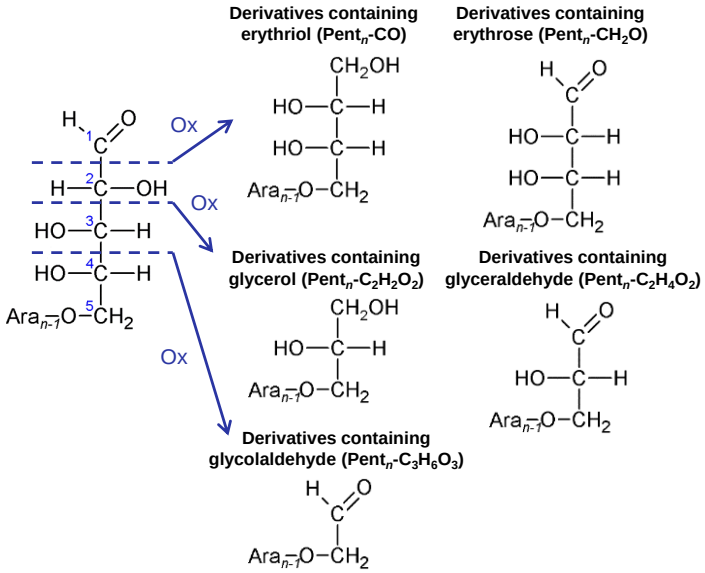

Figure III.7. Neutral oxidation products formed via oxidative cleavage of a carbon-carbon bond at the reducing end of the corresponding non-modified oligosaccharide.

Two ion series were observed with less 28 and 30 Da than the $[Pent_n+Na]^+$, assigned respectively as $[Pent_n-CO+Na]^+$ and $[Pent_n-CH_2O+Na]^+$ (n=3-2). These products, also formed during roasting of Ara_3 (Section III.1.1), were identified as resulting from the cleavage of the C1-C2 bond of the reducing end residue with formation, respectively, of erythriol and erythrose moieties (Figure III.7). The products containing an erythrose moiety ($Pent_n-CH_2O$) can be formed by the decarboxylation of the arabinonic acid moiety, and/or can result from the cleavage of ROO• derivatives (Tudella et al., 2011). Series with additional –OH, –OOH, and keto groups were also detected: $[Pent_n-CO+O+Na]^+$, $[Pent_n-CH_2O+O+Na]^+$, $[Pent_n-CH_2O+O-2H+Na]^+$, and $[Pent_n-CH_2O+2O+Na]^+$ (n=3-2).

Two ion series were observed with less 58 and 60 Da than the $[Pent_n+Na]^+$, assigned respectively as $[Pent_n-C_2H_2O_2+Na]^+$ and $[Pent_n-C_2H_4O_2+Na]^+$ (n=3-2). These products, also formed during roasting of Ara_3 (Section III.1.1), were identified as resulting from the cleavage of the C2-C3 bond of the sugar residue located at reducing end of the corresponding non-modified oligosaccharide with formation, respectively, of glycerol and glyceraldehyde moieties (Figure III.7). Series with more 16 and 14 Da, due to the additional –OH and keto groups, were also detected: $Pent_n-C_2H_2O_2+O$ (n=2), $Pent_n-C_2H_4O_2+O$ (n=3-2), and $Pent_n-C_2H_4O_2+O-2H$ (n=2). Some of these neutral derivatives were also identified after oxidation of mannosyl and galactomannosyl oligosaccharides (Tudella et al., 2011).

An ion series with less 90 Da than the $[Pent_n+Na]^+$ series, assigned to $[Pent_n-C_3H_6O_3+Na]^+$ (n=3-2), was observed. This type of products, also identified after roasting of Ara_3 (Section III.1.1), can result from the cleavage of the C3-C4 bond of the reducing end residue with formation of a glycolaldehyde moiety (Figure III.7). An ion series was also observed with more 16 Da, assigned as $[Pent_n-C_3H_6O_3+O+Na]^+$ (n=3-2). Other series with less 72 Da than the $[Pent_n+Na]^+$ series, $[Pent_n-C_3H_4O_3+O+Na]^+$ (n=3-2), was identified as resulting from the cleavage of the C3-C4 bond of the reducing end residue, also identified after roasting of Ara_3 (Section III.1.1).

Two ion series, $[Pent_n-C_4H_6O_4+O+Na]^+$ and $[Pent_n-C_4H_6O_4+O-2H+Na]^+$ (n=3-2), were identified as resulting from the cleavage of the C4-C5 bond of the reducing end residue. Interestingly, a series of ions with 104 Da less than those of the $[Pent_n+Na]^+$ series, possibly corresponding to $[Pent_n-C_4H_6O_4+O-2H+Na]^+$ ions, was also identified after roasting of Ara_3 (Section III.1.1).

The ions at *m/z* 365 (Figure S2 in Appendix B), 347 (Figure S3), and 203, assigned to neutral derivatives of Ara_3 resulting from the cleavage of the C3-C4 bond (Table III.4), are isobaric with $[M+Na]^+$ ions observed in the ESI-MS spectra of samples containing hexoses: $[Hex_2+Na]^+$, $[Hex_2-H_2O+Na]^+$, and $[Hex+Na]^+$, respectively (Moreira et al., 2011; Nunes et al., 2006). These ions highlight the importance of the MS/MS analysis confirming the proposed assignments for the ions observed by MS. For this reason, ESI-CID-MS/MS spectra were acquired for oxidation products of Ara_3 identified by ESI-MS analysis of both acidic and neutral fractions obtained by LEX/SEC.

Differentiation of isobaric acidic and neutral oxidation products by ESI-CID-MS/MS

Under ESI-(CID)-MS/MS conditions, the $[M+Na]^+$ ions of oligosaccharides showed product ions resulting from glycosidic cleavages (-132 Da due the loss of a $Pent_{res}$), loss of water, and cross-ring cleavages (cleavage of two bonds within the sugar ring) (Asam & Glish, 1997; Simões et al., 2007). The cross-ring ions of Ara_3 (non-oxidized sample) mainly result from neutral losses of $C_2H_4O_2$ (-60 Da) and $C_3H_6O_3$ (-90 Da) (Figure S4 in Appendix B). As the cross-ring cleavages of Ara_3 occur predominantly at the reducing terminal residue with loss of the anomeric oxygen (Section III.1.1), the cross-ring ions were respectively assigned as $^{0,2}A_3^+$ and $^{0,3}A_3^+$ ions, using the nomenclature proposed by Domon & Costello (1988).

Generally, the ESI-MS/MS spectra of the $[M+Na]^+$ ions of the oxidation products of Ara_3 allowed confirming the proposed assignments (Tables III.3 and III.4), and obtaining additional details about the structures. For clarity reasons, not all of these ESI-MS/MS spectra are shown. Also, the fragmentation pattern of the same type of acidic and neutral products was previously described in detail for oxidation products of mannosyl and galactomannosyl oligosaccharides (Tudella et al., 2011), and in the case of oxidation products of Ara_3 formed during thermal treatment (Section III.1.1). The ESI-MS/MS spectra selected to be presented in this paper demonstrate the capability of the MS/MS analysis for the differentiation of isobaric, or isomeric species of acidic and neutral character, previously separated by LEX-SEC chromatography.

Figure III.8 shows the ESI-MS/MS spectra acquired for the ion at *m/z* 321 ($[Pent_2+O+Na]^+$). The ions of the acidic and neutral isomers showed a clearly distinct fragmentation pattern. The MS/MS spectrum acquired from the acidic fraction (Figure III.8A) showed a predominant ion at *m/z* 189 ($[Pent+O+Na]^+$, or specifically [Arabinonic acid+Na]$^+$), resulting from the loss of a $Pent_{res}$. Also, the ion at *m/z* 155 ($[Pent_{res}+Na]^+$) was formed. Considering previous studies on the fragmentation of ^{18}O-labelled oligosaccharides showing that glycosidic cleavages occur preferentially between the anomeric carbon and the glycosidic oxygen (Asam & Glish, 1997; Hofmeister et al., 1991), the occurrence of these product ions indicates that a non-modified $Pent_{res}$ is located at the non-reducing end. The ion at *m/z* 275, formed by loss of 46 Da (HCOOH), corroborates the presence of a carboxylic acid group. The specific presence of arabinonic acid is corroborated by the ions at *m/z* 245, 229, 227, 217, 215 199, 197, 187, 185, 169, 167, 159, 157, and 129, as illustrated in Figure III.8A. The spectrum acquired from a neutral fraction (Figure III.8B) also showed the ions at *m/z* 189 and 155, indicating the presence of a non-modified $Pent_{res}$ located at the non-reducing end. The ion at *m/z* 245, resulting from cross-ring cleavage with loss of 76 Da ($C_2H_4O_2$+O), suggests that the additional –OH group is located at the C1 or C2 of the reducing end residue. However, the glycosidic ions at *m/z* 173 and 171, formed respectively by loss of an oxidized residue ($Pent_{res}$+O) and a Pent unit (-150 Da), and the cross-ring ions at *m/z* 261 and 231, formed respectively by loss of $C_2H_4O_2$ and $C_3H_6O_3$, suggest the presence of a positional isomer with the additional –OH group in the sugar residue located at the non-reducing end.

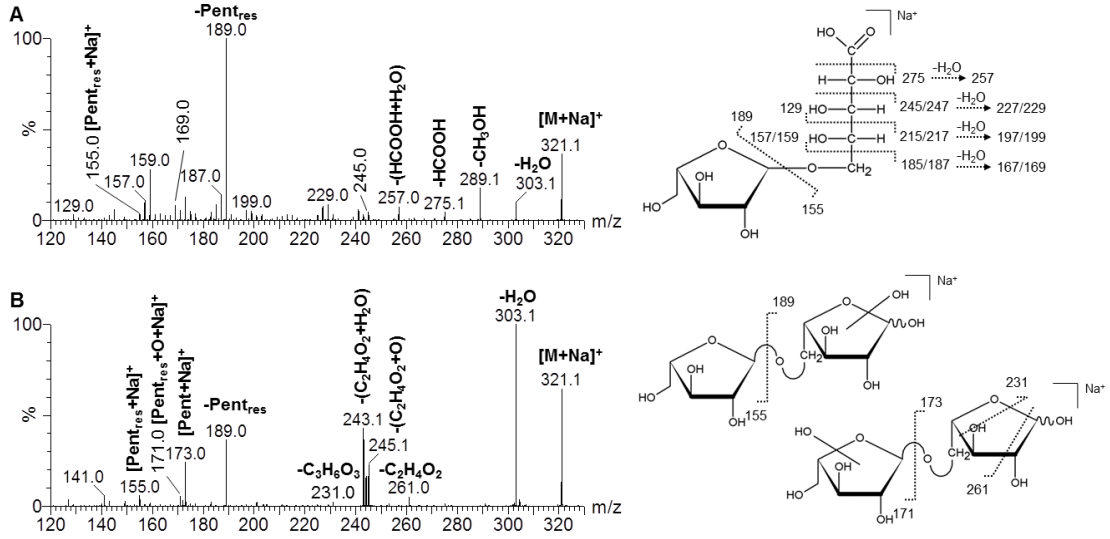

Figure III.8. ESI-MS/MS spectra acquired for the ion at *m/z* 321 ($[Pent_2+O+Na]^+$) from (A) acidic and (B) neutral fractions.

Figure III.9 shows the ESI-MS/MS spectra acquired for the ion at *m/z* 291 ($[Pent_2-CH_2O+O+Na]^+$). The most abundant product ion at *m/z* 159 ([Erythronic acid+Na]$^+$) and that at *m/z* 155 ($[Pent_{res}+Na]^+$) in the spectrum acquired from the acidic fraction (Figure III.9A) indicate the presence of a non-modified $Pent_{res}$ located at the non-reducing end. Also, the ion at *m/z* 245 (-HCOOH) corroborates the presence of a carboxylic acid group. The ions at *m/z* 185 and 129 are due to the cleavage of the C3-C4 bond of the erythronic acid. In the ESI-MS/MS spectrum acquired from a neutral fraction (Figure III.9B), the glycosidic ions at *m/z* 159 ($[Pent-CH_2O+O+Na]^+$) and 155 ($[Pent_{res}+Na]^+$) and the cross-ring ions at *m/z* 143 ($^{0,1}A_1^+$) and 113 ($^{0,2}A_1^+$) corroborate the presence of a non-modified $Pent_{res}$ located at the non-reducing end, and the modification of the sugar residue at the reducing end. Considering that the glycosidic cleavages occur preferentially between the anomeric carbon and the glycosidic oxygen, the ion at *m/z* 173 ($[Pent+Na]^+$) suggests that an isomer with a non-modified $Pent_{res}$ located at the reducing end is also present. However, it is also possible that the modification of the reducing end residue has promoted the glycosidic cleavage from the reducing end. Also, the presence of another isomer containing two modified residues cannot be completely excluded due the presence of the ion at *m/z* 171, assigned to $[Pent_{res}+O+Na]^+$. In summary, the ion at *m/z* 291 in the ESI-MS spectra of neutral fractions can be the contribution of three isomeric oxidation products. Similarly to the ion at *m/z* 291, the ESI-MS/MS spectrum acquired for the ion at *m/z* 261 ($[Pent_2-C_2H_4O_2+O+Na]^+$) from the acidic fraction corroborates the presence of a disaccharide product containing a glyceric acid moiety resulting from the cleavage

of the sugar residue at the reducing end, whereas that acquired from a neutral fraction corroborates the presence of different isomers (Figure S5 in Appendix B).

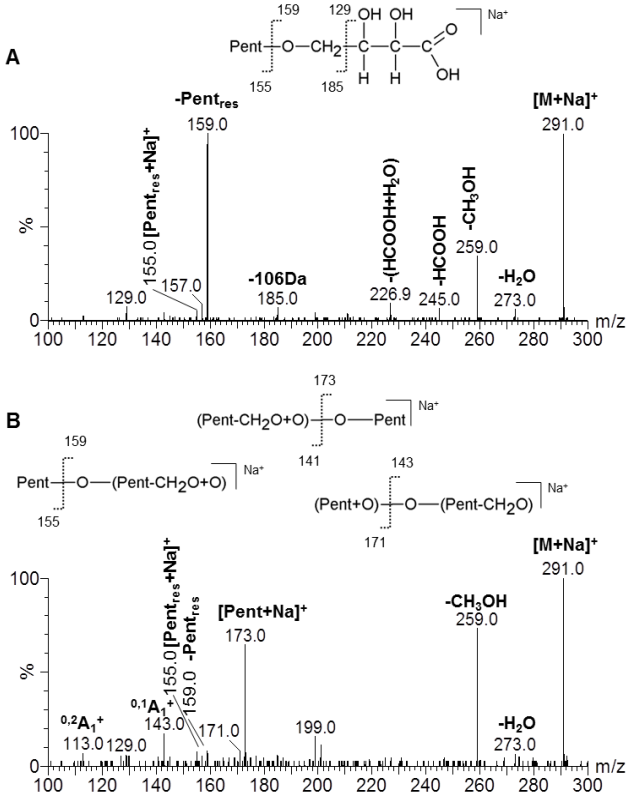

Figure III.9. ESI-MS/MS spectra acquired for the ion at *m/z* 291 ([$Pent_2$-CH_2O+O+Na]$^+$) from (A) acidic and (B) neutral fractions.

Figure III.10 shows the ESI-MS/MS spectra acquired for the ion at *m/z* 305, assigned respectively as [$Pent_2$-CH_2O+2O-2H+Na]$^+$ and [$Pent_2$+Na]$^+$ in acidic and neutral fractions. In the ESI-MS/MS spectrum acquired from the acidic fraction (Figure III.10A), the ion at *m/z* 259 (-HCOOH) corroborates the presence of a carboxylic acid group. The glycosidic ions at *m/z* 173 ([Pent+Na]$^+$) and 155 ([$Pent_{res}$+Na]$^+$) suggest the presence of a non-modified $Pent_{res}$ located at the non-reducing end, and the modification of the reducing end residue. However, the ions at *m/z* 169 ([$Pent_{res}$+O-2H+Na]$^+$) and 159 ([Pent-CH_2O+O+Na]$^+$, or specifically [Erythronic acid+Na]$^+$, suggest that an isomer containing two modified residues is also present. The ESI-MS/MS spectrum acquired from a neutral fraction (Figure III.10B) is consistent with that was previously obtained from a commercial standard of (α1→5)-L-arabinobiose (da Costa et al., 2012).

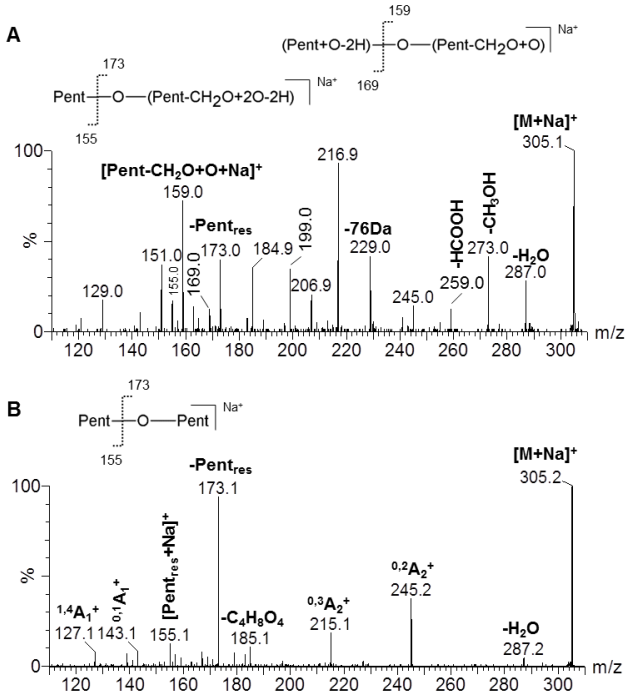

Figure III.10. ESI-MS/MS spectra acquired for the ion at *m/z* 305 from (A) acidic ($[Pent_2\text{-}CH_2O\text{+}2O\text{-}2H\text{+}Na]^+$) and (B) neutral ($[Pent_2\text{+}Na]^+$) fractions.

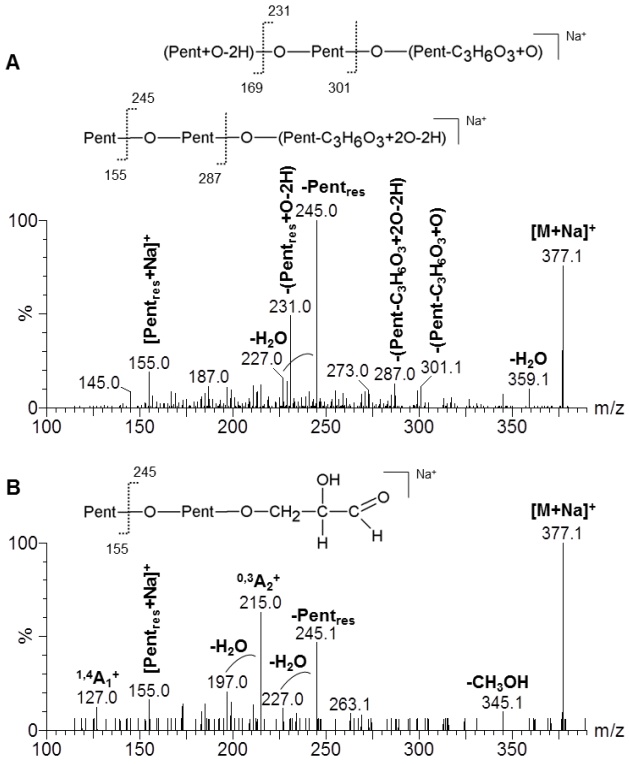

Figure III.11. ESI-MS/MS spectra acquired for the ion at *m/z* 377 from (A) acidic ($[Pent_3\text{-}C_3H_6O_3\text{+}2O\text{-}2H\text{+}Na]^+$) and (B) neutral ($[Pent_3\text{-}C_2H_4O_2\text{+}Na]^+$) fractions.

In Figure III.11 are shown the ESI-MS/MS spectra acquired for the ion at *m/z* 377, assigned respectively as $[Pent_3\text{-}C_3H_6O_3+2O\text{-}2H+Na]^+$ and $[Pent_3\text{-}C_2H_4O_2+Na]^+$ in acidic and neutral fractions. The spectrum acquired from the acidic fraction (Figure III.11A) showed the ions at *m/z* 245 ($\text{-}Pent_{res}$), 287 ($[PentPent_{res}+Na]^+$), and 155 ($[Pent_{res}+Na]^+$), corroborating the presence of two non-modified $Pent_{res}$, and the modification of the reducing end residue. However, the ions at *m/z* 301, formed by loss of 76 Da ($Pent_{res}\text{-}C_3H_6O_3+O$); 231, formed by loss of 146 Da ($Pent_{res}+O\text{-}2H$); and 169, assigned as $[Pent_{res}+O\text{-}2H+Na]^+$, suggest that an isomer containing two modified residues is also present. The spectrum acquired from a neutral fraction (Figure III.11B) also showed the ions at *m/z* 245 and 155. The cross-ring ions at *m/z* 215 ($^{0,3}A_2^+$) and 127 ($^{1,4}A_1^+$) also confirm the presence of a non-modified $Pent_{res}$ at the non-reducing end, and the modification of the reducing end residue, probably due to the cleavage of the C2-C3 bond with formation of a glyceraldehyde moiety.

In summary, the ESI-MS/MS spectra of acidic products showed the characteristic fragmentation pattern of acidic oligosaccharides, including the neutral loss of HCOOH (-46 Da), confirming the existence of a carboxylic acid group. Due the location of the acidic residue at the reducing end of the corresponding non-modified oligosaccharide, the typical fragmentation of the glycosidic linkages yielded the $[Acid\ unit+Na]^+$ ion. Specifically, the ions of arabinonic acid series showed the $[Arabinonic\ acid+Na]^+$ ion at *m/z* 189 (Figure III.8A), erythronic acid series showed the $[Erythronic\ acid+Na]^+$ ion at *m/z* 159 (Figure III.9A), and glyceric acid series showed the $[Glyceric\ acid+Na]^+$ ion at *m/z* 129 (Figure S5A in Appendix B). For the glycolic acid series, it was not possible to observed the $[Glycolic\ acid+Na]^+$ ion at *m/z* 99 because the ESI-MS/MS spectra were acquired from *m/z* value 100. Some of the ESI-MS/MS spectra acquired from both acidic and neutral fractions suggested the coexistence of positional isomers. It is also important to note that the ESI-MS/MS spectra of neutral oxidation products containing two additional oxygen atoms, assigned as hydroperoxy derivatives, showed an abundant product ion resulting from the loss of O_2 (-32 Da) (Figure S6A in Appendix B). The loss of O_2, as well as loss of $HOO^{\bullet}$ (-33 Da) or HOOH (-34 Da), was previously identified as a typical fragmentation pathway of hydroperoxy oligosaccharides (Tudella et al., 2011), and other biomolecules as phospholipids (Reis et al., 2007; Reis, Domingues, Ferrer-Correia & Domingues, 2004b), and amino acids/peptides (Domingues et al., 2003; Fonseca et al., 2009). Also, the ESI-MS/MS spectra acquired for the ion at *m/z* 453 ($[Pent_3+O+Na]^+$) from the acidic and neutral fractions showed a completely different

fragmentation pattern than that acquired from the control mixture ($[Ara_3+K]^+$, Figure S7 in Appendix B), which highlights again the importance of the ESI-MS/MS analysis for the differentiation of isobaric and isomeric compounds.

III.1.2.3. Concluding remarks

This study allowed identifying several acidic and neutral oxidation products of (α1→5)-L-arabinotriose (Ara_3) formed under conditions of Fenton reaction. All these products were identified by ESI-MS analysis of fractions recovered by LEX-SEC chromatography. In acidic fraction, the most abundant ions were assigned as products containing an acidic residue formed by oxidative scission of the sugar residue located at the reducing end of the corresponding non-modified oligosaccharide. In neutral fractions, it was possible to identify oxidation products containing keto, hydroxyl, and hydroperoxy moieties, and also resulting from ring scission. Beyond trisaccharides, depolymerized products were observed, namely as Ara_2 and Ara. The characterization of all tri- and disaccharide derivatives by ESI-MS/MS allowed the differentiation of isobaric and isomeric species of neutral and acidic character.

Considering the diversity of reactions that take place inside of coffee beans during roasting, the results of ESI-MS and ESI-MS/MS obtained in this study may help in identifying specifically structural modifications induced by free radicals on the arabinose side chains of coffee arabinogalactans. Some of the oxidation products of Ara_3 were previously reported when this oligosaccharide was submitted to dry thermal treatments mimicking coffee roasting conditions (Section III.1.1). This observation supports the hypothesis of the occurrence of oxidation reactions during coffee roasting involving the arabinogalactans. It is also important to note that the oxidation of carbohydrates, especially the pentose derived carbohydrates, has not been studied so far. However, it is recognized that this reaction occurs in several industrial processes. In this context, this study may be useful in many areas, not just in coffee research. For example, Gum Arabic, mainly composed by arabinogalactans, is commonly used as a stabilizer, thickening agent, and emulsifier in food industry (Ali, Ziada & Blunden, 2009), but their oxidation products were never investigated. Further studies are also needed to ascertain the influence of the type of sugar in the oxidation products formed.

III.2. UNDERSTANDING NON-ENZYMATIC TRANSGLYCOSYLATION REACTIONS

The dry thermal processing of oligosaccharides structurally related to coffee galactomannans, namely (β1→4)-D-mannotriose (Man_3) (Section I.2.1), and (α1→5)-L-arabinotriose (Ara_3), which is structurally related to the arabinose (Ara) side chains of coffee arabinogalactans (Section III.1.1), promoted the formation of new oligosaccharides. These oligosaccharides have a higher number of monosaccharide units, and new types of glycosidic linkages, evidencing the occurrence of non-enzymatic transglycosylation reactions with the transfer of glycosyl units to the hydroxyl groups of other glycosides induced by roasting.

In order to deeper understand the non-enzymatic transglycosylation reactions induced by roasting, model mixtures containing different proportions of Ara_3 and Man_3 were submitted to dry thermal treatments (Sections III.2.1 and III.2.2).

CHAPTER III. RESULTS AND DISCUSSION

III.2. UNDERSTANDING NON-ENZYMATIC TRANSGLYCOSYLATION REACTIONS

III.2.1. TRANSGLYCOSYLATION REACTIONS BETWEEN GALACTOMANNANS AND ARABINOGALACTANS DURING DRY HEAT TREATMENT

The results and discussion presented in this section were integrally published as follow:

Moreira, A. S. P., Simões, J., Pereira, A. T., Passos, C. P., Nunes, F. M., Domingues, M. R. M., & Coimbra, M. A. (2014). Transglycosylation reactions between galactomannans and arabinogalactans during dry thermal treatment. *Carbohydrate Polymers, 112*, 48-55.

III.2.1.1. Background and aim of the study

The use of oligosaccharides structurally related to coffee galactomannans, namely (β1→4)-D-mannotriose (Man_3) (Moreira et al., 2011), and an oligosaccharide structurally related to arabinose (Ara) side chains of coffee arabinogalactans, (α1→5)-L-arabinotriose (Ara_3) (Section III.1.1), which were individually submitted to dry thermal treatments, allowed to obtain evidences of the occurrence of roasting-induced non-enzymatic transglycosylation reactions involving a type of polysaccharide.

Aiming to investigate the possible occurrence of transglycosylation reactions between galactomannans and Ara side chains of arabinogalactans during coffee roasting, mixtures of Man_3 and Ara_3 were subjected to dry thermal treatments at 200 °C, using the experimental conditions used in the previous studies with each individual oligosaccharide. As the distribution of galactomannans and arabinogalactans in the coffee bean cell wall was shown to be heterogeneous (Sutherland et al., 2004), and these polysaccharides have a different vulnerability to roasting-induced degradation (Simões et al., 2014), three mixtures with different molar proportions of Man_3 and Ara_3 were used to mimic possible regions within the cell wall with a distinct polysaccharide composition. The compounds formed during thermal treatment of each oligosaccharide mixture were analysed by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS), and also according to the sugar and glycosidic linkage compositions. In order to support the formation of hybrid polysaccharide structures of arabinogalactans and galactomannans upon roasting, roasted coffee polysaccharide-rich fractions were also treated with specific glycosidases, and further analysed by neutral sugar or glycosidic linkage analysis.

III.2.1.2. Results and discussion

Identification of roasting-induced products by MALDI-MS

To identify the structural modifications promoted by dry thermal treatment, untreated and thermally treated oligosaccharide (Man_3 and Ara_3) mixtures were analysed by MALDI-MS in the positive-ion mode.

The MALDI-MS spectra of the thermally untreated mixtures (A50M50 in Figure III.12A) showed the ions at *m/z* 527 and 543, attributed to $[Man_3+Na]^+$ and $[Man_3+K]^+$,

respectively. As previously reported for MALDI-MS analysis of Man_3 (Moreira et al., 2011), the $[M+Na]^+$ ion showed higher abundance than $[M+K]^+$ ion. Also, the ion at *m/z* 453 ($[Ara_3+K]^+$) was observed. It was not possible to observe the corresponding sodium adduct at *m/z* 437 ($[Ara_3+Na]^+$), because the spectra were only acquired from *m/z* 450 due the presence of background matrix ions at lower *m/z* values.

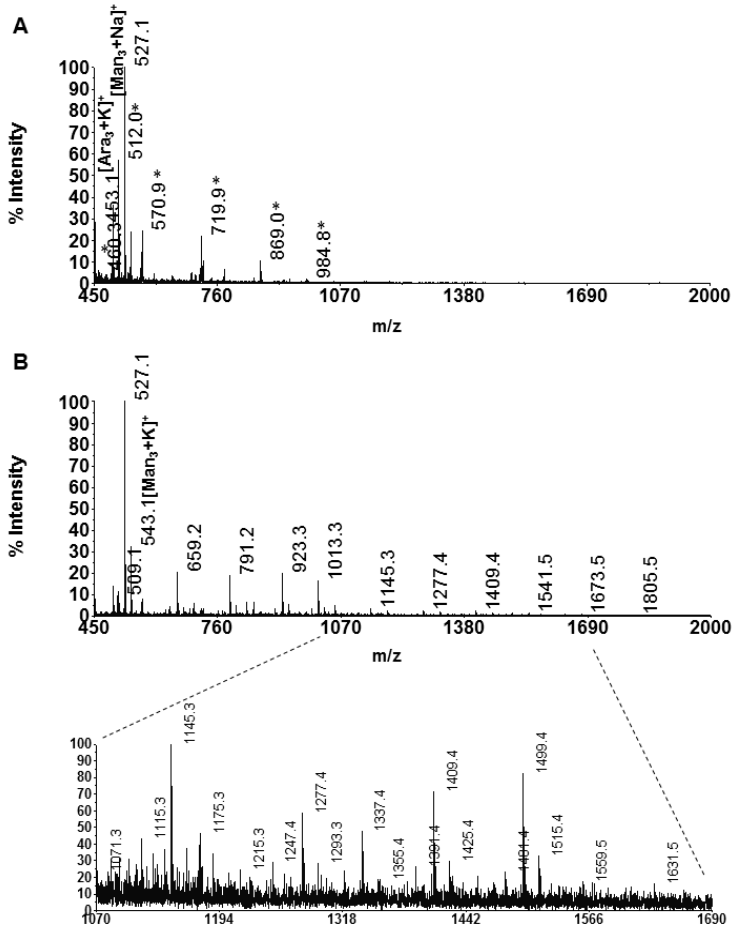

Figure III.12. MALDI-MS spectra of the A50M50 mixture (A) before and (B) after thermal treatment T1 (*m/z* 1070-1690 in inset). The ions marked with an asterisk (*) were attributed to impurities.

In addition to the aforementioned ions, new ions were observed in the MALDI-MS spectra of the thermally treated oligosaccharide mixtures. The spectrum of the A50M50 mixture heated to 200 ºC (T1) is shown in Figure III.12B as an example. Considering all the MALDI-MS spectra acquired from the three different mixtures (A25M75, A50M50, and A75M25) submitted to the three different treatments (T1, T2, and T3), it was possible to observe several ions at *m/z* up to 2147, which is attributable to the sodium adduct of a polysaccharide composed by 13 hexose (Hex) residues. As part of $[Hex_m+Na]^+$ series were also identified the ions at *m/z* 689, 851, 1013, 1175, 1337, and 1499 (*m*=4-9), as

well as the ion at *m/z* 1823 (*m*=11). These spectra also showed several ions at *m/z* up to 2813, which is attributable to the sodium adduct of a polysaccharide composed by 21 pentose (Pent) residues. As part of $[Pent_n+Na]^+$ series were also observed the ions at *m/z* 569, 701, 833, 965, 1097, 1229, 1361, 1493, and 1625 (*n*=4-12). These results are in accordance with the data previously reported for Man_3 (Moreira et al., 2011) and Ara_3 (Section III.1.1) when individually subjected to the same thermal treatments, confirming the occurrence of transglycosylation reactions involving the same type of sugar residues, and the formation of polysaccharides from the oligosaccharide mixtures submitted to dry heating conditions.

The MALDI-MS spectra of the thermally treated mixtures also showed several new ions attributable to $[M+Na]^+$ adducts of polysaccharides composed by both Pent and Hex residues, with *m/z* up to 2603, attributable to $PentHex_{15}$ (Table III.5). As part of the same series were also identified the ions at *m/z* 497, 659, 821, 983, 1145, 1307, 1469, and 1631, attributable to $PentHex_{2-9}$, and at *m/z* 1955 ($PentHex_{11}$). Similarly, combinations of structures containing 2 to 15 Pent residues and single to 13 Hex residues were identified (Table III.5). These results support the hypothesis of the occurrence of transglycosylation reactions involving sugar residues from different origins, and the formation of hybrid polysaccharides from the oligosaccharide mixtures when submitted to dry heating conditions.

Table III.5. Oligo- and polysaccharides composed by pentose and hexose residues ($Pent_nHex_m$) identified in the MALDI-MS spectra of the thermally treated mixtures as $[M+Na]^+$ ions with the indication of the *m/z* value and the proposed composition.

n	*m*											
	1	2	3	4	5	6	7	8	9	11	13	15
1		497	659	821	983	1145	1307	1469	1631	1955		2603
2	467	629	791	953	1115	1277	1439				2411	
3	599	761	923	1085	1247	1409			1895			
4	731	893	1055	1217	1379	1541	1703	1865	2027			
5	863	1025	1187	1349	1511	1673						
6	995	1157	1319	1481	1643	1805	1967					
7	1127	1289	1451									
8	1259	1421	1583	1745					2555			
9	1391	1553	1715									
11	1655	1817										
12		1949										
13		2081										
15			2507									

As previously observed, when the Man_3 (Moreira et al., 2011) and Ara_3 (Section III.1.1) were individually subjected to the same thermal treatments, or when Man_3 was roasted at 200 °C during 2 h (Simões et al., 2014), the formation of dehydrated derivatives was observed, particularly in the mixtures subjected to the longer thermal treatments (T2 and T3). Some of these derivatives have the same nominal mass (calculated by adding the mass of the predominant isotope of each element contributing to the molecule rounded to the nearest integer value) of non-modified oligo- or polysaccharides, as for example the ion observed at *m/z* 1229, identified as $[Pent_9+Na]^+$, but also attributable to $[Pent_3Hex_5\text{-}H_2O+Na]^+$. Thus, for the oligosaccharide mixtures subjected to the T2 or T3 treatment, the coexistence of different compounds with same nominal mass cannot be excluded. In order to gain additional structural information about the intact carbohydrate polymers formed under the dry heating conditions, the sugar compositions and their glycosidic linkages were analysed.

Sugar and glycosidic linkage compositions

The results of sugar and glycosidic linkage analyses obtained for both untreated and thermally treated mixtures are shown in Table III.6.

For thermally untreated ones (T0), the molar percentages of arabinose (Ara) and mannose (Man) (28.5 and 69.2% for A25M75; 56.7 and 41.8% for A50M50; and 75.9 and 22.9% for A75M25, respectively) obtained by GC-FID analysis of the alditol acetate derivatives are in line with the proportions of Ara_3 and Man_3 used in the preparation of the oligosaccharide mixtures. It should be taken into account that the percentages of T-Ara*f*, 5-Ara*f*, T-Man*p*, and 4-Man*p*, determined by direct quantification of the areas of the PMAA derivatives given by GC-MS, are not in accordance with such proportions, possibly due to the differences in the response factors of the partially methylated hexitol acetates and partially methylated pentitol acetates.

For all thermally treated mixtures, new types of glycosidic linkages (absent in the untreated ones) were identified, corroborating the occurrence of transglycosylation reactions inferred by MALDI-MS (Table III.6). The new Ara*f* linkages are in accordance with the previous observations of 2-, 3-, 2,5-, and 3,5-Ara*f* formed when Ara_3 was subjected to the same thermal treatments (Section III.1.1). The formation of 2-, 6-, 2,3-, 2,4-, 2,6-, 3,4-, 3,6-, 2,3,6-, 2,4,6-, and 3,4,6-Man*p* was also observed for the thermally treated oligosaccharide mixtures. The transglycosylation reactions are exemplified in

Figure III.13A for the formation of 6-linked Man*p* by reaction of a reducing Ara*f* with a Man*p*. Revisiting the results obtained when Man_3 was submitted to the same thermal treatments (Moreira et al., 2011), beyond the 2- and 6-Man*p* previously identified, the other new linkages were also identified. These new Man*p* linkages, except 2,4- and 3,4-linkages, were likewise observed for the roasted coffee polysaccharide data (Simões et al., 2014).

Table III.6. Sugar and glycosidic linkage compositions of the oligosaccharide mixtures before (T0) and after thermal treatments (T1, T2, and T3).

Linkages	A25M75				A50M50				A75M25			
	T0	T1	T2	T3	T0	T1	T2	T3	T0	T1	T2	T3
T-Ara*f*	5.4	11.6	12.9	15.0	12.5	13.8	17.3	14.6	27.6	24.7	38.7	23.9
2-Ara*f*			0.8	1.7		0.4	2.2	2.3		0.7	3.0	4.1
3-Ara*f*			0.9	1.3		0.4	3.1	3.7		0.9	4.3	7.2
5-Ara*f*	7.4	8.5	7.6	7.8	20.3	20.4	14.9	15.2	38.0	35.3	28.6	25.0
2,5-Ara*f*			0.5	0.8		0.6	2.4	3.3		1.0	2.2	6.3
3,5-Ara*f*			0.4	0.5		0.4	2.1	3.4		1.0	2.0	6.2
Total	12.8 (28.5)[a]	20.1 (23.7)	23.1 (28.8)	27.1 (29.8)	32.8 (56.7)	36.0 (50.1)	42.0 (55.4)	42.5 (57.3)	65.6 (75.9)	63.6 (74.9)	78.8 (76.5)	72.7 (76.4)
T-Rib*f*			0.3	0.4			0.3	0.4			0.6	0.7
5-Rib*f*	0.2	0.2	0.8	0.3	0.4	0.6	0.4	0.7	1.2	1.1	0.6	1.8
Total	0.2 (t)	0.2 (0.5)	1.1 (0.8)	0.7 (0.8)	0.4 (0.2)	0.6 (1.2)	0.7 (1.2)	1.1 (1.3)	1.2 (0.4)	1.1 (1.7)	1.2 (2.0)	2.5 (3.7)
T-Pent*p*[b]	0.2	0.2	0.8	1.0	0.3	0.2	0.9	1.4	0.9	0.4	1.5	2.1
4-Pent*p*	0.1	0.3	0.4	0.4	0.3	0.6	0.4	0.2	0.2	0.8	0.9	0.3
Total	0.3 (t)	0.5 (0.6)	1.2 (0.3)	1.4 (0.2)	0.6 (t)	0.8 (1.1)	1.3 (1.2)	1.6 (0.2)	1.1 (t)	1.2 (0.2)	2.4 (0.2)	2.4 (0.2)
T-Man*p*	34.3	46.8	35.5	33.6	27.6	26.9	17.6	14.9	14.6	14.9	9.5	6.7
2-Man*p*			1.8	1.8		0.3	1.6	1.8		0.2	0.1	0.6
4-Man*p*	46.7	29.3	16.2	12.4	33.2	28.3	12.6	9.9	15.6	14.9	5.6	4.2
6-Man*p*		1.0	11.5	10.6		2.5	9.9	11.0		1.5	1.5	4.9
2,3-Man*p*			0.2	0.4			0.3	0.3				
2,4-Man*p*			0.4	0.4		0.4	1.1	0.8				0.3
2,6-Man*p*			0.5	0.8			1.2	2.1				0.7
3,4-Man*p*			0.1	0.2			0.2	0.2				
3,6-Man*p*				0.2			0.4	0.6				0.3
4,6-Man*p*	2.0	0.4	4.8	4.9	1.7	2.4	7.3	8.2	0.6	1.2	0.5	3.4
2,3,6- and 2,4,6-Man*p*			0.2	0.7			1.4	2.1				
3,4,6-Man*p*								0.2				
Total	83.0 (69.2)	77.5 (72.5)	71.2 (65.4)	66.0 (64.9)	62.5 (41.8)	60.8 (44.8)	53.6 (40.1)	52.1 (38.7)	30.8 (22.9)	32.7 (22.0)	17.2 (18.2)	21.1 (17.9)
T-Gal*p*	1.8	0.9	0.5	0.3	1.4	0.9	0.4	0.4	0.7	0.5	t[c]	0.1
6-Gal*p*							0.1	0.1				
Total	1.8 (2.0)	0.9 (1.8)	0.5 (2.7)	0.3 (2.9)	1.4 (1.1)	0.9 (1.4)	0.5 (1.4)	0.5 (1.6)	0.7 (0.7)	0.5 (0.6)	t (0.9)	0.1 (0.8)
T-Glc*p*	0.6	0.1	0.8	0.8	0.8	0.9	0.1	0.2	0.2	0.2	t	t
4-Glc*p*	1.3	0.7	2.0	3.6	1.5	0.1	2.0	2.3	0.4	0.7	0.4	1.1
Total	1.9 (0.3)	0.8 (1.0)	2.8 (2.0)	4.4 (1.4)	2.3 (0.2)	1.0 (1.4)	2.1 (0.7)	2.5 (0.9)	0.6 (0.1)	0.9 (0.6)	0.4 (2.2)	1.1 (1.0)

[a]Values in brackets are the molar percentages obtained by GC-FID analysis of the alditol acetate derivatives. The other values were obtained by GC-MS analysis of the PMAA derivatives. [b]Pent*p*=Lyx*p* or Xyl*p*. [c]traces (<0.1%).

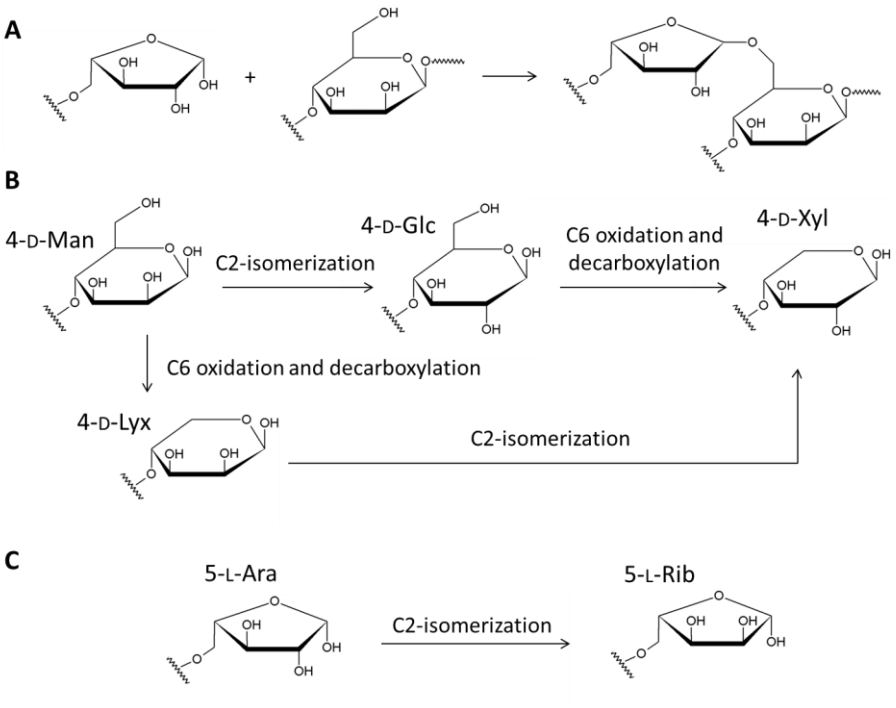

Figure III.13. Reactions proposed to occur during the roasting of the carbohydrates, originating new carbohydrate structures. (A) Transglycosylation reaction involving arabinose and mannose residues forming a 6-Man linkage; (B) Reactions of isomerization and oxidative decarboxylation resulting in the formation of glucose, xylose, and lyxose; (C) Reaction of isomerization of arabinose with concomitant formation of ribose.

An increase of glucose (Glc*p*), namely of 4-Glc*p*, was also observed after thermal treatment of the oligosaccharide mixtures. D-Glc*p* can be formed from D-Man*p* via isomerization at C2 (Figure III.13B) (Moreira et al., 2011; Simões et al., 2014). Similarly, L-Rib*f* should be formed from L-Ara*f* via isomerization at C2 (Figure III.13C). To disclose this hypothesis, the alditol acetates of all untreated and thermally treated oligosaccharide mixtures were analysed by GC-FID and GC-MS, allowing identifying as ribose a peak eluting immediately before the arabinose. Accordingly, a peak characteristic of 5-Pent*f* was observed eluting close to 5-Ara*f* in the GC-MS chromatograms of the PMAA derivatives of all samples (Table III.6), hypothesized as 5-Rib*f*, resulting from the isomerization of 5-Ara*f*. This isomerization can be promoted by the alkali extraction of the arabinan (calcium hydroxide solution at 90°C, according to Megazyme) used to obtain the commercial Ara_3 standard, explaining the presence of 5-Rib*f* in the unroasted samples. Thus, although the relative proportion of this linkage tends to increase upon roasting when compared to the unroasted samples, the presence of 5-Rib*f* cannot be attributed to a roasting promoted reaction. However, a peak characteristic of terminally linked Pent*f* was

also observed in heated samples but not in the untreated ones. This result should evidence the transglycosylation reactions that can be also expected to occur with 5-Rib*f*, giving origin to T-Rib*f* (Table III.6). Accordingly, the peaks corresponding to T- and 5-Rib*f* were also found revisiting the linkage analysis chromatograms of the thermally treated Ara_3 (Section III.1.1).

In the GC-MS chromatograms of the PMAA derivatives of the thermally treated mixtures, two peaks were observed with the retention times and mass spectra similar to those obtained for T-Pent*p* and 4-Pent*p*. These peaks can be attributed to T-Xyl*p* and 4-Xyl*p*, using (β1→4)-D-xylobiose as standard. The 4-Xyl*p* can be formed by the oxidative decarboxylation at C6 of the 4-Glc*p* at the reducing end, formed from 4-Man*p* isomerization, as suggested in Figure III.13B. T-Xyl*p* can be formed from 4-Xyl*p* via transglycosylation reactions, similarly as described for T-Rib*f*, formed from 5-Rib*f*. However, the T-Pent*p* residues can also be attributed to terminally linked lyxose (Lyx) residues, resulting from the oxidative decarboxylation at C6 of the Man located at the non-reducing end. Because the lyxitol acetate used as standard has the same structure of the arabinitol acetate, they could not be distinguished by the sugar analysis performed. Also, when the D-lyxose used as standard was methylated, giving the derivative T-Lyx*p*, it was observed the same retention time of T-Xyl*p* even using different chromatographic columns (DB-1 and FFAP). Because 4-Lyx*p* can also be formed from 4-Man*p* by oxidative decarboxylation, as suggested in Figure III.13B, T- and 4-Pent*p* were the designations used in Table III.6 to refer the Xyl and Lyx derivatives.

Evidence of transglycosylation reactions in roasted coffee polysaccharides

Revisiting the methylation analysis chromatograms of the galactomannan-rich spent coffee grounds (SCG) fractions submitted to dry heating at 160 to 240 ºC for 3 h (Simões et al., 2014), it was possible to observe, although in trace amounts, 2- and 3-Ara*f* linkages, as well as the 2,4- and 3,4-Man*p* linkages, not previously identified (Figure III.14).

In accordance with the results obtained for the model oligosaccharides, it was also possible to find in these samples T- and 4-Pent*p* residues, which should correspond mainly to T- and 4-Lyx*p*, respectively, formed upon the dry heat treatment of the SCG polysaccharides. In contrast to the oligosaccharide samples, there is a low probability for the formation of Xyl*p* from 4-Glc*p* formed via isomerization of reducing 4-Man*p*, since

a very low amount of 4-Man*p* residues in the SCG samples are present as reducing ends. In this work, the occurrence of transglycosylation reactions, particularly between galactomannans and arabinogalactans, was also corroborated with enzymatic treatment of coffee polysaccharide-rich fractions obtained from SCG as well as from instant coffee.

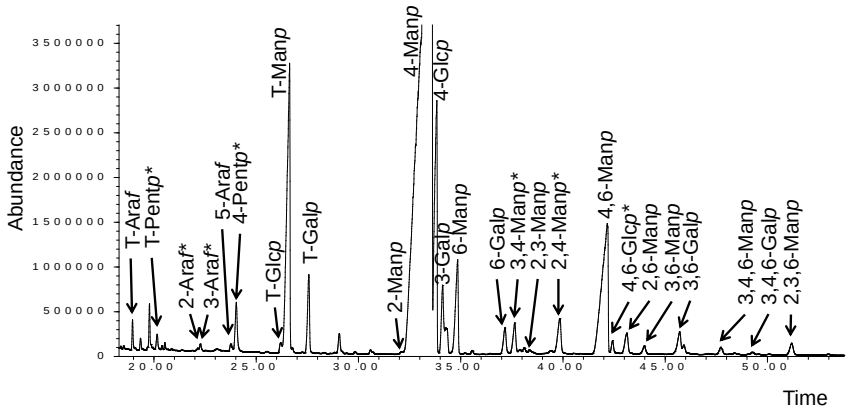

Figure III.14. GC-MS chromatogram of the PMAA derivatives obtained from the SCG galactomannans submitted to dry thermal treatment at 220 ºC for 3h. The peaks marked with an asterisk (*) were identified by revisiting the data previously obtained (Simões et al., 2014).

Table III.7 shows the glycosidic-linkage composition of a polysaccharide fraction soluble in cold water obtained from dry thermal treatment of a water insoluble SCG galactomannan-rich fraction (Simões et al., 2013). It was characterized by the predominance of 4-Man*p* (63.2%), the presence of 4,6- and T-linked Man*p* (4.5 and 9.8%, respectively), as well as T-Gal*p* (4.9%). In accordance with the occurrence of new glycosidic linkages formed by thermal processing, it was also found 6-, 2,3-, 2,4-, 3,4-, 2,3,6-, and 3,4,6-Man*p*. In the this sample, it was also possible to detect small amounts of 3-Gal*p* (3.9%), 3,6-Gal*p* (3.9%), and 5-Ara*f* (0.4%), which are diagnostic linkages for the presence of arabinogalactans. However, according to the extraction procedures used to isolate the galactomannans, it was not expected to obtain arabinogalactans derived from a water insoluble galactomannan-rich fraction. The occurrence of 2- and 3-Ara*f* residues is also in accordance with the formation of new linkages by transglycosylation promoted by the roasting. When this fraction was selectively hydrolysed with a α-galactosidase, which is known to cleave the α-Gal side chains from galactomannans, the resultant glycosidic linkage composition showed the decrease of the relative content of T-Gal*p* and 4,6-Man*p* when compared with the non-hydrolysed polysaccharides. These results confirm the occurrence of T-α-Gal*p* residues linked to the 4,6-Man*p* residues, which is a

structural characteristic of galactomannans. Also, the relative amount of the linkages diagnostic for the presence of arabinogalactans (T-, 3-, 6-, and 3,6-Gal*p*) decreased when the polysaccharide was hydrolysed with β-galactosidase, an enzyme that cleaves the β-Gal residues from the arabinogalactans. These results confirm the presence of arabinogalactan and galactomannan structural features, and new glycosidic linkages, possibly resultant from transglycosylation reactions between galactomannans and arabinogalactans promoted by the dry heat process.

Table III.7. Glycosidic linkage composition of the SCG polysaccharide fraction before and after enzymatic hydrolyses with α-galactosidase and β-galactosidase.

Linkages	SCG polysaccharide fraction		
	non-hydrolysed	α-galactosidase	β-galactosidase
T-Ara*f*	0.7	1.1	0.3
2-Ara*f*	t[a]	t	t
3-Ara*f*	0.1	0.1	t
5-Ara*f*	0.4	0.5	0.5
T-Pent*p*	0.2	0.2	
4-Pent*p*	0.7	1.2	2.1
Pent*p* = Lyx*p* or Xyl*p*			
T-Man*p*	9.8	6.9	10.3
4-Man*p*	63.2	65.8	67.5
6-Man*p*	1.0	0.7	0.8
2,3-Man*p*	0.1	0.1	
2,4-Man*p*	0.9	0.6	0.7
3,4-Man*p*	0.5	0.4	0.7
4,6-Man*p*	4.5	0.8	5.2
2,3,6-Man*p*	t	t	
3,4,6-Man*p*	t		
T-Gal*p*	4.9	3.0	2.1
3-Gal*p*	3.9	6.5	2.0
6-Gal*p*	1.6	2.5	1.5
3,6-Gal*p*	3.9	4.5	2.2
3,4,6-Gal*p*	0.1	0.2	t
T-Glc*p*	0.2	0.4	0.1
4-Glc*p*	3.2	4.1	3.8
4,6-Glc*p*	0.1	0.2	0.1

[a]traces (<0.1%)

Transglycosylation reactions between galactomannans and arabinogalactans were also recognized in instant coffee polysaccharides. Table III.8 shows the sugar composition of a commercial sample of instant coffee. The main sugar residue was Gal (53.6 mol%), followed by Man (36.9 mol%), and Ara (8.2 mol%). In agreement with the literature for instant coffee (Passos et al., 2014a; Wolfrom & Anderson, 1967), the higher percentage of Gal allowed inferring that the arabinogalactans were the most abundant

polysaccharides present. Upon cooling to 4 °C, an insoluble fraction, rich in Man (62.3 mol%) and containing Gal, was recovered, suggesting that the galactomannans were the most abundant polysaccharides present.

Table III.8. Sugar composition of the initial instant coffee sample, precipitates, and supernatants.

	Yield	Sugar composition (mol%)						Total
	(mass%)	Rha	Ara	Xyl	Man	Gal	Glc	sugar (%)
Initial sample of instant coffee		0.4	8.2	0.9	36.9	53.6		50
Precipitation by cooling to 4 °C, decantation, and centrifugation								
Supernatant	79	0.9	9.5	5.3	26.6	55.7		47
Precipitate	18	0.5	6.1	0.8	62.3	30.2		66
Partial enzymatic hydrolysis of the cold water insoluble material								
Mannanase_Sn	32	0.9	10.1	0.4	32.4	53.7	2.5	54
Mannanase_Ppt	68	0.3	3.5		77.0	18.5	0.7	71
Fractionation of the mannanase supernatant by graded ethanol precipitation								
Et50	7	0.3	3.9	1.5	74.2	0.7		80
Et75	14	1.1	8.1	0.5	3.9	85.2	1.1	77
Sn	78	1.2	17.6		48.3	28.3	4.6	22

The partial hydrolysis with an *endo*-(β1→4)-D-mannanase, which cleaves (β1→4) linkages between the mannose residues in the mannan backbone, allowed a partial solubilization of these carbohydrates, giving fraction Mannanase_Sn, with a sugar composition rich in Gal (53.7 mol%), whereas Man and Ara accounted for 32.4 and 10.1 mol%, respectively. The solubilization of arabinogalactan-derived carbohydrates upon treatment with the *endo*-(β1→4)-D-mannanase suggested the presence of hybrid polysaccharides derived from both arabinogalactans and galactomannans, possibly formed via the transglycosylation reactions, as previously reported to occur in roasted oligosaccharides and polysaccharides. The size-exclusion chromatography using BioGel P30 of the Mannanase_Sn fraction showed that the material was split in two distinct fractions, one at the void volume, and another at the exclusion volume, allowing to infer the presence of low and high molecular weight material. Also, the graded ethanol precipitation of the Mannanase_Sn fraction allowed to obtain a small fraction rich in galactomannans (Et50) and a fraction, accounting for 14% of the Mannanase_Sn material, rich in arabinogalactans (Et75), in accordance with previous studies (Nunes & Coimbra, 2001, 2002a). The fraction soluble in 75% ethanol (Sn) was mainly composed by Man (48.3 mol%), and also Gal and Ara (Table III.8), a composition explained by the action of the *endo*-(β1→4)-D-mannanase on galactomannan moieties, forming low molecular weight compounds rich in Man and an arabinogalactan fraction precipitated in 75%

ethanol, obtained from the water insoluble hybrid polysaccharides. Instant coffee is prepared using highly processed thermally treated coffee beans when compared with regular coffee used to prepare coffee brews. However, revisiting the data already published on coffee brew polysaccharides (Nunes & Coimbra, 2007; Nunes et al., 2006), it was observed that it was not possible to separate the galactomannan structural features from those of the arabinogalactans even after exhaustive fractionations, namely, by combined graded ethanol fractionation, anion exchange chromatography on Q-Sepharose FF, and phenylboronic acid Sepharose affinity chromatography or copper affinity chromatography. Based on the data of the present study, it can be inferred that hybrid coffee brew polysaccharides could also be formed by transglycosylation reactions.

III.2.1.3. Concluding remarks

The dry thermal treatment of the mixtures containing arabinosyl and mannosyl oligosaccharides (Ara_3 and Man_3) promoted the formation of polysaccharides composed by both hexose and pentose residues, mainly Man and Ara, corroborating the hypothesis of the occurrence of transglycosylation reactions between galactomannans and arabinose side chains of arabinogalactans during coffee roasting. This hypothesis was also corroborated by the finding of hybrid polysaccharide structures of arabinogalactans and galactomannans obtained from roasted coffee matrices and presenting the same characteristics found for the standards used as models.

III.2. UNDERSTANDING NON-ENZYMATIC TRANSGLYCOSYLATION REACTIONS

III.2.2. NON-ENZYMATIC TRANSGLYCOSYLATION REACTIONS INDUCED BY ROASTING: NEW INSIGHTS FROM MODELS MIMICKING COFFEE BEAN REGIONS WITH DISTINCT POLYSACCHARIDE COMPOSITION

The results and discussion presented in this section were integrally published as follow:

Moreira, A. S. P., Simões, J., Nunes, F. M., Evtuguin, D. V., Domingues, P., Coimbra, M. A., & Domingues, M. R. M. (2016). Nonenzymatic transglycosylation reactions induced by roasting: new insights from models mimicking coffee bean regions with distinct polysaccharide composition. *Journal of Agricultural and Food Chemistry*, *64*, 1831-1840.

III.2.2.1. Background and aim of the study

The possible occurrence of non-enzymatic transglycosylation reactions involving galactomannans and arabinogalactan side chains during coffee roasting was investigated using as model mixtures of (β1→4)-D-mannotriose (Man_3) and (α1→5)-L-arabinotriose (Ara_3). Three mixtures containing different molar proportions of each oligosaccharide (A25M75, A50M50 and A75M25) were used to mimic possible regions within the cell walls with distinct polysaccharide composition. To mimic coffee roasting conditions, the oligosaccharide mixtures were submitted to dry thermal treatments at 200 °C. As observed by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) analysis, the dry thermal processing of these mixtures promoted the formation of arabinan-mannan hybrid structures. This outcome supported the hypothesis of the occurrence of non-enzymatic transglycosylation during coffee roasting involving galactomannans and arabinogalactan side chains, which was corroborated by the finding of the same type of hybrid structures in roasted coffee polysaccharide-rich fractions (Section III.2.1).

In this study we have analysed untreated and thermally treated mixtures of Man_3 and Ara_3 for additional details, aiming to better understand the non-enzymatic transglycosylation reactions occurring under dry heating conditions, namely between galactomannans and arabinogalactan side chains during roasting of green coffee beans. Electrospray ionization mass spectrometry (ESI-MS) and electrospray ionization collision-induced dissociation tandem mass spectrometry (ESI-CID-MS^n) analyses were performed, by infusing labelled (^{18}O) and unlabelled samples. In order to verify if the structures formed during the thermal processing of the model mixtures have the same structural features than those formed during coffee roasting, or additional roasting treatments beyond the roasting of the green coffee beans, a galactomannan-rich fraction isolated from spent coffee grounds (SCG), submitted to additional roasting treatments, was treated with an *endo*-(β1→4)-D-mannanase. The hydrolysed material was further fractioned by ligand exchange/size-exclusion chromatography (LEX-SEC) and analysed by ESI-MS and ESI-CID-MS^n. The elemental composition of the ions identified in the ESI-MS spectra of both oligosaccharide mixtures and SCG sample was obtained by high resolution and high mass accuracy measurements using an Orbitrap-based mass spectrometer, equipped with an ESI source.

III.2.2.2. Results and discussion

Colour and water-solubility of each oligosaccharide mixture upon dry heating

Three different samples of each mixture were heated to 200 °C and maintained at 200 ºC for different periods: T1 (0 min), T2 (30 min), and T3 (60 min). After dry thermal processing, the mixtures (white or off-white powders) acquired a brown coloration, which was more intense for the longer treatments. However, the coloration of the compounds resulting from the longer treatments (T2 and T3) of the mixtures with higher proportion of Ara_3 (A50M50 and A75M25) was darker than that observed when the A25M75 mixture was submitted to the same treatments. These observations are in line with the lower thermal stability of Ara_3 (Section III.1.1) when compared with Man_3 (Moreira et al., 2011), which is also corroborated by the percentages of mass loss in each treatment from the temperature at 150 °C, excluding the initial mass loss due to the loss of adsorbed water molecules (Table S1 in Appendix C). Considering the different mixtures, the mass loss percentages from 150 °C were 0.7-3.5% for T1, 11.0-17.8% for T2, and 15.0-19.1% for T3. Also, independently of the starting mixture, the compounds resulting from T1 were completely dissolved in water. However, as observed for Ara_3 (but not for Man_3), the compounds resulting from T2 treatment of the A75M25 mixture and T3 treatment of all mixtures were only partially dissolved in water. The percentage of water-soluble material was 65.2% for T2 treatment of the A75M25 mixture. For materials resulting from T3, the percentages of water-soluble material were 68.8% for A25M75, 40.7% for A50M50, and 44.4% for A75M25 (Table S1). These values are higher than those obtained with materials resulting from T2 and T3 treatments of Ara_3 (35.3 and 24.6%, respectively) (Section III.1.1), which shows the formation of a higher amount of hydrophobic compounds when the amount of Ara_3 is higher.

In summary, the observed differences in the coloration and water-solubility of the thermally treated mixtures suggest that the structural modifications occurred to higher extent with increasing molar proportion of Ara_3 in the starting mixture and the time at 200 °C. This is in agreement with the higher diversity of ions observed in the MALDI-MS spectra of the oligosaccharide mixtures subjected to the longer treatments (T2 and T3) (Section III.2.1). To obtain a deeper insight into the structural modifications induced by dry thermal processing, both untreated samples and water-soluble fractions recovered from thermally treated samples were analysed by ESI-MS and ESI-MS^n.

Identification of hybrid and non-hybrid compounds upon dry heating

In preliminary testing, we have found that, under ESI-MS conditions, neutral oligosaccharides ionize better in positive than in negative ion mode. For this reason, positive ion mode was preferred for ESI-MS analysis of both untreated and thermally treated mixtures.

As typical of neutral oligosaccharides (Moreira et al., 2011; Reis et al., 2004a), the oligosaccharides in the starting mixtures (Man_3 and Ara_3), and also the roasting-induced compounds, were mainly detected as sodium adduct ions ($[M+Na]^+$). Accordingly, the most abundant ions in the ESI-MS spectra of the untreated mixtures (A75M25 in Figure III.15A) were observed at *m/z* 437 ($[Ara_3+Na]^+$) and 527 ($[Man_3+Na]^+$).

After thermal processing, new ions, not observed in the ESI-MS spectra of the untreated mixtures, were identified (Figure III.15B-D). Independent of the starting mixture (A25M75, A50M50, or A75M25), the diversity of ions was higher in the ESI-MS spectra obtained after the longer treatments (T2 and T3). Due to the complexity of these ESI-MS spectra, only the ions observed with a relative abundance ≥ 15% in at least two ESI-MS spectra acquired from the thermally treated mixtures were considered (Table III.9). The assignment of these ions was supported on the basis of their fragmentation pattern under ESI-CID-MS^n conditions, as will be later described, and corroborated by the elemental composition obtained from high resolution and high mass accuracy measurements using a hybrid quadrupole-Orbitrap mass spectrometer (data relating to the A50M50 mixture submitted to the T3 treatment in Table S2 in Appendix C). Accordingly, the ions identified in the ESI-MS spectra of the thermally treated mixtures were attributed to $[M+Na]^+$ ions of non-hybrid compounds composed by hexose (Hex) or pentose (Pent) residues and derivatives, and hybrid compounds composed by both Hex and Pent residues and derivatives (Table III.9). As supported by sugar and glycosidic linkage analyses (Section III.2.1), Pent and Hex were mainly arabinose (Ara) and mannose (Man), respectively. However, new sugar residues, although in minor amounts, were formed during thermal processing, namely ribose (Rib), xylose (Xyl), and lyxose (Lyx) that are isomers of Ara, and thus they are not distinguishable by MS. Therefore, in Table III.9, Pent represents mainly Ara, but also Rib, Xyl, and Lyx, which are present in minor amounts. Similarly, Hex represents mainly Man, but also isomeric sugars (glucose and galactose) that are present in minor amounts.

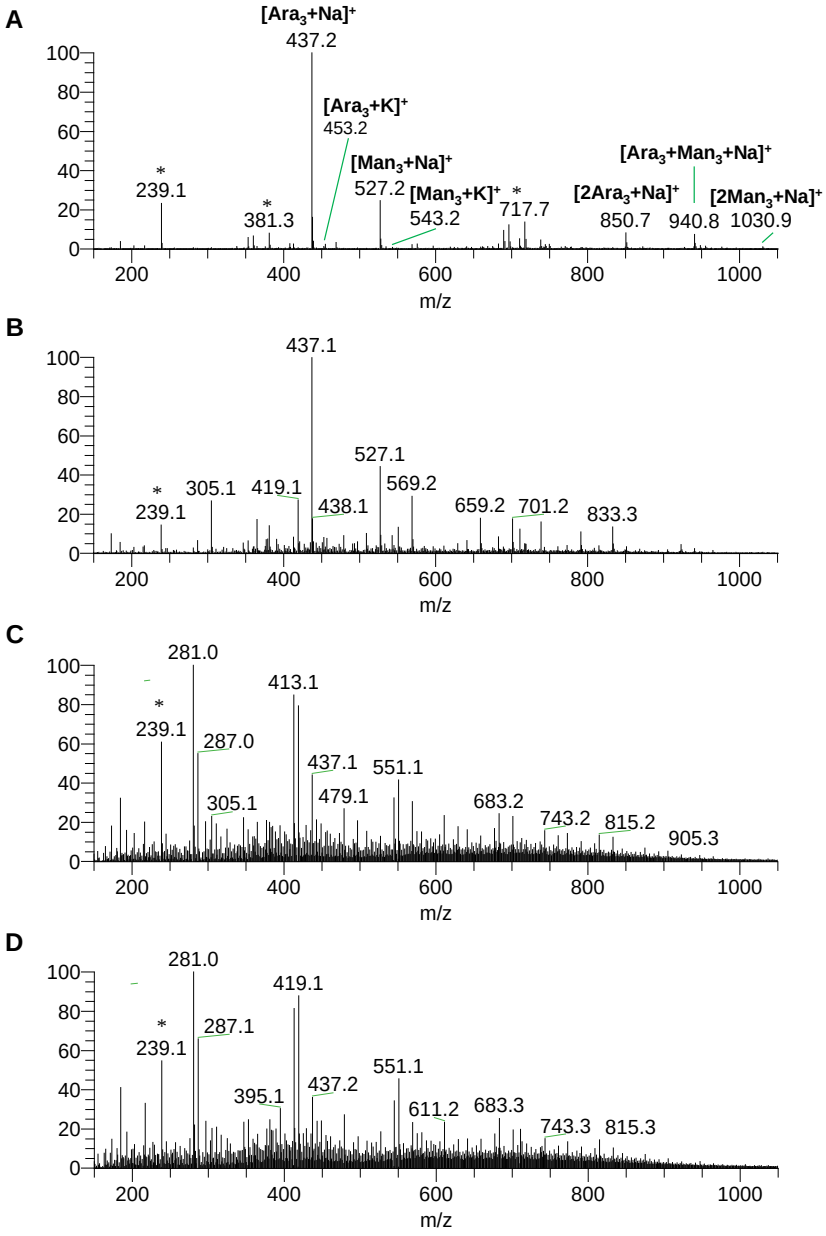

Figure III.15. Positive ion ESI-LIT-MS spectra acquired from the A75M25 mixture, (A) before (T0) and after thermal treatments: (B) T1, (C) T2, and (D) T3. Ions marked with an asterisk (*) are attributed to impurities.

As observed for Ara_3 (Section III.1.1), the non-hybrid compounds composed by Pent residues and derivatives identified after thermal treatment of the oligosaccharide mixtures include Pent oligosaccharides ($Pent_n$) with lower (n=2) and higher (n=4-5) degrees of polymerization than that of the Pent oligosaccharide in the starting mixtures (Ara_3). The respective monosaccharide was also observed at *m/z* 173 ($[Pent+Na]^+$). Also, $Pent_n$ (n=2-6) derivatives resulting from the formation of a keto group (-2 Da), and dehydration with loss of one, two and three water molecules, as well as resulting from the

oxidative scission of a furanose ring with loss of $C_2H_4O_2$, $C_3H_4O_2$, $C_2H_6O_3$, $C_3H_4O_3$, $C_3H_6O_3$, and $C_2H_8O_5$ were identified.

Table III.9. Summary of the $[M+Na]^+$ ions identified in the ESI-LIT-MS spectra acquired from the thermally treated mixtures.

Proposed assignments		Number (*n*) of pentose (Pent) units					
		1	2	3	4	5	6
Non-hybrid compounds composed by pentose residues and derivatives							
$[Pent_n+Na]^+$		173	305	†437[b]	†569	†701	
$[Pent_n-2H+Na]^+$	(-2)[a]			435			
$[Pent_n-H_2O+Na]^+$	(-18)		287	†419	†551	†683	
$[Pent_n-2H_2O+Na]^+$	(-36)			401			
$[Pent_n-3H_2O+Na]^+$	(-54)			383			
$[Pent_n-C_2H_4O_2+Na]^+$	(-60)			377	†509	†641	
$[Pent_n-C_3H_4O_2+Na]^+$	(-72)			†365	†497	†629	
$[Pent_n-C_2H_6O_3+Na]^+$	(-78)		227				
$[Pent_n-C_3H_4O_3+Na]^+$	(-88)		217				
$[Pent_n-C_3H_6O_3+Na]^+$	(-90)			†347	†479	†611	‡743
$[Pent_n-C_2H_8O_5+Na]^+$	(-112)		193	325			
Non-hybrid compounds composed by hexose residues and derivatives							
$[Hex_n+Na]^+$		203	†365	†527			
$[Hex_n-H_2O+Na]^+$	(-18)	185	†347	†509			
$[Hex_n-3H_2O+Na]^+$	(-54)	149	311	473			
$[Hex_n-4H_2O+Na]^+$	(-72)		293	455			
$[Hex_n-5H_2O+Na]^+$	(-90)			†437			
$[Hex_n-6H_2O+Na]^+$	(-108)		257	†419	†581	‡743	
Pentose-hexose hybrid oligosaccharides and derivatives							
$[PentHex_n+Na]^+$			†497	†659			
$[PentHex_n-H_2O+Na]^+$	(-18)	317	†479	†641			
$[PentHex_n-3H_2O+Na]^+$	(-54)	281	443	605			
$[PentHex_n-4H_2O+Na]^+$	(-72)		425	587			
$[PentHex_n-5H_2O+Na]^+$	(-90)			†569			
$[PentHex_n-6H_2O+Na]^+$	(-108)		389	†551	†713		
$[Pent_2Hex_n+Na]^+$			†629				
$[Pent_2Hex_n-H_2O+Na]^+$	(-18)	449	†611				
$[Pent_2Hex_n-3H_2O+Na]^+$	(-54)	413	575				
$[Pent_2Hex_n-4H_2O+Na]^+$	(-72)	395					
$[Pent_2Hex_n-5H_2O+Na]^+$	(-90)			†701			
$[Pent_2Hex_n-6H_2O+Na]^+$	(-108)		521	†683			
$[Pent_3Hex_n-H_2O+Na]^+$	(-18)	†581	‡743				
$[Pent_3Hex_n-3H_2O+Na]^+$	(-54)	545					
$[Pent_3Hex_n-4H_2O+Na]^+$	(-72)	†527					
$[Pent_4Hex_n-H_2O+Na]^+$	(-18)	†713					
$[Pent_4Hex_n-3H_2O+Na]^+$	(-54)	677					
$[Pent_4Hex_n-4H_2O+Na]^+$	(-72)	†659					

Only the ions observed with a relative abundance ≥ 15% in at least two of the ESI-LIT-MS spectra were considered. [a]*m/z* values. [b]Ions attributed to different compounds having the same nominal mass are marked with a symbol: † for two and ‡ for three compounds. [c]Values in parentheses are the *m/z* value differences compared to the $[M+Na]^+$ ions of the corresponding non-modified oligosaccharide.

As observed for Man_3 (Moreira et al., 2011), the non-hybrid compounds composed by Hex residues and derivatives identified after thermal treatment of the model mixtures include Hex mono- and oligosaccharides (Hex_n, n=1-3). The corresponding dehydrated derivatives formed by loss of one and from three to six water molecules were also observed ($[Hex_n\text{-}xH_2O+Na]^+$; n=1-5; x=1, 3-6). Revisiting the ESI-MS spectra previously acquired from Man_3 subjected to T2 or T3 treatment, beyond the derivatives resulting from the loss of one and three water molecules previously reported (Moreira et al., 2011), it was possible observing the ions attributed to dehydrated Hex oligosaccharides resulting from the loss of four, five, and six water molecules.

In agreement to what was previously observed by MALDI-MS (Section III.2.1), hybrid oligosaccharides composed by both Hex and Pent units were also identified as $[M+Na]^+$ ions in the ESI-MS spectra of the thermally treated mixtures, namely $PentHex_2$ (*m/z* 497), $PentHex_3$ (*m/z* 659) and $Pent_2Hex_2$ (*m/z* 629). The observation of these hybrids corroborates the hypothesis of the occurrence of non-enzymatic transglycosylation reactions involving galactomannans and arabinose side chains of arabinogalactans during coffee roasting. Also, several ions (Table III.9) were attributed to dehydrated derivatives formed by loss of one and three up to six water molecules from hybrid oligosaccharides ($[Pent_mHex_n\text{-}xH_2O+Na]^+$; m, n=1-4; x=1, 3-6).

As highlighted in Table III.9, among the compounds identified in the ESI-MS spectra of the thermally treated mixtures as $[M+Na]^+$ ions, there are compounds with the same nominal mass (calculated by adding the mass of the predominant isotope of each element contributing to the molecule rounded to the nearest integer value). Some of these compounds are isobaric compounds, having the same nominal mass but different elemental composition, and thus, different accurate mass (Table S2 in Appendix C); other are isomers, having the same elemental composition, and thus, same exact mass. Isobaric and isomeric compounds were differentiated on the basis of specific fragmentation seen in the ESI-CID-MS^n spectra, as will be detailed and exemplified in the following sections.

Diagnostic neutral losses and product ions observed under ESI-CID-MS^n conditions

The fragmentation of $[M+Na]^+$ ions of neutral and reducing oligosaccharides under ESI-CID-MS^n conditions results from glycosidic linkage cleavages, cross-ring cleavages (cleavage of two bonds within a sugar ring), and loss of water. As inferred from ^{18}O-

labelling of the anomeric oxygen of standard oligosaccharides (da Costa et al., 2012; Hofmeister et al., 1991), including Ara_3 (Section III.1.1), glycosidic cleavages mainly occur between the anomeric carbon and the glycosidic linkage oxygen, whereas cross-ring cleavages and loss of water occur mainly at the reducing end residue with loss of the anomeric oxygen. The cross-ring cleavages at the reducing end residue depend on the oligosaccharide structure, giving rise to neutral losses of CH_2O (-30 Da), $C_2H_4O_2$ (-60 Da), $C_3H_6O_3$ (-90 Da), and $C_4H_8O_4$ (-120 Da). On the basis of the knowledge not only of the typical fragmentation pathways of non-modified oligosaccharides, but also of derivatives formed when Man_3 (Moreira et al., 2011) and Ara_3 (Section III.1.1) were individually submitted to the thermal treatments at 200 °C, it was possible identify neutral losses and product ions diagnostic of different ion series observed in the ESI-MS spectra of the thermally treated mixtures, which are summarized in Table III.10.

Table III.10. Diagnostic neutral losses and product ions observed under ESI-CID-MS^n conditions.

Neutral losses and product ions	Assignment(s)
Neutral loss (Da)	
96	$(Pent\text{-}2H_2O)_{res}$; $(Pent\text{-}3H_2O)$
114	$(Pent\text{-}H_2O)_{res}$; $(Pent\text{-}2H_2O)$
126	$(Hex\text{-}3H_2O)$; $(Hex\text{-}2H_2O)_{res}$
132	$Pent_{res}$; $(Pent\text{-}H_2O)$
144	$(Hex\text{-}H_2O)_{res}$; $(Hex\text{-}2H_2O)$
148	(Pent-2H)
150	Pent
162	Hex_{res}; $(Hex\text{-}H_2O)$; Combined loss of $Pent_{res}$ and CH_2O (30 Da) from cross-ring cleavage
180	Hex
294	Combined loss of $Pent_{res}$ and Hex_{res}
Product ion (*m/z*)	
149	$[Hex\text{-}3H_2O+Na]^+$
155	$[Pent_{res}+Na]^+$; $[Pent\text{-}H_2O+Na]^+$
167	$[(Hex\text{-}H_2O)_{res}+Na]^+$; $[Hex\text{-}2H_2O+Na]^+$
171	$[Pent\text{-}2H+Na]^+$
173	$[Pent+Na]^+$
185	$[Hex_{res}+Na]^+$; $[Hex\text{-}H_2O+Na]^+$
203	$[Hex+Na]^+$
335	$[PentHex+Na]^+$

The ESI-MS^n fragmentation of $[M+Na]^+$ ions of Hex oligosaccharides produces product ions resulting from glycosidic cleavages with neutral losses of a Hex residue (Hex_{res}) (162 Da) and a Hex (180 Da), and diagnostic product ions at *m/z* 203 ($[Hex+Na]^+$) and 185 ($[Hex_{res}+Na]^+$) (Figure III.16A). Similarly, the ESI-MS^n spectra of Pent oligosaccharides show product ions resulting from glycosidic cleavages with loss a Pent residue ($Pent_{res}$) (132 Da) and Pent (150 Da), and product ions at *m/z* 173 ($[Pent+Na]^+$) and 155 ($[Pent_{res}+Na]^+$) (Figure III.17A). Product ions resulting from the neutral loss of 162 Da observed with relative abundance ≤10% in the ESI-MS^n spectra of Pent oligosaccharides, as $Pent_5$ and derivatives (Table III.11), are not due to the loss of Hex_{res} but to the combined loss of $Pent_{res}$ and CH_2O (30 Da) from cross-ring cleavage. Neutral losses and product ions observed in the ESI-MS^n spectra of non-hybrid oligosaccharides, composed by Pent or Hex residues, are also observed in the ESI-MS^n spectra of hybrid oligosaccharides, composed by both Pent and Hex residues. However, the neutral loss of 294 Da, due to the combined loss of $Pent_{res}$ and Hex_{res}, is specific for Pent-Hex hybrid oligosaccharides, as well as the product ion at *m/z* 335 ($[PentHex+Na]^+$) (Figure III.18A).

When sugar residues are modified, they yield new diagnostic neutral losses and product ions, which can be specific or not of one type of modification. The ESI-MS^n fragmentation of an oligosaccharide bearing a (Hex-$3H_2O$) produces a product ion resulting from the neutral loss of 126 Da, corresponding to the loss of (Hex-$3H_2O$), and a product ion at *m/z* 149 ($[Hex-3H_2O+Na]^+$). The loss of 144 Da is observed in the ESI-MS^n spectra of either (Hex-H_2O) derivatives, resulting from the loss of $(Hex-H_2O)_{res}$, or (Hex-$2H_2O$) derivatives, resulting from the loss of (Hex-$2H_2O$). The product ion at *m/z* 167 is also observed in the ESI-MS^n spectra of both (Hex-H_2O) and (Hex-$2H_2O$) derivatives, corresponding to $[(Hex-H_2O)_{res}+Na]^+$ and $[(Hex-2H_2O+Na]^+$, respectively. In the case of (Hex-H_2O) derivatives, the product ion at *m/z* 185 can also be due to $[Hex-H_2O+Na]^+$, and not exclusively to $[Hex_{res}+Na]^+$. Similarly, the neutral loss of 114 Da is observed either in the ESI-MS^n spectra of (Pent-H_2O) derivatives, resulting from the loss of $(Pent-H_2O)_{res}$, or in the ESI-MS^n spectra of (Pent-$2H_2O$) derivatives, resulting from the loss of (Pent-$2H_2O$). In the case of (Pent-H_2O) derivatives, the product ion at *m/z* 155 can also be due to $[Pent-H_2O+Na]^+$, and exclusively to $[Pent_{res}+Na]^+$. The neutral loss of 132 and 162 Da can also be due to the presence a mono-dehydrated sugar, (Pent-H_2O) and (Hex-H_2O), respectively. Also, the neutral loss of 96 Da can be due either to the loss of $(Pent-2H_2O)_{res}$, or (Pent-$3H_2O$). In the case of oligosaccharides bearing a (Pent-2H), their fragmentation produces a product ion resulting from neutral loss of 148 Da, due to the

loss of (Pent-2H), and the product ion at *m/z* 171 ([Pent-2H+Na]$^+$). The identification of neutral losses and product ions diagnostic of different ion series (Table III.10) was essential to disclose the presence of isobaric and isomeric compounds.

Table III.11. Summary of the pairs of isobaric/isomeric compounds and one set of three compounds having the same nominal mass identified as [M+Na]$^+$ ions in the ESI-MS spectra of the thermally treated mixtures.

Ion (*m/z*)	Proposed assignments	Formula	Characteristic neutral losses			
			132 Da	162 Da	144 Da	126 Da
347	[Pent$_3$-C$_3$H$_6$O$_3$+Na]$^+$ [Hex$_2$-H$_2$O+Na]$^+$	C$_{12}$H$_{20}$NaO$_{10}$	Pent$_{res}$	Hex$_{res}$/(Hex-H$_2$O)	(Hex-H$_2$O)$_{res}$	
365	[Pent$_3$-C$_3$H$_4$O$_2$+Na]$^+$ [Hex$_2$+Na]$^+$	C$_{12}$H$_{22}$NaO$_{11}$	Pent$_{res}$	(Pent$_{res}$+CH$_2$O)^a Hex$_{res}$		
419	[Pent$_3$-H$_2$O+Na]$^+$ [Hex$_3$-6H$_2$O+Na]$^+$	C$_{15}$H$_{24}$NaO$_{12}$ C$_{18}$H$_{20}$NaO$_{10}$	Pent$_{res}$/(Pent-H$_2$O)	Hex$_{res}$		(Hex-3H$_2$O)
437	[Pent$_3$+Na]$^+$ [Hex$_3$-5H$_2$O+Na]$^+$	C$_{15}$H$_{26}$NaO$_{13}$ C$_{18}$H$_{22}$NaO$_{11}$	Pent$_{res}$	Hex$_{res}$	(Hex-2H$_2$O)	(Hex-3H$_2$O)/ (Hex-2H$_2$O)$_{res}$
479	[Pent$_4$-C$_3$H$_6$O$_3$+Na]$^+$ [PentHex$_2$-H$_2$O+Na]$^+$	C$_{17}$H$_{28}$NaO$_{14}$	Pent$_{res}$ Pent$_{res}$	(Pent$_{res}$+CH$_2$O)^a Hex$_{res}$/(Hex-H$_2$O)	(Hex-H$_2$O)$_{res}$	
497	[Pent$_4$-C$_3$H$_4$O$_2$+Na]$^+$ [PentHex$_2$+Na]$^+$	C$_{17}$H$_{30}$NaO$_{15}$	Pent$_{res}$ Pent$_{res}$	(Pent$_{res}$+CH$_2$O)^a Hex$_{res}$		
509	[Pent$_4$-C$_2$H$_4$O$_2$+Na]$^+$ [Hex$_3$-H$_2$O+Na]$^+$	C$_{18}$H$_{30}$NaO$_{15}$	Pent$_{res}$	(Pent$_{res}$+CH$_2$O)^a Hex$_{res}$/(Hex-H$_2$O)	(Hex-H$_2$O)$_{res}$	
527	[Hex$_3$+Na]$^+$ [Pent$_3$Hex-4H$_2$O+Na]$^+$	C$_{18}$H$_{32}$NaO$_{16}$ C$_{21}$H$_{28}$NaO$_{14}$	Pent$_{res}$	Hex$_{res}$		(Hex-3H$_2$O)
551	[Pent$_4$-H$_2$O+Na]$^+$ [PentHex$_3$-6H$_2$O+Na]$^+$	C$_{20}$H$_{32}$NaO$_{16}$ C$_{23}$H$_{28}$NaO$_{14}$	Pent$_{res}$/(Pent-H$_2$O) Pent$_{res}$	Hex$_{res}$		(Hex-3H$_2$O)
569	[Pent$_4$+Na]$^+$ [PentHex$_3$-5H$_2$O+Na]$^+$	C$_{20}$H$_{34}$NaO$_{17}$ C$_{23}$H$_{30}$NaO$_{15}$	Pent$_{res}$ Pent$_{res}$	Hex$_{res}$	(Hex-2H$_2$O)	(Hex-3H$_2$O)/ (Hex-2H$_2$O)$_{res}$
581	[Hex$_4$-6H$_2$O+Na]$^+$ [Pent$_3$Hex-H$_2$O+Na]$^+$	C$_{24}$H$_{30}$NaO$_{15}$ C$_{21}$H$_{34}$NaO$_{17}$	Pent$_{res}$	Hex$_{res}$ (Hex-H$_2$O)	(Hex-H$_2$O)$_{res}$	(Hex-3H$_2$O)
611	[Pent$_5$-C$_3$H$_6$O$_3$+Na]$^+$ [Pent$_2$Hex$_2$-H$_2$O+Na]$^+$	C$_{22}$H$_{36}$NaO$_{18}$	Pent$_{res}$ Pent$_{res}$	(Pent$_{res}$+CH$_2$O)^a Hex$_{res}$/(Hex-H$_2$O)	(Hex-H$_2$O)$_{res}$	
629	[Pent$_5$-C$_3$H$_4$O$_2$+Na]$^+$ [Pent$_2$Hex$_2$+Na]$^+$	C$_{22}$H$_{38}$NaO$_{19}$	Pent$_{res}$ Pent$_{res}$	(Pent$_{res}$+CH$_2$O)^a Hex$_{res}$		
641	[Pent$_5$-C$_2$H$_4$O$_2$+Na]$^+$ [PentHex$_3$-H$_2$O+Na]$^+$	C$_{23}$H$_{38}$NaO$_{19}$	Pent$_{res}$ Pent$_{res}$	(Pent$_{res}$+CH$_2$O)^a Hex$_{res}$/(Hex-H$_2$O)	(Hex-H$_2$O)$_{res}$	
659	[PentHex$_3$+Na]$^+$ [Pent$_4$Hex-4H$_2$O+Na]$^+$	C$_{23}$H$_{40}$NaO$_{20}$ C$_{26}$H$_{36}$NaO$_{18}$	Pent$_{res}$ Pent$_{res}$	Hex$_{res}$		(Hex-3H$_2$O)
683	[Pent$_5$-H$_2$O+Na]$^+$ [Pent$_2$Hex$_3$-6H$_2$O+Na]$^+$	C$_{25}$H$_{40}$NaO$_{20}$ C$_{28}$H$_{36}$NaO$_{18}$	Pent$_{res}$/(Pent-H$_2$O) Pent$_{res}$	Hex$_{res}$		(Hex-3H$_2$O)
701	[Pent$_5$+Na]$^+$ [Pent$_2$Hex$_3$-5H$_2$O+Na]$^+$	C$_{25}$H$_{42}$NaO$_{21}$ C$_{28}$H$_{38}$NaO$_{19}$	Pent$_{res}$ Pent$_{res}$	(Pent$_{res}$+CH$_2$O)^a Hex$_{res}$	(Hex-2H$_2$O)	(Hex-3H$_2$O)/ (Hex-2H$_2$O)$_{res}$
713	[PentHex$_4$-6H$_2$O+Na]$^+$ [Pent$_4$Hex-H$_2$O+Na]$^+$	C$_{29}$H$_{38}$NaO$_{19}$ C$_{26}$H$_{42}$NaO$_{21}$	Pent$_{res}$ Pent$_{res}$	Hex$_{res}$ (Hex-H$_2$O)	(Hex-H$_2$O)$_{res}$	(Hex-3H$_2$O)
743	[Hex$_5$-6H$_2$O+Na]$^+$ [Pent$_6$-C$_3$H$_6$O$_3$+Na]$^+$ [Pent$_3$Hex$_2$-H$_2$O+Na]$^+$	C$_{30}$H$_{40}$NaO$_{20}$ C$_{27}$H$_{44}$NaO$_{22}$	Pent$_{res}$ Pent$_{res}$	Hex$_{res}$ Hex$_{res}$/(Hex-H$_2$O)	(Hex-H$_2$O)$_{res}$	(Hex-3H$_2$O)

[a]The corresponding product ion was observed with a relative abundance ≤10%.

Differentiation of isobaric and isomeric compounds by ESI-CID-MS^n

The presence of isobaric/isomeric compounds was inferred from the observation of specific product ions in the ESI-MS^2 spectra acquired on the LIT mass spectrometer (Table III.11). In most of the cases, the presence of isobaric/isomeric compounds was confirmed by ESI-MS^n, *n*=3-4. As an example, it is shown the tandem MS-based strategy to discriminate between the pairs Hex_3/($Pent_3Hex$-$4H_2O$) (*m/z* 527) and $Pent_3$/(Hex_3-$5H_2O$) (*m/z* 437).

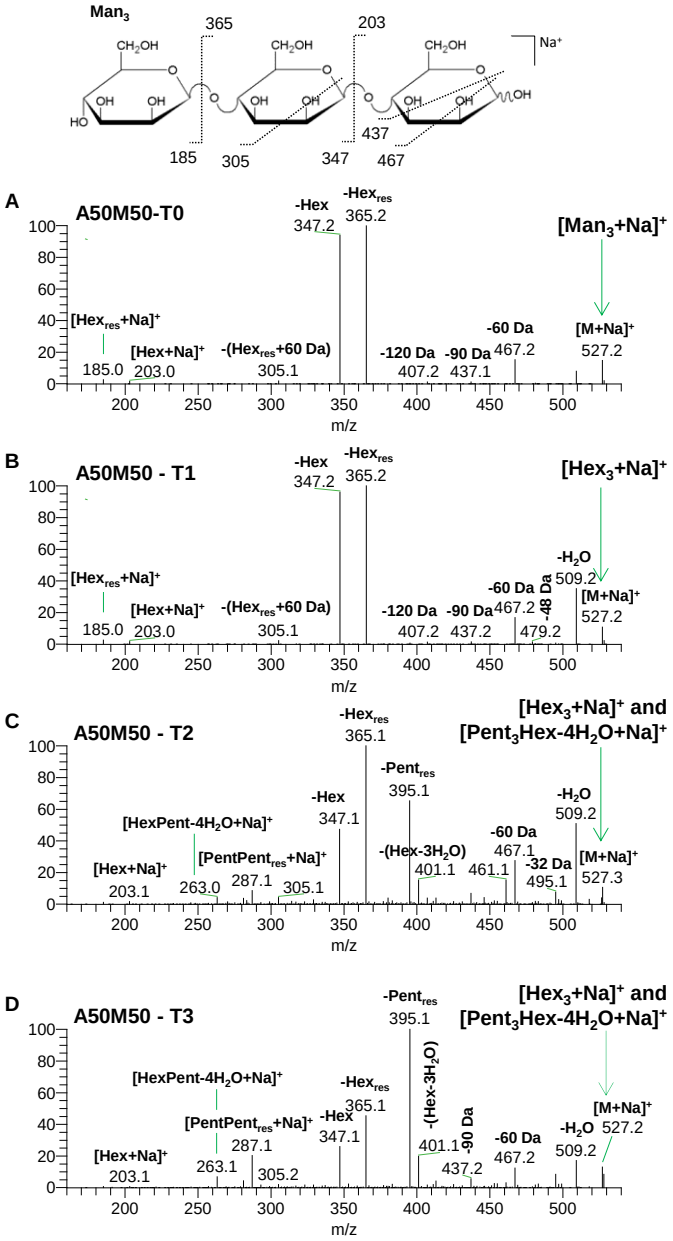

Figure III.16. ESI-MS^2 spectra of the ion at *m/z* 527 acquired from the (A) untreated A50M50 mixture (T0), and from samples of the A50M50 mixture subjected to the (B) T1, (C) T2, and (D) T3 treatments.

In the MS analysis of carbohydrate-rich samples, the ion (or product ion) at *m/z* 527 is usually indicative of Hex_3. The ESI-MS^2 spectrum acquired from the A50M50 mixture submitted to the T1 treatment (Figure III.16B) is similar to that acquired from the untreated mixture (Figure III.16A), suggesting the exclusive presence of the precursor ion $[Hex_3+Na]^+$, as corroborated by high resolution and high mass accuracy MS data. However, new product ions, absent in the ESI-MS^2 spectrum acquired after T1 (Figure III.16B), namely the product ions at *m/z* 401 (-126 Da) and 395 (-132 Da), identified as resulting from the loss of (Hex-$3H_2O$) and $Pent_{res}$, and the product ions at *m/z* 287 and 263, attributed to $[PentPent_{res}+Na]^+$ and $[HexPent-4H_2O+Na]^+$, respectively, were observed in the ESI-MS^2 spectra acquired after the longer treatments, T2 (C) and T3 (D). These new ions suggest the presence of the precursor ion $[Pent_3Hex-4H_2O+Na]^+$ beyond $[Hex_3+Na]^+$, as corroborated by high resolution and high mass accuracy MS data. The relative abundance of the new product ions at *m/z* 401 and 395 increases with increasing treatment time, whereas that at *m/z* 365 (-162 Da, $-Hex_{res}$) decreases. These changes suggest the increase of the proportion of $[Pent_3Hex-4H_2O+Na]^+$ and the decrease of $[Hex_3+Na]^+$ ions with increasing treatment time.

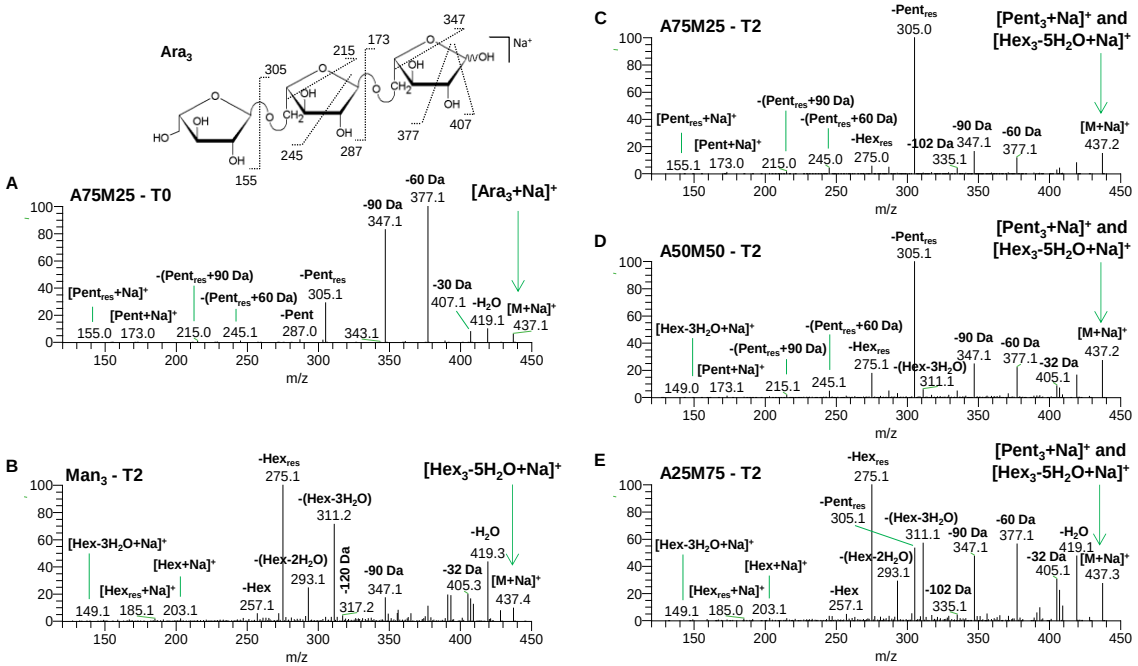

Figure III.17. ESI-MS^2 spectra of the ion at *m/z* 437 acquired from the (A) untreated A75M25 mixture and (B) the Man_3 sample subjected to the T2 treatment, and from mixtures subjected to the T2 treatment: (C) A75M25, (D) A50M50, and (E) A25M75.

Panels C, D, and E of Figure III.17 show the ESI-MS^2 spectra of the ion at *m/z* 437 acquired from each sample of the A75M25, A50M50, and A25M75 mixtures, respectively, subjected to the T2 treatment. The product ions at *m/z* 305 (-$Pent_{res}$) and 173 (-2$Pent_{res}$), also observed in the ESI-MS^2 spectrum of the Ara_3 (Figure III.17A), corroborate the presence of $[Pent_3+Na]^+$ as precursor ion. The coexistence of another precursor ion, $[Hex_3-5H_2O+Na]^+$, is supported by the product ions at *m/z* 311 (-126 Da), 293 (-144 Da), and 275 (-162 Da), formed respectively by loss of (Hex-$3H_2O$), (Hex-$2H_2O$), and Hex_{res}, as well as those at *m/z* 203 ($[Hex+Na]^+$), 185 ($[Hex_{res}+Na]^+$), and 149 ($[Hex-3H_2O+Na]^+$). These product ions were also observed in the ESI-MS^2 spectrum of (Hex_3-5H_2O) formed when Man_3 was individually subjected to the treatment T2 (Figure III.17B) (Moreira et al., 2011). The relative abundance of the product ions at *m/z* 311, 293, and 275 suggests the increase of the proportion of $[Hex_3-5H_2O+Na]^+$ ions and the decrease of $[Pent_3+Na]^+$ ions when the proportion of Man_3 increases (Figure III.17C-E).

As illustrated by the ESI-MS^2 spectra of the ions at *m/z* 527 and 437 in Figures III.16 and III.17, the formation of each isobaric compound, and also of each isomeric compound, was dependent on the thermal treatment and the starting mixture. This suggests that a variety of structures may be formed during the green coffee roasting, depending on the roasting conditions and distribution of the polysaccharides in the coffee beans.

ESI-MS^n analysis after labelling with oxygen-18

Aiming to gain more information about the structure of the compounds formed during thermal processing, thermally treated mixtures were dissolved in ^{18}O-enriched water before MS analysis. On the basis of the principles of nucleophilic addition reactions of aldehydes and ketones, they react with water to yield geminal diols (hydrates). The hydration reaction is reversible, and a geminal diol can eliminate water to regenerate an aldehyde or ketone (McMurry, 2008). Accordingly, the ^{18}O-labelling of not only the anomeric oxygen of non-modified reducing sugar residues, but also new carbonyl groups, either of ketones or aldehydes, formed by dry thermal processing, was observed.

ESI-MS^n of ^{18}O-labelled Pent-Hex hybrid oligosaccharides

The ESI-MS^2 spectrum of the ion at *m/z* 659 acquired from the A50M50 mixture subjected to the T1 treatment, and that acquired after ^{18}O-labelling are shown in Figures

III.18A and B. In accordance with what was previously presented, this species could have the contribution of two isobaric compounds, $PentHex_3$ and ($Pent_4Hex$-$4H_2O$). As corroborated by the ESI-MS^2 spectrum before ^{18}O-labelling (Figure III.18A), namely by the absence of the product ion at *m/z* 533, resulting from the neutral loss of (Hex-$3H_2O$), only $PentHex_3$ was present.

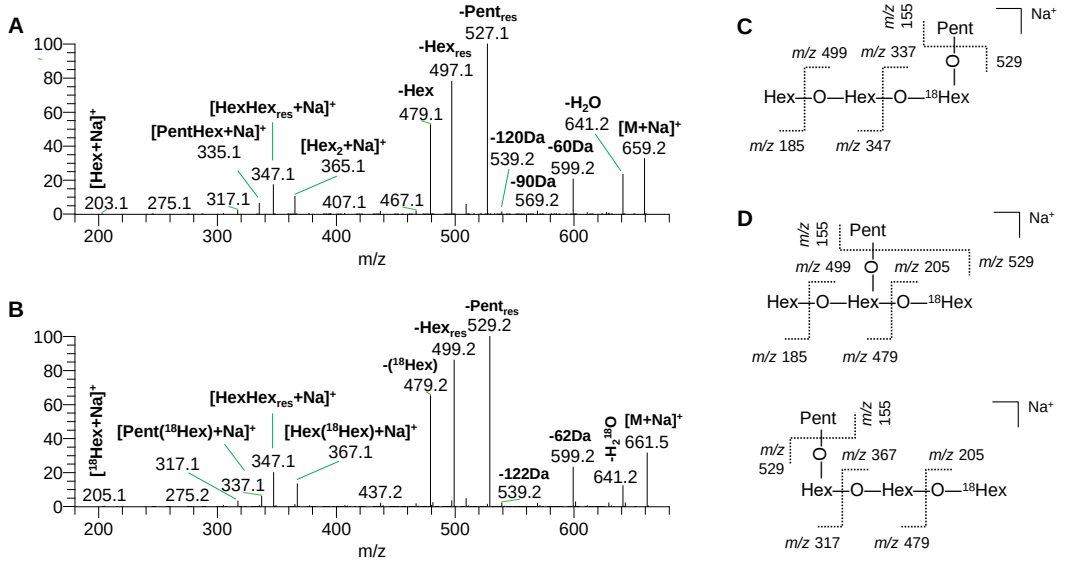

Figure III.18. ESI-MS^2 spectrum of the ion at *m/z* 659 ([Hex_3Pent+$Na]^+$) acquired from the A50M50 mixture subjected to the (A) T1 treatment, and (B) the corresponding ESI-MS^2 spectrum acquired after labelling with oxygen-18. The different proposed structures are represented in panels (C) and (D).

The ESI-MS^2 spectrum of ^{18}O-labelled $PentHex_3$ (Figure III.18B) shows the product ion at *m/z* 529, resulting from the loss of unlabelled $Pent_{res}$, and the product ion at *m/z* 205, attributed to $[M+Na]^+$ of a ^{18}O-labelled Hex. These product ions indicate that one of the three Hex units, and not the Pent, was ^{18}O-labelled. This means that a Hex is located at the reducing end of $PentHex_3$. The observation of the product ion at *m/z* 529 also indicates that the three Hex units are linked together, having several possible binding sites for the Pent. The product ion at *m/z* 337 (-$2Hex_{res}$) suggests that the Pent is linked to the Hex located at the reducing end (Figure III.18C). However, the product ion at *m/z* 479, formed by loss of the ^{18}O-labelled Hex, supports the presence of other structures with the Pent not linked to the Hex located at the reducing end (Figure III.18D).

In any of the possible structures, it is of note that the Pent-Hex glycosidic linkage involves the anomeric carbon of the Pent. The formation of the Pent-Hex glycosidic linkage involving the anomeric carbon of the Pent, and not of the Hex, can be favoured

by the higher reactivity of pentoses compared to hexoses. In the case of the thermally treated mixtures, the most abundant pentoses and hexoses are Ara and Man, respectively. As also previously observed by glycosidic linkage analysis (Section III.2.1), new types of Man glycosidic linkages, absent in the untreated mixtures, namely (1→2), (1→6), (1→2,3), (1→2,6), (1→3,6), and (1→2,3,6) linkages, were formed during thermal processing of the Ara_3-Man_3 mixtures. Accordingly, these new Man linkages were identified when a coffee galactomannan-rich fraction was submitted to different dry thermal treatments (Simões et al., 2014). In both cases, the formation of (1→6)-Man linkages was favoured over that of other types of Man linkages. As reported for British gums produced by dry heating of starch (Tomasik, Wiejak & Pałasiński, 1989), the formation of (1→6) linkages can be favoured over others in terms of stereochemical and thermodynamic aspects. This can be due to the primary hydroxyl group located at the C6-position, whereas the other hydroxyl groups of Man are secondary. According to Tomasik et al. (1989), anhydrosugars are possible intermediates in the formation of new glycosidic linkages, but the mechanisms involved in the non-enzymatic transglycosylation reactions induced by dry heating are far to be elucidated.

ESI-MS^n of ^{18}O-labelled oligosaccharides containing a tri-dehydrated hexose

Both non-hybrid and hybrid oligosaccharides containing a (Hex-$3H_2O$) were labelled by dissolving selected thermally treated mixtures in $H_2{}^{18}O$. In Figure III.19 are shown, as example, the ESI-MS^2 spectrum of the ion at *m/z* 413 ($[Pent_2Hex\text{-}3H_2O+Na]^+$) acquired from the A50M50 mixture subjected to the T2 treatment (A), and that acquired after ^{18}O-labelling (B). The product ion observed after ^{18}O-labelling (B) at *m/z* 287 (-128 Da), resulting from the loss of ^{18}O-labelled (Hex-$3H_2O$), and that observed at *m/z* 151, formed by loss of two unlabelled $Pent_{res}$, attributed to a ^{18}O-labelled (Hex-$3H_2O$), confirm the presence of ($Pent_2Hex\text{-}3H_2O$) structures with the (Hex-$3H_2O$) located at the reducing end of the corresponding non-modified oligosaccharide.

The occurrence of ^{18}O-labelling at the (Hex-$3H_2O$) suggests that this moiety has a carbonyl group. As previously demonstrated by deuterium-labelling of oligosaccharides bearing a (Hex-$3H_2O$) formed from Man_3 (Moreira et al., 2011), (Hex-$3H_2O$) may be isomaltol, which has a keto group. Due to the reversibility of the hydration reaction of ketones (McMurry, 2008), the ^{18}O-labelling of the keto group of isomaltol moieties present in the thermally treated mixtures dissolved in $H_2{}^{18}O$ could be expected. However,

because conditions used in this work for ^{18}O-labelling had previously been used only to label the carbonyl group of aldehydes, the ketone 3-octanone was used as standard, confirming the ^{18}O-labelling of a keto group under the conditions used (Figures S1 and S2 in Appendix C).

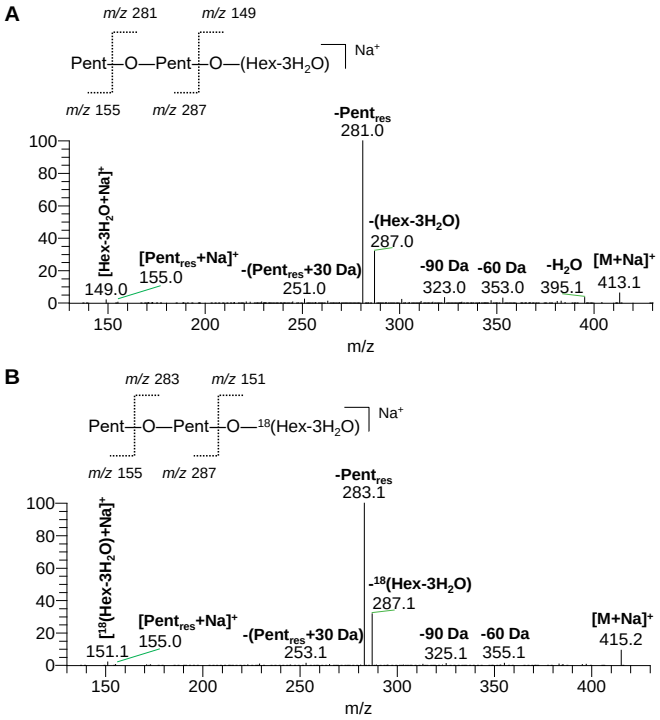

Figure III.19. (A) ESI-MS^2 spectrum of the ion at *m/z* 413 ($[Pent_2Hex\text{-}3H_2O+Na]^+$) acquired from the A50M50 mixture subjected to the T2 treatment, and (B) the corresponding ESI-MS^2 spectrum acquired after labelling with oxygen-18.

Analysis of a coffee galactomannan-rich fraction

To check if the structures formed during thermal processing of the model mixtures have the same structural features than those formed during coffee roasting, or additional roasting treatments beyond the roasting of the green coffee beans, a galactomannan-rich fraction isolated from spent coffee grounds (SCG) and submitted to two additional roasting treatments of 1 h at 200 °C was analysed. First, this galactomannan-rich fraction was treated with an *endo*-(*β*1→4)-D-mannanase to selectively cleave the (*β*1→4)-D-mannan backbone between adjacent (β1→4)-linked Man residues. According to the known mechanism of action of this enzyme, the hydrolysis of the (*β*1→4)-D-mannan backbone is hindered by the presence of substituted Man residues, non-Man residues interspersed in the mannan backbone, modified Man residues, or other Man glycosidic linkages that not (β1→4), yielding oligosaccharides that contain these structural details.

Thus, the hydrolysed material was further fractionated by LEX-SEC and then analysed by ESI-MS and ESI-MS^n.

In accord with previous studies (Tudella et al., 2011; Section III.1.2), the LEX-SEC fraction that eluted between 16 and 17 min (Figure S3 in Appendix C) was assigned to the neutral oligosaccharides with the higher molecular weight that resulted from the enzymatic hydrolysis. Due to the enzyme action, these oligosaccharides contain structural details of the coffee galactomannans under study. In the ESI-MS spectrum of this fraction (Figure S4 in Appendix C), it was possible to identify the ions at *m/z* 497, 509, 527, 641, and 659, also identified in the thermal treated mixtures, and other ions at *m/z* 671, 689, 803, 833, and 851. On the basis of their ESI-CID-MS^n fragmentation patterns and the elemental composition obtained from high resolution and high mass accuracy measurements using a hybrid quadrupole-Orbitrap mass spectrometer, the identification of these ions is as follows: 497, $[PentHex_2+Na]^+$; 509, $[Hex_3-H_2O+Na]^+$; 527, $[Hex_3+Na]^+$; 641, $[PentHex_3-H_2O+Na]^+$; 659, $[PentHex_3+Na]^+$; 671, $[Hex_4-H_2O+Na]^+$; 689, $[Hex_4+Na]^+$; 803, $[PentHex_4-H_2O+Na]^+$; 833, $[Hex_5-H_2O+Na]^+$; 851, $[Hex_5+Na]^+$.

According to the glycosidic linkage composition of the coffee galactomannan-rich fraction under study (Simões et al., 2014), the observation of Hex_{3-5} after enzymatic hydrolysis can be related to the presence of substituted Man residues, namely by single Gal residues as occur in green coffee galactomannans; non-mannose residues (Glc and Gal) interspersed in the mannan backbone; and new glycosidic linkages resistant to the enzyme action, such as (1→6) Man linkages, formed during roasting. Because single arabinose residues occur as side chains in green coffee galactomannans (Nunes et al., 2005), $PentHex_2$ and $PentHex_3$ can result from the cleavage of the original mannan backbone, without any modification promoted by roasting. Contrarily, the dehydrated derivatives are the result of dehydration reactions occurring during roasting. The ESI-MS^n fragmentation of $[M+Na]^+$ ions of ($PentHex_3-H_2O$) and ($PentHex_4-H_2O$) suggests that loss of a water molecule occurred at a Hex unit, as implied by the loss of an intact $Pent_{res}$ (132 Da) and $(Hex-H_2O)_{res}$ (144 Da). This is in accordance with that was observed with the Ara_3-Man_3 mixtures, reinforcing the validity of the models used. The absence of hybrid domains formed by non-enzymatic transglycosylation between galactomannans and arabinogalactans during roasting in the galactomannan-rich fraction analysed in this study can be related to its original location in the green coffee beans and the roasting conditions used.

III.2.2.3. Concluding remarks

The analysis of the model mixtures containing different molar proportions of Ara_3 and Man_3, maintained at 200 °C for different periods, showed that different structures can be formed during coffee roasting, depending on the distribution of the polysaccharides in the beans and the roasting conditions. Furthermore, the diversity of isobaric and isomeric compounds formed highlights the importance of a detailed structural characterization when real roasted carbohydrate-rich matrices such as coffee are analysed.

III.3. UNDERSTANDING COFFEE MELANOIDIN FORMATION

The coffee roasting process leads to the formation of high molecular weight nitrogenous brown-coloured compounds, known as melanoidins. These compounds were estimated to account for up to around 25% (w/w) of the dry weight of roasted coffee beans, and coffee brews, prepared by hot water extraction from ground and roasted coffee beans, are considered one of the main sources of melanoidins in human diet. Although several biological activities, such as antioxidant, antimicrobial, anticariogenic, anti-inflammatory, antihypertensive, and antiglycative activities, have been attributed to coffee melanoidins, their precise structures and mechanisms involved in their formation are far to be elucidated (Section I.2.3).

In order to better understand the structures and mechanisms involved in the formation of coffee melanoidins, selected model systems, composed by an oligosaccharide mixed with a chlorogenic acid and/or a peptide, were submitted to dry thermal treatments, mimicking the coffee roasting process (Sections III.3.1 and III.3.2).

III.3. UNDERSTANDING COFFEE MELANOIDIN FORMATION

III.3.1. CHLOROGENIC ACID–ARABINOSE HYBRID DOMAINS IN COFFEE MELANOIDINS: EVIDENCES FROM A MODEL SYSTEM

The results and discussion presented in this section were integrally published as follow:

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., Passos, C. P., Santos, S. A. O., Silvestre, A. J. D., Silva, A. M. N., Rangel, M., & Domingues, M. R. M. (2015). Chlorogenic acid–arabinose hybrid domains in coffee melanoidins: evidences from a model system. *Food Chemistry*, *185*, 135-144.

III.3.1.1. Background and aim of the study

The studies undertaken to date dealing with the structural characterization of coffee melanoidins have shown that polysaccharides, galactomannans and type II arabinogalactans, together with proteins, chlorogenic acids (CGAs) and sucrose, are involved in coffee melanoidin formation. However, it is still unclear how these different constituents (or their derivatives) are linked in the melanoidin structures (Section I.2.3).

Arabinose (Ara) from arabinogalactan side chains was hypothesized as a possible binding site for CGAs in coffee melanoidins (Bekedam et al., 2008b). However, no evidences have been reported of the presence of CGAs covalently linked to the Ara side chains of the arabinogalactans incorporated in coffee melanoidin structures. In order to investigate this hypothesis, an equimolar mixture of 5-*O*-caffeoylquinic acid (5-CQA), the most abundant CGA in green coffee beans, and (α1→5)-L-arabinotriose (Ara_3), an oligosaccharide structurally related to Ara side chains of arabinogalactans, was submitted to dry thermal treatments, mimicking coffee roasting conditions. The compounds formed during thermal processing were identified by direct electrospray ionization mass spectrometry (ESI-MS) analysis. The identification of these compounds was confirmed by determination of elemental compositions using high resolution MS and their fragmentation pattern under tandem MS (ESI-MS^n). High-performance liquid chromatography (HPLC) with photodiode array detection (PDA) online coupled to ESI-MS and ESI-MS^n was also used to investigate the presence of structures having the same elemental composition (isomers), and thus not able to be differentiated by direct MS analysis. In order to support the formation of chlorogenic acid-arabinose hybrid structures during coffee roasting, fractions previously recovered from spent coffee grounds (SCG) were also analysed by ESI-MS and ESI-MS^n.

III.3.1.2. Results and discussion

Thermal stability of the 5-CQA and mixture with Ara_3

To optimize the thermal conditions used in this work, the thermal stability of the 5-*O*-caffeoylquinic acid (5-CQA; for simplicity, also abbreviated in this work as CQA) and mixture with (α1→5)-L-arabinotriose (Ara_3) was investigated. The respective thermogravimetric (TG) and first derivative thermogravimetric (DTG) curves are shown

as Supplementary Material (Figure S1 in Appendix D), together with that previously obtained for Ara_3 (Section III.1.1). Considering that the loss of weight until around 100 °C is due to the loss of adsorbed water, it can be observed that 5-CQA is thermally stable until around 200 ºC, as reported in previous studies (Owusu-Ware, Chowdhry, Leharne & Antonijević, 2013; Sharma et al., 2002). More specifically, under the conditions used in this work, the first decomposition process of 5-CQA has a peak temperature at 234 °C. Distinctly, the degradation of the Ara_3 and the mixture begins below 200 ºC. In fact, a huge diversity of new compounds was identified when Ara_3 was individually heated from room temperature to 200 ºC (Section III.1.1).

In this work, the model mixture containing Ara_3 and 5-CQA was also heated from room temperature to 200 ºC (200T1). Aiming to obtain a lower degradation extent and, if possible, to identify intermediate degradation products, dry thermal treatments at a lower temperature were also performed. The mixture was heated to 175 ºC and maintained at this temperature for two different periods: 0 (175T1) and 30 min (175T2). The total mass loss percentages were 5.5% for 175T1, 9.9% for 175T2, and 8.1% for 200T1. The solid mixture of 5-CQA and Ara_3 (a yellowish-white powder) acquired a dark brown coloration and appearance of a brittle caramel when submitted to the thermal treatments at 175 ºC and 200 ºC, similar to Ara_3 when individually submitted at 200 ºC. The caramel resulting from the treatment 175T1 had a slightly less intense colour than those from 175T2 and 200T1, suggesting that the lowest degradation was promoted by the treatment 175T1, as corroborated by the lowest total mass loss. No visual colour change was observed when the 5-CQA was individually submitted to the same thermal treatments, corroborating that the development of the brown coloration during dry thermal processing of the mixture resulted from the transformation of sugar moieties.

Identification of chlorogenic acid-arabinose hybrids and other structures

To evaluate the reactivity between Ara_3 and 5-CQA when the model mixture was submitted to the different thermal treatments, both untreated and thermally treated samples, completely solubilized in water, were analysed by ESI-MS. Under ESI-MS conditions, neutral oligosaccharides ionize preferentially in positive mode as $[M+Na]^+$ ions (Moreira et al., 2011; Zaia, 2004), whereas chlorogenic acids ionize preferentially in negative mode as $[M-H]^-$ ions due to the presence of the carboxylic acid group (Clifford, Johnston, Knight & Kuhnert, 2003). For this reason, in order to obtain a clearer picture

of the different compounds formed during thermal processing, ESI-MS spectra were acquired in both negative and positive ion modes.

The negative ion ESI-MS spectrum of the untreated mixture showed as base peak the ion at *m/z* 353, attributed to $[CQA\text{-}H]^-$, and the second most abundant ion at *m/z* 413, attributed to $[Ara_3\text{-}H]^-$ (Figure III.20A). The ions observed at *m/z* 767 and 827 were attributed to $[CQA+Ara_3\text{-}H]^-$ and $[2Ara_3\text{-}H]^-$, respectively.

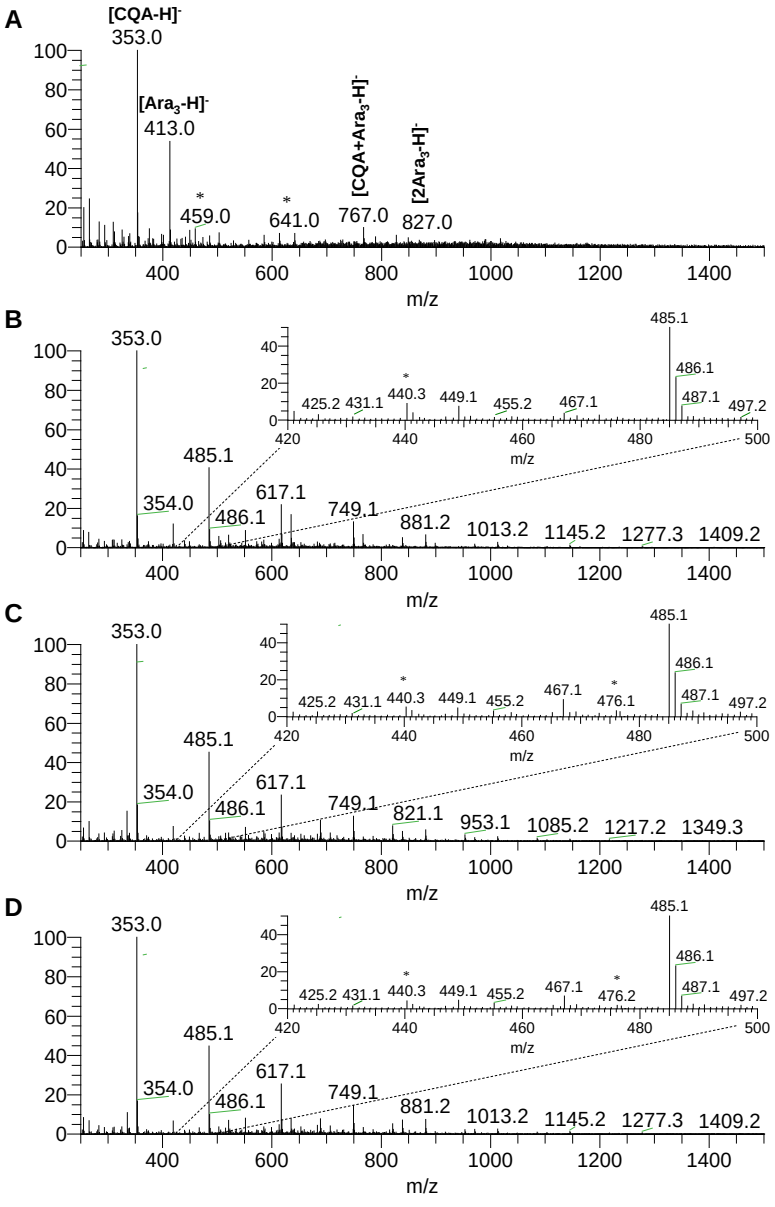

Figure III.20. Negative ion ESI-LIT-MS spectra of the (A) untreated mixture and after thermal treatments: (B) 175T1, (C) 175T2, and (D) 200T1. Ions marked with an asterisk (*) are attributed to solvent impurities.

Table III.12. Summary of the $[M-H]^-$ ions identified in the negative ESI-MS spectra of the thermally treated mixtures of Ara_3 and 5-CQA with the indication of the *m/z* value and the proposed assignment.

n	1	2	3	4	5	6	7	8	9	10	11	12
Pent$_n$CQA series												
$[Pent_nCQA-H]^-$	485	617	749	881	1013	1145	1277	1409	1541[a]	1673[a]	1805[a]	1937[a]
$[Pent_nCQA-H_2O-H]^-$	467	599	731	863	995	1127	1259	1391	1523[a]	1655[a]		
$[Pent_nCQA-2H_2O-H]^-$	449	581	713	845	977	1109	1241					
$[Pent_nCQA-3H_2O-H]^-$	431	563	695	827	959	1091						
Pent$_n$(CQA)$_2$ series												
$[Pent_n(CQA)_2-H]^-$	821	953	1085	1217	1349	1481						
$[Pent_n(CQA)_2-H_2O-H]^-$	803	935										
Pent$_n$QA series												
$[Pent_nQA-H]^-$	323	455	587	719	851							
$[Pent_nQA-H_2O-H]^-$	305	437	569									
Pent$_n$CA series												
$[Pent_nCA-H]^-$		443	575									
$[Pent_nCA-H_2O-H]^-$		425	557									
CQA and derivatives without a sugar moiety												
$[(CQA)_n-H]^-$	353	689										
$[(CQA)_n-H_2O-H]^-$	335											
$[(CQA)_nQA-H]^-$	527											
$[(CQA)_nCA-H]^-$	515											
$[(CQA)_nCA-H_2O-H]^-$	497											

[a]Ions observed exclusively in the ESI-MS spectrum acquired on the LTQ-Orbitrap in the *m/z* range 150-2000. The other ions were also observed in the ESI-MS spectra acquired on the LIT mass spectrometer in the *m/z* range 150-1500.

Independently of the thermal treatment, the ion at *m/z* 353 ($[CQA-H]^-$) remained as the base peak in the negative ESI-MS spectra of the thermally treated mixtures (Figure III.20B-D). Also, several new ions, not observed in the ESI-MS spectrum of the untreated mixture (Figure III.20A), were identified. These ions, summarized in Table III.12 with the indication of the *m/z* value and the proposed assignment, were assigned as $[M-H]^-$ ions of hybrid compounds, derived from Ara_3 and 5-CQA, and 5-CQA derivatives not bearing a sugar moiety formed during thermal processing of the model mixture, which will be described in detail later. The assignment of each ion was supported by the elemental composition obtained from high resolution and exact mass measurements using a LTQ-Orbitrap hybrid mass spectrometer (Table S1 in Appendix D). Note that the sugar moiety of each hybrid compound is composed by pentose (Pent) units. According to the

sugar and glycosidic linkage analyses performed when Ara_3 was individually submitted to dry thermal treatments at 200 °C (Sections III.1.1 and III.2.2), the Pent units are mainly arabinose, although other Pent units, formed by isomerization, oxidation and decarboxylation reactions, were identified.

As result of the different ionization preferences of 5-CQA and Ara_3, the positive ion ESI-MS spectrum of the untreated mixture (Figure III.21A), in contrast to the corresponding negative ion ESI-MS spectrum (Figure III.20A), showed as base peak the ion at *m/z* 437 ($[Ara_3+Na]^+$), and the second most abundant ion at *m/z* 377 ($[CQA+Na]^+$). Also, the $[M+H]^+$ and $[M+K]^+$ ions of Ara_3 (*m/z* 415 and 453) and $[CQA+H]^+$ (*m/z* 355) were observed, but with relative abundances lower than 3.5%.

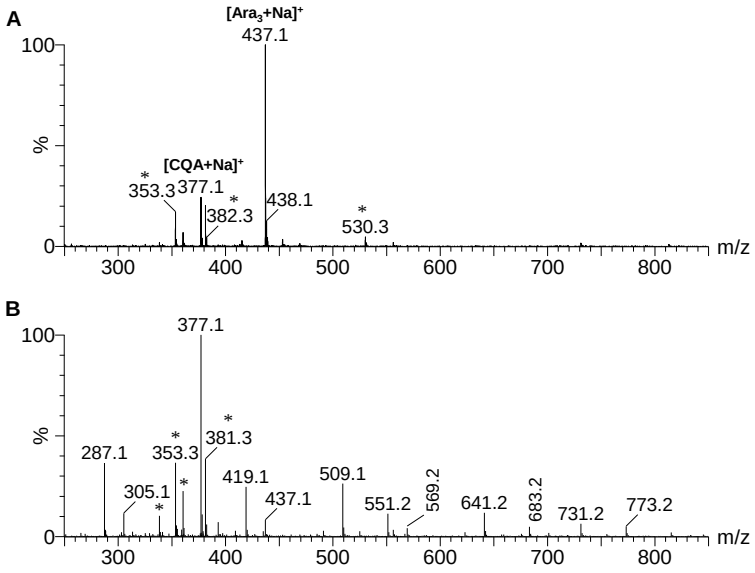

Figure III.21. Positive ion ESI-QTOF-MS spectra of the (A) untreated mixture and (B) the mixture heated at 175 °C for 30 min (175T2). Ions marked with an asterisk (*) are attributed to solvent impurities.

In respect to the positive ion ESI-MS spectra of the thermally treated mixtures, it is important to note that, despite the ionization of oligosaccharides occurs preferentially in positive ion mode, the abundance of the ion at *m/z* 437 decreased and the ion at *m/z* 377 became the base peak after the thermal treatments, corroborating that the Ara_3 was more degraded than 5-CQA, as expected considering the respective thermogravimetric (TG) curves (Figure S1 in Appendix D). As an example, the positive ion ESI-MS spectrum of the mixture heated at 175 °C for 30 min (175T2) is shown in Figure III.21B. As previously described for thermally treated Ara_3 (Section III.1.1), the ions observed at *m/z* 305 and 569 correspond to $[Pent_2+Na]^+$ and $[Pent_4+Na]^+$, supporting the occurrence of

depolymerization and polymerization (transglycosylation) reactions, respectively. The ions at *m/z* 287, 419, 551, and 683 correspond to $[M+Na]^+$ ions of dehydrated Pent oligosaccharides ($Pent_n$-H_2O, *n*=2-5) (Section III.1.1). In accordance with the respective ESI-MS spectrum acquired in negative ion mode (Figure III.20C and Table III.12), the ions at *m/z* 509, 641, and 773 are attributable to $[M+Na]^+$ ions of $Pent_nCQA$ (*n*=1-3) hybrid compounds, which will be described below. The ion observed at *m/z* 731 correspond to $[2CQA+Na]^+$, in accordance with other noncovalently-linked dimers observed in the negative ESI-MS.

In order to confirm the proposed assignments and gain additional information about their structures, the hybrid compounds and 5-CQA derivatives not bearing a sugar moiety formed during thermal processing of the model mixture were characterized by $ESI\text{-}MS^n$. For their characterization, the negative ion mode was preferred because they ionize better in negative than in positive mode.

Likewise $[Ara_3+Na]^+$ ions (Section III.1.1), the $ESI\text{-}MS^2$ fragmentation of $[Ara_3\text{-}H]^-$ ions (*m/z* 413) (Figure S2A in Appendix D) yielded product ions resulting from glycosidic cleavages, loss of water, and cross-ring cleavages with neutral losses of $C_2H_4O_2$ (-60 Da) and $C_3H_6O_3$ (-90 Da). The $ESI\text{-}MS^2$ spectrum of $[M\text{-}H]^-$ ions of 5-CQA (*m/z* 353) (Figure S2B), in accordance with previous studies (Clifford et al., 2003; Fang, Yu & Prior, 2002), showed as base peak the product ion at *m/z* 191, corresponding to $[QA\text{-}H]^-$, and the product ion at *m/z* 179 with a low relative abundance, corresponding to $[CA\text{-}H]^-$. Note that the negative charge is preferentially retained at the QA moiety due the existence of a free carboxyl group (–COOH). Also, it is possible to observe the product ions at *m/z* 173 and 161, formed by loss of CA (-180 Da) and QA (-192 Da) moieties, respectively. Similar to the nomenclature used for the product ions of oligosaccharides resulting from glycosidic cleavages (Section III.1.1), and avoiding the confusion with the dehydration induced by dry thermal processing, these product ions are assigned as deprotonated acid residues, respectively, $[QA_{res}\text{-}H]^-$ and $[CA_{res}\text{-}H]^-$, instead as $[QA\text{-}H_2O\text{-}H]^-$ and $[CA\text{-}H_2O\text{-}H]^-$ designations used in previous studies by other authors (Clifford et al., 2003; Fang et al., 2002). The product ions observed at *m/z* 309 (-44 Da) and 135 (-218 Da) were respectively formed by loss of CO_2 and by combined loss of the QA_{res} (-174 Da) and CO_2 (-44 Da). The knowledge of the typical fragmentation pathways of Ara_3 and 5-CQA was essential to understand the fragmentation pattern of the new ions identified after thermal processing of the mixture, corresponding to hybrid compounds,

derived from Ara_3 and 5-CQA, and 5-CQA derivatives not bearing a sugar moiety. All these ions and their fragmentation patterns are described in the following sections.

a) $Pent_nCQA$ hybrids

Several hybrid compounds composed by a CQA covalently linked with pentose (Pent) units were identified in the negative ESI-MS spectra of the thermally treated mixtures, corroborating the hypothesis of linkages between chlorogenic acids and arabinose in coffee melanoidin structures. For all treated mixtures, the one with the highest relative abundance was observed at *m/z* 485, corresponding to $[M-H]^-$ of a compound formed by the reaction of a Pent and a CQA molecule with the release of a water molecule, assigned as $[PentCQA-H]^-$. As part of the same ion series were also observed the ions at *m/z* 617, 749, 881, 1013, 1145, 1277, 1409, 1541, 1673, 1805 and 1937, assigned as $[Pent_nCQA-H]^-$ ions, n = 2-12. The $Pent_nCQA$ (n=1-12) compounds were also observed as $[M-2H]^{2-}$ ions at *m/z* 242, 308, 374, 440, 506, 572, 638, 704, 770, 836, 902 and 968, but having lower relative abundance than the corresponding $[M-H]^-$ ions. The positive ESI-MS spectra of the thermally treated mixtures, as previously mentioned, showed the ions with *m/z* 509, 641, and 773, assigned as $[Pent_nCQA+Na]^+$ ions (n = 1-3). The identification of $Pent_nCQA$ compounds bearing a lower (n=2) and higher (n=4-12) number of sugar units than that of the oligosaccharide (Ara_3) in the starting mixture also corroborates the occurrence of depolymerization and polymerization (transglycosylation) reactions.

All the ESI-MS^2 spectra acquired from ions assigned as $[Pent_nCQA-H]^-$ support the presence of covalently linked Pent to CQA moieties, allowing to confirm the proposed assignments. As example, the ESI-MS^2 spectra of $Pent_{1-3}CQA$ are shown in Figure III.22.

Figure III.22A shows the ESI-MS^2 spectrum of the ion observed at *m/z* 485, attributed to $[PentCQA-H]^-$. The product ion at *m/z* 353, formed by loss of a $Pent_{res}$ and attributed to $[CQA-H]^-$, confirms the presence of a CQA linked to a Pent. The product ion at *m/z* 323 (base peak), formed by loss of a CA_{res} and attributed to $[PentQA-H]^-$, suggests that the Pent unit was linked to the QA moiety. However, the product ion at *m/z* 293, formed by loss of a QA and attributed to $[(PentCA)_{res}-H]^-$, suggests the presence of other structures with the Pent unit linked to the CA moiety.

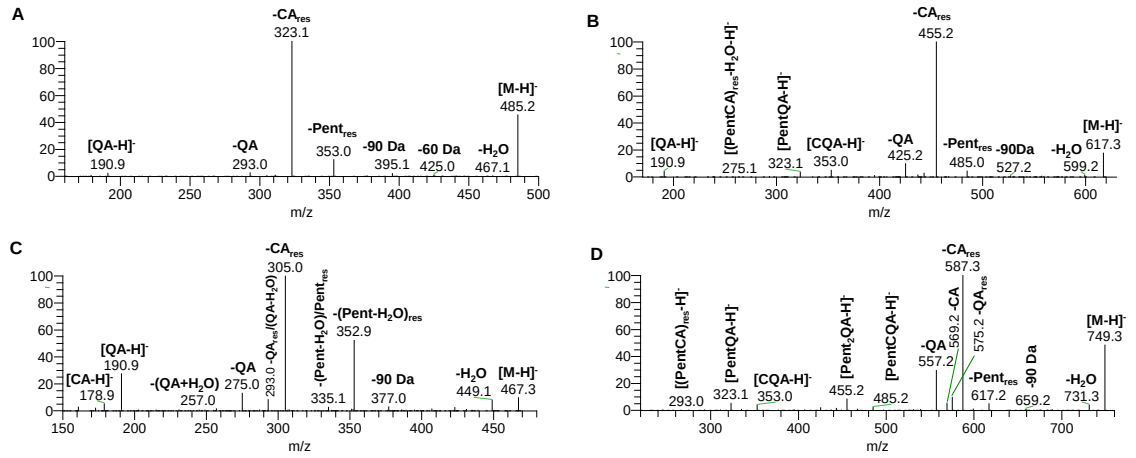

Figure III.22. ESI-MS^2 spectra of the ions observed at *m/z* (A) 485 ([PentCQA-H]$^-$), (B) 617 ([Pent$_2$CQA-H]$^-$), (C) 467 ([PentCQA-H$_2$O-H]$^-$) and (D) 749 ([Pent$_3$CQA-H]$^-$), acquired from the mixture heated to 175 °C (175T1).

Similar to the [PentCQA-H]$^-$ ions, the ESI-MS^2 fragmentation of [M-H]$^-$ ions of Pent$_2$CQA (*m/z* 617, Figure III.22B) and Pent$_3$CQA (*m/z* 749, Figure III.22D) did not yield products ions formed by loss of CQA or (CQA)$_{res}$, not allowing to confirm the presence of structures with the Pent residues linked together. For these compounds, the complementary study of the fragmentation of the correspondent [M+Na]$^+$ ions allowed to obtain additional structural information. The ESI-MS/MS fragmentation of [Pent$_2$CQA+Na]$^+$ (*m/z* 641) and [Pent$_3$CQA+Na]$^+$ (*m/z* 773) (Figure S3 in Appendix D) yielded, respectively, product ions at *m/z* 305 ([Pent$_2$+Na]$^+$) and 437 ([Pent$_3$+Na]$^+$), confirming the presence of structures with the Pent units linked together. The ESI-MS^2 fragmentation of [Pent$_{2\text{-}3}$CQA-H]$^-$ ions (Figure III.22B and 22D) also produced product ions formed by loss of CA$_{res}$ and QA, suggesting the coexistence of isomers having the sugar moiety (Pent$_2$ or Pent$_3$) linked either to the QA or CA moiety of the CQA. In fact, the presence of isomers for each Pent$_n$CQA (*n*=1-12) compound was expected, considering the different linkage possibilities, even in the simplest structure (PentCQA). In this case, considering that the anomeric oxygen of the Pent is involved in a glycosidic linkage, there are five free hydroxyl groups (three in QA and two in CA) as possible binding sites in the CQA. Also, α- and β-anomers may be formed. Moreover, the free carboxylic acid group of the QA moiety may be involved in the formation of an ester linkage.

b) $Pent_n(CQA)_2$ hybrids

In the negative ESI-MS spectra of the thermally treated mixtures, another ion series was observed at *m/z* 821, 953, 1085, 1217, 1349 and 1481 (Table III.12), attributed to $[M-H]^-$ ions of compounds bearing two CQAs covalently linked with a variable number of Pent units (1-6), with the release of a water molecule for each linkage ($[Pent_n(CQA)_2-H]^-$, n=1-6).

All the ESI-MS^2 spectra acquired from ions assigned as $[Pent_n(CQA)_2-H]^-$ support the presence of two CQA molecules and one or more Pent units. As an example, the ESI-MS^2 spectrum of the ion observed at *m/z* 821, attributed to $[Pent(CQA)_2-H]^-$, is shown in Supplementary (Figure S4A in Appendix D). The product ion at *m/z* 689, formed by loss of a $Pent_{res}$, suggests that the two CQA molecules (or their derivatives) are linked together. On the other hand, the product ion at *m/z* 323, attributed to $[PentQA-H]^-$, suggests that the Pent unit is linked to a QA moiety.

c) $Pent_nQA$ and $Pent_nCA$ hybrids

Also, $[M-H]^-$ ions of compounds composed exclusively by quinic (QA) or caffeic (CA) acid moieties, derived from a CQA, covalently linked with a variable number of Pent units were identified in the negative ESI-MS spectra of the thermally treated mixtures and assigned as $[Pent_nQA-H]^-$ (n=1-5) and $[Pent_nCA-H]^-$ (n=2-3), respectively (Table III.12).

All the ESI-MS^2 spectra acquired from ions assigned as $[Pent_nQA-H]^-$ and $[Pent_nCA-H]^-$ support the presence of one or more Pent units linked to a QA or a CA moiety, respectively. For the compounds bearing two or more Pent units ($n\geq2$), the observation of product ions attributed to $[Pent_n-H]^-$ and $[Pent_{n-1}Pent_{res}-H]^-$ corroborated the presence of structures having all the Pent units linked together.

d) Dehydrated derivatives of $Pent_nCQA$, $Pent_n(CQA)_2$, $Pent_nQA$ and $Pent_nCA$

For all the aforementioned series, $[M-H]^-$ ions of dehydrated derivatives resulting from the loss of another water molecule were also identified in the negative ESI-MS spectra of the thermally treated mixtures, assigned as $[Pent_nCQA-H_2O-H]^-$ (n=1-10), $[Pent_n(CQA)_2-H_2O-H]^-$ (n=1-2), $[Pent_nQA-H_2O-H]^-$ (n=1-3), and $[Pent_nCA-H_2O-H]^-$ (n=2-3). Also, $[M-H]^-$ ions of $Pent_nCQA$ derivatives resulting from the loss of two and

three additional water molecules were identified, assigned as [Pent$_n$CQA-2H$_2$O-H]$^-$ (n=1-7) and [Pent$_n$CQA-3H$_2$O-H]$^-$ (n=1-6), respectively (Table III.12).

As an example of a dehydrated derivative of Pent$_n$CQA compounds, the ESI-MS^2 spectrum of the ion observed at *m/z* 467, attributed to [PentCQA-H$_2$O-H]$^-$, is shown in Figure III.22C. The product ion at *m/z* 353, with a difference of 114 Da (132-18) from the precursor ion, suggests that the dehydration induced by thermal processing occurred at the Pent moiety. Considering the loss of water at the Pent moiety, the product ions at *m/z* 335 (-132 Da) and 293 (-174 Da) can be identified as resulting from the loss of (Pent-H$_2$O) and QA$_{res}$, respectively. However, these product ions can also result, respectively, from the loss of Pent$_{res}$ and (QA-H$_2$O), and therefore the coexistence of other structures bearing an intact Pent and a dehydrated QA moiety cannot be completely excluded.

Similarly, the ESI-MS^2 spectrum of the ion observed at *m/z* 803 (Figure S4B in Appendix D), attributed to [Pent(CQA)$_2$-H$_2$O-H]$^-$, showed the product ion at *m/z* 689 (-114 Da), suggesting the presence of a dehydrated Pent moiety. However, the product ion at *m/z* 485, with a difference of 318 Da (336-18) from the precursor ion, formed by loss of (CQA-H$_2$O)$_{res}$, suggests the coexistence of other structures bearing an intact Pent and a dehydrated CQA. All the ESI-MS^2 spectra acquired from ions assigned as [Pent$_n$QA-H$_2$O-H]$^-$ and [Pent$_n$CA-H$_2$O-H]$^-$ also showed a product ion with a difference of 114 Da from the precursor ion, corroborating the presence of structures bearing a dehydrated Pent unit.

e) CQA derivatives without a sugar moiety

After thermal processing (175T1, 175T2 and 200T1) of the model mixture, the MS^2 fragmentation pattern of the ion observed at *m/z* 353 ([CQA-H]$^-$) was not changed. After coffee roasting at 230 °C for 5-6 min, it was observed the decrease of 5-CQA content, while the content of isomers, namely 3-CQA and 4-CQA, increased (Farah et al., 2005). Considering that [M-H]$^-$ ions of 3-, 4- and 5-CQAs produce distinct ESI-MS^2 spectra (Clifford et al., 2003; Fang et al., 2002), changes in the fragmentation pattern of the ion at *m/z* 353 could be indicative of 5-CQA isomerization, not observed in this study.

In accordance with previous studies reporting the dehydration of CQAs during coffee roasting (Farah et al., 2005; Jaiswal et al., 2012; Jaiswal et al., 2014), a dehydrated derivative of CQA was observed at *m/z* 335 ([CQA-H$_2$O-H]$^-$) in the negative ESI-MS spectra of the thermally treated mixtures. The respective ESI-MS^2 spectrum is shown in

Figure S5 (Appendix D). The product ion at *m/z* 179, with a difference of 156 Da (174-18) from the precursor ion, formed by loss of $(QA\text{-}H_2O)_{res}$ and attributed to $[CA\text{-}H]^-$, corroborates that the loss of the water molecule occurred at the QA moiety. This is also corroborated with the product ions at *m/z* 173, 161 (base peak), and 135, attributed to $[QA\text{-}H_2O\text{-}H]^-$, $[CA_{res}\text{-}H]^-$, and $[CA\text{-}CO_2\text{-}H]^-$, respectively. According to previous studies, the loss of a single water molecule from the QA moiety of CQAs during coffee roasting can produce either caffeoyl-1,5-quinides (lactones, abbreviated as CQLs) (Farah et al., 2005), or caffeoylshikimic acids (CSAs) (Jaiswal et al., 2012; Jaiswal et al., 2014). Considering the ESI-MS^2 fragmentation reported for both CQLs and CSAs (Jaiswal, Matei, Ullrich & Kuhnert, 2011), the possible coexistence of CQLs and CSAs formed during thermal processing of the mixture cannot be excluded.

Other compounds derived from CQA, not bearing a sugar moiety, were also identified in the negative ESI-MS spectra of the thermally treated mixtures as $[M\text{-}H]^-$ ions at *m/z* 689, 527, 515, and 497, assigned as $[(CQA)_2\text{-}H]^-$, $[(CQA)QA\text{-}H]^-$, $[(CQA)CA\text{-}H]^-$, and $[(CQA)CA\text{-}H_2O\text{-}H]^-$, respectively (Table III.12). In accordance with the proposed assignments, the ESI-MS^2 spectra of the ions observed at *m/z* 515, 527 and 689 (Figure S6 in Appendix D) support the presence of a CQA covalently linked with a CA, a QA or another CQA molecule, respectively. However, the product ion observed at *m/z* 395 (-132 Da) in the ESI-MS^2 spectrum of the ion at *m/z* 527, attributed to $[(CQA)QA\text{-}H]^-$, suggests the coexistence of $[Pent_4\text{-}H_2O\text{-}H]^-$ precursor ions, which is corroborated by the observation of the corresponding $[M+Na]^+$ ions (*m/z* 551) in the positive ion ESI-MS spectra of the thermally treated mixtures (Figure III.21B). The ESI-MS^n spectra (*n*=2-3) acquired from the ion observed at *m/z* 497, attributed to $[(CQA)CA\text{-}H_2O\text{-}H]^-$, corroborate the loss of a water molecule at a QA moiety, not excluding the possibility of the formation of either a lactone or a shikimic acid moiety (Figure S7).

The $[M\text{-}H]^-$ ions of CQA derivatives without a sugar moiety were also identified in samples of only 5-CQA submitted to the thermal treatments 175T1, 175T2 and 200T1 (data not shown). As in the ESI-MS spectra of the thermally treated mixtures, these ions showed a low relative abundance (≤ 1.5%) and the ion at *m/z* 353 ($[CQA\text{-}H]^-$) remained as the base peak after thermal processing, corroborating the thermal stability of 5-CQA until around 200 ºC, as evidenced by TG analysis.

Differentiation of PentCQA isomers

To unveil possible isomeric structures, in particular of Pent$_n$CQA compounds, the mixture heated to 175 ºC (175T1) was further analysed by HPLC-PDA-ESI-MS and HPLC-PDA-ESI-MS^n. The Ara$_3$ that did not react and other oligosaccharides formed during the thermal processing of the mixture were not retained by the C18 column, confirming the covalent linkage between CQA and sugar moieties of the Pent$_n$CQA hybrid compounds, which were retained by the column.

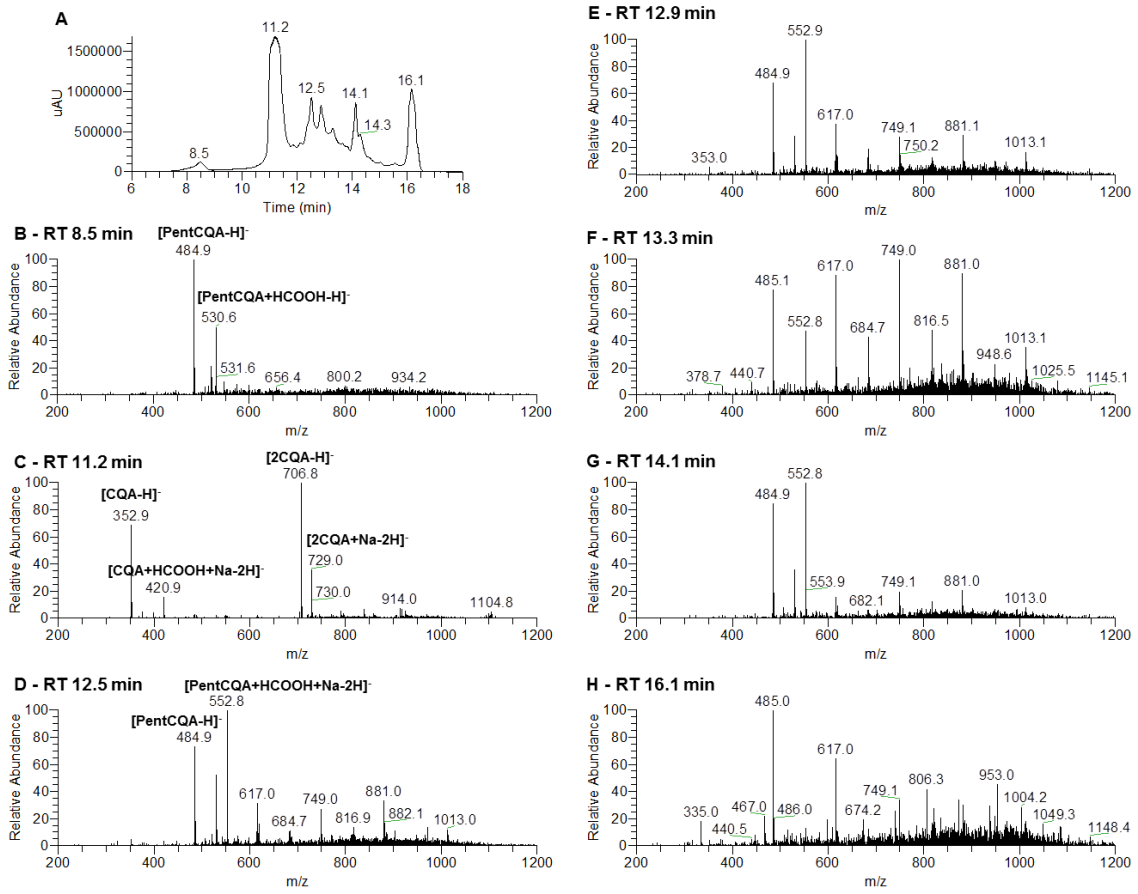

Figure III.23. (A) HPLC-UV chromatogram recorded at 325 nm obtained from mixture heated to 175 °C (175T1), and (B-H) HPLC-ESI-MS spectra associated with the major peaks with retention times (RTs) between 8.5-16.1 min.

Figure III.23A shows the HPLC-UV chromatogram recorded at 325 nm, a characteristic absorption wavelength of CQAs. According to this chromatogram, 5-CQA that did not react and the compounds bearing a CQA moiety formed during the thermal processing of the mixture, including the Pent$_n$CQA compounds, eluted between 7.5-16.5 min. On the other hand, the HPLC-MS spectra associated with the major chromatogram

peaks (Figure III.23B-H) show that the chromatographic separation of each Pent$_n$CQA compound having a distinct number of Pent units (n=1-12) was not achieved, but isomers of these compounds were separated, eluting at different retention times (RTs). However, it was not possible to achieve a perfect separation of all the isomeric structures of each Pent$_n$CQA compound. Since a more reliable separation of the isomers was obtained for the simplest hybrid compound (PentCQA), observed as [M-H]$^-$ at *m/z* 485, the respective reconstructed ion chromatogram (RIC) is shown in Figure III.24A.

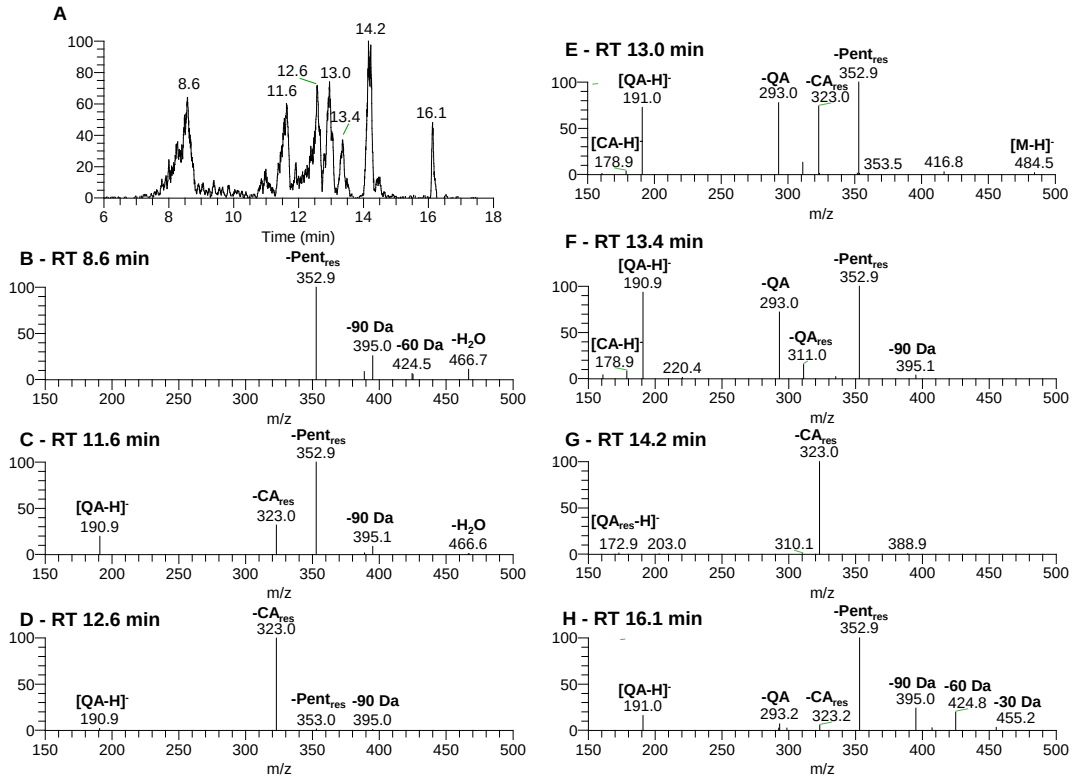

Figure III.24. Differentiation of PentCQA isomers in the mixture heated to 175 °C (175T1) by HPLC-ESI-MS and HPLC-ESI-MS^2: (A) Reconstructed ion chromatogram of the ion with *m/z* 485 ([PentCQA-H]$^-$), and (B-H) the respective HPLC-ESI-MS^2 spectra acquired at different retention times (RTs).

According to the RIC chromatogram (Figure III.24A), PentCQA eluted in seven major peaks with RTs ranging from 8.6 to 16.1 min. This suggests the presence of at least seven isomeric structures, as corroborated by the distinct HPLC-MS^2 spectra obtained at each RT (Figure III.24B-H). As there are no standards available, it was not possible to identify the specific isomers giving rise to these HPLC-MS/MS spectra. Nevertheless, the absence of the product ion formed by loss of QA (*m/z* 293) in the MS^2 spectra obtained at RTs 11.6 (C), 12.6 (D) and 14.2 (G) min suggests the presence of isomers having the Pent linked to the CQA through the QA moiety. On other hand, the absence of the product

ion formed by loss of CA_{res} (*m/z* 323) in the MS^2 spectrum obtained at RT 13.4 min (F) suggests an isomer having the Pent linked to the CA moiety of the CQA. In fact, the possible reaction of the anomeric carbonyl group of the Pent with any one of the five free hydroxyl groups in the CQA, giving rise to α- and/or β-anomers, as well as the possible reactivity of the acid group of the QA moiety justify the diversity of isomers formed.

Identification of Pent$_{1-2}$CQA in fractions recovered from spent coffee grounds

In order to validate the strategy used to identify possible hybrid structures formed from chlorogenic acids and arabinose side chains of arabinogalactans during coffee roasting, fractions recovered from spent coffee grounds (SCG) were analysed by ESI-MS. In both negative ion ESI-MS spectra of MAE3_PptEt and MAE4_PptEt fractions, the most abundant ions were observed at *m/z* 191 and 353, attributed to $[QA-H]^-$ and $[CQA-H]^-$, respectively. They also showed, although with a low relative abundance, the ions at *m/z* 485 and 617, attributed to $[Pent_{1-2}CQA-H]^-$, as well as the ion at *m/z* 335, attributed to $[CQA-H_2O-H]^-$. These assignments were corroborated with respective negative ESI-MS^n (*n*=2-3) spectra, showing the typical product ions identified from the fragmentation of the ions with the same *m/z* value identified after thermal processing of the model mixture. Also in accordance with the data obtained from the thermal treated mixture, the Pent$_{1-2}$CQA compounds identified in the SCG fractions may have been formed during coffee roasting. Accordingly, this type of compounds may have also been incorporated into the coffee melanoidin structures, also formed during roasting.

III.3.1.3. Concluding remarks

The dry thermal processing of a model mixture composed by equimolar amounts of arabinotriose (Ara_3) and 5-*O*-caffeoylquinic acid (5-CQA) promoted the formation of several hybrid compounds composed by one or two CQAs covalently linked with a variable number of pentose residues (mainly arabinose), corroborating the hypothesis of arabinose from arabinogalactan side chains as a possible binding site for chlorogenic acid derivatives in coffee melanoidin structures. The further analysis by HPLC-MS and HPLC-MS^n allowed demonstrating the presence of isomeric hybrid structures, namely PentCQA isomers. These results highlight the structural complexity of compounds that can be formed between chlorogenic acids and carbohydrates during coffee roasting. Also,

the formation of these chlorogenic acid-carbohydrate hybrid structures with functionalization of the carbohydrate moiety by inclusion of carboxylic groups can increase their reactivity and constitute the starting point for the incorporation of carbohydrates in coffee melanoidins through the reaction of the chlorogenic acids present.

The identification of Pent$_n$CQA compounds from the model mixture, as well as the knowledge of their fragmentation pattern under ESI-MS^n conditions, made possible their identification in fractions recovered from spent coffee grounds, opening new perspectives for their identification in coffee melanoidin structures, but also in melanoidins from other sources.

The presence of covalently linked chlorogenic acids to the melanoidin structures may contribute to their antioxidant activity. Having this in mind, the roasting of oligosaccharides or polysaccharides used as functional ingredients in the presence of chlorogenic acids may be used as a method to improve the antioxidant activity of food products. Future work is needed to assess biological activities of hybrid compounds formed from oligo- or polysaccharides and chlorogenic acids, as well as studies with synthetic standards are needed to identify the specific fragmentation pattern of the different isomers formed.

III.3. UNDERSTANDING COFFEE MELANOIDIN FORMATION

III.3.2. ACID CATALYSES, AND MAILLARD REACTION PREVENTS, NON-ENZYMATIC TRANSGLYCOSYLATION REACTIONS DURING COFFEE ROASTING

III.3.2.1. Background and aim of the study

Melanoidins are by definition high molecular weight nitrogenous and brown-coloured compounds, end products of the Maillard reaction (MR). They are formed during the roasting of green coffee beans, but also during the heat processing of other food products like bread, meat, or malt (Section I.2.3).

The different chemical composition of melanoidin populations isolated from coffee brews, namely those containing covalently-linked CGAs, strongly suggest that the classical MR mechanism is not the only mechanism operating during coffee melanoidin formation. As coffee melanoidins are reported as beneficial to health, it is necessary to understand their structural features and formation mechanisms, and fully exploit their beneficial effects (Section I.2.3).

In this study, aiming to better understand the roasting-induced compounds formed by MR between coffee polysaccharides and proteins, and their relation with the non-enzymatic transglycosylation reactions, model mixtures containing equimolar amounts of (β1→4)-D-mannotriose (Man_3), an oligosaccharide structurally related to coffee galactomannans (Nunes et al., 2005), and a dipeptide composed by tyrosine (Y) and leucine (L), amino acids present in green coffee beans (Illy et al., 1995), were submitted to thermal treatment under dry conditions, mimicking the coffee roasting process. To evaluate the influence in the compounds formed of the amino acid's position in the peptide chain, two isomeric peptides, tyrosine-leucine (YL) and leucine-tyrosine (LY), were used. Also, the influence in the non-enzymatic transglycosylation and MR of 5-O-caffeoylquinic acid (CQA), the most abundant chlorogenic acid in green coffee beans, was investigated using mixtures with equimolar amounts of Man_3 and CQA and Man_3, YL, and CQA. The reactivity between proteins and chlorogenic acids during coffee roasting was also investigated using a mixture with equimolar amounts of YL and CQA. For comparison with the results obtained from the mixtures, a sample of Man_3 was also submitted to the same thermal/roasting conditions. All the unroasted and roasted samples were further analysed according to the glycosidic linkage composition, and by high-performance liquid chromatography coupled with electrospray ionization tandem mass spectrometry (HPLC-ESI-MS^n). The elemental composition of the ions identified in the roasted mixtures Man_3-CQA and Man_3-YL was obtained by high resolution and high mass accuracy measurements using an Orbitrap-based mass spectrometer.

III.3.2.2. Results and discussion

Mass losses and changes in colour and water-solubility after roasting

In this study, (β1→4)-D-mannotriose (Man_3) and mixtures with tyrosine-leucine (YL), leucine-tyrosine (LY), 5-*O*-caffeoylquinic acid (CQA), and both YL and CQA, as well as a mixture YL-CQA were submitted to thermal treatment under dry conditions, mimicking the coffee roasting process. The energy requirements for Maillard reaction (MR) are usually lesser than those required for caramelization, i.e. the reactions occurring when polyhydroxylcarbonyl compounds, namely reducing sugars, are exposed to relatively high temperatures in the absence of amino compounds (Hodge, 1953). Thus, attending to the oligosaccharide thermal stability (Moreira et al., 2011) to avoid an extensive degradation in the presence of a peptide, the mixtures Man_3-YL, Man_3-LY, and Man_3-YL-CQA were heated from room temperature to 175 ºC (175T1). For comparison with the results obtained with these mixtures, the other mixtures under study, Man_3-CQA and YL-CQA, and also a sample of Man_3 were submitted to the same thermal/roasting conditions.

During the roasting process, the mixtures Man_3-YL, Man_3-LY, Man_3-CQA, and Man_3-YL-CQA (off-white powders) acquired a dark brown coloration. The mass loss (ML) between 150-175 °C, excluding the initial ML due to the loss of adsorbed water molecules, was 2.2% for Man_3-CQA, 4.3% for Man_3-YL-CQA, 5.2% for Man_3-YL, and 6.5% for Man_3-LY. For Man_3 (a white power), no change in colour was observed after roasting. The total ML percentage was 7.1%, with no ML between 150-175 °C, allowing to infer that a higher degradation extent occurred in the mixtures containing a peptide. Accordingly, a water-insoluble fraction was recovered after roasting from the mixtures Man_3-YL-CQA, Man_3-YL, and Man_3-LY, whereas the mixture Man_3-CQA remained completely water-soluble at the concentration used (5 mg/mL). In respect to the mixture YL-CQA (an off-white power), it acquired a light brown coloration, and had a ML between 150-175 °C of 1.0%. In order to gain information about the structural modifications induced by roasting, all unroasted (T0) and roasted (175T1) samples (water-soluble fractions) were analysed according to the glycosidic linkage composition, and by high-performance liquid chromatography (HPLC, or simply LC) coupled with electrospray ionization mass spectrometry (ESI-MS) and tandem MS ($ESI\text{-}MS^n$).

Glycosidic linkage composition

Table III.13 shows the glycosidic linkage composition of both unroasted (T0) and roasted (175T1) samples. Terminally linked mannopyranose (T-Man*p*; 40.3-43.2%) and (1→4)-linked mannopyranose (4-Man*p*; 52.2-53.3%) were the most abundant residues identified in the unroasted samples of Man_3 and mixtures Man_3-YL, Man_3-LY, Man_3-CQA, and Man_3-YL-CQA. Also, in small amounts were identified 4,6-Man*p* (1.2-2.0%), T-Gal*p* (2.0-2.6%), T-Glc*p* (0.2-0.9%) and 4-Glc*p* (0.4-0.8%). In the unroasted Man_3-CQA mixture, small amounts of 6-Man*p* (0.7%), 3,4-Man*p* (0.3%) and 2,4-Man*p* (0.3%) were also identified.

Table III.13. Glycosidic linkage composition (percentage area) of unroasted (T0) and roasted (175T1) samples of Man_3 and mixtures Man_3-YL, Man_3-LY, Man_3-CQA, Man_3-MalA, Man_3-CitA, and Man_3-YL-CQA.

Linkage	Man_3		Man_3-YL		Man_3-LY		Man_3-CQA		Man_3-MalA		Man_3-CitA		Man_3-YL-CQA	
	T0	175T1	T0	175T1	T0	175T1	T0	175T1	T0	175T1	T0	175T1	T0	175T1
T-Man*p*	42.1	43.6	42.1	60.4	42.8	62.2	40.3	42.5	39.2	41.9	41.3	44.5	43.2	60.5
2-Man*p*				0.7		1.0		2.9		3.0		3.1		1.0
4-Man*p*	52.7	51.4	53.3	27.8	53.2	26.3	52.7	19.5	57.3	24.8	54.9	10.3	52.2	23.1
6-Man*p*				3.2		2.8	0.7	14.1		12.6		18.9		3.7
3,4-Man*p*		0.1		0.3		0.4	0.3	1.1		0.8		0.4		0.5
2,3-Man*p*								0.8		0.2		0.4		
2,4-Man*p*		0.2		0.9		0.6	0.3	1.0		1.4		1.1		0.9
4,6-Man*p*	1.6	1.7	1.7	1.7	1.2	1.3	2.0	6.2	0.6	7.5	1.0	7.1	1.5	2.8
2,6-Man*p*				0.6		0.6		2.6		0.3		2.4		
3,6-Man*p*								0.8		0.7		2.2		
3,4,6-Man*p*								0.7		0.4		0.7		
2,3,6/2,4,6-Man*p*								2.3		0.4		1.5		
2,3,4,6-Man*p*								0.9				0.3		
T-Gal*p*	2.1	2.2	2.1	3.0	2.0	2.5	2.6	1.2	1.8	1.1	1.9	0.8	2.1	4.1
6-Gal*p*								1.8		1.7		2.8		
T-Glc*p*	0.9	0.2	0.2	0.2	0.2	0.3	0.4	0.6	0.5		0.4		0.2	0.2
4-Glc*p*	0.7	0.7	0.6	1.2	0.4	2.0	0.7	1.0	0.5	3.2	0.5	3.6	0.8	3.1

When Man_3 was roasted by heating from room temperature up to 175 °C (175T1), little change in the glycosidic linkage composition was observed. This is in accordance with previous data showing the thermal stability of Man_3 up to temperatures of approximately 200 °C (Moreira et al., 2011). In terms of new glycosidic linkages, only small amounts of 3,4-Man*p* (0.1%) and 2,4-Man*p* (0.2%) were identified. When the

mixture Man_3-YL was submitted to the same roasting conditions, more new Man*p* linkages, especially 6-Man*p* (3.2%), were formed, allowing to infer an increase in the extent of non-enzymatic transglycosylation reactions induced by roasting. However, a marked increase of T-Man*p* residues (42.1 to 60.4%) was observed, indicating the occurrence of depolymerization. As a similar behaviour was observed for the mixture Man_3-LY, the glycosidic linkage composition seems not to be influenced by the amino acid's position in the peptide chain.

After roasting of the mixture Man_3-CQA, it was observed a higher diversity and percentage of new glycosidic linkages than what was observed after roasting of the mixtures Man_3-peptide. On the other hand, it was not observed a marked increase of T-Man*p* residues (40.3 to 42.5%). These results suggest that the presence of CQA may act as a catalyst of non-enzymatic transglycosylation reactions. In order to disclose the possibility of an acid-catalysed transglycosylation, mixtures of Man_3 with malic acid (MalA) and citric acid (CitA), two acids present in green coffee beans (Wei et al., 2010), were submitted to the same roasting conditions. Similar to what was observed after roasting of the mixture Man_3-CQA, a huge diversity of new glycosidic linkages, absent in the respective unroasted sample, was identified after roasting of the mixtures Man_3-MalA and Man_3-CitA (Table III.13), corroborating that acids catalyse non-enzymatic transglycosylation reactions induced by roasting.

The roasting-induced changes in the glycosidic composition of the mixture Man_3-YL-CQA were similar to what was observed for the mixtures Man_3-YL and Man_3-LY, with a marked increase of T-Man*p* (43.2 to 60.5%), and formation of the same new Man*p* linkages, with a similar percentage, except 2,6-Man*p* that was not identified in the roasted mixture Man_3-YL-CQA. However, the diversity of new Man*p* linkages and their percentage was lower than that was observed for the mixture Man_3-CQA (Table III.13). These findings indicate that the extent of non-enzymatic transglycosylation is inhibited by the peptide, even in the presence of an acid like CQA. According to the classical MR mechanism, it can be inferred that in the presence of a peptide, few reducing sugar residues were free to participate in non-enzymatic transglycosylation reactions, because most of them have been reacted with peptide amino groups by MR. This can contribute to a lower extent of non-enzymatic transglycosylation reactions during coffee roasting, particularly in coffee bean regions rich in proteins. Accordingly, the reducing terminal residues of coffee polysaccharides react preferentially with the amino groups of coffee proteins, decreasing the number of reducing residues that could be directly involved in

the non-enzymatic transglycosylation reactions. This fact can explain the high content of melanoidins formed during roasting, although polysaccharides account for about 50% of green coffee beans' dry weight.

Analysis of roasting-induced compounds by HPLC-ESI-MS and HPLC-ESI-MS^n

The roasting-induced compounds were identified and characterized by HPLC-ESI-MS and HPLC-ESI-MS^n (for simplify, also abbreviated as LC-MS and LC-MS^n, respectively). For these analyses, positive mode was preferred because better signals were obtained in positive than in negative mode, due to the high proton affinity of peptides, and the easier facility of sugars to form sodium adduct ions ($[M+Na]^+$).

Regarding to the LC-MS analysis of the unroasted samples, as typical of neutral oligosaccharides (Moreira et al., 2011), Man_3 was mainly identified as $[M+Na]^+$ ions (*m/z* 527). As typical of peptides (Fonseca et al., 2009), YL and LY were mainly identified as $[M+H]^+$ ions (*m/z* 295). CQA was mostly detected as $[M+Na]^+$ ions (*m/z* 377), but also $[M+H]^+$ (*m/z* 355) ions of CQA were identified, having a relative abundance ≥30%. Other related ions having a lower relative abundance than those previously described were also identified in the unroasted samples, namely $[Man_3+K]^+$ ions (*m/z* 543), $[2Peptide+H]^+$ (*m/z* 589), $[2CQA+Na]^+$ (*m/z* 731), $[Peptide+Hex_3+H]^+$ (*m/z* 799), and ($[2Man_3+Na]^+$ (*m/z* 1031). After roasting, new ions, absent in the unroasted samples, were identified by LC-MS analysis. In Table III.14 are summarized $[M+Na]^+$ and $[M+H]^+$ ions identified after roasting of Man_3 and mixtures Man_3-YL, Man_3-CQA, Man_3-YL-CQA, and YL-CQA, and in Supplementary Table S1 (Appendix E) are summarized $[M+Na]^+$ and $[M+H]^+$ ions identified after roasting of Man_3-LY, with the indication of the *m/z* values and the proposed assignments. As will be detailed later, the new ions identified from all roasted samples were assigned to different types of roasting-induced compounds, grouped as follow: a) hexose (Hex) and oligosaccharides composed up to 9 Hex units (Hex_n, n=1-9) and dehydrated derivatives, b) dehydrated and oxidative deaminated derivatives of peptides, c) Amadori compounds formed by condensation of the N-terminal amino group of a peptide (YL or LY) with the reducing residue of Hex_n (Hex_nYL and Hex_nLY), d) dehydrated derivatives of Hex_nYL and Hex_nLY, e) compounds bearing two peptide molecules covalently linked with a variable number of Hex units ($Hex_n(YL)_2$ and $Hex_n(LY)_2$), f) derivatives of CQA with and without a Hex_n moiety, g) a compound formed by reaction of a peptide and a CQA molecule (YL(CQA)), and h) compounds

formed by reaction of a peptide, a CQA molecule, and a sugar composed by a variable number of Hex units (Hex_n(YL)CQA).

Table III.14. Summary of $[M+Na]^+$ and $[M+H]^+$ ions identified by LC-MS analysis after roasting of the Man_3 and mixtures Man_3-YL, Man_3-CQA, Man_3-YL-CQA, and YL-CQA, with the indication of the *m/z* values and the proposed assignments.

n	1	2	3	4	5	6	7	8	9
Roasted Man_3									
$[Hex_n+Na]^+$		365	527	689					
Roasted mixture Man_3-YL									
$[Hex_n+Na]^+$	203	365	527	689	851	1013			
$[Hex_n-H_2O+Na]^+$		347	509						
$[Hex_n-2H_2O+Na]^+$		329	491						
$[Hex_n-3H_2O+Na]^+$		311	473						
$[(YL)_n+H]^+$	295								
$[(YL)_n-H_2O+H]^+$	277								
$[(YL)_n-NH_3+O+H]^+$	294								
$[Hex_nYL+H]^+$	457	619	781	943	1105	1267			
$[Hex_nYL-H_2O+H]^+$	439	601	763	925	1087	1249			
$[Hex_nYL-2H_2O+H]^+$	421	583	745	907					
$[Hex_nYL-3H_2O+H]^+$	403	565	727	889					
$[Hex_nYL-4H_2O+H]^+$	385	547	709	871					
$[Hex_n(YL)_2+H]^+$	733	895	1057						
Roasted mixture Man_3-CQA[a]									
$[Hex_n+Na]^+$		365	527	689	851	1013	1175	1337	1499
$[Hex_n-H_2O+Na]^+$		†347	†509	671	833	995	1157	1319	
$[Hex_n-2H_2O+Na]^+$		329	491	653					
$[Hex_n-3H_2O+Na]^+$		311	473	635					
$[(CQA)_n+Na]^+$	377†	713							
$[(CQA)_nCA+Na]^+$	539‡								
$[(CQA)_nCA-H_2O+Na]^+$	521‡								
$[Hex_nCQA+Na]^+$	539‡	701†	863	1025	1187	1349			
$[Hex_nCQA-H_2O+Na]^+$	521‡	683†							
$[Hex_nCQA-3H_2O+Na]^+$	485	647							
$[Hex_n(CQA)_2+Na]^+$	875								
$[Hex_nQA+Na]^+$	377†	539‡	701†						
$[Hex_nQA-H_2O+Na]^+$		521‡	683†						
$[Hex_nCA-H_2O+Na]^+$	†347	†509							
Roasted mixture Man_3-YL-CQA[b]									
$[Hex_n+Na]^+$		365	527	689					
$[Hex_n-H_2O+Na]^+$		347	509						
$[Hex_n-2H_2O+Na]^+$		329	491						
$[Hex_n-3H_2O+Na]^+$		311	473						
$[YL_n+H]^+$	295								
$[YL_n-H_2O+H]^+$	277								
$[YL_n-NH_3+O+H]^+$	294								
$[(CQA)_n+Na]^+$	377								
$[YL_n(CQA)+H]^+$	631								
$[Hex_nYL+H]^+$	457	619	781	943	1105				
$[Hex_nYL-H_2O+H]^+$	439	601	763	925	1087				
$[Hex_nYL-2H_2O+H]^+$	421	583	745	907	1069				
$[Hex_nYL-3H_2O+H]^+$	403	565	727	889					
$[Hex_nYL-4H_2O+H]^+$	385	547	709	871					
$[Hex_n(YL)_2+H]^+$	733	895	1057						
$[Hex_nYLCQA+H]^+$	793	955	1117						
Roasted mixture YL-CQA									
$[YL_n+H]^+$	295	571							
$[(CQA)_n+Na]^+$	377	713							
$[YL_n(CQA)+H]^+$	631								

[a]According to the accurate masses found by LTQ-Orbitrap, the ions marked with the symbol † or ‡ were attributed to different isobaric compounds: † for two and ‡ for three possible compounds. [b]Similar to the mixture Man_3-CQA, some of the ions observed after roasting of the Man_3-YL-CQA can be the contribution of different isobaric compounds. The ion assignment was made considering the most abundant isobaric compounds identified in the roasted mixture Man_3-CQA.

In following sections, the ion series identified from each roasted sample are described, as well as details on their fragmentation pattern under LC-MS^n conditions are presented, corroborating the proposed assignments. The assignment of the ions identified after roasting of the mixtures Man$_3$-CQA and Man$_3$-YL was also supported by elemental composition obtained from high resolution and high mass accuracy measurements using an Orbitrap-based mass spectrometer (Tables S2 and S3 in Appendix E). Due to the different ionization preferences/efficiencies of oligosaccharides and chlorogenic acids/peptides, the MS data obtained by direct analysis on the LTQ-Orbitrap, without HPLC coupling, were acquired in negative ion mode, as previously made in the analysis of the roasted mixture arabinotriose-CQA (Section III.3.1). Additionally, for roasted mixture Man$_3$-YL, the LC retention times (RTs), and the most abundant product ions observed in the LC-MS^2 spectra, with the indication of the *m/z* values, mass differences relative to the precursor ion, and the identification of the most informative product ions, are shown in Supplementary Table S4 (Appendix E). Note that the analysis of the reconstructed ion chromatograms (RICs) suggests the presence of isomeric compounds, i.e. compounds with the same elemental composition but different structures, eluting at different RTs. However, the exact structural differences were not possible to be inferred based on the respective LC-MS^n spectra (n=2-3) because they were very similar, most probably due to the presence of positional isomers. In the case of the compounds bearing a sugar moiety, the structural differences of the isomers can be related to different structures of the sugar moiety, differing on glycosidic linkage positions, and anomeric configuration.

a) Hex$_n$ and dehydrated derivatives

After roasting of the mixtures Man$_3$-LY (Table S1 in Appendix E) and Man$_3$-YL (Table III.14), hexose (Hex) monosaccharides, but also Hex oligosaccharides with a degree of polymerization (DP, i.e. number of monosaccharide units) from 2 up to 5 and 6, respectively, were identified. In the LC-MS analysis of the roasted Man$_3$ sample (Table III.14), besides Hex$_3$, only Hex$_2$ and Hex$_4$ were identified, and the RICs of the respective $[M+Na]^+$ ions had a low signal intensity. As inferred by glycosidic linkage analysis, these results corroborate an increase in the extent of non-enzymatic transglycosylation reactions during roasting of the mixtures Man$_3$-peptide (YL or LY), in comparison to the Man$_3$ sample. However, the extent of non-enzymatic transglycosylation reactions was

even greater when the mixture Man_3-CQA was roasted, as corroborated by the identification of Hex oligosaccharides with a DP up to 9. As observed for roasted Man_3, only Hex of oligosaccharides with a DP from 2 up to 4 were identified by LC-MS analysis of the roasted mixture Man_3-YL-CQA, corroborating the inhibition of non-enzymatic transglycosylation reactions due to the presence of the peptide.

Dehydrated Hex_n derivatives, resulting from loss of one, two, and three water molecules, were identified after roasting of the mixtures Man_3-YL, Man_3-CQA, and Man_3-YL-CQA (Table III.14), and the mixture Man_3-LY (Table S1 in Appendix E). Under the same roasting conditions, no dehydrated derivatives were formed from Man_3, indicating that the total energy required for the formation of dehydrated Hex_n derivatives was decreased due to the presence of CQA and/or a peptide. This type of dehydrated derivatives was previously observed when Man_3 (alone) was roasted by heating from room temperature up to 200 °C, maintained at 200 °C for additional 30 and 60 min (Moreira et al., 2011).

According to that was previously reported (Moreira et al., 2011), the LC-MS^2 fragmentation of $[M+Na]^+$ ions of Hex_n and dehydrated derivatives yielded product ions resulting from glycosidic cleavages, loss of water, and cross-ring cleavages, corroborating the proposed ion assignments.

b) Dehydrated and oxidative deaminated derivatives of peptides

After roasting of the mixtures Man_3-YL, Man_3-LY, Man_3-YL-CQA, and YL-CQA, non-modified peptide was identified. As previously reported (Fonseca et al., 2009), the LC-MS^2 fragmentation of the respective $[M+H]^+$ ions (*m/z* 295) (Figure S1 in Appendix E) gave rise to the base peak product ion at *m/z* 136 (-159 Da) in the case of YL, and *m/z* 182 (-113 Da) in the case of LY, assigned, respectively, as a_1 and y_1, according to the nomenclature originally proposed by Roepstorff & Fohlman (1984), and later modified by Biemann (1988).

In the roasted mixtures Man_3-YL, Man_3-LY, and Man_3-YL-CQA, a dehydrated peptide, (YL-H_2O) or (LY-H_2O), was also identified as $[M+H]^+$ ions (*m/z* 277). According to the respective LC-MS^2 spectra, the peptide dehydration involves the C-terminal carboxyl group, but an exact structure cannot be deduced. For YL (Table S4 in Appendix E), the involvement of the C-terminal carboxyl group in the peptide dehydration was corroborated by the observation in the respective LC-MS^2 spectrum of

the base peak product ion at *m/z* 136 (a_1), resulting from the neutral loss of 141 Da, instead of 159 Da observed for the respective non-modified peptide (Figure S1A in Appendix E). The LY dehydration led to a greater change in the LC-MS^2 fragmentation, compared to that of the non-modified peptide. In particular, the involvement of the C-terminal carboxyl group in the LY dehydration was corroborated by the product ion observed at *m/z* 171 (-106 Da), resulting from the neutral loss of the 4-methylphenol moiety of the tyrosine residue (Y_{res}), indicating that the hydroxyl group of the Y side chain was not involved in the LY dehydration. The involvement of the C-terminal carboxyl group in peptide dehydration, namely by C- to N-terminal cyclization, was previously reported when angiotensin II (solid) was submitted to 220 °C for 10 s (Liu, Topchiy, Lehmann & Basile, 2015).

Also in the roasted mixtures Man_3-YL, Man_3-LY, and Man_3-YL-CQA, a peptide derivative formed during roasting by oxidative deamination, i.e. by release of ammonia (-17 Da) at the N-terminal amino acid residue with formation of a ketone group (+16 Da), assigned as (Peptide-NH_3+O), was observed as $[M+H]^+$ ions at *m/z* 294. For both peptides, the respective LC-MS^2 spectra corroborate the occurrence of oxidative deamination at the N-terminal amino acid residue. In the case of (YL-NH_3+O) (Table S4 in Appendix E), the oxidative deamination at the N-terminal Y residue (Y_{res}, 163 Da) is corroborated by the product ion observed at *m/z* 132 (-162 Da, -$(Y\text{-}NH_3+O)_{res}$). In the case of (LY-NH_3+O), the product ion observed at *m/z* 182 (-112 Da), resulting from the loss of $(L\text{-}NH_3+O)_{res}$, corroborates the oxidative deamination at the N-terminal L_{res}. The oxidative deamination of free amino acids was previously reported as occurring during roasting of cocoa beans (Pinto & Chichester, 1966).

In the roasted mixture YL-CQA, it was also observed an ion at *m/z* 571, attributed to $[M+H]^+$ ions of a compound formed by the reaction of two YL molecules with the release of a water molecule, assigned as $(YL)_2$. Accordingly, the respective LC-MS^2 spectrum showed a product ion at *m/z* 295, attributed to $[YL+H]^+$, formed by loss of the another YL_{res}.

c) Hex_nYL and Hex_nLY series

In the roasted mixtures Man_3-YL and Man_3-LY, a series of ions was observed at *m/z* 457, 619, 781, 943, 1105, and 1267. These ions were identified as $[M+H]^+$ ions of Amadori compounds formed by condensation of the N-terminal amino group of the

peptide, YL or LY, with reducing hexose (Hex) sugars, assigned as Hex_nYL and Hex_nLY, respectively. The sugar moiety ranged from a Hex monosaccharide (n=1) to an oligosaccharide composed by 6 Hex units (n=6). Similarly, $[M+H]^+$ ions of Hex_nYL (n=1-5) were identified in the roasted mixture Man_3-YL-CQA, indicating that the formation of the Amadori compounds derived from YL was not affected by the presence of the CQA.

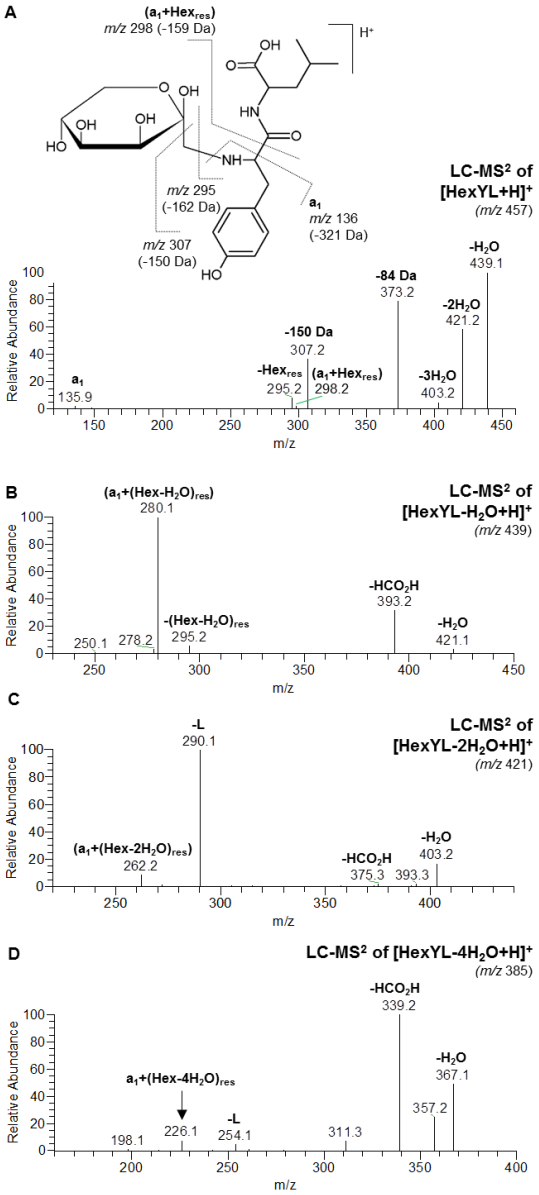

Figure III.25. LC-MS^2 spectra of $[M+H]^+$ ions of (A) HexYL, (B) (HexYL-H_2O), (C) (HexYL-$2H_2O$), and (D) (HexYL-$4H_2O$), acquired from the roasted mixture Man_3-YL.

According to the proposed fragmentation pathways of hexose-derived Amadori compounds (Jerić, Versluis, Horvat & Heck, 2002; Wang, Lu, Liu & He, 2008), the MS^2 spectra of $[M+H]^+$ ions (*m/z* 457) of HexYL (Figure III.25A) and HexLY (Figure III.26A)

show the product ions formed by loss of H_2O (-18 Da), $2H_2O$ (-36 Da) and $3H_2O$ (-54 Da), and by combined loss of $3H_2O$ and CH_2O (-84 Da). The product ions observed at *m/z* 295 (-162 Da, $-Hex_{res}$) and 307 (-150 Da) derived of the cleavage between the sugar and peptide moieties. For HexYL (Figure III.25A), the product ion observed at *m/z* 298 (-159 Da), corresponding to the peptide fragment a_1 bearing a Hex_{res}, corroborates the linkage of the Hex to the Y. Similarly, the product ion observed at *m/z* 248 (-209 Da) in the MS^2 spectrum of $[M+H]^+$ ions of HexLY (Figure III.26A) corroborates the linkage of the Hex to the L. In the case of Hex_nYL and Hex_nLY compounds with $n \geq 2$, the LC-MS^2 fragmentation of the respective $[M+H]^+$ ions gave rise to additional product ions resulting from successive losses of Hex residues (Table S4 in Appendix E).

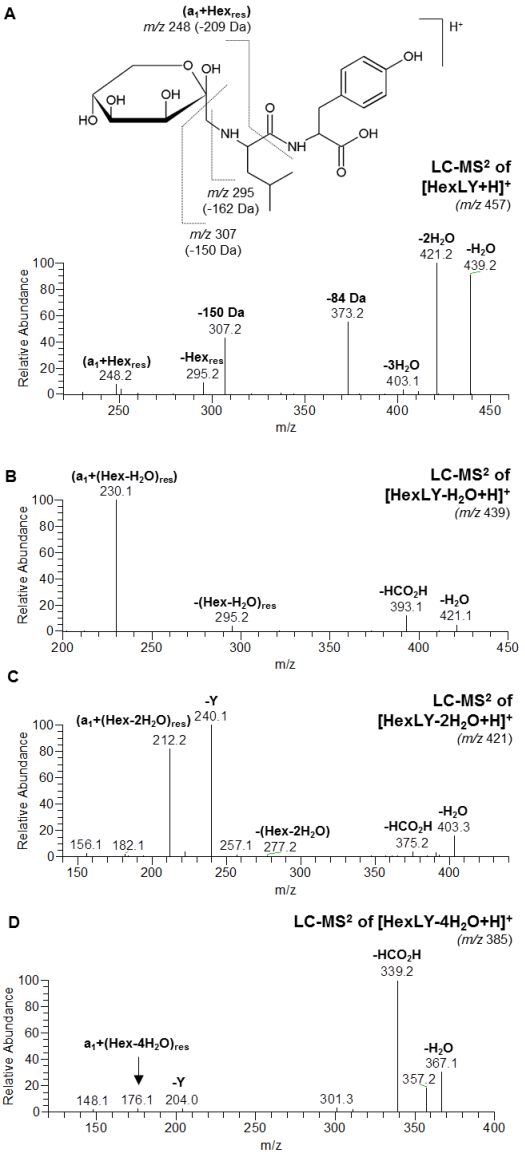

Figure III.26. LC-MS^2 spectra of $[M+H]^+$ ions of (A) HexLY, (B) (HexLY-H_2O), (C) (HexLY-$2H_2O$), and (D) (HexLY-$4H_2O$), acquired from the roasted mixture Man_3-LY.

d) Dehydrated derivatives of Hex_nYL and Hex_nLY

In the roasted mixtures Man_3-YL, Man_3-LY, and Man_3-YL-CQA were also identified different ion series assigned to $[M+H]^+$ ions of dehydrated derivatives of Hex_nYL, or Hex_nLY in the case of the mixture Man_3-LY. More detailed in Table III.14, with the indication of the respective *m/z* values, these derivatives have less one (-18 Da) up to four (-72 Da) water molecules than the corresponding non-modified Hex_nYL, or Hex_nLY.

As shown in Figures III.25B and III.26B, the LC-MS^2 fragmentation of $[M+H]^+$ ions of (HexYL-H_2O) and (HexLY-H_2O) (*m/z* 439) produced a product ion at *m/z* 295 ($[Peptide+H]^+$), resulting from the neutral loss of $(Hex\text{-}H_2O)_{res}$ (-144 Da), which corroborates the presence of a mono-dehydrated Hex in the structure. On the other hand, the linkage of the (Hex-H_2O) to the N-terminal amino acid was corroborated by a product ion observed at *m/z* 280 from (HexYL-H_2O), and at *m/z* 230 from (HexYL-H_2O), corresponding to the peptide fragment a_1 bearing a $(Hex\text{-}H_2O)_{res}$. In the case of the (HexYL-$2H_2O$) and (HexLY-$2H_2O$) compounds formed during roasting of the mixtures, the LC-MS^2 fragmentation of the respective $[M+H]^+$ ions (*m/z* 421) produced an abundant product ion resulting from the loss of the C-terminal amino acid, respectively, L (-131 Da, *m/z* 290) (Figure III.25C) and Y (-181 Da, *m/z* 240) (Figure III.26C). The presence of a di-dehydrated Hex linked to the N-terminal amino acid is corroborated by a product ion observed, respectively, at *m/z* 262 (-159 Da) and 212 (-209 Da), corresponding to $(a_1+Hex\text{-}2H_2O)_{res}$. Similarly, the LC-$MS^2$ spectra of $[M+H]^+$ ions of (HexYL-$3H_2O$) (Table S4 in Appendix E) and (HexLY-$3H_2O$) showed a product ion resulting from the loss of the C-terminal amino acid, observed at *m/z* 272 (-131 Da, -L) and 222 (-181 Da, -Y), respectively. Also, a product ion assigned as $(a_1+Hex\text{-}3H_2O)_{res}$ was observed, respectively, at *m/z* 244 (-159 Da) and 194 (-209 Da), corroborating the linkage of the tri-dehydrated Hex to the N-terminal amino acid.

As observed for Hex_nYL and Hex_nLY compounds, additional product ions resulting from successive losses of Hex residues were observed in the LC-MS^2 spectra of $[M+H]^+$ ions of ($Hex_nYL\text{-}H_2O$), ($Hex_nLY\text{-}H_2O$), ($Hex_nYL\text{-}2H_2O$), and ($Hex_nLY\text{-}2H_2O$) compounds with $n \geq 2$. It is of note that the LC-MS^2 spectra of $[M+H]^+$ ions of ($Hex_2YL\text{-}3H_2O$) (*m/z* 565), eluting between 19.6-21.2 min from the roasted Man_3-YL mixture, showed a product ion at *m/z* 403 (-162 Da), resulting from the loss of an intact Hex_{res}, suggesting the presence of structures with (Hex-$3H_2O$) linked to the N-terminal amino

acid. However, a product ion was observed at *m/z* 439 (-126 Da), probably resulting from the loss of (Hex-3H_2O). This product ion suggest the presence of another structures with (Hex-3H_2O) linked to an intact Hex unit, and not linked directly to the N-terminal amino acid (Table S4 in Appendix E).

As shown in the Figure III.25D, the LC-MS^2 fragmentation of $[M+H]^+$ ions of (HexYL-4H_2O) (*m/z* 385) gave rise to a product ion at *m/z* 254 (-131 Da, -L), and another one at *m/z* 226 (-159 Da). Similarly, the MS^2 spectra of $[M+H]^+$ ions of (HexLY-4H_2O) (*m/z* 385) (Figure III.26D) showed a product ion at *m/z* 204 (-181 Da, -Y), and and another one at *m/z* 176 (-209 Da). The product ions observed at *m/z* 226 and 176, assigned as (a_1+(Hex-4H_2O)$_{res}$), suggest that the loss of four water molecules occurred at the Hex unit linked to the N-terminal amino acid. Distinctly, the LC-MS^2 fragmentation of (Hex$_n$YL-4H_2O) (Table S4 in Appendix E) and (Hex$_n$LY-4H_2O) compounds with $n\geq2$ gave rise to product ions resulting from neutral loss of 126 and 144 Da, probably due to the neutral losses of (Hex-3H_2O) and (Hex-H_2O)$_{res}$, corroborating that the loss of three water molecules occurred at one of the Hex units, and the another water loss occurred at another Hex unit. On the other hand, the presence of an intact peptide moiety in the structures is corroborated by the observation of product ions resulting from a neutral loss of 277 Da from the precursor ion, due to the cleavage of the bond between the amino group and α carbon of the N-terminal amino acid of a non-modified peptide.

e) Hex$_n$(YL)$_2$ and Hex$_n$(LY)$_2$ series

In the LC-MS analysis of the roasted mixtures Man$_3$-YL and Man$_3$-YL-CQA (Table III.14), and Man$_3$-LY (Table S1 in Appendix E), another ion series was observed at *m/z* 733, 895 and 1057, attributed to $[M+H]^+$ ions of compounds bearing two peptide molecules covalently linked with a variable number of Hex units (*n*), with the release of a water molecule for each linkage. These compounds were assigned as Hex$_n$(YL)$_2$ and Hex$_n$(LY)$_2$ (*n*=1-3). The LC-MS^2 spectra of the respective $[M+H]^+$ ions corroborate the presence of two peptide molecules linked to one or more Hex units. In particular, the observation of the product ion at *m/z* 571 (Table S4 in Appendix E), attributed to $[(Peptide)_2+H]^+$, suggests the presence of structures with the two peptide molecules linked together.

f) Derivatives of CQA with and without a Hex_n moiety

The repeating sugar unit of Man_3, mannose ($C_6H_{12}O_6$), and the caffeic acid (CA) moiety ($C_9H_8O_4$) of 5-CQA have the same nominal mass (calculated by adding the mass of the predominant isotope of each element contributing to the molecule rounded to the nearest integer value). For this reason, and attending to the roasting-induced compounds formed from a mixture of 5-CQA and an arabinosyl oligosaccharide (Ara_3) (Section III.3.1), it could be expected the formation of isobaric compounds during roasting of the mixtures containing Man_3 and 5-CQA. Isobaric compounds have the same nominal mass, but different elemental composition, and thus, different accurate mass. For this reason, the assignment of the ions observed after roasting of the mixture Man_3-CQA was supported by the elemental composition obtained from high resolution and high mass accuracy measurements using an Orbitrap-based mass spectrometer (Supplementary Table S2 in Appendix E).

According to that was observed for the roasted mixture Ara_3-CQA (Section III.3.1), non-modified 5-CQA, detected mainly as $[M+Na]^+$ ions at *m/z* 377, was identified after roasting of the mixture Man_3-CQA. Also, derivatives of CQA not bearing a sugar moiety were identified as $[M+Na]^+$ ions, assigned as $[CQA_2+Na]^+$ (*m/z* 713), $[(CQA)CA+Na]^+$ (*m/z* 539), and $[(CQA)CA-H_2O+Na]^+$ (*m/z* 521), respectively. Also, $[M+Na]^+$ ions (*m/z* values in Table III.14) of compounds composed by one or two CQAs covalently linked with a variable number of Hex units (*n*), assigned as $[Hex_nCQA+Na]^+$ (*n*=1-6) and $[Hex_n(CQA)_2+Na]^+$ (*n*=1), as well as Hex_nCQA dehydrated derivatives resulting from the loss of one and three water molecules, assigned as $[Hex_nCQA-H_2O+Na]^+$ and $[Hex_nCQA-3H_2O+Na]^+$, were also identified after roasting of the mixture Man_3-CQA. Furthermore, $[M+Na]^+$ ions of CQA derivatives bearing a sugar moiety but composed exclusively by the quinic or caffeic acid moiety of CQAs, or their dehydrated derivatives, were also identified, assigned as $[Hex_nQA+Na]^+$ (*n*=1-3), $[Hex_nQA-H_2O+Na]^+$ (*n*=2-3), and $[Hex_nCA-H_2O+Na]^+$ (*n*=1-2). As highlighted in Table III.14, some of these compounds identified in the roasted Man_3-CQA are isobaric compounds, and thus only differentiable based on the accurate masses found by LTQ-Orbitrap (Table S2 in Appendix E).

Distinctly to that was observed for the roasted mixture Man_3-CQA, Hex_nCQA compounds, or their dehydrated derivatives, were not identified from the roasted mixture Man_3-YL-CQA. This can be related with the high reactivity between carbonyl and amino groups. Probably, most the carbonyl groups of reducing sugars reacted with the amino

groups of the YL peptide, and thus few carbonyl groups were free to react with CQA molecules to yield Hex_nCQA compounds. However, it can not be excluded the possibility of HPLC-MS signal suppression effect, since Hex_nCQA and Hex_nYL compounds had similar retention times under the HPLC conditions used.

g) YL(CQA) compound

In the roasted mixture YL-CQA, beyond $(YL)_n$ and $(CQA)_n$ (n=1-2), it was also identified a roasting-induced compound formed by the reaction a peptide and a CQA molecule with release of a water molecule, assigned as YL(CQA). It was detected mainly as $[M+H]^+$ ions at *m/z* 631, eluting at RT 19.5 and 27.8 min. For both RTs, the respective ESI-MS^2 spectrum (RT 19.5 min in Figure III.27A) showed a product ion at *m/z* 337, resulting from the neutral loss of YL (-294 Da), and another at *m/z* 277, resulting from the loss of CQA (-354 Da), corroborating the linkage between an intact peptide (YL) and an intact CQA molecule. On the other hand, the observation of a product ion at *m/z* 326 (-305 Da), probably resulting from the combined loss of a QA_{res} (-174 Da) and L (-131 Da), suggested that the linkage between YL and CQA involved the N-terminal amino acid (Y) of YL and the CA moiety of CQA. Assuming that the structure have an intact peptide (YL) and an intact CQA that are linked by a bond between the Y amino acid and CA moiety, the product ion observed at *m/z* 469, probably resulting from the loss of CA_{res} (-162 Da), suggested that this is a cyclic structure, or that the product ion at *m/z* 469 is resulted from an internal fragmentation with an arrangement of the structure. Instead, it can not be excluded the possibility of different structures have been eluted at the same RT. The roasting-induced compound YL(CQA) was also identified from the roasted mixture Man_3-YL-CQA, observed as $[M+H]^+$ ions (*m/z* 631) at RT 28.0 min.

h) Hex_n(YL)CQA series

In the roasted mixture Man_3-YL-CQA, it was also observed an ion series at *m/z* 793, 955, and 1117, attributed to $[M+H]^+$ ions of compounds formed by the reaction of a peptide (YL), a CQA molecule, and a sugar composed by one, two, or three Hex units (n), with the release of a water molecule for each linkage. These compounds were assigned as Hex_n(YL)CQA (n=1-3). The interpretation of the respective LC-MS^n spectra is hampered by the fact of the Hex monosaccharide and the CA moiety of CQA have the

same nominal mass. In specific, product ions with the same nominal *m/z* value can result from the loss of Hex_{res} and CA_{res} (-162 Da), and the loss of Hex and CA (-180 Da).

In Figure III.27B is shown the LC-MS^2 spectrum of Hex(YL)CQA (*m/z* 793). The product ions observed at *m/z* 634 (-159 Da), attributed to the peptide fragment a_1 bearing a $(HexCQA)_{res}$ moiety, at *m/z* 516 (-277 Da), resulting from the loss of (YL-NH_3), and at *m/z* 295, attributed to $[YL+H]^+$, suggested that the peptide is only linked through the amino group of the N-terminal amino acid (Y) to the HexCQA moiety. Considering the reactivity between carbonyl and amino groups, it can be expected that the Y amino group has been reacted with the carbonyl group of Hex. On the other hand, the product ion observed at *m/z* 342, attributed to $[(CAHex)_{res}+NH_3+H]^+$, suggested that the CQA molecule is linked through the CA moiety to the Hex unit. However, the presence of other isomeric structures can not be excluded. The product ions observed at *m/z* 631 and 613 can be resulted from neutral losses of Hex_{res}/CA_{res} (-162 Da) and Hex/CA (-180 Da), respectively.

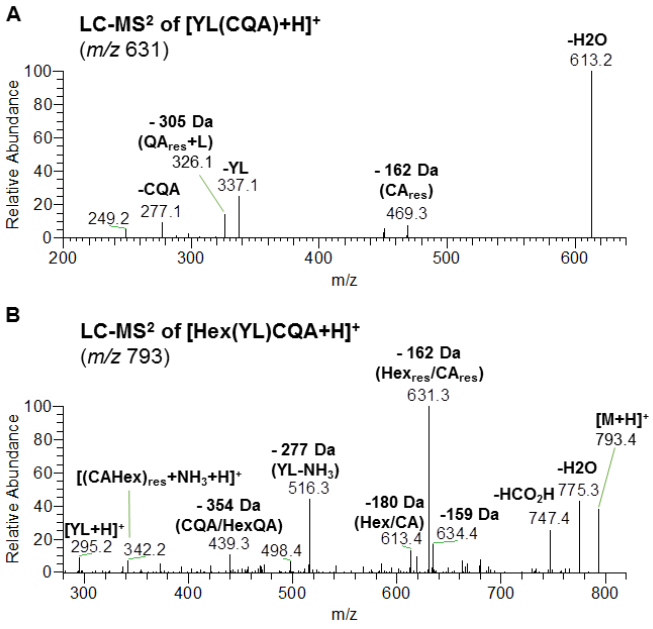

Figure III.27. LC-MS^2 spectra of $[M+H]^+$ ions of (A) (YL)CQA acquired from the roasted mixture YL-CQA, and of (B) Hex(YL)CQA acquired from the roasted mixture Man_3-YL-CQA.

III.3.2.3. Concluding remarks

The dry thermal treatment of the model mixture containing (β1→4)-D-mannotriose (Man_3), an oligosaccharide structurally related to the coffee galactomannans, and 5-*O*-

caffeoylquinic acid (5-CQA), the most abundant chlorogenic acid in green coffee beans, promoted the formation of hexose oligosaccharides with a degree of polymerization up to at least 9, and a diversity of new glycosidic linkages, not observed when Man_3 was individually submitted to the same thermal conditions. Similarly, a diversity of new glycosidic linkages was observed after thermal treatment of the model mixtures containing Man_3 and one of two acids, malic and citric, present in green coffee beans. These results led to the conclusion that acids catalyse non-enzymatic transglycosylation reactions induced by dry thermal treatment.

The diversity of new glycosidic linkages identified after thermal treatment of the model mixtures containing Man_3 and one of two isomeric peptides, YL and LY, used as models of the proteins, as well as the mixture containing Man_3, 5-CQA and YL were lower than those observed after thermal treatment of the mixture Man_3-CQA. This led to the conclusion that Maillard reaction (MR) prevents non-enzymatic transglycosylation reactions induced by dry thermal treatment. According to the classical MR mechanism, in the presence of peptides/proteins, reducing sugar residues react with amino groups of peptides/proteins, decreasing the number of reducing residues available to be directly involved in the non-enzymatic transglycosylation reactions. This can explain a low extent of non-enzymatic transglycosylation reactions during coffee roasting, particularly in coffee bean regions rich in proteins, although polysaccharides account for about 50% of green coffee beans' dry weight.

CHAPTER IV. GENERAL CONCLUDING REMARKS

Polysaccharides are the major components of green coffee beans, comprising about 50% of their dry weight. The most abundant coffee polysaccharides are galactomannans and type II arabinogalactans. During the roasting process, galactomannans and arabinogalactans undergo structural modifications, and participate in reactions involving other coffee bean components, such as proteins and chlorogenic acids. On the basis of the current knowledge, it can be stated that these modifications are far to be completely elucidated due to the diversity and the higher complexity of the new structures formed. Galactomannans and arabinogalactans when react with proteins, chlorogenic acids, and sucrose originate high molecular weight nitrogenous brown-coloured compounds, known as melanoidins. Despite their exact structures remain unknown, several biological activities and beneficial health effects have been attributed to coffee melanoidins. In this context, the present work was conducted with the main aim of obtaining an overall view of the structural modifications in coffee polysaccharides promoted by roasting, contributing to the elucidation of the structures and formation mechanisms of melanoidins.

To achieve the proposed aim, due to the complexity of the coffee matrix, simple model systems were submitted to dry thermal treatment, mimicking the coffee roasting process. In terms of model compounds, two oligosaccharides were used: (β1→4)-D-mannotriose (Man_3), structurally related to the backbone of galactomannans, and (α1→5)-L-arabinotriose (Ara_3), structurally related to the arabinose (Ara) side chains of arabinogalactans. These oligosaccharides were subjected to dry thermal treatment alone or in mixtures, namely mixtures with 5-*O*-caffeoylquinic acid (5-CQA), the most abundant chlorogenic acid in green coffee beans, and one of two dipeptides composed by tyrosine (Y) and leucine (L), used as models of the proteins. As hydroxyl radicals ($HO^{\bullet}$) may be involved in the oxidation of coffee arabinogalactans, the $HO^{\bullet}$-induced oxidation of Ara_3 was also studied. The identification of the modifications induced by thermal treatment (dry) or oxidation induced by $HO^{\bullet}$ (in solution) was performed using mostly mass spectrometry (MS)-based analytical strategies. In order to validate the conclusions achieved with the model systems, roasted coffee polysaccharide-rich samples were also analysed.

The detailed study of the compounds formed when Ara_3 was submitted in different experiments to dry thermal treatments at 200 ºC, and oxidation induced by $HO^{\bullet}$ formed under conditions of Fenton reaction (Fe^{2+}/H_2O_2) suggests that a huge diversity of compounds can be formed during the roasting from the Ara side chains of coffee

arabinogalactans. Also, identical compounds were formed under the different conditions studied, supporting the hypothesis of the occurrence of HO•-induced oxidation reactions during coffee roasting involving the arabinogalactans.

When Ara_3 was mixed with Man_3 in different molar proportions, and the mixtures submitted to dry thermal treatments at 200 ºC, oligo- and polysaccharides composed by both pentose and hexose residues were formed, corroborating the hypothesis of the occurrence of non-enzymatic transglycosylation reactions between arabinogalactans and galactomannans during coffee roasting. This hypothesis was also supported by data obtained from spent coffee grounds (SCG) and instant coffee samples, suggesting the presence of hybrid polysaccharides derived from arabinogalactans and galactomannans. On the other hand, the use of mixtures containing different proportions of each oligosaccharide, submitted to different extent of thermal treatment led to the conclusion that, depending on the roasting conditions and the distribution of the polysaccharides in the bean cell walls, different non-hybrid and hybrid structures can be formed during coffee roasting. This may explain to the heterogeneity of melanoidin structures formed during coffee roasting.

The hypothesis that Ara derived from arabinogalactan side chains is a possible binding site for the chlorogenic acid derivatives in melanoidin structures was corroborated by using of a model mixture containing equimolar amounts of Ara_3 and 5-CQA. The dry thermal treatment of this mixture, either at 175 ºC or 200 ºC, promoted the formation of hybrid structures composed by CQA molecules covalently linked to a variable number of sugar residues. The same type of structures, probably formed during the roasting process, was identified in samples obtained from SCG, opening new perspectives for the identification of chlorogenic acid-Ara hybrid domains in melanoidin structures.

On the other hand, the dry thermal treatment at 175 °C of the mixture Ara_3-CQA, but also of the mixture Man_3-CQA, led to the formation of oligosaccharides having a higher degree of polymerization than that of the oligosaccharide in the starting mixture, and new types of glycosidic linkages. These new oligosaccharides are indicatives of the occurrence of non-enzymatic transglycosylation reactions, not occurring when Man_3 was individually submitted to the same thermal treatment at 175 °C. These results allowed to infer that 5-CQA acts as a catalyst of non-enzymatic transglycosylation reactions. However, when model mixtures containing a peptide, even also containing 5-CQA, were subjected to the same treatment at 175 °C, the extent of transglycosylation reactions

decreased. This outcome can explain the low extent of non-enzymatic transglycosylation reactions during coffee roasting, although polysaccharides are the major components of the coffee beans. Accordingly, the reducing terminal residues of coffee polysaccharides react preferentially with the amino groups of proteins by Maillard reaction, decreasing the number of reducing residues that could be directly involved in the non-enzymatic transglycosylation reactions. This type of structures formed by Maillard reaction of polysaccharides and proteins can constitute the skeleton of coffee melanoidins. Probably, the chemical composition of this skeleton varies with the original location in the coffee bean of the polysaccharides and proteins involved in the formation of the melanoidins, and also with the roasting conditions. This may contribute to the heterogeneity of melanoidins structures formed during coffee roasting. However, this diversity still requires to be exploited concerning their structural features and biological activities.

The identification of the structural modifications in coffee polysaccharides promoted by roasting pave the way to the understanding of the mechanisms of formation of melanoidins and structure-activity relationships of these compounds. However, the complete structural characterization of coffee melanoidins, as well as the understanding of the formation mechanisms and structure-activity relationships remain a big challenge. To date, it is known that they do not have a repetitive structure. There are coffee melanoidin populations with distinct chemical composition. As supported with the date obtained in this work, this heterogeneity may be related with the original location in the beans of the different coffee components involved in the formation of the melanoidins, and the roasting conditions used. In future studies on the structural characterization of coffee melanoidins, it is needed, first, to develop/optimize procedures for isolation and purification of different melanoidin populations. Practical considerations must be taken into account for analysis of the purified coffee melanoidins through analytical approaches based on MS using soft ionization methods, either electrospray ionization (ESI) or matrix assisted laser desorption/ionization (MALDI). Due to their high molecular weight, entire melanoidins are not suitable to be analysed. Melanoidins need to be cleaved into lower weight melanoidin fragments prior MS analysis. Attending that carbohydrates, proteins, and chlorogenic acids are involved in the formation of coffee melanoidins, it is need to develop a method for specific controlled cleavages as, for example, enzyme cocktails consisting of different type of enzymes, namely glycosidases and proteases, to cleave coffee melanoidins into lower weight melanoidin fragments able to be analysed by MS.

From a more general point of view, the results obtained in this work provide new information on the modification of oligo- and polysaccharides subjected to dry thermal treatment, but also to oxidation induced by $HO^{\bullet}$, facilitating their identification in complex matrixes like coffee, but also in other carbohydrate-rich samples. For example, the results obtained can be extended to starch submitted to the bakery processing. The reliability of the models used in this work was corroborated by the identification of the same type of modifications in coffee polysaccharide samples.

CHAPTER V. REFERENCES

Ali, B. H., Ziada, A., & Blunden, G. (2009). Biological effects of gum arabic: A review of some recent research. *Food and Chemical Toxicology, 47*, 1-8.

Arts, S. J. H. F., Mombarg, E. J. M., van Bekkum, H., & Sheldon, R. A. (1997). Hydrogen peroxide and oxygen in catalytic oxidation of carbohydrates and related compounds. *Synthesis, 6*, 597-613.

Asam, M. R., & Glish, G. L. (1997). Tandem mass spectrometry of alkali cationized polysaccharides in a quadrupole ion trap. *Journal of the American Society for Mass Spectrometry, 8*, 987-995.

Aspinall, G. O. (1959). Structural chemistry of the hemicelluloses. In L. W. Melville (Ed.), *Advances in Carbohydrate Chemistry*, vol. 14 (pp. 429-468): Academic Press.

Bähre, F., & Maier, H. (1996). Electrophoretic clean-up of organic acids from coffee for the GC/MS analysis. *Fresenius' Journal of Analytical Chemistry, 355*, 190-193.

Bauer, S. (2012). Mass spectrometry for characterizing plant cell wall polysaccharides. *Frontiers in Plant Science, 3*, 45.

Bekedam, E. K., De Laat, M. P. F. C., Schols, H. A., Van Boekel, M. A. J. S., & Smit, G. (2007). Arabinogalactan proteins are incorporated in negatively charged coffee brew melanoidins. *Journal of Agricultural and Food Chemistry, 55*, 761-768.

Bekedam, E. K., Loots, M. J., Schols, H. A., Van Boekel, M. A. J. S., & Smit, G. (2008a). Roasting effects on formation mechanisms of coffee brew melanoidins. *Journal of Agricultural and Food Chemistry, 56*, 7138-7145.

Bekedam, E. K., Schols, H. A., van Boekel, M. A. J. S., & Smit, G. (2006). High molecular weight melanoidins from coffee brew. *Journal of Agricultural and Food Chemistry, 54*, 7658-7666.

Bekedam, E. K., Schols, H. A., Van Boekel, M. A. J. S., & Smit, G. (2008b). Incorporation of chlorogenic acids in coffee brew melanoidins. *Journal of Agricultural and Food Chemistry, 56*, 2055-2063.

Berger, O., McBride, R., Razi, N., & Paulson, J. (2008). Symbol notation extension for pathogen polysaccharides (proposal) (http://glycomics.scripps.edu/coreD/PGAnomenclature.pdf). In): The Scripps Research Institute, Consortium for Functional Glycomics, Core D.

Bhattacharya, S. (2015). *Conventional and advanced food processing technologies*: John Wiley & Sons, Ltd.

Biemann, K. (1988). Contributions of mass spectrometry to peptide and protein structure. *Biomedical and Environmental Mass Spectrometry, 16*, 99-111.

Bradbury, A. G. W., & Halliday, D. J. (1990). Chemical structures of green coffee bean polysaccharides. *Journal of Agricultural and Food Chemistry, 38*, 389-392.

Casal, E., Lebrón-Aguilar, R., Moreno, F. J., Corzo, N., & Quintanilla-López, J. E. (2010). Selective linkage detection of *O*-sialoglycan isomers by negative electrospray ionization ion trap tandem mass spectrometry. *Rapid Communications in Mass Spectrometry, 24*, 885-893.

Ciucanu, I., & Kerek, F. (1984). A simple and rapid method for the permethylation of carbohydrates. *Carbohydrate Research, 131*, 209-217.

Clifford, M. N. (2000). Chlorogenic acids and other cinnamates – nature, occurrence, dietary burden, absorption and metabolism. *Journal of the Science of Food and Agriculture, 80*, 1033-1043.

Clifford, M. N., Johnston, K. L., Knight, S., & Kuhnert, N. (2003). Hierarchical scheme for LC-MS^n identification of chlorogenic acids. *Journal of Agricultural and Food Chemistry, 51*, 2900-2911.

Coelho, C., Ribeiro, M., Cruz, A. C. S., Domingues, M. R. M., Coimbra, M. A., Bunzel, M., & Nunes, F. M. (2014). Nature of phenolic compounds in coffee melanoidins. *Journal of Agricultural and Food Chemistry, 62*, 7843-7853.

da Costa, E. V., Moreira, A. S. P., Nunes, F. M., Coimbra, M. A., Evtuguin, D. V., & Domingues, M. R. M. (2012). Differentiation of isomeric pentose disaccharides by electrospray ionization tandem mass spectrometry and discriminant analysis. *Rapid Communications in Mass Spectrometry, 26*, 2897-2904.

Daglia, M., Papetti, A., Aceti, C., Sordelli, B., Spini, V., & Gazzani, G. (2007). Isolation and determination of α-dicarbonyl compounds by RP-HPLC-DAD in green and roasted Coffee. *Journal of Agricultural and Food Chemistry, 55*, 8877-8882.

Dass, C. (2007). *Fundamentals of contemporary mass spectrometry*. Hoboken: John Wiley & Sons.

de Lederkremer, R. M., & Marino, C. (2003). Acids and other products of oxidation of sugars. *Advances in Carbohydrate Chemistry and Biochemistry, 58*, 199-306.

Dey, P. M. (1978). Biochemistry of plant galactomannans. *Advances in Carbohydrate Chemistry and Biochemistry, 35*, 341-376.

Domingues, M. R. M., Domingues, P., Reis, A., Fonseca, C., Amado, F. M. L., & Ferrer-Correia, A. J. V. (2003). Identification of oxidation products and free radicals of

tryptophan by mass spectrometry. *Journal of the American Society for Mass Spectrometry, 14*, 406-416.

Domon, B., & Costello, C. E. (1988). A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates *Glycoconjugate Journal, 5*, 397-409.

Duan, J., & Kasper, D. L. (2011). Oxidative depolymerization of polysaccharides by reactive oxygen/nitrogen species. *Glycobiology, 21*, 401-409.

Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., & Smith, F. (1956). Colorimetric method for determination of sugars and related substances. *Analytical Chemistry, 28*, 350-356.

Ekman, R., Silberring, J., Westman-Brinkmalm, A. M., Kraj, A., Desiderio, D. M., & Nibbering, N. M. (2009). *Mass Spectrometry: Instrumentation, Interpretation, and Applications*. New Jersey: John Wiley & Sons, Inc.

Eskin, N. A. M., Ho, C.-T., & Shahidi, F. (2013). Chapter 6 - Browning reactions in foods. In N. A. M. Eskin & F. Shahidi (Eds.), *Biochemistry of Foods (Third Edition)*, (pp. 245-289). San Diego: Academic Press.

Fang, N., Yu, S., & Prior, R. L. (2002). LC/MS/MS characterization of phenolic constituents in dried plums. *Journal of Agricultural and Food Chemistry, 50*, 3579-3585.

Fang, T. T., & Bendiak, B. (2007). The stereochemical dependence of unimolecular dissociation of monosaccharide-glycolaldehyde anions in the gas phase: a basis for assignment of the stereochemistry and anomeric configuration of monosaccharides in oligosaccharides by mass spectrometry via a key discriminatory product ion of disaccharide fragmentation, *m/z* 221. *Journal of the American Chemical Society, 129*, 9721-9736.

Fang, T. T., Zirrolli, J., & Bendiak, B. (2007). Differentiation of the anomeric configuration and ring form of glucosyl-glycolaldehyde anions in the gas phase by mass spectrometry: isomeric discrimination between *m/z* 221 anions derived from disaccharides and chemical synthesis of *m/z* 221 standards. *Carbohydrate Research, 342*, 217-235.

Farah, A., de Paulis, T., Moreira, D. P., Trugo, L. C., & Martin, P. R. (2006). Chlorogenic acids and lactones in regular and water-decaffeinated Arabica coffees. *Journal of Agricultural and Food Chemistry, 54*, 374-381.

Farah, A., de Paulis, T., Trugo, L. C., & Martin, P. R. (2005). Effect of roasting on the formation of chlorogenic acid lactones in coffee. *Journal of Agricultural and Food Chemistry, 53*, 1505-1513.

Faure, A. M., Andersen, M. L., & Nyström, L. (2012). Ascorbic acid induced degradation of beta-glucan: Hydroxyl radicals as intermediates studied by spin trapping and electron spin resonance spectroscopy. *Carbohydrate Polymers, 87*, 2160-2168.

Faure, A. M., Werder, J., & Nyström, L. (2013). Reactive oxygen species responsible for beta-glucan degradation. *Food Chemistry, 141*, 589-596.

Fiddler, W., Parker, W. E., Wasserman, A. E., & Doerr, R. C. (1967). Thermal decomposition of ferulic acid. *Journal of Agricultural and Food Chemistry, 15*, 757-761.

Fischer, M., Reimann, S., Trovato, V., & Redgwell, R. J. (2001). Polysaccharides of green Arabica and Robusta coffee beans. *Carbohydrate Research, 330*, 93-101.

Fogliano, V., & Morales, F. J. (2011). Estimation of dietary intake of melanoidins from coffee and bread. *Food & Function, 2*, 117-123.

Fonseca, C., Domingues, M. R. M., Simões, C., Amado, F., & Domingues, P. (2009). Reactivity of Tyr–Leu and Leu–Tyr dipeptides: identification of oxidation products by liquid chromatography–tandem mass spectrometry. *Journal of Mass Spectrometry, 44*, 681-693.

Frank, O., Blumberg, S., Kunert, C., Zehentbauer, G., & Hofmann, T. (2007). Structure determination and sensory analysis of bitter-tasting 4-vinylcatechol oligomers and their identification in roasted coffee by means of LC-MS/MS. *Journal of Agricultural and Food Chemistry, 55*, 1945-1954.

Fry, S. C. (1998). Oxidative scission of plant cell wall polysaccharides by ascorbate-induced hydroxyl radicals. *Biochemical Journal, 332*, 507-515.

Galli, V., & Barbas, C. (2004). Capillary electrophoresis for the analysis of short-chain organic acids in coffee. *Journal of Chromatography A, 1032*, 299-304.

Ginz, M., Balzer, H. H., Bradbury, A. G. W., & Maier, H. G. (2000). Formation of aliphatic acids by carbohydrate degradation during roasting of coffee. *European Food Research and Technology, 211*, 404-410.

Gniechwitz, D., Reichardt, N., Meiss, E., Ralph, J., Steinhart, H., Blaut, M., & Bunzel, M. (2008a). Characterization and fermentability of an ethanol soluble high molecular weight coffee fraction. *Journal of Agricultural and Food Chemistry, 56*, 5960-5969.

Gniechwitz, D., Reichardt, N., Ralph, J., Blaut, M., Steinhart, H., & Bunzel, M. (2008b). Isolation and characterisation of a coffee melanoidin fraction. *Journal of the Science of Food and Agriculture, 88*, 2153-2160.

Golon, A., & Kuhnert, N. (2012). Unraveling the chemical composition of caramel. *Journal of Agricultural and Food Chemistry, 60*, 3266-3274.

Gonçalves, V. M. F., Evtuguin, D. V., & Domingues, M. R. M. (2008). Structural characterization of the acetylated heteroxylan from the natural hybrid *Paulownia elongata/Paulownia fortunei*. *Carbohydrate Research, 343*, 256-266.

Goodman, B. A., Pascual, E. C., & Yeretzian, C. (2011). Real time monitoring of free radical processes during the roasting of coffee beans using electron paramagnetic resonance spectroscopy. *Food Chemistry, 125*, 248-254.

Guay, D. F., Cole, B. J. W., Fort, R. C., Hausman, M. C., Genco, J. M., Elder, T. J., & Overly, K. R. (2001). Mechanisms of oxidative degradation of carbohydrates during oxygen delignification. II. Reaction of photochemically generated hydroxyl radicals with methyl β-cellobioside. *Journal of Wood Chemistry and Technology, 21*, 67-79.

Gutiérrez, C., Ortolá, M. D., Chiralt, A., & Fito, P. (1993). SEM analysis of roasted coffee porosity (Spanish). In *Proceedings of 15th International Colloquium on the Chemistry of Coffee*, (pp. 661-671). Paris: ASIC - Association for Science and Information on Coffee.

Gutteridge, J. M. (1987). Ferrous-salt-promoted damage to deoxyribose and benzoate. The increased effectiveness of hydroxyl-radical scavengers in the presence of EDTA. *Biochemical Journal, 243*, 709-714.

Hegele, J., Münch, G., & Pischetsrieder, M. (2009). Identification of hydrogen peroxide as a major cytotoxic component in Maillard reaction mixtures and coffee. *Molecular Nutrition & Food Research, 53*, 760-769.

Hodge, J. E. (1953). Dehydrated foods, chemistry of browning reactions in model systems. *Journal of Agricultural and Food Chemistry, 1*, 928-943.

Hofmeister, G. E., Zhou, Z., & Leary, J. A. (1991). Linkage position determination in lithium-cationized disaccharides: tandem mass spectrometry and semiempirical calculations. *Journal of the American Chemical Society, 113*, 5964-5970.

Illy, A., Illy, E., Macrae, R., Petracco, M., Sondahl, M. R., Valussi, S., & Viani, R. (1995). *Espresso coffee: the chemistry of quality*. London: Academic Press.

Isbell, H. S., & Frush, H. L. (1987). Mechanisms for hydroperoxide degradation of disaccharides and related compounds. *Carbohydrate Research, 161*, 181-193.

IUPAC (1976). Nomenclature of cyclitols. *Biochemical Journal, 153*, 23-31.

Jaiswal, R., Matei, M. F., Golon, A., Witt, M., & Kuhnert, N. (2012). Understanding the fate of chlorogenic acids in coffee roasting using mass spectrometry based targeted and non-targeted analytical strategies. *Food & Function, 3*, 976-984.

Jaiswal, R., Matei, M. F., Subedi, P., & Kuhnert, N. (2014). Does roasted coffee contain chlorogenic acid lactones or/and cinnamoylshikimate esters? *Food Research International, 61*, 214-227.

Jaiswal, R., Matei, M. F., Ullrich, F., & Kuhnert, N. (2011). How to distinguish between cinnamoylshikimate esters and chlorogenic acid lactones by liquid chromatography–tandem mass spectrometry. *Journal of Mass Spectrometry, 46*, 933-942.

Jerić, I., Versluis, C., Horvat, Š., & Heck, A. J. R. (2002). Tracing glycoprotein structures: electrospray ionization tandem mass spectrometric analysis of sugar–peptide adducts. *Journal of Mass Spectrometry, 37*, 803-811.

Kane, R. W., & Timpa, J. D. (1992). A high-performance liquid chromatography study of D-cellobiose degradation under Fenton conditions. *Journal of Carbohydrate Chemistry, 11*, 779-797.

Karas, M., & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry, 60*, 2299-2301.

Konda, C., Bendiak, B., & Xia, Y. (2012). Differentiation of the stereochemistry and anomeric configuration for 1-3 linked disaccharides via tandem mass spectrometry and ^{18}O-labeling. *Journal of the American Society for Mass Spectrometry, 23*, 347-358.

Kroh, L. W. (1994). Caramelisation in food and beverages. *Food Chemistry, 51*, 373-379.

Kumazawa, K., & Masuda, H. (2003). Investigation of the change in the flavor of a coffee drink during heat processing. *Journal of Agricultural and Food Chemistry, 51*, 2674-2678.

Leloup, V., Louvrier, A., & Liardon, R. (1995). Degradation mechanisms of chlorogenic acids during roasting. In *Proceedings of the 16th ASIC Colloquium (Kyoto)*, (pp. 192-198). Paris: ASIC - Association for Science and Information on Coffee.

Levěque, P., Godechal, Q., & Gallez, B. (2008). EPR spectroscopy and imaging of free radicals in food. *Israel Journal of Chemistry, 48*, 19-26.

Liu, C., Topchiy, E., Lehmann, T., & Basile, F. (2015). Characterization of the dehydration products due to thermal decomposition of peptides by liquid chromatography-tandem mass spectrometry. *Journal of Mass Spectrometry, 50*, 625-632.

Maier, H. G., Diemair, W., & Ganssmann, J. (1968). Isolation and characterization of brown roast compounds of coffee (In German). *Zeitschrift für Lebensmitteluntersuchung und -Forschung A, 137*, 287-292.

Marraccini, P., Rogers, W. J., Caillet, V., Deshayes, A., Granato, D., Lausanne, F., Lechat, S., Pridmore, D., & Pétiard, V. (2005). Biochemical and molecular characterization of α-D-galactosidase from coffee beans. *Plant Physiology and Biochemistry, 43*, 909-920.

McMurry, J. (2008). *Organic chemistry - International student edition* (7th ed.). Belmont, CA: Thomson Brooks/Cole.

Mischnick, P. (2012). Mass spectrometric characterization of oligo- and polysaccharides and their derivatives. In M. Hakkarainen (Ed.), *Mass Spectrometry of Polymers – New Techniques*, vol. 248 (pp. 105-174). Heidelberg: Springer

Montavon, P., Mauron, A.-F., & Duruz, E. (2003). Changes in green coffee protein profiles during roasting. *Journal of Agricultural and Food Chemistry, 51*, 2335-2343.

Moody, G. J. (1963). The action of Fenton's reagent on carbohydrates. *Tetrahedron, 19*, 1705-1710.

Moon, J.-K., & Shibamoto, T. (2009). Role of roasting conditions in the profile of volatile flavor chemicals formed from coffee beans. *Journal of Agricultural and Food Chemistry, 57*, 5823-5831.

Moon, J.-K., Yoo, H. S., & Shibamoto, T. (2009). Role of roasting conditions in the level of chlorogenic acid content in coffee beans: correlation with coffee acidity. *Journal of Agricultural and Food Chemistry, 57*, 5365-5369.

Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., Simões, J., & Domingues, M. R. M. (2011). Evaluation of the effect of roasting on the structure of coffee galactomannans using model oligosaccharides. *Journal of Agricultural and Food Chemistry, 59*, 10078-10087.

Moreira, A. S. P., Nunes, F. M., Domingues, M. R., & Coimbra, M. A. (2012). Coffee melanoidins: structures, mechanisms of formation and potential health impacts. *Food & Function, 3*, 903-915.

Mueller, U., Sauer, T., Weigel, I., Pichner, R., & Pischetsrieder, M. (2011). Identification of H_2O_2 as a major antimicrobial component in coffee. *Food & Function, 2*, 265-272.

Murray, K. K., Boyd, R. K., Eberlin, M. N., Langley, G. J., Li, L., & Naito, Y. (2013). Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013). *Pure and Applied Chemistry, 85*, 1515-1609.

Navarini, L., Gilli, R., Gombac, V., Abatangelo, A., Bosco, M., & Toffanin, R. (1999). Polysaccharides from hot water extracts of roasted *Coffea arabica* beans: isolation and characterization. *Carbohydrate Polymers, 40*, 71-81.

Nebesny, E., & Budryn, G. (2006). Changes in free radicals content in coffee beans throughout convective roasting and microwaving and during storage. *Deutsche Lebensmittel-Rundschau, 102*, 526-530.

Nunes, F. M., & Coimbra, M. A. (2001). Chemical characterization of the high molecular weight material extracted with hot water from green and roasted arabica coffee. *Journal of Agricultural and Food Chemistry, 49*, 1773-1782.

Nunes, F. M., & Coimbra, M. A. (2002a). Chemical characterization of galactomannans and arabinogalactans from two arabica coffee infusions as affected by the degree of roast. *Journal of Agricultural and Food Chemistry, 50*, 1429-1434.

Nunes, F. M., & Coimbra, M. A. (2002b). Chemical characterization of the high molecular weight material extracted with hot water from green and roasted robusta coffees as affected by the degree of roast. *Journal of Agricultural and Food Chemistry, 50*, 7046-7052.

Nunes, F. M., & Coimbra, M. A. (2007). Melanoidins from coffee infusions. Fractionation, chemical characterization, and effect of the degree of roast. *Journal of Agricultural and Food Chemistry, 55*, 3967-3977.

Nunes, F. M., & Coimbra, M. A. (2010). Role of hydroxycinnamates in coffee melanoidin formation. *Phytochemistry Reviews, 9*, 171-185.

Nunes, F. M., Cruz, A. C. S., & Coimbra, M. A. (2012). Insight into the mechanism of coffee melanoidin formation using modified "in bean" models. *Journal of Agricultural and Food Chemistry, 60*, 8710-8719.

Nunes, F. M., Domingues, M. R., & Coimbra, M. A. (2005). Arabinosyl and glucosyl residues as structural features of acetylated galactomannans from green and roasted coffee infusions. *Carbohydrate Research, 340*, 1689-1698.

Nunes, F. M., Reis, A., Domingues, M. R. M., & Coimbra, M. A. (2006). Characterization of galactomannan derivatives in roasted coffee beverages. *Journal of Agricultural and Food Chemistry, 54*, 3428-3439

Nunes, F. M., Reis, A., Silva, A. M. S., Domingues, M. R. M., & Coimbra, M. A. (2008). Rhamnoarabinosyl and rhamnoarabinoarabinosyl side chains as structural features of coffee arabinogalactans. *Phytochemistry, 69*, 1573-1585.

Oosterveld, A., Coenen, G. J., Vermeulen, N. C. B., Voragen, A. G. J., & Schols, H. A. (2004). Structural features of acetylated galactomannans from green *Coffea arabica* beans. *Carbohydrate Polymers, 58*, 427-434.

Oosterveld, A., Harmsen, J. S., Voragen, A. G. J., & Schols, H. A. (2003a). Extraction and characterization of polysaccharides from green and roasted *Coffea arabica* beans. *Carbohydrate Polymers, 52*, 285-296.

Oosterveld, A., Voragen, A. G. J., & Schols, H. A. (2003b). Effect of roasting on the carbohydrate composition of *Coffea arabica* beans. *Carbohydrate Polymers, 54*, 183-192.

Ovalle, R., Soll, C. E., Lim, F., Flanagan, C., Rotunda, T., & Lipke, P. N. (2001). Systematic analysis of oxidative degradation of polysaccharides using PAGE and HPLC–MS. *Carbohydrate Research, 330*, 131-139.

Owusu-Ware, S. K., Chowdhry, B. Z., Leharne, S. A., & Antonijević, M. D. (2013). Quantitative analysis of overlapping processes in the non-isothermal decomposition of chlorogenic acid by peak fitting. *Thermochimica Acta, 565*, 27-33.

Pascual, E. C., Goodman, B. A., & Yeretzian, C. (2002). Characterization of free radicals in soluble coffee by electron paramagnetic resonance spectroscopy. *Journal of Agricultural and Food Chemistry, 50*, 6114-6122.

Passos, C. P., Cepeda, M. R., Ferreira, S. S., Nunes, F. M., Evtuguin, D. V., Madureira, P., Vilanova, M., & Coimbra, M. A. (2014a). Influence of molecular weight on *in vitro* immunostimulatory properties of instant coffee. *Food Chemistry, 161*, 60-66.

Passos, C. P., & Coimbra, M. A. (2013). Microwave superheated water extraction of polysaccharides from spent coffee grounds. *Carbohydrate Polymers, 94*, 626-633.

Passos, C. P., Moreira, A. S. P., Domingues, M. R. M., Evtuguin, D. V., & Coimbra, M. A. (2014b). Sequential microwave superheated water extraction of mannans from spent coffee grounds. *Carbohydrate Polymers, 103*, 333-338.

Perrone, D., Farah, A., & Donangelo, C. M. (2012). Influence of coffee roasting on the incorporation of phenolic compounds into melanoidins and their relationship with

antioxidant activity of the brew. *Journal of Agricultural and Food Chemistry, 60*, 4265-4275.

Perrone, D., Farah, A., Donangelo, C. M., de Paulis, T., & Martin, P. R. (2008). Comprehensive analysis of major and minor chlorogenic acids and lactones in economically relevant Brazilian coffee cultivars. *Food Chemistry, 106*, 859-867.

Pettolino, F. A., Walsh, C., Fincher, G. B., & Bacic, A. (2012). Determining the polysaccharide composition of plant cell walls. *Nature Protocols, 7*, 1590-1607

Pinto, A., & Chichester, C. O. (1966). Changes in the content of free amino acids during roasting of cocoa beans. *Journal of Food Science, 31*, 726-732.

Ponder, G. R., & Richards, G. N. (1993). Pyrolysis of some ^{13}C-labeled glucans: a mechanistic study. *Carbohydrate Research, 244*, 27-47.

Prajapati, V. D., Jani, G. K., Moradiya, N. G., Randeria, N. P., Nagar, B. J., Naikwadi, N. N., & Variya, B. C. (2013). Galactomannan: A versatile biodegradable seed polysaccharide. *International Journal of Biological Macromolecules, 60*, 83-92.

Prozil, S. O., Costa, E. V., Evtuguin, D. V., Cruz Lopes, L. P., & Domingues, M. R. M. (2012). Structural characterization of polysaccharides isolated from grape stalks of Vitis vinifera L. *Carbohydrate Research, 356*, 252-259.

Purlis, E. (2010). Browning development in bakery products – A review. *Journal of Food Engineering, 99*, 239-249.

Rawel, H. M., Rohn, S., & Kroll, J. (2005). Characterisation of 11s protein fractions and phenolic compounds from green coffee beans under special consideration of their interactions: A review. *Deutsche Lebensmittel-Rundschau, 101*, 148-160.

Redgwell, R., Curti, D., Rogers, J., Nicolas, P., & Fischer, M. (2003). Changes to the galactose/mannose ratio in galactomannans during coffee bean (*Coffea arabica* L.) development: implications for in vivo modification of galactomannan synthesis. *Planta, 217*, 316-326.

Redgwell, R. J., Curti, D., Fischer, M., Nicolas, P., & Fay, L. B. (2002a). Coffee bean arabinogalactans: acidic polymers covalently linked to protein. *Carbohydrate Research, 337*, 239-253.

Redgwell, R. J., Trovato, V., Curti, D., & Fischer, M. (2002b). Effect of roasting on degradation and structural features of polysaccharides in Arabica coffee beans. *Carbohydrate Research, 337*, 421-431.

Reis, A., Coimbra, M. A., Domingues, P., Ferrer-Correia, A. J., & Domingues, M. R. M. (2004a). Fragmentation pattern of underivatised xylo-oligosaccharides and their

alditol derivatives by electrospray tandem mass spectrometry. *Carbohydrate Polymers, 55*, 401-409.

Reis, A., Domingues, M. R. M., Amado, F. M. L., Ferrer-Correia, A. J., & Domingues, P. (2007). Radical peroxidation of palmitoyl-lineloyl-glycerophosphocholine liposomes: Identification of long-chain oxidised products by liquid chromatography–tandem mass spectrometry. *Journal of Chromatography B, 855*, 186-199.

Reis, A., Domingues, P., Ferrer-Correia, A. J. V., & Domingues, M. R. M. (2004b). Tandem mass spectrometry of intact oxidation products of diacylphosphatidylcholines: evidence for the occurrence of the oxidation of the phosphocholine head and differentiation of isomers. *Journal of Mass Spectrometry, 39*, 1513-1522.

Reis, A., Pinto, P., Evtuguin, D. V., Neto, C. P., Domingues, P., Ferrer-Correia, A. J., & Domingues, M. R. M. (2005). Electrospray tandem mass spectrometry of underivatised acetylated xylo-oligosaccharides. *Rapid Communications in Mass Spectrometry, 19*, 3589-3599.

Rodrigues, J. A., Taylor, A. M., Sumpton, D. P., Reynolds, J. C., Pickford, R., & Thomas-Oates, J. (2007). Mass spectrometry of carbohydrates: newer aspects. *Advances in Carbohydrate Chemistry and Biochemistry, 61*, 59-141.

Roepstorff, P., & Fohlman, J. (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomedical Mass Spectrometry, 11*, 601.

Rufián-Henares, J. A., & Morales, F. J. (2007). Angiotensin-I converting enzyme inhibitory activity of coffee melanoidins. *Journal of Agricultural and Food Chemistry, 55*, 1480-1485.

Ruiz-Matute, A. I., Hernández-Hernández, O., Rodríguez-Sánchez, S., Sanz, M. L., & Martínez-Castro, I. (2011). Derivatization of carbohydrates for GC and GC–MS analyses. *Journal of Chromatography B, 879*, 1226-1240.

Ryan, T. T., & Aust, S. D. (1992). The role of iron in oxygen-mediated toxicities. *Critical Reviews in Toxicology, 22*, 119-141.

Sassaki, G. L., & Souza, L. M. d. (2013). Mass spectrometry strategies for structural analysis of carbohydrates and glycoconjugates. In A. V. Coelho & C. F. Franco (Eds.), *Tandem Mass Spectrometry - Molecular Characterization*). http://www.intechopen.com/books/tandem-mass-spectrometry-molecular-characterization: InTech.

Scheirs, J., Camino, G., & Tumiatti, W. (2001). Overview of water evolution during the thermal degradation of cellulose. *European Polymer Journal, 37*, 933-942.

Schenker, S., Handschin, S., Frey, B., Perren, R., & Escher, F. (2000). Pore structure of coffee beans affected by roasting conditions. *Journal of Food Science, 65*, 452-457.

Schenker, S., Heinemann, C., Huber, M., Pompizzi, R., Perren, R., & Escher, R. (2002). Impact of roasting conditions on the formation of aroma compounds in coffee beans. *Journal of Food Science, 67*, 60-66.

Scholz, B., & Maier, H. (1990). Isomers of quinic acid and quinide in roasted coffee. *Zeitschrift für Lebensmittel-Untersuchung und Forschung, 190*, 132-134.

Schweikert, C., Liszkay, A., & Schopfer, P. (2000). Scission of polysaccharides by peroxidase-generated hydroxyl radicals. *Phytochemistry, 53*, 565-570.

Sewell, P. A. (2000). Chromatography: Liquid | Theory of Liquid Chromatography. In I. D. Wilson (Ed.), *Encyclopedia of Separation Science*, (pp. 779-787). Oxford: Academic Press.

Sharma, R. K., Fisher, T. S., & Hajaligol, M. R. (2002). Effect of reaction conditions on pyrolysis of chlorogenic acid. *Journal of Analytical and Applied Pyrolysis, 62*, 281-296.

Shen, D. K., & Gu, S. (2009). The mechanism for thermal decomposition of cellulose and its main products. *Bioresource Technology, 100*, 6496-6504.

Showalter, A. M. (2001). Arabinogalactan-proteins: structure, expression and function. *Cellular and Molecular Life Sciences, 58*, 1399-1417.

Simões, J., Domingues, P., Reis, A., Nunes, F. M., Coimbra, M. A., & Domingues, M. R. M. (2007). Identification of anomeric configuration of underivatized reducing glucopyranosyl-glucose disaccharides by tandem mass spectrometry and multivariate analysis. *Analytical Chemistry, 79*, 5896-5905.

Simões, J., Maricato, É., Nunes, F. M., Domingues, M. R., & Coimbra, M. A. (2014). Thermal stability of spent coffee ground polysaccharides: Galactomannans and arabinogalactans. *Carbohydrate Polymers, 101*, 256-264.

Simões, J., Nunes, F. M., Domingues, M. R., & Coimbra, M. A. (2013). Extractability and structure of spent coffee ground polysaccharides by roasting pre-treatments. *Carbohydrate Polymers, 97*, 81-89.

Simões, J., Nunes, F. M., Domingues, M. R. M., & Coimbra, M. A. (2010). Structural features of partially acetylated coffee galactomannans presenting immunostimulatory activity. *Carbohydrate Polymers, 79*, 397-402.

Simões, J., Nunes, F. M., Domingues, P., Coimbra, M. A., & Domingues, M. R. (2012). Mass spectrometry characterization of an *Aloe vera* mannan presenting immunostimulatory activity. *Carbohydrate Polymers, 90*, 229-236.

Šoltés, L., Mendichi, R., Kogan, G., Schiller, J., Stankovská, M., & Arnhold, J. (2006). Degradative action of reactive oxygen species on hyaluronan. *Biomacromolecules, 7*, 659-668.

Sparkman, O. D., Penton, Z. E., & Kitson, F. G. (2011). Chapter 2 - Gas chromatography. In O. D. Sparkman, Z. E. Penton & F. G. Kitson (Eds.), *Gas Chromatography and Mass Spectrometry (Second edition)*, (pp. 15-83). Amsterdam: Academic Press.

Stadler, R. H., & Fay, L. B. (1995). Antioxidative reactions of caffeine: Formation of 8-oxocaffeine (1,3,7-trimethyluric acid) in coffee subjected to oxidative stress. *Journal of Agricultural and Food Chemistry, 43*, 1332-1338.

Stadler, R. H., Welti, D. H., Stämpfli, A. A., & Fay, L. B. (1996). Thermal decomposition of caffeic acid in model systems: identification of novel tetraoxygenated phenylindan isomers and their stability in aqueous solution. *Journal of Agricultural and Food Chemistry, 44*, 898-905.

Stauder, M., Papetti, A., Mascherpa, D., Schito, A. M., Gazzani, G., Pruzzo, C., & Daglia, M. (2010). Antiadhesion and antibiofilm activities of high molecular weight coffee components against *Streptococcus mutans*. *Journal of Agricultural and Food Chemistry, 58*, 11662-11666.

Sun, D.-W. (2012). *Thermal food processing: new technologies and quality issues* (Second ed.). Boca Raton Taylor & Francis.

Sutherland, P. W., Hallett, I. C., MacRae, E., Fischer, M., & Redgwell, R. J. (2004). Cytochemistry and immunolocalisation of polysaccharides and proteoglycans in the endosperm of green Arabica coffee beans. *Protoplasma, 223*, 203-211.

Takenaka, M., Sato, N., Asakawa, H., Wen, X., Murata, M., & Homma, S. (2005). Characterization of a metal-chelating substance in coffee. *Bioscience, Biotechnology, and Biochemistry, 69*, 26-30.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., & Matsuo, T. (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.

Tomasik, P., Wiejak, S., & Pałasiński, M. (1989). The thermal decomposition of carbohydrates. Part II.* The decomposition of starch. In R. S. Tipson & H. Derek

(Eds.), *Advances in Carbohydrate Chemistry and Biochemistry*, vol. Volume 47 (pp. 279-343): Academic Press.

Totlani, V. M., & Peterson, D. G. (2007). Influence of epicatechin reactions on the mechanisms of Maillard product formation in low moisture model systems. *Journal of Agricultural and Food Chemistry, 55*, 414-420.

Tudella, J., Nunes, F. M., Paradela, R., Evtuguin, D. V., Domingues, P., Amado, F., Coimbra, M. A., Barros, A. I. R. N. A., & Domingues, M. R. M. (2011). Oxidation of mannosyl oligosaccharides by hydroxyl radicals as assessed by electrospray mass spectrometry. *Carbohydrate Research, 346*, 2603-2611.

Uchiyama, H., Dobashi, Y., Ohkouchi, K., & Nagasawa, K. (1990). Chemical change involved in the oxidative reductive depolymerization of hyaluronic acid. *Journal of Biological Chemistry, 265*, 7753-7759.

Varela, O. (2003). Oxidative reactions and degradations of sugars and polysaccharides. *Advances in Carbohydrate Chemistry and Biochemistry, 58*, 307-369.

Verzelloni, E., Tagliazucchi, D., Del Rio, D., Calani, L., & Conte, A. (2011). Antiglycative and antioxidative properties of coffee fractions. *Food Chemistry, 124*, 1430-1435.

Vitaglione, P., Morisco, F., Mazzone, G., Amoruso, D. C., Ribecco, M. T., Romano, A., Fogliano, V., Caporaso, N., & D'Argenio, G. (2010). Coffee reduces liver damage in a rat model of steatohepatitis: The underlying mechanisms and the role of polyphenols and melanoidins. *Hepatology, 52*, 1652-1661.

Wang, J., Lu, Y.-M., Liu, B.-Z., & He, H.-Y. (2008). Electrospray positive ionization tandem mass spectrometry of Amadori compounds. *Journal of Mass Spectrometry, 43*, 262-264.

Wei, F., Furihata, K., Hu, F., Miyakawa, T., & Tanokura, M. (2010). Complex mixture analysis of organic compounds in green coffee bean extract by two-dimensional NMR spectroscopy. *Magnetic Resonance in Chemistry, 48*, 857-865.

Wei, F., Furihata, K., Koda, M., Hu, F., Miyakawa, T., & Tanokura, M. (2012). Roasting process of coffee beans as studied by nuclear magnetic resonance: time course of changes in composition. *Journal of Agricultural and Food Chemistry, 60*, 1005-1012.

Wei, F., & Tanokura, M. (2015). Chemical changes in the components of coffee beans during roasting. In V. R. Preedy (Ed.), *Coffee in health and disease prevention*, (pp. 83-91). San Diego: Academic Press.

Welch, K. D., Davis, T. Z., & Aust, S. D. (2002). Iron autoxidation and free radical generation: effects of buffers, ligands, and chelators. *Archives of Biochemistry and Biophysics, 397*, 360-369.

Wolfrom, M. L., & Anderson, L. E. (1967). Polysaccharides from instant coffee powder. *Journal of Agricultural and Food Chemistry, 15*, 685-687.

Wolfrom, M. L., Laver, M. L., & Patin, D. L. (1961). Carbohydrates of the coffee bean. II. Isolation and characterization of a mannan. *The Journal of Organic Chemistry, 26*, 4533-4535.

Wolfrom, M. L., & Patin, D. L. (1965). Carbohydrates of the coffee bean. IV. An arabinogalactan. *The Journal of Organic Chemistry, 30*, 4060-4063.

Yamashita, M., & Fenn, J. B. (1984). Electrospray ion source. Another variation on the free-jet theme. *The Journal of Physical Chemistry, 88*, 4451-4459.

Yeretzian, C., Jordan, A., Badoud, R., & Lindinger, W. (2002). From the green bean to the cup of coffee: investigating coffee roasting by on-line monitoring of volatiles. *European Food Research and Technology, 214*, 92-104.

Yeretzian, C., Jordan, A., & Lindinger, W. (2003). Analysing the headspace of coffee by proton-transfer-reaction mass-spectrometry. *International Journal of Mass Spectrometry, 223–224*, 115-139.

Zaia, J. (2004). Mass spectrometry of oligosaccharides. *Mass Spectrometry Reviews, 23*, 161-227.

APPENDIX A. SUPPLEMENTARY MATERIAL OF SECTION III.1.1

SUPPLEMENTARY TABLES

Supplementary Table S1. Summary of the ions observed in the ESI-MS spectrum of Ara_3 heated to 200 °C, maintained at 200 °C for 0 min (T1), 30 min (T2), and 60 min (T3).

n		1	2	3	4	5	6	7	8	9
Products of depolymerization and polymerization										
		173[a]	**305**	**437**	**569**	**701**	**833**	**965**	**1097**	**1229**
$[Pent_n+Na]^+$	**T1**	6.3 ± 3.5[b]	28.5 ± 2.7	100%	42.8 ± 1.5	30.9 ± 5.2	23.0 ± 3.9	11.5 ± 1.4	5.3 ± 0.6	2.2 ± 0.6
	T2	2.5 ± 0.4	17.5 ± 2.2	47.9 ± 0.7	29.5 ± 2.6	24.8 ± 1.5	17.4 ± 1.2	9.8 ± 0.8	4.8 ± 0.4	1.7 ± 0.2
	T3	2.4 ± 0.6	12.7 ± 0.4	45.6 ± 10.6	20.4 ± 2.1	20.9 ± 3.5	13.3 ± 2.1	7.0 ± 0.1	3.2 ± 0.6	
Dehydration products										
			287	**419**	**551**	**683**	**815**	**947**	**1079**	**1211**
$[Pent_n-18Da+Na]^+$	**T1**		8.3 ± 0.8	28.5 ± 1.1	19.5 ± 0.3	12.2 ± 1.1	9.9 ± 0.4	5.6 ± 1.2	1.3 ± 1.0	1.0 ± 0.8
	T2		52.4 ± 4.8	100%	56.0 ± 1.9	39.1 ± 1.4	25.7 ± 1.2	13.2 ± 0.8	5.6 ± 0.7	2.0 ± 0.5
	T3		54.1 ± 3.6	100%	54.9 ± 1.2	36.4 ± 1.1	22.0 ± 0.7	11.0 ± 0.7	4.4 ± 0.5	1.3 ± 0.4
			269	**401**	**533**	**665**	**797**	**929**	**1061**	
$[Pent_n-36Da+Na]^+$	**T1**		1.4 ± 0.2	5.5 ± 0.7	5.7 ± 0.2	4.6 ± 1.2	4.0 ± 0.6	2.6 ± 0.4	1.3 ± 0.8	
	T2		5.4 ± 0.5	12.2 ± 1.5	8.6 ± 0.3	9.0 ± 1.4	7.0 ± 0.5	3.6 ± 0.4	1.2 ± 0.1	
	T3		4.5 ± 0.1	9.6 ± 0.2	7.9 ± 0.7	7.7 ± 1.6	7.8 ± 3.0	3.9 ± 1.4	1.0 ± 0.3	
			251	**383**	**515**	**647**	**779**	**911**	**1043**	
$[Pent_n-54Da+Na]^+$	**T1**			2.0 ± 0.5	3.6 ± 0.6	4.2 ± 0.6	3.7 ± 1.3	2.6 ± 0.8	1.3 ± 1.1	
	T2		4.6 ± 0.4	15.2 ± 2.6	12.3 ± 1.3	14.6 ± 1.6	8.5 ± 1.0	4.3 ± 0.5	1.6 ± 0.3	
	T3		2.8 ± 0.2	11.9 ± 1.3	10.1 ± 0.8	10.8 ± 2.6	8.0 ± 2.2	4.2 ± 1.4	1.4 ± 0.5	
Oxidation products										
			321	**453**	**585**	**717**	**849**	**981**	**1113**	
$[Pent_n+16Da+Na]^+$	**T1**		4.7 ± 1.1	14.4 ± 5.8	6.9 ± 1.8	6.4 ± 1.3	5.8 ± 1.9	3.7 ± 1.5	1.6 ± 1.0	
(and $[Pent_n+K]^+$)	**T2**		3.2 ± 0.2	9.0 ± 1.0	5.7 ± 0.2	6.6 ± 0.1	4.1 ± 0.1	2.3 ± 0.3		
	T3		2.4 ± 0.1	13.5 ± 4.3	5.6 ± 0.4	7.3 ± 2.6	5.6 ± 2.3	3.0 ± 1.3		
				435	**567**	**699**	**831**	**963**		
$[Pent_n-2Da+Na]^+$	**T1**			8.3 ± 1.5	5.8 ± 1.7	5.9 ± 1.7	6.0 ± 2.9	3.5 ± 2.7		
	T2			16.9 ± 0.7	10.6 ± 0.1	11.5 ± 1.0	6.6 ± 0.2	3.2 ± 0.1		
	T3			21.9 ± 4.3	11.8 ± 2.3	12.6 ± 2.9	8.0 ± 1.1	3.9 ± 0.4		
Products of carbon-carbon bond cleavage										
			277	**409**	**541**	**673**	**805**	**937**		
$[Pent_n-28Da+Na]^+$	**T1**		1.2 ± 0.5	3.3 ± 0.5	2.8 ± 0.8	2.6 ± 0.8	2.6 ± 1.5	1.6 ± 0.9		
	T2		1.6 ± 0.1	6.4 ± 0.1	3.3 ± 0.4	5.4 ± 0.2	3.4 ± 0.4	1.7 ± 0.2		
	T3		1.3 ± 1.2	5.8 ± 2.5	3.1 ± 0.3	5.7 ± 2.7	4.3 ± 1.6	2.1 ± 0.8		
				407	**539**	**671**	**803**	**935**		
$[Pent_n-30Da+Na]^+$	**T1**			3.7 ± 0.5	3.0 ± 0.9	3.4 ± 1.5	6.4 ± 5.9	2.8 ± 1.6		
	T2			8.0 ± 1.1	4.0 ± 0.1	5.7 ± 0.5	4.2 ± 0.5	2.2 ± 0.1		
	T3			5.5 ± 0.9	3.2 ± 0.4	5.7 ± 2.8	4.7 ± 1.6	5.0 ± 1.6		
			261	**393**	**525**	**657**	**789**	**921**	**1053**	
$[Pent_n-44Da+Na]^+$	**T1**			4.0 ± 1.9	3.4 ± 1.2	3.4 ± 1.3	3.0 ± 1.0	2.5 ± 1.1	1.2 ± 0.8	
	T2		2.8 ± 0.3	6.9 ± 0.4	5.7 ± 0.3	6.5 ± 0.2	5.4 ± 0.1	3.2 ± 0.1	1.1 ± 0.1	
	T3		1.8 ± 0.1	5.5 ± 0.4	4.6 ± 0.3	5.7 ± 0.9	5.8 ± 2.2	3.5 ± 1.6	1.1 ± 0.4	
			259	**391**	**523**	**655**	**787**	**919**		
$[Pent_n-46Da+Na]^+$	**T1**		2.4 ± 1.5	8.1 ± 6.2	4.3 ± 2.0	3.7 ± 1.2	3.7 ± 1.6	2.5 ± 1.5		
	T2		3.3 ± 0.1	9.0 ± 1.1	4.8 ± 0.2	5.5 ± 0.1	4.4 ± 0.3	2.3 ± 0.1		
	T3		1.9 ± 0.1	5.9 ± 0.7	3.3 ± 0.1	4.2 ± 0.8	4.2 ± 1.4	2.2 ± 0.7		
			247	**379**	**511**	**643**	**775**	**907**		
$[Pent_n-58Da+Na]^+$	**T1**		4.5 ± 1.3	8.3 ± 0.6	5.2 ± 0.9	3.8 ± 0.8	3.8 ± 1.1	2.7 ± 0.8		
	T2		8.3 ± 1.0	5.1 ± 0.6	4.1 ± 0.1	5.2 ± 0.7	4.6 ± 0.5	2.7 ± 0.3		
	T3		4.9 ± 1.3	3.4 ± 0.4	3.1 ± 0.3	4.5 ± 1.2	5.5 ± 2.6	2.9 ± 0.3		

Supplementary Table S1 (Continued). Summary of the ions observed in the ESI-MS spectrum of Ara_3 heated to 200 °C, maintained at 200 °C for 0 min (T1), 30 min (T2), and 60 min (T3).

n		1	2	3	4	5	6	7	8	9
			245	**377**	**509**	**641**	**773**	**905**	**1037**	
$[Pent_n$-60Da+Na$]^+$	**T1**		2.3 ± 0.6	7.0 ± 1.5	9.0 ± 0.3	6.0 ± 4.0	6.3 ± 4.8	5.6 ± 2.4	3.4 ± 2.6	
	T2		8.4 ± 0.6	13.0 ± 0.5	11.2 ± 0.5	10.0 ± 0.5	8.2 ± 0.3	4.3 ± 0.2	1.5 ± 0.1	
	T3		4.0 ± 0.9	7.9 ± 0.1	6.6 ± 0.3	7.1 ± 0.6	7.5 ± 2.9	3.6 ± 1.0	1.2 ± 0.3	
			233	**365**	**497**	**629**	**761**	**893**	**1025**	
$[Pent_n$-72Da+Na$]^+$	**T1**			1.8 ± 0.4	3.4 ± 0.8	3.8 ± 1.1	4.0 ± 1.1	3.4 ± 1.1	1.8 ± 1.1	
	T2		1.5 ± 0.0	7.2 ± 1.0	8.0 ± 0.4	8.2 ± 0.6	7.3 ± 0.4	4.4 ± 0.1	1.8 ± 0.1	
	T3		1.7 ± 0.5	6.1 ± 0.6	7.0 ± 0.3	7.6 ± 0.1	7.5 ± 2.4	4.6 ± 1.1	1.6 ± 0.3	
			231	**363**	**495**	**627**	**759**	**891**	**1023**	
$[Pent_n$-74Da+Na$]^+$	**T1**			3.9 ± 1.0	4.4 ± 0.5	4.1 ± 1.1	4.3 ± 1.3	3.4 ± 1.5	1.9 ± 1.4	
	T2		4.5 ± 0.3	12.8 ± 2.3	7.3 ± 0.2	7.6 ± 0.1	6.8 ± 0.4	4.3 ± 0.2	2.0 ± 0.1	
	T3		2.2 ± 0.6	8.0 ± 3.4	5.0 ± 0.9	5.6 ± 0.4	6.6 ± 1.9	4.0 ± 0.9	1.6 ± 0.5	
			229	**361**	**493**	**625**	**757**	**889**	**1021**	
$[Pent_n$-76Da+Na$]^+$	**T1**		2.0 ± 0.5	3.5 ± 0.7	6.0 ± 2.6	4.7 ± 1.0	4.1 ± 1.0	3.0 ± 1.3	1.6 ± 1.0	
	T2		6.4 ± 0.4	9.4 ± 2.4	8.3 ± 0.7	8.1 ± 1.0	6.1 ± 0.5	3.5 ± 0.2	1.4 ± 0.1	
	T3		3.8 ± 0.3	6.7 ± 2.1	5.8 ± 0.1	6.6 ± 1.1	6.1 ± 1.7	3.6 ± 1.4	1.4 ± 0.5	
			227	**359**	**491**	**623**	**755**	**887**	**1019**	
$[Pent_n$-78Da+Na$]^+$	**T1**			3.6 ± 1.2	6.5 ± 1.0	5.1 ± 1.5	5.2 ± 2.5	3.8 ± 1.8	2.0 ± 1.4	
	T2		2.4 ± 0.2	6.1 ± 0.2	7.6 ± 0.6	8.7 ± 0.9	6.9 ± 0.3	3.9 ± 0.3	1.6 ± 0.1	
	T3		1.9 ± 0.4	5.7 ± 2.0	7.9 ± 2.1	9.4 ± 3.8	9.0 ± 4.4	7.0 ± 4.9	2.4 ± 1.4	
			217	**349**	**481**	**613**	**745**	**877**	**1009**	
$[Pent_n$-88Da+Na$]^+$	**T1**		1.6 ± 0.6	1.7 ± 0.6	4.0 ± 1.5	3.7 ± 1.1	3.3 ± 0.8	2.8 ± 1.2	1.8 ± 1.4	
	T2		3.3 ± 0.6	5.1 ± 0.4	7.2 ± 0.3	7.9 ± 0.4	8.9 ± 0.1	3.9 ± 0.3	1.7 ± 0.3	
	T3		3.0 ± 1.0	7.8 ± 4.6	7.6 ± 2.1	7.1 ± 1.6	9.0 ± 5.0	4.1 ± 1.4	1.7 ± 0.4	
			215	**347**	**479**	**611**	**743**	**875**	**1007**	
$[Pent_n$-90Da+Na$]^+$	**T1**		2.6 ± 0.7	6.2 ± 1.4	10.5 ± 1.3	7.0 ± 1.1	6.9 ± 3.1	4.9 ± 2.3	2.8 ± 1.8	
	T2		3.0 ± 0.1	16.6 ± 2.5	15.1 ± 0.8	13.8 ± 0.6	10.6 ± 1.0	6.1 ± 0.4	2.7 ± 0.4	
	T3		2.6 ± 0.2	10.5 ± 2.5	11.3 ± 0.1	10.8 ± 0.1	9.8 ± 2.6	5.6 ± 1.3	2.2 ± 0.2	
				333	**465**	**597**	**729**	**861**	**993**	**1125**
$[Pent_n$-104Da+Na$]^+$	**T1**			2.9 ± 0.4	5.6 ± 0.1	4.6 ± 0.1	4.4 ± 0.8	4.0 ± 1.3	2.2 ± 1.2	1.3 ± 1.5
	T2			3.0 ± 0.4	7.8 ± 0.5	7.9 ± 0.5	7.0 ± 0.5	4.6 ± 0.3	2.1 ± 0.1	
	T3			2.4 ± 0.3	6.0 ± 0.1	6.3 ± 0.6	6.5 ± 2.1	4.4 ± 1.5	2.0 ± 0.6	
			193	**325**	**457**	**589**	**721**	**853**	**985**	
$[Pent_n$-112Da+Na$]^+$	**T1**		1.1 ± 0.6	3.1 ± 0.6	11.3 ± 1.6	3.8 ± 1.8	4.6 ± 1.6	3.6 ± 1.2	2.3 ± 1.3	**1117**
	T2		7.5 ± 0.3	13.8 ± 1.7	15.8 ± 0.8	10.5 ± 0.7	8.5 ± 0.5	4.6 ± 0.8	1.8 ± 0.2	1.3 ± 1.5
	T3		5.1 ± 0.1	7.9 ± 0.2	10.2 ± 0.5	7.2 ± 0.6	7.2 ± 2.5	4.2 ± 1.3	1.7 ± 0.7	
				319	**451**	**583**	**715**	**847**	**979**	
$[Pent_n$-118Da+Na$]^+$	**T1**			6.4 ± 2.4	8.5 ± 2.0	4.6 ± 1.4	4.2 ± 1.6	3.8 ± 1.8	2.3 ± 1.4	
	T2			21.0 ± 7.5	4.7 ± 0.1	4.7 ± 0.1	4.8 ± 0.2	3.1 ± 0.1	1.5 ± 0.2	
	T3			12.3 ± 7.3	3.8 ± 0.3	3.5 ± 0.2	4.6 ± 1.8	3.1 ± 0.8	1.6 ± 0.5	

Supplementary Table S2. Exact mass measurement of the ions assigned as modified pentose oligosaccharides identified in the ESI-MS spectrum of Ara_3 heated to 200 °C (T1), as well as theoretical mass and mass error for their predicted formula.

Ion series (modification)	*n*	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM)	
					mDa	ppm
$[Pent_n$-18Da+Na$]^+$ (loss of H_2O)	2	$C_{10}H_{16}O_8Na$	287.0744	287.0743	0.1	0.4
	3	$C_{15}H_{24}O_{12}Na$	419.1161	419.1165	-0.4	-1.1
	4	$C_{20}H_{32}O_{16}Na$	551.1583	551.1588	-0.5	-0.9
	5	$C_{25}H_{40}O_{20}Na$	683.1997	683.2011	-1.4	-2.0
	6	$C_{30}H_{48}O_{24}Na$	815.2427	815.2433	-0.6	-0.8
	7	$C_{35}H_{56}O_{28}Na$	947.2849	947.2856	-0.7	-0.7
	8	$C_{40}H_{64}O_{32}Na$	1079.3269	1079.3278	-0.9	-0.9
	9	$C_{45}H_{72}O_{36}Na$	1211.3641	1211.3701	-6.0	-5.0

Supplementary Table S2 (Continued). Exact mass measurement of the ions assigned as modified pentose oligosaccharides identified in the ESI-MS spectrum of Ara_3 heated to 200 °C (T1), as well as theoretical mass and mass error for their predicted formula.

Ion series (modification)	*n*	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM)	
					mDa	ppm
$[Pent_n\text{-}36Da+Na]^+$ (loss of $2H_2O$)	2	$C_{10}H_{14}O_7Na$	269.0710	269.0637	7.3	27.0
	3	$C_{15}H_{22}O_{11}Na$	401.1078	401.1060	1.8	4.5
	4	$C_{20}H_{30}O_{15}Na$	533.1481	533.1482	-0.1	-0.3
	5	$C_{25}H_{38}O_{19}Na$	665.1935	665.1905	3.0	4.5
	6	$C_{30}H_{46}O_{23}Na$	797.2366	797.2328	3.8	4.8
	7	$C_{35}H_{54}O_{27}Na$	929.2814	929.2750	6.4	6.9
	8	$C_{40}H_{62}O_{31}Na$	1061.3175	1061.3173	0.2	0.2
$[Pent_n\text{-}54Da+Na]^+$ (loss of $3H_2O$)	3	$C_{15}H_{20}O_{10}Na$	383.1040	383.0954	8.6	22.4
	4	$C_{20}H_{28}O_{14}Na$	515.1412	515.1377	3.5	6.8
	5	$C_{25}H_{36}O_{18}Na$	647.1888	647.1799	8.9	13.7
	6	$C_{30}H_{44}O_{22}Na$	779.2282	779.2222	6.0	7.7
	7	$C_{35}H_{52}O_{26}Na$	911.2783	911.2645	13.8	15.2
	8	$C_{40}H_{60}O_{30}Na$	1043.3113	1043.3067	4.6	4.4
$[Pent_n\text{+}16Da+Na]^+$ (an extra O atom) and $[Pent_n+K]^+$	2	$C_{10}H_{18}O_{10}Na$	321.0741	321.0798	-5.7	-17.6
		$C_{10}H_{18}O_9K$		321.0588	15.3	47.7
	3	$C_{15}H_{26}O_{14}Na$	453.1050	453.1220	-17.0	-37.6
		$C_{15}H_{26}O_{13}K$		453.1010	4.0	8.7
	4	$C_{20}H_{34}O_{18}Na$	585.1502	585.1643	-14.1	-24.1
		$C_{20}H_{34}O_{17}K$		585.1433	6.9	11.8
	5	$C_{25}H_{42}O_{22}Na$	717.1942	717.2065	-12.3	-17.2
		$C_{25}H_{42}O_{21}K$		717.1856	8.6	12.0
	6	$C_{30}H_{50}O_{26}Na$	849.2372	849.2488	-11.6	-13.7
		$C_{30}H_{50}O_{25}K$		849.2278	9.4	11.0
	7	$C_{35}H_{58}O_{30}Na$	981.2852	981.2911	-5.9	-6.0
		$C_{35}H_{58}O_{29}K$		981.2701	15.1	15.4
	8	$C_{40}H_{66}O_{34}Na$	1113.3254	1113.3333	-7.9	-7.1
		$C_{40}H_{66}O_{33}K$		1113.3123	13.1	11.7
$[Pent_n\text{-}2Da+Na]^+$ (loss of two H atoms)	3	$C_{15}H_{24}O_{13}Na$	435.1069	435.1115	-4.6	-10.5
	4	$C_{20}H_{32}O_{17}Na$	567.1515	567.1537	-2.2	-3.9
	5	$C_{25}H_{40}O_{21}Na$	699.1964	699.1960	0.4	0.6
	6	$C_{30}H_{48}O_{25}Na$	831.2400	831.2382	1.8	2.1
	7	$C_{35}H_{56}O_{29}Na$	963.2861	963.2805	5.6	5.8
$[Pent_n\text{-}28Da+Na]^+$ (loss of C_1O_1)	2	$C_9H_{18}O_8Na$	277.0776	277.0899	-12.3	-44.5
	3	$C_{14}H_{26}O_{12}Na$	409.1212	409.1322	-11.0	-26.9
	4	$C_{19}H_{34}O_{16}Na$	541.1564	541.1745	-18.1	-33.4
	5	$C_{24}H_{42}O_{20}Na$	673.2036	673.2167	-13.1	-19.5
	6	$C_{29}H_{50}O_{24}Na$	805.2450	805.2590	-14.0	-17.4
	7	$C_{34}H_{58}O_{28}Na$	937.2808	937.3012	-20.4	-21.8
$[Pent_n\text{-}30Da+Na]^+$ (loss of $C_1H_2O_1$)	3	$C_{14}H_{24}O_{12}Na$	407.1137	407.1165	-2.8	-7.0
	4	$C_{19}H_{32}O_{16}Na$	539.1554	539.1588	-3.4	-6.3
	5	$C_{24}H_{40}O_{20}Na$	671.1951	671.2011	-6.0	-8.9
	6	$C_{29}H_{48}O_{24}Na$	803.2368	803.2433	-6.5	-8.1
	7	$C_{34}H_{56}O_{28}Na$	935.2878	935.2856	2.2	2.4

Supplementary Table S2 (Continued). Exact mass measurement of the ions assigned as modified pentose oligosaccharides identified in the ESI-MS spectrum of Ara_3 heated to 200 °C (T1), as well as theoretical mass and mass error for their predicted formula.

Ion series (modification)	*n*	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM)	
					mDa	mDa
$[Pent_n$-44Da+Na$]^+$ (loss of $C_2H_4O_1$)	3	$C_{13}H_{22}O_{12}Na$	393.0938	393.1009	-7.1	-18.1
	4	$C_{18}H_{30}O_{16}Na$	525.1403	525.1432	-2.9	-5.4
	5	$C_{23}H_{38}O_{20}Na$	657.1846	657.1854	-0.8	-1.2
	6	$C_{28}H_{46}O_{24}Na$	789.2297	789.2277	2.0	2.6
	7	$C_{33}H_{54}O_{28}Na$	921.2715	921.2699	1.6	1.7
	8	$C_{38}H_{62}O_{32}Na$	1053.3195	1053.3122	7.3	6.9
$[Pent_n$-46Da+Na$]^+$ (loss of $C_1H_2O_2$)	2	$C_9H_{16}O_7Na$	259.0798	259.0794	0.4	1.6
	3	$C_{14}H_{24}O_{11}Na$	391.1192	391.1216	-2.4	-6.2
	4	$C_{19}H_{32}O_{15}Na$	523.1580	523.1639	-5.9	-11.3
	5	$C_{24}H_{40}O_{19}Na$	655.1985	655.2061	-7.6	-11.7
	6	$C_{29}H_{48}O_{23}Na$	787.2372	787.2484	-11.2	-14.2
	7	$C_{34}H_{56}O_{27}Na$	919.2738	919.2907	-16.9	-18.3
$[Pent_n$-58Da+Na$]^+$ (loss of $C_2H_2O_2$)	2	$C_8H_{16}O_7Na$	247.0809	247.0794	1.5	6.2
	3	$C_{13}H_{24}O_{11}Na$	379.1189	379.1216	-2.7	-7.2
	4	$C_{18}H_{32}O_{15}Na$	511.1572	511.1639	-6.7	-13.1
	5	$C_{23}H_{40}O_{19}Na$	643.1982	643.2061	-7.9	-12.4
	6	$C_{28}H_{48}O_{23}Na$	775.2422	775.2484	-6.2	-8.0
	7	$C_{33}H_{56}O_{27}Na$	907.2804	907.2907	-10.3	-11.3
$[Pent_n$-60Da+Na$]^+$ (loss of $C_2H_4O_2$)	2	$C_8H_{14}O_7Na$	245.0654	245.0637	1.7	6.8
	3	$C_{13}H_{22}O_{11}Na$	377.1064	377.1060	0.4	1.1
	4	$C_{18}H_{30}O_{15}Na$	509.1467	509.1482	-1.5	-3.0
	5	$C_{23}H_{38}O_{19}Na$	641.1875	641.1905	-3.0	-4.7
	6	$C_{28}H_{46}O_{23}Na$	773.2276	773.2328	-5.2	-6.7
	7	$C_{33}H_{54}O_{27}Na$	905.2686	905.2750	-6.4	-7.1
	8	$C_{38}H_{62}O_{31}Na$	1037.3121	1037.3173	-5.2	-5.0
$[Pent_n$-72Da+Na$]^+$ (loss of $C_3H_4O_2$)	3	$C_{12}H_{22}O_{11}Na$	365.1021	365.1060	-3.9	-10.6
	4	$C_{17}H_{30}O_{15}Na$	497.1445	497.1482	-3.7	-7.5
	5	$C_{22}H_{38}O_{19}Na$	629.1887	629.1905	-1.8	-2.9
	6	$C_{27}H_{46}O_{23}Na$	761.2299	761.2328	-2.9	-3.8
	7	$C_{32}H_{54}O_{27}Na$	893.2718	893.2750	-3.2	-3.6
	8	$C_{37}H_{62}O_{31}Na$	1025.3112	1025.3173	-6.1	-5.9
$[Pent_n$-74Da+Na$]^+$ (loss of $C_3H_6O_2$)	3	$C_{12}H_{20}O_{11}Na$	363.0927	363.0903	2.4	6.5
	4	$C_{17}H_{28}O_{15}Na$	495.1341	495.1326	1.5	3.0
	5	$C_{22}H_{36}O_{19}Na$	627.1782	627.1748	3.4	5.3
	6	$C_{27}H_{44}O_{23}Na$	759.2208	759.2171	3.7	4.9
	7	$C_{32}H_{52}O_{27}Na$	891.2602	891.2594	0.8	0.9
	8	$C_{37}H_{60}O_{31}Na$	1023.3016	1023.3016	0.0	0.0
$[Pent_n$-76Da+Na$]^+$ (loss of $C_2H_4O_3$)	2	$C_8H_{14}O_6Na$	229.0711	229.0688	2.3	10.0
	3	$C_{13}H_{22}O_{10}Na$	361.1099	361.1111	-1.2	-3.2
	4	$C_{18}H_{30}O_{14}Na$	493.1494	493.1533	-3.9	-8.0
	5	$C_{23}H_{38}O_{18}Na$	625.1907	625.1956	-4.9	-7.8
	6	$C_{28}H_{46}O_{22}Na$	757.2330	757.2378	-4.8	-6.4
	7	$C_{33}H_{54}O_{26}Na$	889.2742	889.2801	-5.9	-6.6
	8	$C_{38}H_{62}O_{30}Na$	1021.3151	1021.3224	-7.3	-7.1

Supplementary Table S2 (Continued). Exact mass measurement of the ions assigned as modified pentose oligosaccharides identified in the ESI-MS spectrum of Ara_3 heated to 200 °C (T1), as well as theoretical mass and mass error for their predicted formula.

Ion series (modification)	*n*	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM)	
					mDa	mDa
$[Pent_n$-78Da+Na$]^+$ (loss of $C_2H_6O_3$)	3	$C_{13}H_{20}O_{10}Na$	359.0914	359.0954	-4.0	-11.2
	4	$C_{18}H_{28}O_{14}Na$	491.1326	491.1377	-5.1	-10.3
	5	$C_{23}H_{36}O_{18}Na$	623.1732	623.1799	-6.7	-10.8
	6	$C_{28}H_{44}O_{22}Na$	755.2189	755.2222	-3.3	-4.4
	7	$C_{33}H_{52}O_{26}Na$	887.2601	887.2645	-4.4	-4.9
	8	$C_{38}H_{60}O_{30}Na$	1019.3089	1019.3067	2.2	2.1
$[Pent_n$-88Da+Na$]^+$ (loss of $C_3H_4O_3$)	2	$C_7H_{14}O_6Na$	217.0723	217.0688	3.5	16.1
	3	$C_{12}H_{22}O_{10}Na$	349.1063	349.1111	-4.8	-13.7
	4	$C_{17}H_{30}O_{14}Na$	481.1400	481.1533	-13.3	-27.7
	5	$C_{22}H_{38}O_{18}Na$	613.1761	613.1956	-19.5	-31.8
	6	$C_{27}H_{46}O_{22}Na$	745.2230	745.2378	-14.8	-19.9
	7	$C_{32}H_{54}O_{26}Na$	877.2666	877.2801	-13.5	-15.4
	8	$C_{37}H_{62}O_{30}Na$	1009.3070	1009.3224	-15.4	-15.2
$[Pent_n$-90Da+Na$]^+$ (loss of $C_3H_6O_3$)	2	$C_7H_{12}O_6Na$	215.0554	215.0532	2.2	10.4
	3	$C_{12}H_{20}O_{10}Na$	347.0974	347.0954	2.0	5.7
	4	$C_{17}H_{28}O_{14}Na$	479.1373	479.1377	-0.4	-0.8
	5	$C_{22}H_{36}O_{18}Na$	611.1777	611.1799	-2.2	-3.7
	6	$C_{27}H_{44}O_{22}Na$	743.2144	743.2222	-7.8	-10.5
	7	$C_{32}H_{52}O_{26}Na$	875.2609	875.2645	-3.6	-4.1
	8	$C_{37}H_{60}O_{30}Na$	1007.3001	1007.3067	-6.6	-6.6
$[Pent_n$-104Da+Na$]^+$ (loss of $C_4H_8O_3$)	3	$C_{11}H_{18}O_{10}Na$	333.0852	333.0798	5.4	16.3
	4	$C_{16}H_{26}O_{14}Na$	465.1295	465.1220	7.5	16.1
	5	$C_{21}H_{34}O_{18}Na$	597.1728	597.1643	8.5	14.3
	6	$C_{26}H_{42}O_{22}Na$	729.2144	729.2065	7.9	10.8
	7	$C_{31}H_{50}O_{26}Na$	861.2560	861.2488	7.2	8.4
	8	$C_{36}H_{58}O_{30}Na$	993.2977	993.2911	6.6	6.7
	9	$C_{41}H_{66}O_{34}Na$	1125.3328	1125.3333	-0.5	-0.5
$[Pent_n$-112Da+Na$]^+$ (loss of $C_2H_8O_5$)	2	$C_8H_{10}O_4Na$	193.0526	193.0477	4.9	25.5
	3	$C_{13}H_{18}O_8Na$	325.0951	325.0899	5.2	15.9
	4	$C_{18}H_{26}O_{12}Na$	457.1333	457.1322	1.1	2.4
	5	$C_{23}H_{34}O_{16}Na$	589.1779	589.1745	3.4	5.8
	6	$C_{28}H_{42}O_{20}Na$	721.2209	721.2167	4.2	5.8
	7	$C_{33}H_{50}O_{24}Na$	853.2670	853.2590	8.0	9.4
	8	$C_{38}H_{58}O_{28}Na$	985.3044	985.3012	3.2	3.2
	9	$C_{43}H_{66}O_{32}Na$	1117.3475	1117.3435	4.0	3.6
$[Pent_n$-118Da+Na$]^+$ (loss of $C_4H_6O_4$)	3	$C_{11}H_{20}O_9Na$	319.0992	319.1005	-1.3	-4.1
	4	$C_{16}H_{28}O_{13}Na$	451.1396	451.1428	-3.2	-7.0
	5	$C_{21}H_{36}O_{17}Na$	583.1777	583.1850	-7.3	-12.6
	6	$C_{26}H_{44}O_{21}Na$	715.2191	715.2273	-8.2	-11.4
	7	$C_{31}H_{52}O_{25}Na$	847.2591	847.2695	-10.4	-12.3
	8	$C_{36}H_{60}O_{29}Na$	979.3014	979.3118	-10.4	-10.6

SUPPLEMENTARY FIGURES

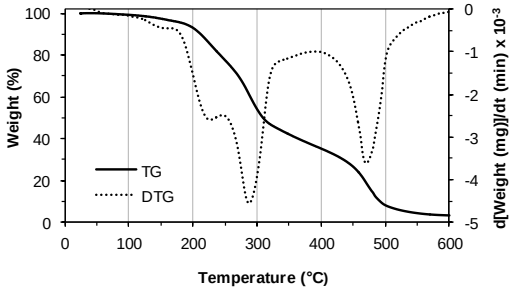

Figure S1. TG and DTG curves of Ara_3 obtained from room temperature to 600 °C.

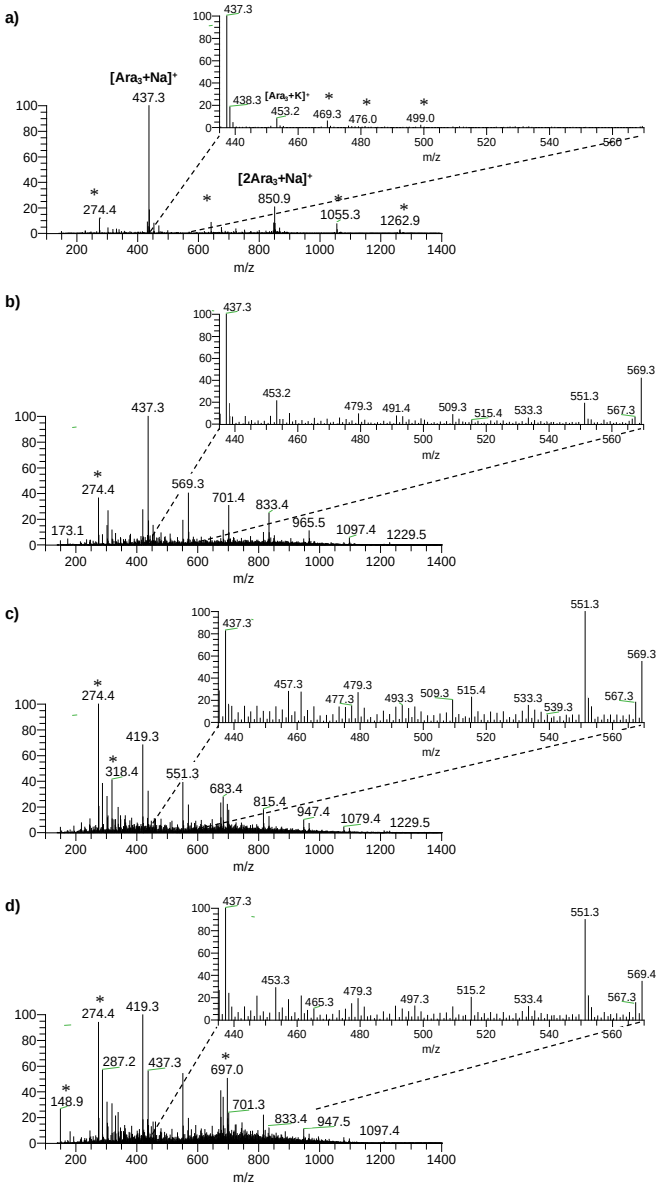

Figure S2. ESI-MS spectra obtained for Ara_3 (A) before (T0) and after heating from room temperature to 200 °C, maintained at 200 °C for (B) 0 min (T1), (C) 30 min (T2), and (D) 60 min (T3). Ions marked with an asterisk are attributed to impurities.

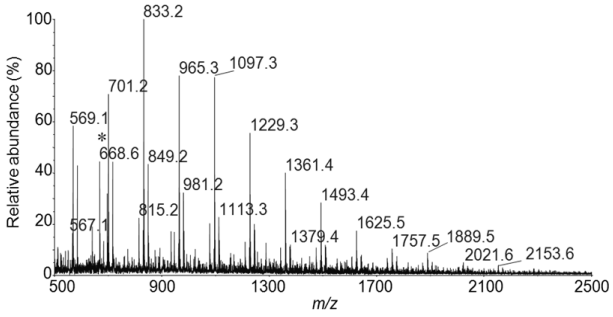

Figure S3. MALDI-MS spectrum of Ara_3 heated to 200 °C (T1). *Ion attributed to an impurity.

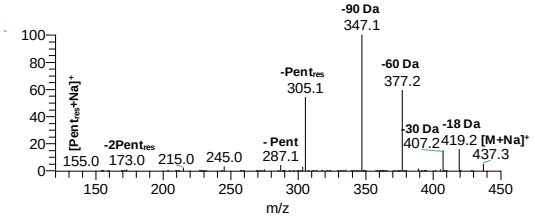

Figure S4. ESI-MS^2 spectrum of the ion at *m/z* 437 ($[Pent_3+Na]^+$) acquired from Ara_3 after T1.

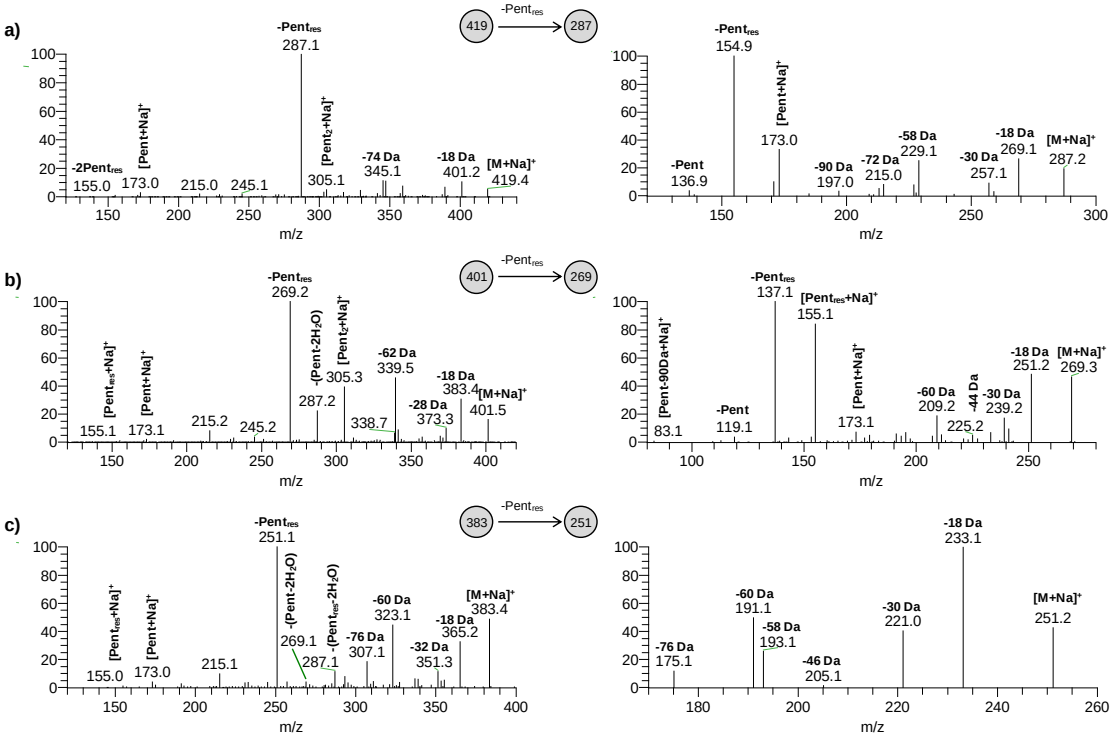

Figure S5. ESI-MS^2 spectra of the ions at *m/z* (A) 419 ($[Pent_3-H_2O+Na]^+$), (B) 401 ($[Pent_3-2H_2O+Na]^+$), and (C) 383 ($[Pent_3-3H_2O+Na]^+$) acquired from Ara_3 after T1, and MS^3 spectra of their product ion formed by loss of a $Pent_{res}$.

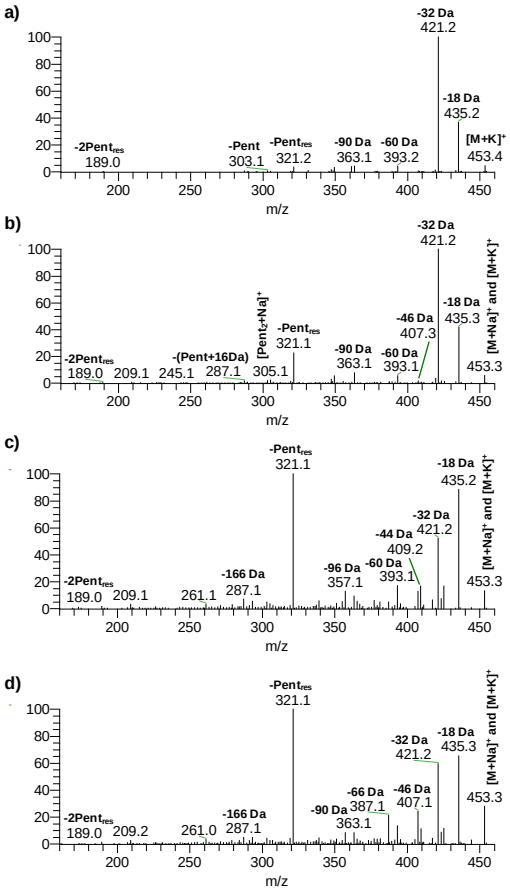

Figure S6. ESI-MS^2 spectra of the ion at *m/z* 453 ([Pent$_3$+K]$^+$ and/or [Pent$_3$+O+Na]$^+$) obtained for Ara$_3$ (A) before (T0) and after heating from room temperature to 200 °C, maintained at 200 °C for (B) 0 min (T1), (C) 30 min (T2), and (D) 60 min (T3).

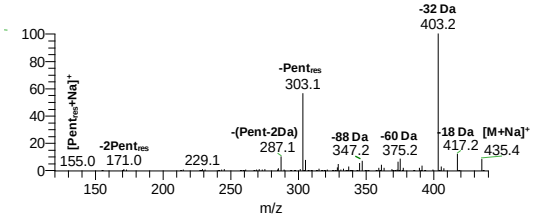

Figure S7. ESI-MS^2 spectrum of the ion at *m/z* 435 ([Pent$_3$-2Da+Na]$^+$) acquired from Ara$_3$ after T1.

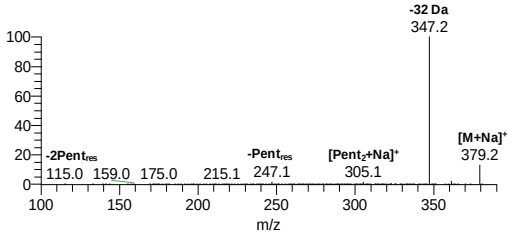

Figure S8. ESI-MS^2 spectrum of the ion at *m/z* 379 ([Pent$_3$-58Da+Na]$^+$) acquired from Ara$_3$ after T1.

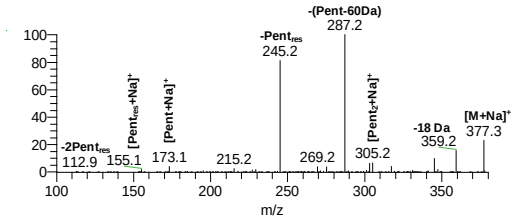

Figure S9. ESI-MS^2 spectrum of the ion at *m/z* 377 ([Pent$_3$-60Da+Na]$^+$) acquired from Ara$_3$ after T1.

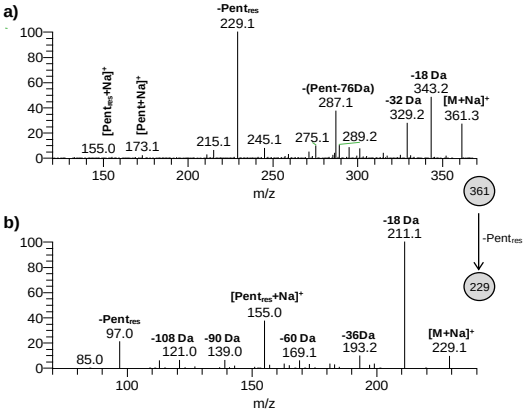

Figure S10. (A) ESI-MS^2 spectrum of the ion at *m/z* 361 ([Pent$_3$-76Da+Na]$^+$) acquired from Ara$_3$ after T1, and (B) MS^3 spectrum of its product ion formed by loss of a Pent$_{res}$.

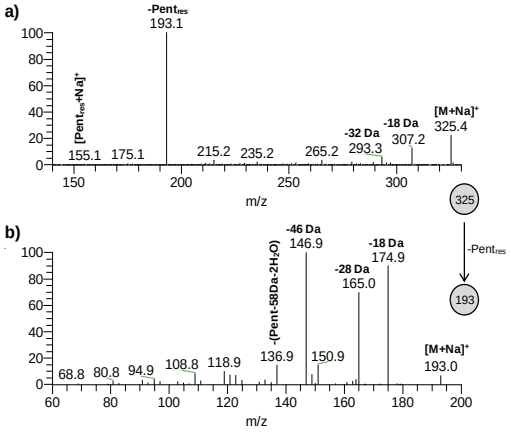

Figure S11. (A) ESI-MS^2 spectrum of the ion at *m/z* 325 ([Pent$_3$-112Da+Na]$^+$) acquired from Ara$_3$ after T1, and (B) MS^3 spectrum of its product ion formed by loss of a Pent$_{res}$.

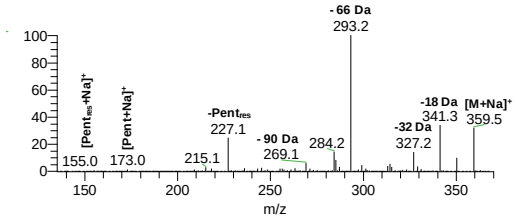

Figure S12. ESI-MS^2 spectrum of the ion at *m/z* 359 ([Pent$_3$-78Da+Na]$^+$) acquired from Ara$_3$ after T1.

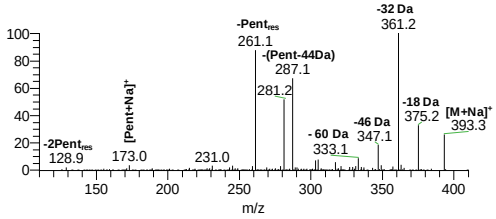

Figure S13. ESI-MS^2 spectrum of the ion at *m/z* 393 ([Pent$_3$-44Da+Na]$^+$) acquired from Ara$_3$ after T1.

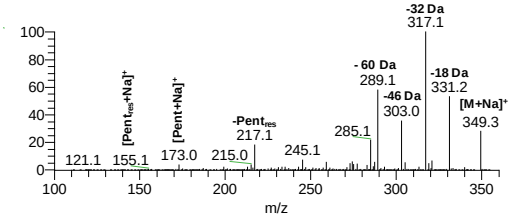

Figure S14. ESI-MS^2 spectrum of the ion at *m/z* 349 ([Pent$_3$-88Da+Na]$^+$) acquired from Ara$_3$ after T1.

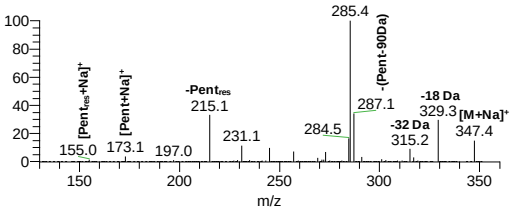

Figure S15. ESI-MS^2 spectrum of the ion at *m/z* 347 ([Pent$_3$-90Da+Na]$^+$) acquired from Ara$_3$ after T1.

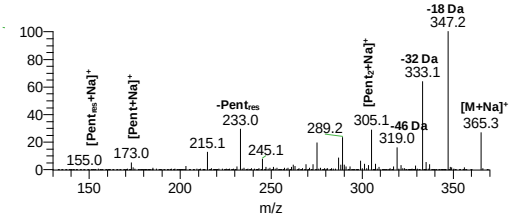

Figure S16. ESI-MS^2 spectrum of the ion at *m/z* 365 ([Pent$_3$-72Da+Na]$^+$) acquired from Ara$_3$ after T1.

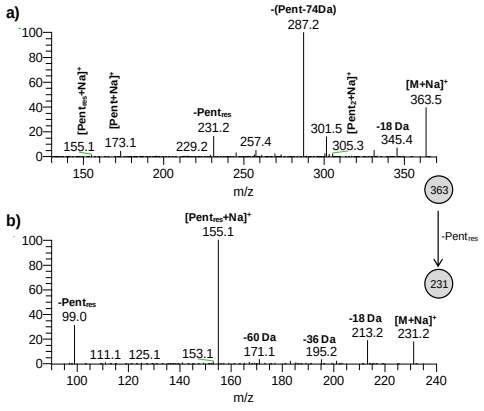

Figure S17. (A) ESI-MS^2 spectrum of the ion at *m/z* 363 ([Pent$_3$-74Da+Na]$^+$) acquired from Ara$_3$ after T1, and (B) MS^3 spectrum of its product ion formed by loss of a Pent$_{res}$.

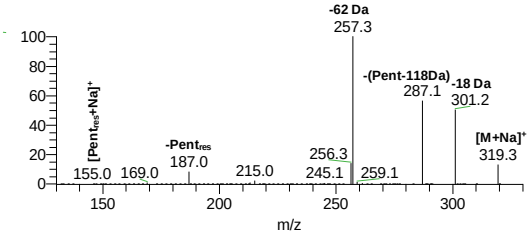

Figure S18. ESI-MS^2 spectrum of the ion at *m/z* 319 ([Pent$_3$-118Da+Na]$^+$) acquired from Ara$_3$ after T1.

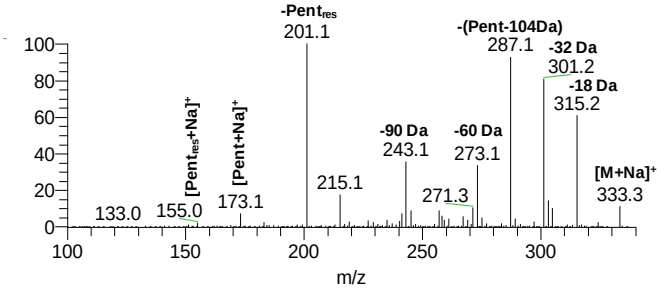

Figure S19. ESI-MS^2 spectrum of the ion at *m/z* 333 ([Pent$_3$-104Da+Na]$^+$) acquired from Ara$_3$ after T1.

APPENDIX B. SUPPLEMENTARY MATERIAL OF SECTION III.1.2

SUPPLEMENTARY TABLES

Table S1. Exact mass measurement of the oxidation products observed in the ESI-MS spectrum of the acidic fraction (F1), as well as theoretical mass and mass error for the predicted formula. The lock mass was the calculated monoisotopic mass of $[M+Na]^+$ ions of the (β1→4)-D-mannotriose used as internal standard.

Ions (*m/z* value, proposed assignment)	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM)	
				mDa	ppm
187, $[Pent_1+O-2H+Na]^+$	$C_5H_8O_6Na$	187.0262	187.0219	4.3	23.2
189, $[Pent_1+O+Na]^+$	$C_5H_{10}O_6Na$	189.0428	189.0375	5.3	28.0
203, $[Pent_1+2O-2H+Na]^+$	$C_5H_8O_7Na$	203.0155	203.0168	-1.3	-6.3
205, $[Pent_1+2O+Na]^+$	$C_5H_{10}O_7Na$	205.0308	205.0324	-1.6	-7.9
231, $[Pent_2-C_3H_6O_3+O+Na]^+$	$C_7H_{12}O_7Na$	231.0465	231.0481	-1.6	-6.8
245, $[Pent_2-C_3H_6O_3+2O-2H+Na]^+$	$C_7H_{10}O_8Na$	245.0344	245.0273	7.1	28.8
247, $[Pent_2-C_3H_6O_3+2O+Na]^+$	$C_7H_{12}O_8Na$	247.0500	247.0430	7.0	28.4
259, $[Pent_2-C_2H_4O_2+O-2H+Na]^+$	$C_8H_{12}O_8Na$	259.0489	259.0430	5.9	22.8
261, $[Pent_2-C_2H_4O_2+O+Na]^+$	$C_8H_{14}O_8Na$	261.0584	261.0586	-0.2	-0.9
275, $[Pent_2-C_2H_4O_2+2O-2H+Na]^+$	$C_8H_{12}O_9Na$	275.0471	275.0379	9.2	33.4
277, $[Pent_2-C_2H_4O_2+2O+Na]^+$	$C_8H_{14}O_9Na$	277.0529	277.0536	-0.7	-2.4
289, $[Pent_2-CH_2O+O-2H+Na]^+$	$C_9H_{14}O_9Na$	289.0585	289.0536	4.9	17.1
291, $[Pent_2-CH_2O+O+Na]^+$	$C_9H1_6O_9Na$	291.0607	291.0692	-8.5	-29.2
303, $[Pent_2-CH_2O+2O-4H+Na]^+$	$C_9H_{12}O_{10}Na$	303.0334	303.0328	0.6	1.9
305, $[Pent_2-CH_2O+2O-2H+Na]^+$	$C_9H_{14}O_{10}Na$	305.0574	305.0485	8.9	29.3
307, $[Pent_2-CH_2O+2O+Na]^+$	$C_9H_{16}O_{10}Na$	307.0652	307.0641	1.1	3.5
317, $[Pent_2+O-4H+Na]^+$	$C_{10}H_{14}O_{10}Na$	317.0548	317.0485	6.3	20.0
319, $[Pent_2+O-2H+Na]^+$	$C_{10}H_{16}O_{10}Na$	319.0668	319.0641	2.7	8.4
321, $[Pent_2+O+Na]^+$	$C_{10}H_{18}O_{10}Na$	321.0817	321.0798	1.9	6.0
333, $[Pent_2+2O-4H+Na]^+$	$C_{10}H_{14}O_{11}Na$	333.0483	333.0434	4.9	14.8
335, $[Pent_2+2O-2H+Na]^+$	$C_{10}H_{16}O_{11}Na$	335.0594	335.0590	0.4	1.1
337, $[Pent_2+2O+Na]^+$	$C_{10}H_{18}O_{11}Na$	337.0649	337.0747	-9.8	-29.0
351, $[Pent_2+3O-2H+Na]^+$	$C_{10}H_{16}O_{12}Na$	351.0527	351.0539	-1.2	-3.5
361, $[Pent_3-C_3H_6O_3+O-2H+Na]^+$	$C_{12}H_{18}O_{11}Na$	361.0783	361.0747	3.6	10.0
363, $[Pent_3-C_3H_6O_3+O+Na]^+$	$C_{12}H_{20}O_{11}Na$	363.0908	363.0903	0.5	1.3
377, $[Pent_3-C_3H_6O_3+2O-2H+Na]^+$	$C_{12}H_{18}O_{12}Na$	377.0758	377.0696	6.2	16.5
379, $[Pent_3-C_3H_6O_3+2O+Na]^+$	$C_{12}H_{20}O_{12}Na$	379.0928	379.0852	7.6	19.9
391, $[Pent_3-C_2H_4O_2+O-2H+Na]^+$	$C_{13}H_{20}O_{12}Na$	391.0877	391.0852	2.5	6.3
393, $[Pent_3-C_2H_4O_2+O+Na]^+$	$C_{13}H_{22}O_{12}Na$	393.1002	393.1009	-0.7	-1.8
407, $[Pent_3-C_2H_4O_2+2O-2H+Na]^+$	$C_{13}H_{20}O_{13}Na$	407.0841	407.0802	3.9	9.7
409, $[Pent_3-C_2H_4O_2+2O+Na]^+$	$C_{13}H_{22}O_{13}Na$	409.0974	409.0958	1.6	3.9
421, $[Pent_3-CH_2O+O-2H+Na]^+$	$C_{14}H_{22}O_{13}Na$	421.0979	421.0958	2.1	5.0
423, $[Pent_3-CH_2O+O+Na]^+$	$C_{14}H_{24}O_{13}Na$	423.1055	423.1115	-6.0	-14.1
435, $[Pent_3-CH_2O+2O-4H+Na]^+$	$C_{14}H_{20}O_{14}Na$	435.0679	435.0751	-7.2	-16.5
437, $[Pent_3-CH_2O+2O-2H+Na]^+$	$C_{14}H_{22}O_{14}Na$	437.0916	437.0907	0.9	2.0
439, $[Pent_3-CH_2O+2O+Na]^+$	$C_{14}H_{24}O_{14}Na$	439.1119	439.1064	5.5	12.6
449, $[Pent_3+O-4H+Na]^+$	$C_{15}H_{24}O_{14}Na$	449.0822	449.0907	-8.5	-19.0
451, $[Pent_3+O-2H+Na]^+$	$C_{15}H_{24}O_{14}Na$	451.1078	451.1064	1.4	3.2
453, $[Pent_3+O+Na]^+$	$C_{15}H_{26}O_{14}Na$	453.1166	453.1220	-5.4	-12.0
[a]453, $[Pent_3+K]^+$	$C_{15}H_{26}O_{13}K$		453.1010	15.6	34.3
465, $[Pent_3+2O-4H+Na]^+$	$C_{15}H_{22}O_{15}Na$	465.0856	465.0856	0.0	-0.1
467, $[Pent_3+2O-2H+Na]^+$	$C_{15}H_{24}O_{15}Na$	467.1032	467.1013	1.9	4.1
469, $[Pent_3+2O+Na]^+$	$C_{15}H_{26}O_{15}Na$	469.1017	469.1169	-15.2	-32.5
483, $[Pent_3+3O-2H+Na]^+$	$C_{14}H_{22}O_{14}Na$	483.1054	483.0962	9.2	19.0
485, $[Pent_3+3O+Na]^+$	$C_{14}H_{24}O_{14}Na$	485.1250	485.1119	13.1	27.1
497, $[Pent_3+4O-2H+Na]^+$	$C_{15}H_{22}O_{17}Na$	497.0750	497.0755	-0.5	-0.9
499, $[Pent_3+4O-2H+Na]^+$	$C_{15}H_{24}O_{17}Na$	499.1056	499.0911	14.5	29.0

[a]Theoretical mass and mass error considering the proposed assignment in the control mixture.

Table S2. Exact mass measurement of the oxidation products observed in the ESI-MS spectrum of the neutral fraction F2, as well as theoretical mass and mass error for the predicted formula. The lock mass was the calculated monoisotopic mass of the $[Pent_3+Na]^+$ ions.

Ions (*m/z*, proposed assignment)	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM)	
				mDa	ppm
365, $[Pent_3-C_3H_4O_3+O+Na]^+$	$C_{12}H_{22}O_{11}Na$	365.0977	365.1060	-8.3	-22.7
[a]365, $[Pent_2+4O-4H+Na]^+$	$C_{10}H_{14}O_{13}Na$		365.0332	64.5	176.6
407, $[Pent_3-CH_2O+Na]^+$	$C_{14}H_{24}O_{12}Na$	407.1252	407.1165	8.7	21.3
[a]407, $[Pent_3-C_2H_4O_2+2O-2H+Na]^+$	$C_{13}H_{20}O_{13}Na$		407.0802	45.0	110.6
425, $[Pent_3-CO+O+Na]^+$	$C_{14}H_{26}O_{13}Na$	425.1235	425.1271	-3.6	-8.5
433, $[Pent_3-4H+Na]^+$	$C_{15}H_{22}O_{13}Na$	433.1021	433.0958	6.3	14.5
435, $[Pent_3-2H+Na]^+$	$C_{15}H_{24}O_{13}Na$	435.1160	435.1115	4.5	10.4
[a]435, $[Pent_3-CH_2O+2O-4H+Na]^+$	$C_{14}H_{20}O_{14}Na$		435.0751	40.9	94.1
449, $[Pent_3+O-4H+Na]^+$	$C_{15}H_{22}O_{13}Na$	449.0840	449.0907	-6.7	-15.0
451, $[Pent_3+O-2H+Na]^+$	$C_{15}H_{24}O_{14}Na$	451.1097	451.1064	3.3	7.4
453, $[Pent_3+O+Na]^+$	$C_{15}H_{26}O_{14}Na$	453.1208	453.1220	-1.2	-2.7
467, $[Pent_3+2O-2H+Na]^+$	$C_{15}H_{24}O_{15}Na$	467.0916	467.1013	-9.7	-20.7
469, $[Pent_3+2O+Na]^+$	$C_{15}H_{26}O_{15}Na$	469.1037	469.1169	-13.2	-28.2
481, $[Pent_3+3O-4H+Na]^+$	$C_{15}H_{22}O_{16}Na$	481.0833	481.0806	2.7	5.7
483, $[Pent_3+3O-2H+Na]^+$	$C_{14}H_{22}O_{14}Na$	483.0877	483.0962	-8.5	-17.6
485, $[Pent_3+3O+Na]^+$	$C_{14}H_{24}O_{14}Na$	485.1019	485.1119	-10.0	-20.5

[a]Theoretical mass and mass error considering the most probable assignment obtained in the acidic fraction.

Table S3. Exact mass measurement of the oxidation products (excluding that used as lock mass) observed in the ESI-MS spectrum of the neutral fraction F3, as well as theoretical mass and mass error for the predicted formula. The lock mass was the calculated monoisotopic mass of the $[Pent_2+Na]^+$ ions.

Ions (*m/z*, proposed assignment)	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM)	
				mDa	ppm
303, $[Pent_2-2H+Na]^+$	$C_{10}H_{16}O_9Na$	303.0612	303.0692	-8.0	-26.4
[a]303, $[Pent_2-CH_2O+2O-4H+Na]^+$	$C_9H_{12}O_{10}Na$		303.0328	28.4	93.7
321, $[Pent_2+O+Na]^+$	$C_{10}H_{18}O_{10}Na$	321.0751	321.0798	-4.7	-14.5
333, $[Pent_3-C_4H_6O_4+O-2H+Na]^+$	$C_{11}H_{18}O_{10}Na$	333.0807	333.0798	0.9	2.8
[a]333, $[Pent_2+2O-4H+Na]^+$	$C_{10}H_{14}O_{11}Na$		333.0434	37.3	112.0
335, $[Pent_3-C_4H_6O_4+O+Na]^+$	$C_{11}H_{20}O_{10}Na$	335.0968	335.0954	1.4	4.1
[a]335, $[Pent_2+2O-2H+Na]^+$	$C_{10}H_{16}O_{11}Na$		335.0590	37.8	112.7
365, $[Pent_3-C_3H_4O_3+O+Na]^+$	$C_{12}H_{22}O_{11}Na$	365.1092	365.1060	3.2	8.8
377, $[Pent_3-C_2H_4O_2+Na]^+$	$C_{13}H_{22}O_{11}Na$	377.1106	377.1060	4.6	12.2
[a]377, $[Pent_3-C_3H_6O_3+2O-2H+Na]^+$	$C_{12}H_{18}O_{12}Na$		377.0696	41.0	108.7
379, $[Pent_3-C_2H_2O_2+Na]^+$	$C_{13}H_{24}O_{11}Na$	379.1139	379.1216	-7.7	-20.4
[a]379, $[Pent_3-C_3H_6O_3+2O+Na]^+$	$C_{12}H_{20}O_{12}Na$		379.0852	28.7	75.6
393, $[Pent_3-C_2H_4O_2+O+Na]^+$	$C_{13}H_{22}O_{12}Na$	393.1091	393.1009	8.2	20.9
409, $[Pent_3-CO+Na]^+$	$C_{14}H_{26}O_{12}Na$	409.1300	409.1322	-2.2	-5.4
[a]409, $[Pent_3-C_2H_4O_2+2O+Na]^+$	$C_{13}H_{22}O_{13}Na$		409.0958	34.2	83.6
435, $[Pent_3-2H+Na]^+$	$C_{15}H_{24}O_{13}Na$	435.1187	435.1115	7.2	16.6
[a]435, $[Pent_3-CH_2O+2O-4H+Na]^+$	$C_{14}H_{20}O_{14}Na$		435.0751	43.6	100.3
437, $[Pent_3+Na]^+$	$C_{15}H_{26}O_{13}Na$	437.1359	437.1271	8.8	20.1
[a]437, $[Pent_3-CH_2O+2O-2H+Na]^+$	$C_{14}H_{22}O_{14}Na$		437.0907	45.2	103.3
451, $[Pent_3+O-2H+Na]^+$	$C_{15}H_{24}O_{14}Na$	451.1024	451.1064	-4.0	-8.8
453, $[Pent_3+O+Na]^+$	$C_{15}H_{26}O_{14}Na$	453.124	453.1220	2.1	4.6

[a]Theoretical mass and mass error considering the most probable assignment obtained in the acidic fraction.

Table S4. Exact mass measurement of the oxidation products (excluding that used as lock mass) observed in the ESI-MS spectrum of the neutral fraction F4, as well as theoretical mass and mass error for the predicted formula. The lock mass was the calculated monoisotopic mass of the $[Pent_2+Na]^+$ ions.

Ions (*m/z*, proposed assignment)	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM)	
				mDa	ppm
187, $[Pent_1+O-2H+Na]^+$	$C_5H_8O_6Na$	187.0239	187.0219	2.0	10.9
201, $[Pent_2-C_4H_6O_4+O-2H+Na]^+$	$C_6H_{10}O_6Na$	201.0450	201.0375	7.5	37.3
215, $[Pent_2-C_3H_6O_3+Na]^+$	$C_7H_{12}O_6Na$	215.0609	215.0532	7.7	36.0
233, $[Pent_2-C_3H_4O_3+O+Na]^+$	$C_7H_{14}O_7Na$	233.0645	233.0637	0.8	3.3
245, $[Pent_2-C_2H_4O_2+Na]^+$	$C_8H_{14}O_7Na$	245.0665	245.0637	2.8	11.3
[a]245, $[Pent_2-C_3H_6O_3+2O-2H+Na]^+$	$C_7H_{10}O_8Na$		245.0273	39.2	159.8
247, $[Pent_2-C_2H_2O_2+Na]^+$	$C_8H_{16}O_7Na$	247.0841	247.0794	4.7	19.1
[a]247, $[Pent_2-C_3H_6O_3+2O+Na]^+$	$C_7H_{12}O_8Na$		247.0430	41.1	166.4
263, $[Pent_2-C_2H_2O_2+O+Na]^+$	$C_8H_{16}O_8Na$	263.0773	263.0743	3.0	11.5
275, $[Pent_2-CH_2O+Na]^+$	$C_9H_{16}O_8Na$	275.0695	275.0743	-4.8	-17.4
[a]275, $[Pent_2-C_2H_4O_2+2O-2H+Na]^+$	$C_8H_{12}O_9Na$		275.0379	31.6	114.9
277, $[Pent_2-CO+Na]^+$	$C_9H_{18}O_8Na$	277.0853	277.0899	-4.6	-16.7
[a]277, $[Pent_2-C_2H_4O_2+2O+Na]^+$	$C_8H_{14}O_9Na$		277.0536	31.7	114.6
291, $[Pent_2-CH_2O+O+Na]^+$	$C_9H_{16}O_9Na$	291.0670	291.0692	-2.2	-7.6
293, $[Pent_2-CO+O+Na]^+$	$C_9H_{18}O_9Na$	293.0840	293.0849	-0.9	-2.9
303, $[Pent_2-2H+Na]^+$	$C_{10}H_{16}O_9Na$	303.0691	303.0692	-0.1	-0.3
[a]303, $[Pent_2-CH_2O+2O-4H+Na]^+$	$C_9H_{12}O_{10}Na$		303.0328	36.3	119.7
319, $[Pent_2+O-2H+Na]^+$	$C_{10}H_{16}O_{10}Na$	319.0536	319.0641	-10.5	-33.0
321, $[Pent_2+O+Na]^+$	$C_{10}H_{18}O_{10}Na$	321.0767	321.0798	-3.1	-9.6
333, $[Pent_3-C_4H_6O_4+O+Na]^+$	$C_{11}H_{18}O_{10}Na$	333.0809	333.0798	1.1	3.4
[a]333, $[Pent_2+2O-4H+Na]^+$	$C_{10}H_{14}O_{11}Na$		333.0434	37.5	112.6
335, $[Pent_3-C_4H_6O_4+O+Na]^+$	$C_{11}H_{20}O_{10}Na$	335.0930	335.0954	-2.4	-7.2
[a]335, $[Pent_2+2O-2H+Na]^+$	$C_{10}H_{16}O_{11}Na$		335.0590	34.0	101.4
337, $[Pent_2+2O+Na]^+$	$C_{10}H_{18}O_{11}Na$	337.0862	337.0747	11.5	34.2
347, $[Pent_3-C_3H_6O_3+Na]^+$	$C_{12}H_{20}O_{10}Na$	347.0991	347.0954	3.7	10.6
363, $[Pent_3-C_3H_6O_3+O+Na]^+$	$C_{12}H_{20}O_{11}Na$	363.0869	363.0903	-3.4	-9.5
365, $[Pent_3-C_3H_4O_3+O+Na]^+$	$C_{12}H_{22}O_{11}Na$	365.1169	365.1060	10.9	29.9
421, $[Pent_3-CH_2O+O-2H+Na]^+$	$C_{14}H_{22}O_{13}Na$	421.0995	421.0958	3.7	8.8
435, $[Pent_3-2H+Na]^+$	$C_{15}H_{24}O_{13}Na$	435.1113	435.1115	-0.2	-0.4
[a]435, $[Pent_3-CH_2O+2O-4H+Na]^+$	$C_{14}H_{20}O_{14}Na$		435.0751	36.2	83.3
437, $[Pent_3+Na]^+$	$C_{15}H_{26}O_{13}Na$	437.1271	437.1271	0.0	0.0
[a]437, $[Pent_3-CH_2O+2O-2H+Na]^+$	$C_{14}H_{22}O_{14}Na$		437.0907	36.4	83.2
451, $[Pent_3+O-2H+Na]^+$	$C_{15}H_{24}O_{14}Na$	451.1053	451.1064	-1.1	-2.4
453, $[Pent_3+O+Na]^+$	$C_{15}H_{26}O_{14}Na$	453.1188	453.1220	-3.2	-7.1

[a]Theoretical mass and mass error considering the most probable assignment obtained in the acidic fraction.

Table S5. Exact mass measurement of the oxidation products observed in the ESI-MS spectrum of the neutral fraction F5 with a relative abundance higher than 20%, as well as theoretical mass and mass error for the predicted formula. The lock mass was the calculated monoisotopic mass of $[M+Na]^+$ ions of the (β1→4)-D-mannotriose used as internal standard.

Ions (*m/z*, proposed assignment)	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM)	
				mDa	ppm
171, $[Pent_1+O+Na]^+$	$C_5H_8O_5Na$	171.0260	171.0269	-0.9	-5.5
173, $[Pent_1+Na]^+$	$C_5H_{10}O_5Na$	173.0432	173.0426	0.6	3.5
189, $[Pent_1+O+Na]^+$	$C_5H_{10}O_6Na$	189.0366	189.0375	-0.9	-4.8
203, $[Pent_2-C_4H_6O_4+O+Na]^+$	$C_6H_{12}O_6Na$	203.0541	203.0532	0.9	4.6
[a]203, $[Pent_1+2O-2H+Na]^+$	$C_5H_8O_7Na$		203.0168	37.3	183.8
231, $[Pent_2-C_3H_6O_3+O+Na]^+$	$C_7H_{12}O_7Na$	231.0499	231.0481	1.8	7.9
233, $[Pent_2-C_3H_4O_3+O+Na]^+$	$C_7H_{14}O_7Na$	233.0703	233.0637	6.6	28.2
245, $[Pent_2-C_2H_2O_2+Na]^+$	$C_8H_{14}O_7Na$	245.0646	245.0637	0.9	3.6
[a]245, $[Pent_2-C_3H_6O_3+2O-2H+Na]^+$	$C_7H_{10}O_8Na$		245.0273	37.3	152.1
259, $[Pent_2-C_2H_4O_2+O-2H+Na]^+$	$C_8H_{12}O_8Na$	259.0448	259.0430	1.8	7.0
261, $[Pent_2-C_2H_4O_2+O+Na]^+$	$C_8H_{14}O_8Na$	261.0668	261.0586	8.2	31.3
263, $Pent_2-C_2H_2O_2+O+Na]^+$	$C_8H_{16}O_8Na$	263.0810	263.0743	6.7	25.5
289, $[Pent_2-CH_2O+O-2H+Na]^+$	$C_9H_{14}O_9Na$	289.0541	289.0536	0.5	1.9
291, $[Pent_2-CH_2O+O+Na]^+$	$C_9H_{16}O_9Na$	291.0736	291.0692	4.4	15.1
303, $[Pent_2-2H+Na]^+$	$C_{10}H_{16}O_9Na$	303.0706	303.0692	1.4	4.6
[a]303, $[Pent_2-CH_2O+2O-4H+Na]^+$	$C_9H_{12}O_{10}Na$		303.0328	37.8	124.7

Table S5 (Continued). Exact mass measurement of the oxidation products observed in the ESI-MS spectrum of the neutral fraction F5 with a relative abundance higher than 20%, as well as theoretical mass and mass error for the predicted formula. The lock mass was the calculated monoisotopic mass of $[M+Na]^+$ ions of the (β1→4)-D-mannotriose used as internal standard.

Ions (*m/z*, proposed assignment)	Predicted formula	Exact mass (EM, Da)	Theoretical mass (TM, Da)	Error (EM-TM) mDa	mDa
305, $[Pent_2+Na]^+$	$C_{10}H_{18}O_9Na$	305.0867	305.0849	1.8	6.1
[a]305, $[Pent_2-CH_2O+2O-2H+Na]^+$	$C_9H_{14}O_{10}Na$		305.0485	38.2	125.3
307, $[Pent_2-CH_2O+2O+Na]^+$	$C_9H_{16}O_{10}Na$	307.0595	307.0641	-4.6	-15.0
317, $[Pent_2+O-4H+Na]^+$	$C_{10}H_{14}O_{10}Na$	317.0475	317.0485	-1.0	-3.0
319, $[Pent_2+O-2H+Na]^+$	$C_{10}H_{16}O_{10}Na$	319.0644	319.0641	0.3	0.9
321, $[Pent_2+O+Na]^+$	$C_{10}H_{18}O_{10}Na$	321.0807	321.0798	0.9	2.9
337, $[Pent_2+2O+Na]^+$	$C_{10}H_{18}O_{11}Na$	337.0683	337.0747	-6.4	-18.9
423, $[Pent_3-CH_2O+O+Na]^+$	$C_{14}H_{24}O_{13}Na$	423.1019	423.1115	-9.6	-22.6
439, $[Pent_3-CH_2O+2O+Na]^+$	$C_{14}H_{24}O_{14}Na$	439.1072	439.1064	0.8	1.9
451, $[Pent_3+O-2H+Na]^+$	$C_{15}H_{24}O_{14}Na$	451.1105	451.1064	4.1	9.1

[a] Theoretical mass and mass error considering the most probable assignment obtained in the acidic fraction.

SUPPLEMENTARY FIGURES

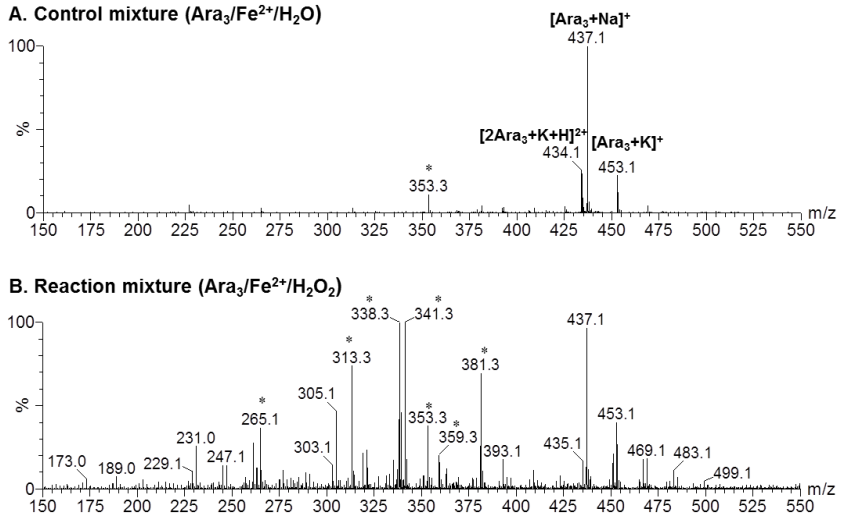

Figure S1. ESI-MS spectra of the (A) control mixture and (B) reaction mixture. *Impurity

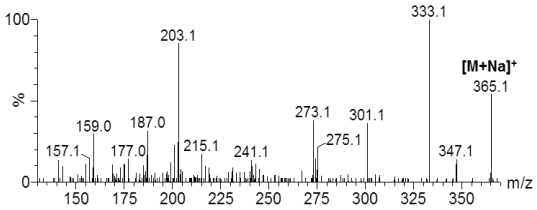

Figure S2. ESI-MS/MS spectrum acquired for the ion at *m/z* 365 ($[Pent_3-C_3H_4O_3+O+Na]^+$) from a neutral fraction.

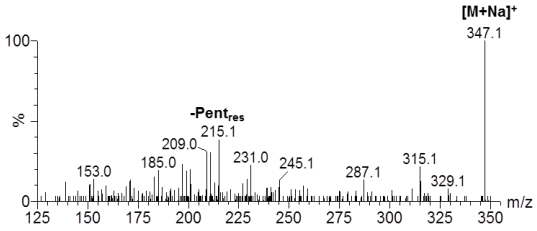

Figure S3. ESI-MS/MS spectrum acquired for the ion at *m/z* 347 ([$Pent_3$-$C_3H_6O_3$+Na]$^+$) from a neutral fraction.

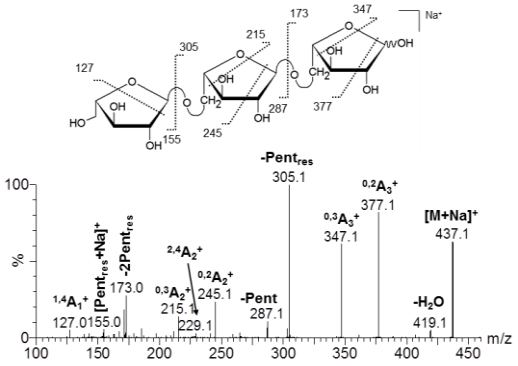

Figure S4. ESI-MS/MS spectrum of [M+Na]$^+$ ions of Ara_3 (*m/z* 437) acquired from the stock solution.

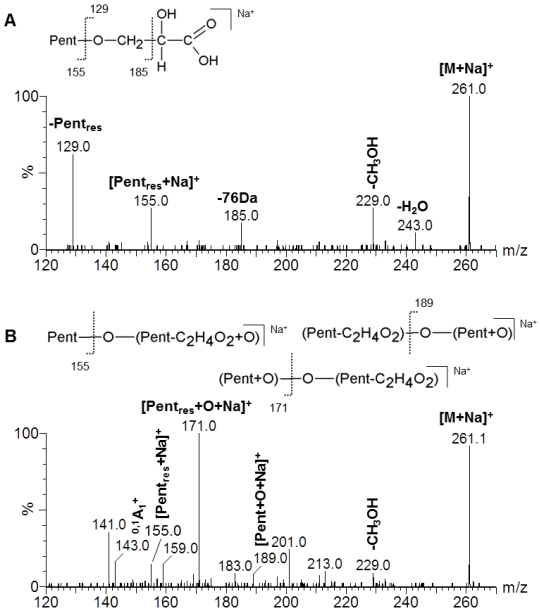

Figure S5. ESI-MS/MS spectra acquired for the ion at *m/z* 261 ([$Pent_2$-$C_2H_4O_2$+O+Na]$^+$) from the (A) acidic and (B) a neutral fraction.

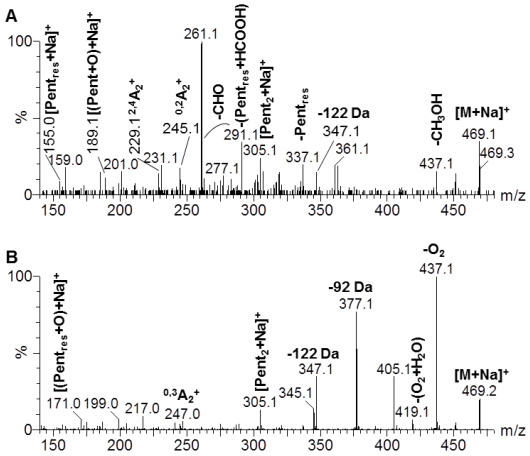

Figure S6. ESI-MS/MS spectra acquired for the ion at *m/z* 469 ([$Pent_3$+2O+$Na]^+$) from the (A) acidic fraction and (B) a neutral fraction.

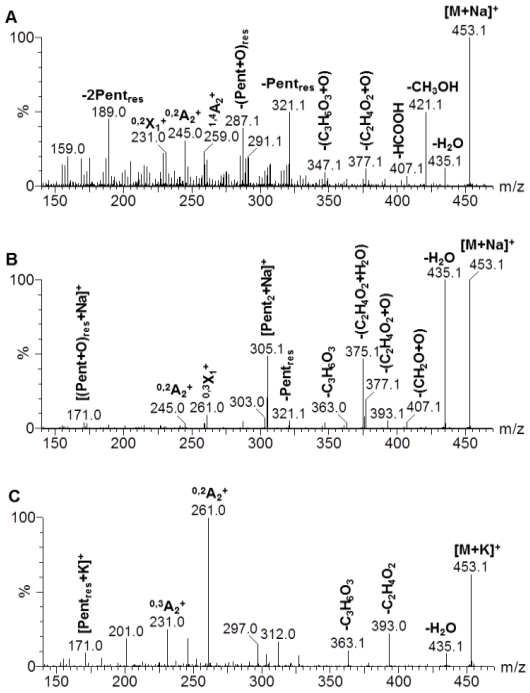

Figure S7. ESI-MS/MS spectra acquired for the ion at *m/z* 453 from the (A) acidic fraction, (B) a neutral fraction, and (C) the control mixture.

APPENDIX C. SUPPLEMENTARY MATERIAL OF SECTION III.2.2

SUPPLEMENTARY TABLES

Table S1. Total mass loss and mass loss from 150 °C (%) during each thermal treatment and water-solubility (%) of the resulting compounds.

Sample	T1		T2		T3	
A25M75	7.1 (0.7)[c]	100[d]	18.1 (11.0)	100	21.8 (15.0)	68.8
A50M50	7.5 (3.5)	100	19.5 (15.1)	100	24.1 (19.1)	40.7
A75M25	6.1 (3.4)	100	21.8 (17.8)	65.2	18.5 (18.0)	44.4
Man_3[a]	7.7 (0.1)	100	15.8 (8.1)	100	20.0 (12.3)	100
Ara_3[b]	6.0 (3.5)	100	19.4 (18.1)	35.3	21.7 (20.2)	24.6

[a]Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., Simões, J., & Domingues, M. R. M. (2011). Evaluation of the effect of roasting on the structure of coffee galactomannans using model oligosaccharides. *Journal of Agricultural and Food Chemistry*, *59*, 10078-10087. [b]Moreira, A. S. P., Coimbra, M. A., Nunes, F. M., & Domingues, M. R. M. (2013). Roasting-induced changes in arabinotriose, a model of coffee arabinogalactan side chains. *Food Chemistry*, *138*, 2291-2299 (Section III.1.1 of this thesis). [c]Percentage of total mass loss. The values in brackets are the percentages considering the mass losses from 150 °C, excluding the initial mass loss due to the water evaporation. [d]Percentage of water-soluble material.

Table S2. Accurate masses found by Q Exactive Orbitrap for the ions identified after heating of the A50M50 mixture submitted to the thermal treatment T3. The theoretical mass and the difference between the theoretical and experimental masses for each predicted formula were obtained from Xcalibur software.

Experimental mass (*m/z*)	Theoretical mass (*m/z*)	Mass error (ppm)	RDB equiv.	Composition	Proposed assignment(s)
149.02100	149.02092	0.57	3.5	$C_6H_6O_3Na$	$[Hex\text{-}3H_2O+Na]^+$
173.04217	173.04204	0.72	0.5	$C_5H_{10}O_5Na$	$[Pent+Na]^+$
185.04212	185.04204	0.41	1.5	$C_6H_{10}O_5Na$	$[Hex\text{-}H_2O+Na]^+$
193.04726	193.04713	0.67	3.5	$C_8H_{10}O_4Na$	$[Pent_2\text{-}C_2H_8O_5+Na]^+$
203.05264	203.05261	0.15	0.5	$C_6H_{12}O_6Na$	$[Hex+Na]^+$
217.06830	217.06826	0.19	0.5	$C_7H_{14}O_6Na$	$[Pent_2\text{-}C_3H_4O_3+Na]^+$
227.05260	227.05261	-0.04	2.5	$C_8H_{12}O_6Na$	$[Pent_2\text{-}C_2H_6O_3+Na]^+$
257.04194	257.04204	-0.41	7.5	$C_{12}H_{10}O_5Na$	$[Hex_2\text{-}6H_2O+Na]^+$
281.06329	281.06317	0.41	4.5	$C_{11}H_{14}O_7Na$	$[PentHex\text{-}3H_2O+Na]^+$
287.07383	287.07374	0.32	2.5	$C_{10}H_{16}O_8Na$	$[Pent_2\text{-}H_2O+Na]^+$
293.06311	293.06317	-0.22	5.5	$C_{12}H_{14}O_7Na$	$[Hex_2\text{-}4H_2O+Na]^+$
305.08438	305.08430	0.25	1.5	$C_{10}H_{18}O_9Na$	$[Pent_2+Na]^+$
311.07367	311.07374	-0.22	4.5	$C_{12}H_{16}O_8Na$	$[Hex_2\text{-}3H_2O+Na]^+$
317.08445	317.08430	0.46	2.5	$C_{11}H_{18}O_9Na$	$[PentHex\text{-}H_2O+Na]^+$
325.08934	325.08939	-0.15	4.5	$C_{13}H_{18}O_8Na$	$[Pent_3\text{-}C_2H_8O_5+Na]^+$
347.09506	347.09487	0.55	2.5	$C_{12}H_{20}O_{10}Na$	$[Pent_3\text{-}C_3H_6O_3+Na]^+$/ $[Hex_2\text{-}H_2O+Na]^+$
365.10562	365.10543	0.51	1.5	$C_{12}H_{22}O_{11}Na$	$[Pent_3\text{-}C_3H_4O_2+Na]^+/[Hex_2+Na]^+$
377.10559	377.10543	0.42	2.5	$C_{13}H_{22}O_{11}Na$	$[Pent_3\text{-}C_2H_4O_2+Na]^+$

Table S2 (Continued). Accurate masses found by Q Exactive Orbitrap for the ions identified after heating of the A50M50 mixture submitted to the thermal treatment T3. The theoretical mass and the difference between the theoretical and experimental masses for each predicted formula were obtained from Xcalibur software.

Experimental mass (*m/z*)	Theoretical mass (*m/z*)	Mass error (ppm)	RDB equiv.	Composition	Proposed assignment(s)
383.09547	383.09487	1.57	5.5	$C_{15}H_{20}O_{10}Na$	$[Pent_3\text{-}3H_2O+Na]^+$
389.08430	389.08430	-0.01	8.5	$C_{17}H_{18}O_9Na$	$[PentHex_2\text{-}6H_2O+Na]^+$
395.09490	395.09487	0.08	6.5	$C_{16}H_{20}O_{10}Na$	$[Pent_2Hex\text{-}4H_2O+Na]^+$
401.1055	401.10543	0.17	4.5	$C_{15}H_{22}O_{11}Na$	$[Pent_3\text{-}2H_2O+Na]^+$
413.10549	413.10543	0.14	5.5	$C_{16}H_{22}O_{11}Na$	$[Pent_2Hex\text{-}3H_2O+Na]^+$
419.09530	419.09487	1.03	8.5	$C_{18}H_{20}O_{10}Na$	$[Hex_3\text{-}6H_2O+Na]^+$
419.11632	419.11600	0.77	3.5	$C_{15}H_{24}O_{12}Na$	$[Pent_3\text{-}H_2O+Na]^+$
425.10582	425.10543	0.91	6.5	$C_{17}H_{22}O_{11}Na$	$[PentHex_2\text{-}4H_2O+Na]^+$
435.11127	435.11091	0.82	3.5	$C_{15}H_{24}O_{13}Na$	$[Pent_3\text{-}2H+Na]^+$
437.10598	437.10543	1.25	7.5	$C_{18}H_{22}O_{11}Na$	$[Hex_3\text{-}5H_2O+Na]^+$
437.12705	437.12656	1.12	2.5	$C_{15}H_{26}O_{13}Na$	$[Pent_3+Na]^+$
443.11643	443.11600	0.98	5.5	$C_{17}H_{24}O_{12}Na$	$[PentHex_2\text{-}3H_2O+Na]^+$
449.12700	449.12656	0.98	3.5	$C_{16}H_{26}O_{13}Na$	$[Pent_2Hex\text{-}H_2O+Na]^+$
455.11645	455.11600	0.99	6.5	$C_{18}H_{24}O_{12}Na$	$[Hex_3\text{-}4H_2O+Na]^+$
473.12717	473.12656	1.29	5.5	$C_{18}H_{26}O_{13}Na$	$[Hex_3\text{-}3H_2O+Na]^+$
479.13740	479.13713	0.57	3.5	$C_{17}H_{28}O_{14}Na$	$[Pent_4\text{-}C_3H_6O_3+Na]^+$/ $[PentHex_2\text{-}H_2O+Na]^+$
497.14795	497.14769	0.52	2.5	$C_{17}H_{30}O_{15}Na$	$[Pent_4\text{-}C_3H_4O_2+Na]^+$/ $[PentHex_2+Na]^+$
509.14804	509.14769	0.69	3.5	$C_{18}H_{30}O_{15}Na$	$[Pent_4\text{-}C_2H_4O_2+Na]^+$/ $[Hex_3\text{-}H_2O+Na]^+$
521.12665	521.12656	0.17	9.5	$C_{22}H_{26}O_{13}Na$	$[Pent_2Hex_2\text{-}6H_2O+Na]^+$
527.13722	527.13713	0.18	7.5	$C_{21}H_{28}O_{14}Na$	$[Pent_3Hex\text{-}4H_2O+Na]^+$
527.15853	527.15826	0.52	2.5	$C_{18}H_{32}O_{16}Na$	$[Hex_3+Na]^+$
545.14783	545.14769	0.25	6.5	$C_{21}H_{30}O_{15}Na$	$[Pent_3Hex\text{-}3H_2O+Na]^+$
551.13721	551.13713	0.15	9.5	$C_{23}H_{28}O_{14}Na$	$[PentHex_3\text{-}6H_2O+Na]^+$
551.15827	551.15826	0.03	4.5	$C_{20}H_{32}O_{16}Na$	$[Pent_4\text{-}H_2O+Na]^+$
569.14744	569.14769	-0.44	8.5	$C_{23}H_{30}O_{15}Na$	$[PentHex_3\text{-}5H_2O+Na]^+$
569.16916	569.16882	0.6	3.5	$C_{20}H_{34}O_{17}Na$	$[Pent_4+Na]^+$
575.15838	575.15826	0.22	6.5	$C_{22}H_{32}O_{16}Na$	$[Pent_2Hex_2\text{-}3H_2O+Na]^+$
581.14752	581.14769	-0.29	9.5	$C_{24}H_{30}O_{15}Na$	$[Hex_4\text{-}6H_2O+Na]^+$
581.16914	581.16882	0.55	4.5	$C_{21}H_{34}O_{17}Na$	$[Pent_3Hex\text{-}H_2O+Na]^+$
587.15813	587.15826	-0.21	7.5	$C_{23}H_{32}O_{16}Na$	$[PentHex_3\text{-}4H_2O+Na]^+$
605.16876	605.16882	-0.1	6.5	$C_{23}H_{34}O_{17}Na$	$[PentHex_3\text{-}3H_2O+Na]^+$
611.17959	611.17939	0.34	4.5	$C_{22}H_{36}O_{18}Na$	$[Pent_5\text{-}C_3H_6O_3+Na]^+$/ $[Pent_2Hex_2\text{-}H_2O+Na]^+$
629.19019	629.18995	0.38	3.5	$C_{22}H_{38}O_{19}Na$	$[Pent_5\text{-}C_3H_4O_2+Na]^+$/ $[Pent_2Hex_2+Na]^+$
641.19043	641.18995	0.75	4.5	$C_{23}H_{38}O_{19}Na$	$[Pent_5\text{-}C_2H_4O_2+Na]^+$/ $[PentHex_3\text{-}H_2O+Na]^+$
659.17853	659.17939	-1.3	8.5	$C_{26}H_{36}O_{18}Na$	$[Pent_4Hex\text{-}4H_2O+Na]^+$
659.20026	659.20051	-0.39	3.5	$C_{23}H_{40}O_{20}Na$	$[PentHex_3+Na]^+$
677.18982	677.18995	-0.19	7.5	$C_{26}H_{38}O_{19}Na$	$[Pent_4Hex\text{-}3H_2O+Na]^+$

Table S2 (Continued). Accurate masses found by Q Exactive Orbitrap for the ions identified after heating of the A50M50 mixture submitted to the thermal treatment T3. The theoretical mass and the difference between the theoretical and experimental masses for each predicted formula were obtained from Xcalibur software.

Experimental mass (*m/z*)	Theoretical mass (*m/z*)	Mass error (ppm)	RDB equiv.	Composition	Proposed assignment(s)
683.17914	683.17939	-0.36	10.5	$C_{28}H_{36}O_{18}Na$	$[Pent_2Hex_3\text{-}6H_2O+Na]^+$
683.20005	683.20051	-0.68	5.5	$C_{25}H_{40}O_{20}Na$	$[Pent_5\text{-}H_2O+Na]^+$
701.18960	701.18995	-0.5	9.5	$C_{28}H_{38}O_{19}Na$	$[Pent_2Hex_3\text{-}5H_2O+Na]^+$
701.21035	701.21108	-1.04	4.5	$C_{25}H_{42}O_{21}Na$	$[Pent_5+Na]^+$
713.18999	713.18995	0.06	10.5	$C_{29}H_{38}O_{19}Na$	$[PentHex_4\text{-}6H_2O+Na]^+$
713.21023	713.21108	-1.19	5.5	$C_{26}H_{42}O_{21}Na$	$[Pent_4Hex\text{-}H_2O+Na]^+$
743.20107	743.20051	0.75	10.5	$C_{30}H_{40}O_{20}Na$	$[Hex_5\text{-}6H_2O+Na]^+$
743.22125	743.22164	-0.53	5.5	$C_{27}H_{44}O_{22}Na$	$[Pent_6\text{-}C_3H_6O_3+Na]^+$/ $[Pent_3Hex_2\text{-}H_2O+Na]^+$

SUPPLEMENTARY FIGURES

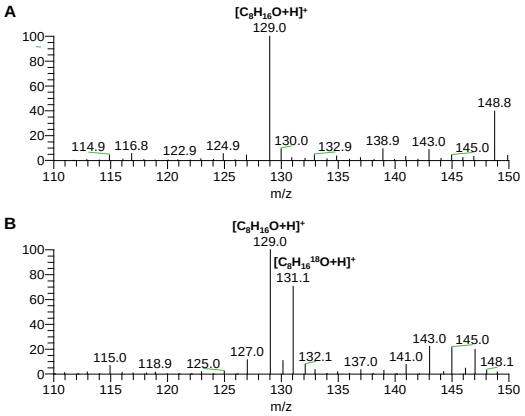

Figure S1. ESI-MS spectra of 3-octanone ($C_8H_{16}O$), (A) before and (B) after ^{18}O-labelling procedure (B).

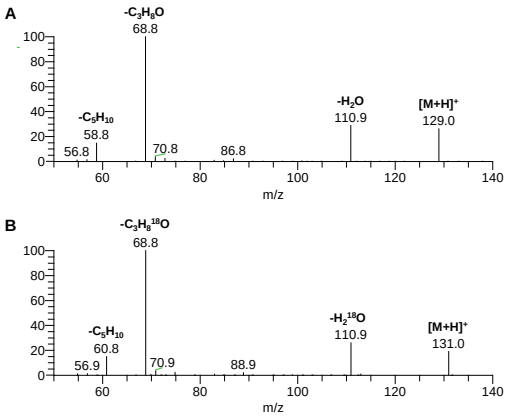

Figure S2. ESI-MS^2 spectra of $[M+H]^+$ ions of (A) unlabelled (*m/z* 129) and (B) ^{18}O-labelled 3-octanone (*m/z* 131).

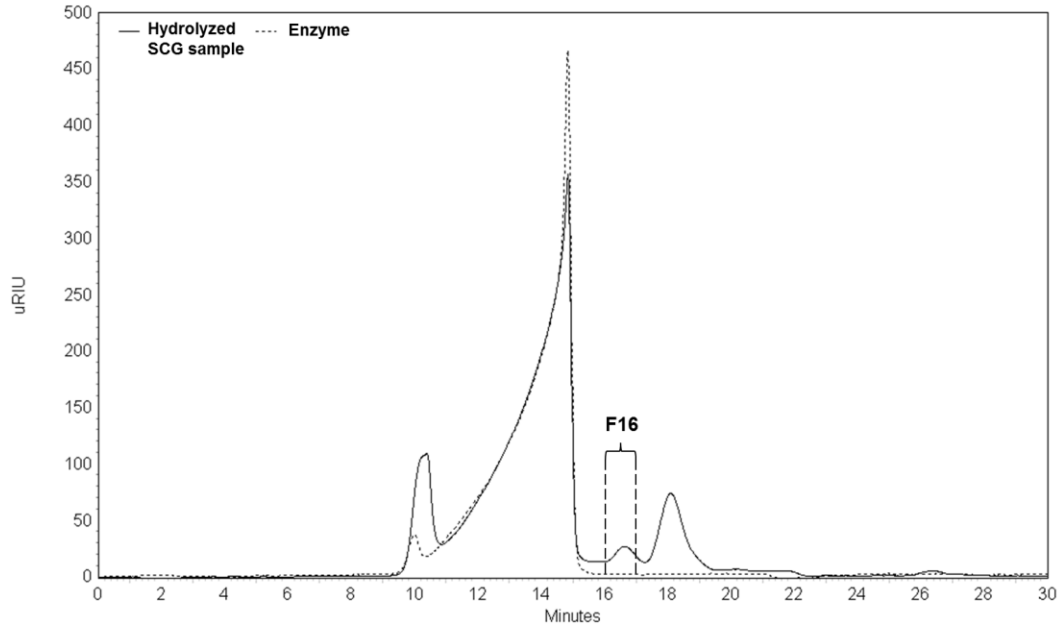

Figure S3. LEX-SEC chromatograms of the hydrolysed material obtained after enzymatic treatment of a SCG galactomannan-rich fraction with an *endo*-(β1→4)-D-mannanase and the respective enzyme preparation, both in buffer. F16 - Fraction that eluted between 16-17 min.

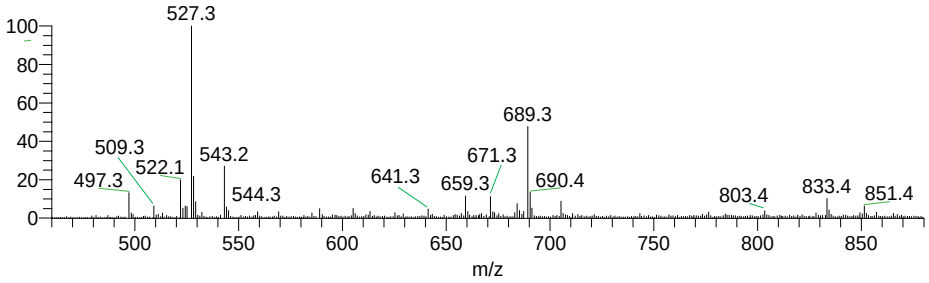

Figure S4. ESI-LIT-MS spectrum of F16 obtained by LEX-SEC of the hydrolysed material obtained after enzymatic treatment of a SCG galactomannan-rich fraction.

APPENDIX D. SUPPLEMENTARY MATERIAL OF SECTION III.3.1

SUPPLEMENTARY TABLES

Table S1. Accurate masses found by LTQ-Orbitrap for the ions identified after heating of the model mixture to 175 ºC. The theoretical mass and the difference between the theoretical and experimental masses for each predicted formula were obtained from Xcalibur software.

Experimental mass (*m/z*)	Theoretical mass (*m/z*)	Mass error (ppm)	Composition	Proposed assignment
305.0866	305.0867	-0.22	$C_{12}H_{17}O_9$	$[PentQA-H_2O-H]^-$
323.0972	323.0973	-0.23	$C_{12}H_{19}O_{10}$	$[PentQA-H]^-$
335.0760	335.0761	-0.43	$C_{16}H_{15}O_8$	$[CQA-H_2O-H]^-$
353.0865	353.0867	-0.59	$C_{16}H_{17}O_9$	$[CQA-H]^-$
425.1075	425.1078	-0.79	$C_{19}H_{21}O_{11}$	$[Pent_2CA-H_2O-H]^-$
431.2270	431.2276	-1.3	$C_{21}H_{35}O_9$	$[PentCQA-3H_2O-H]^-$
437.1286	437.1290	-0.84	$C_{17}H_{25}O_{13}$	$[Pent_2QA-H_2O-H]^-$
443.1186	443.1184	0.45	$C_{19}H_{23}O_{12}$	$[Pent_2CA-H]^-$
449.1057	449.1078	-4.76	$C_{21}H_{21}O_{11}$	$[PentCQA-2H_2O-H]^-$
455.1393	455.1395	-0.51	$C_{17}H_{27}O_{14}$	$[Pent_2QA-H]^-$
467.1181	467.1184	-0.65	$C_{21}H_{23}O_{12}$	$[PentCQA-H_2O-H]^-$
485.1287	485.1290	-0.55	$C_{21}H_{25}O_{13}$	$[PentCQA-H]^-$
497.1079	497.1078	0.04	$C_{25}H_{21}O_{11}$	$[(CQA)CA-H_2O-H]^-$
515.1182	515.1184	-0.37	$C_{25}H_{23}O_{12}$	$[(CQA)CA-H]^-$
527.1393	527.1395	-0.5	$C_{23}H_{27}O_{14}$	$[(CQA)QA-H]^-$
563.1379	563.1395	-2.9	$C_{26}H_{27}O_{14}$	$[Pent_2CQA-3H_2O-H]^-$
569.1707	569.1712	-0.92	$C_{22}H_{33}O_{17}$	$[Pent_3QA-H_2O-H]^-$
575.1606	575.1607	-0.11	$C_{24}H_{31}O_{16}$	$[Pent_3CA-H]^-$
581.1476	581.1501	-4.3	$C_{26}H_{29}O_{15}$	$[Pent_2CQA-2H_2O-H]^-$
587.1812	587.1818	-1.01	$C_{22}H_{35}O_{18}$	$[Pent_3QA-H]^-$
599.1601	599.1607	-0.94	$C_{26}H_{31}O_{16}$	$[Pent_2CQA-H_2O-H]^-$
617.1707	617.1712	-0.85	$C_{26}H_{33}O_{17}$	$[Pent_2CQA-H]^-$
689.1706	689.1712	-0.91	$C_{32}H_{33}O_{17}$	$[CQA_2-H]^-$
695.1801	695.1818	-2.45	$C_{31}H_{35}O_{18}$	$[Pent_3CQA-3H_2O-H]^-$
713.1896	713.1924	-3.86	$C_{31}H_{37}O_{19}$	$[Pent_3CQA-2H_2O-H]^-$
719.2231	719.2240	-1.32	$C_{27}H_{43}O_{22}$	$[Pent_4QA-H]^-$
731.2021	731.2029	-1.12	$C_{31}H_{39}O_{20}$	$[Pent_3CQA-H_2O-H]^-$
749.2126	749.2135	-1.18	$C_{31}H_{41}O_{21}$	$[Pent_3CQA-H]^-$
803.2024	803.2029	-0.65	$C_{37}H_{39}O_{20}$	$[PentCQA_2-H_2O-H]^-$
821.2126	821.2135	-1.08	$C_{37}H_{41}O_{21}$	$[PentCQA_2-H]^-$
827.2223	827.224	-2.15	$C_{36}H_{43}O_{22}$	$[Pent_4CQA-3H_2O-H]^-$
845.2318	845.2346	-3.33	$C_{36}H_{45}O_{23}$	$[Pent_4CQA-2H_2O-H]^-$
851.2654	851.2663	-1.07	$C_{32}H_{51}O_{26}$	$[Pent_5QA-H]^-$
863.2445	863.2452	-0.79	$C_{36}H_{47}O_{24}$	$[Pent_4CQA-H_2O-H]^-$
881.2548	881.2557	-1.07	$C_{36}H_{49}O_{25}$	$[Pent_4CQA-H]^-$

Table S1 (Continued). Accurate masses found by LTQ-Orbitrap for the ions identified after heating of the model mixture to 175 ºC. The theoretical mass and the difference between the theoretical and experimental masses for each predicted formula were obtained from Xcalibur software.

Experimental mass (*m/z*)	Theoretical mass (*m/z*)	Mass error (ppm)	Composition	Proposed assignment
935.2440	935.2452	-1.26	$C_{42}H_{47}O_{24}$	$[Pent_2CQA_2\text{-}H_2O\text{-}H]^-$
953.2548	953.2557	-0.99	$C_{42}H_{49}O_{25}$	$[Pent_2CQA_2\text{-}H]^-$
959.2662	959.2663	-0.07	$C_{41}H_{51}O_{26}$	$[Pent_5CQA\text{-}3H_2O\text{-}H]^-$
977.2740	977.2769	-2.94	$C_{41}H_{53}O_{27}$	$[Pent_5CQA\text{-}2H_2O\text{-}H]^-$
995.28610	995.28740	-1.34	$C_{41}H_{55}O_{28}$	$[Pent_5CQA\text{-}H_2O\text{-}H]^-$
1013.2974	1013.2980	-0.59	$C_{41}H_{57}O_{29}$	$[Pent_5CQA\text{-}H]^-$
1085.2968	1085.2980	-1.11	$C_{47}H_{57}O_{29}$	$[Pent_3CQA_2\text{-}H]^-$
1091.3083	1091.3086	-0.23	$C_{46}H_{59}O_{30}$	$[Pent_6CQA\text{-}3H_2O\text{-}H]^-$
1109.3162	1109.3191	-2.64	$C_{46}H_{61}O_{31}$	$[Pent_6CQA\text{-}2H_2O\text{-}H]^-$
1127.3284	1127.3297	-1.15	$C_{46}H_{63}O_{32}$	$[Pent_6CQA\text{-}H_2O\text{-}H]^-$
1145.3391	1145.3403	-1.01	$C_{46}H_{65}O_{33}$	$[Pent_6CQA\text{-}H]^-$
1217.3388	1217.3403	-1.21	$C_{52}H_{65}O_{33}$	$[Pent_4CQA_2\text{-}H]^-$
1241.3573	1241.3614	-3.29	$C_{51}H_{69}O_{35}$	$[Pent_7CQA\text{-}2H_2O\text{-}H]^-$
1259.3709	1259.3720	-0.84	$C_{51}H_{71}O_{36}$	$[Pent_7CQA\text{-}H_2O\text{-}H]^-$
1277.3809	1277.3825	-1.27	$C_{51}H_{73}O_{37}$	$[Pent_7CQA\text{-}H]^-$
1349.3804	1349.3825	-1.57	$C_{57}H_{73}O_{37}$	$[Pent_5CQA_2\text{-}H]^-$
1391.4124	1391.4142	-1.3	$C_{56}H_{79}O_{40}$	$[Pent_8CQA\text{-}H_2O\text{-}H]^-$
1409.4228	1409.4248	-1.4	$C_{56}H_{81}O_{41}$	$[Pent_8CQA\text{-}H]^-$
1523.4544	1523.4565	-1.33	$C_{61}H_{87}O_{44}$	$[Pent_9CQA\text{-}H_2O\text{-}H]^-$
1541.4655	1541.4670	-0.98	$C_{61}H_{89}O_{45}$	$[Pent_9CQA\text{-}H]^-$
1655.5000	1655.4987	0.78	$C_{66}H_{95}O_{48}$	$[Pent_{10}CQA\text{-}H_2O\text{-}H]^-$
1673.5078	1673.5093	-0.88	$C_{66}H_{97}O_{49}$	$[Pent_{10}CQA\text{-}H]^-$
1805.5493	1805.5516	-1.24	$C_{71}H_{105}O_{53}$	$[Pent_{11}CQA\text{-}H]^-$
1937.5905	1937.5938	-1.71	$C_{76}H_{113}O_{57}$	$[Pent_{12}CQA\text{-}H]^-$

SUPPLEMENTARY FIGURES

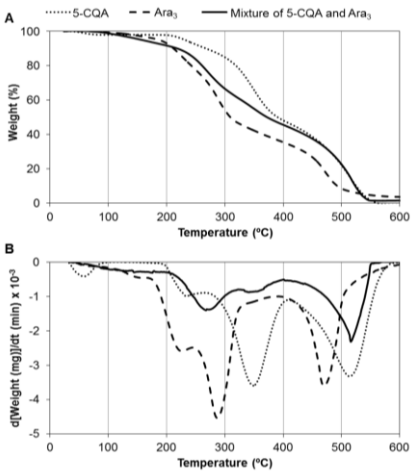

Figure S1. (A) TG and (B) DTG curves of the 5-*O*-caffeoylquinic acid (5-CQA), α-(1→5)-L-arabinotriose (Ara_3), and their equimolar mixture.

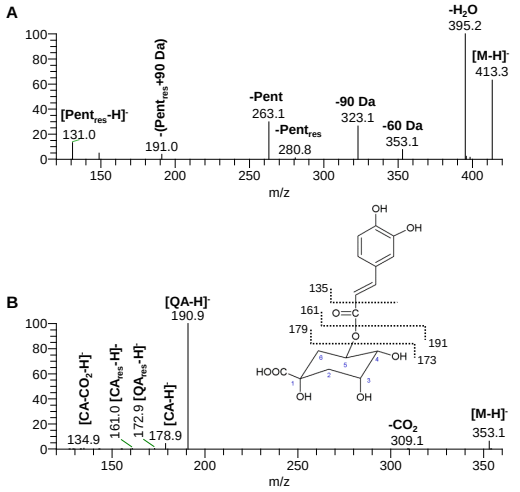

Figure S2. ESI-MS^2 spectra of [M-H]$^-$ ions of (A) Ara$_3$ (*m/z* 413) and (B) 5-CQA (*m/z* 353) acquired from the untreated mixture.

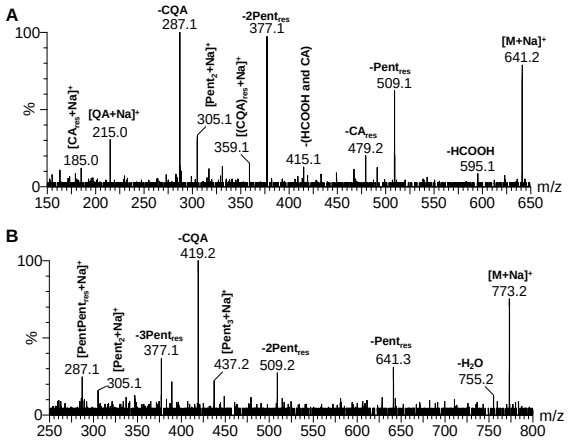

Figure S3. ESI-MS/MS spectra of the ions observed at *m/z* (A) 641 ([Pent$_2$CQA+Na]$^+$) and (B) 773 ([Pent$_3$CQA+Na]$^+$), acquired from the mixture heated at 175 °C for 30 min.

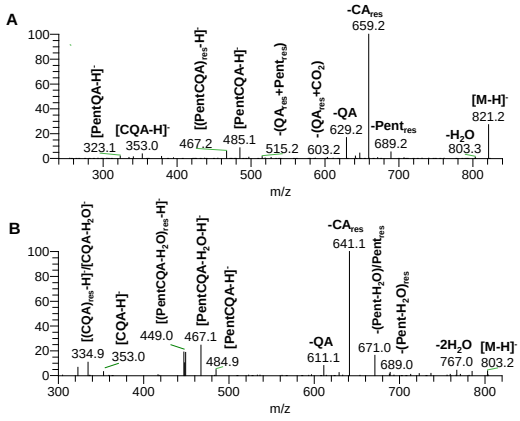

Figure S4. ESI-MS^2 spectra of the ions observed at *m/z* (A) 821 ([Pent(CQA)$_2$-H]$^-$) and (B) 803 ([Pent(CQA)$_2$-H$_2$O-H]$^-$), acquired from the mixture heated at 175 °C for 30 min (175T2).

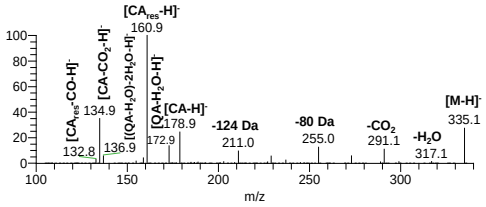

Figure S5. ESI-MS^2 spectrum of the ion observed at *m/z* 335 [CQA-H_2O-H]$^-$, acquired from the mixture heated to 175 °C (175T1).

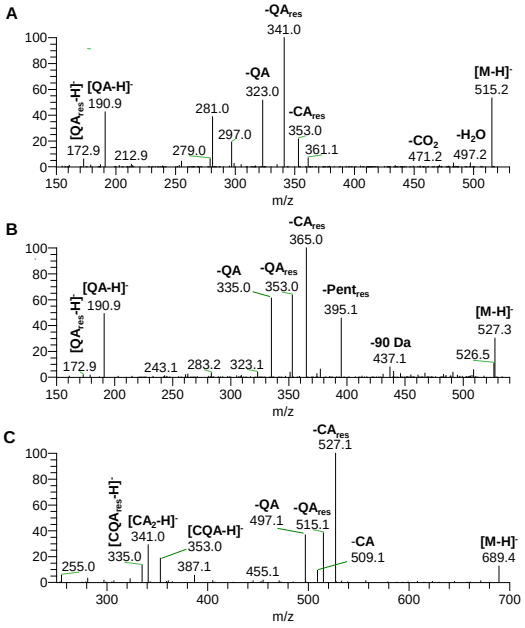

Figure S6. ESI-MS^2 spectra of the ions identified at *m/z* (A) 515 ([(CQA)CA-H]$^-$), (B) 527 ([(CQA)QA-H]$^-$ and [$Pent_4$-H_2O-H]$^-$), and (C) 689 ([$(CQA)_2$-H]$^-$), acquired from thermally treated mixtures (175T1 and 175T2).

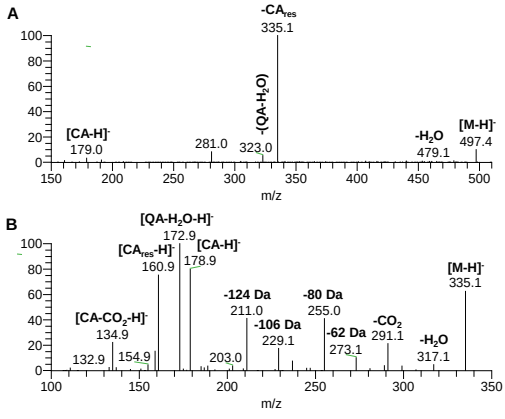

Figure S7. (A) ESI-MS^2 spectrum of the ion observed at *m/z* 497 ([(CQA)CA-H_2O-H]$^-$), and (B) the respective ESI-MS^3 spectrum of the product ion formed by loss of CA_{res} (*m/z* 335), acquired from the mixture heated at 175 °C for 30 min.

APPENDIX E. SUPPLEMENTARY MATERIAL OF SECTION III.3.2

SUPPLEMENTARY TABLES

Table S1. Summary of $[M+Na]^+$ and $[M+H]^+$ ions identified by LC-MS analysis of the roasted mixture Man_3-LY, with the indication of the *m/z* value and the proposed assignment.

n	1	2	3	4	5	6
$[Hex_n+Na]^+$	203	365	527	689	851	
$[Hex_n-H_2O+Na]^+$		347	509	671	833	
$[Hex_n-2H_2O+Na]^+$		329	491			
$[Hex_n-3H_2O+Na]^+$		311	473			
$[(LY)_n+H]^+$	295					
$[(LY)_n-H_2O+H]^+$	277					
$[(LY)_n-NH_3+O+H]^+$	294					
$[Hex_nLY+H]^+$	457	619	781	943	1105	1267
$[Hex_nLY-H_2O+H]^+$	439	601	763	925	1087	1249
$[Hex_nLY-2H_2O+H]^+$	421	583	745	907	1069	
$[Hex_nLY-3H_2O+H]^+$	403	565	727	889		
$[Hex_nLY-4H_2O+H]^+$	385	547	709	871		
$[Hex_n(LY)_2+H]^+$	733	895	1057			

Table S2. Accurate masses found by LTQ-Orbitrap for the ions identified after roasting of the mixture Man_3-CQA. The theoretical mass and the difference between the theoretical and experimental masses for each predicted formula were obtained from Xcalibur software.

Experimental mass (*m/z*)	Theoretical mass (*m/z*)	Mass error (ppm)	RDB equiv.	Composition	Proposed assignment
323.0757	323.0761	-1.28	8.5	$C_{15}H_{15}O_8$	$[HexCA-H_2O-H]^-$
323.097	323.0973	-0.94	3.5	$C_{12}H_{19}O_{10}$	$[Hex_2-H_2O-H]^-$
341.1074	341.1078	-1.17	2.5	$C_{12}H_{21}O_{11}$	$[Hex_2-H]^-$
353.0862	353.0867	-1.41	8.5	$C_{16}H_{17}O_9$	$[CQA-H]^-$
353.1072	353.1078	-1.75	3.5	$C_{13}H_{21}O_{11}$	$[HexQA-H]^-$
461.1071	461.1078	-1.51	12.5	$C_{22}H_{21}O_{11}$	$[HexCQA-3H_2O-H]^-$
497.1071	497.1078	-1.46	15.5	$C_{25}H_{21}O_{11}$	$[(CQA)CA-H_2O-H]^-$
497.1281	497.129	-1.76	10.5	$C_{22}H_{25}O_{13}$	$[HexCQA-H_2O-H]^-$
497.1492	497.1501	-1.86	5.5	$C_{19}H_{29}O_{15}$	$[Hex_2QA-H_2O-H]^-$
503.1599	503.1607	-1.61	3.5	$C_{18}H_{31}O_{16}$	$[Hex_3-H]^-$

Table S2 (Continued). Accurate masses found by LTQ-Orbitrap for the ions identified after roasting of the mixture Man_3-CQA. The theoretical mass and the difference between the theoretical and experimental masses for each predicted formula were obtained from Xcalibur software.

Experimental mass (*m/z*)	Theoretical mass (*m/z*)	Mass error (ppm)	RDB equiv.	Composition	Proposed assignment
515.1182	515.1184	-0.37	14.5	$C_{25}H_{23}O_{12}$	$[(CQA)CA-H]^-$
515.1385	515.1395	-1.96	9.5	$C_{22}H_{27}O_{14}$	$[HexCQA-H]^-$
515.1596	515.1607	-2.02	4.5	$C_{19}H_{31}O_{16}$	$[Hex_2QA-H]^-$
623.1594	623.1607	-2.02	13.5	$C_{28}H_{31}O_{16}$	$[Hex_2CQA-3H_2O-H]^-$
659.1805	659.1818	-1.97	11.5	$C_{28}H_{35}O_{18}$	$[Hex_2CQA-H_2O-H]^-$
659.2029	659.2029	-0.09	6.5	$C_{25}H_{39}O_{20}$	$[Hex_3QA-H_2O-H]^-$
677.1908	677.1924	-2.3	10.5	$C_{28}H_{37}O_{19}$	$[Hex_2CQA-H]^-$
677.2121	677.2135	-1.97	5.5	$C_{25}H_{41}O_{21}$	$[Hex_3QA-H]^-$
689.1703	689.1712	-1.4	16.5	$C_{32}H_{33}O_{17}$	$[(CQA)_2-H]^-$
839.2433	839.2452	-2.29	11.5	$C_{34}H_{47}O_{24}$	$[Hex_3CQA-H]^-$
851.222	851.224	-2.41	17.5	$C_{38}H_{43}O_{22}$	$[Hex(CQA)_2-H]^-$
1001.2961	1001.298	-1.9	12.5	$C_{40}H_{57}O_{29}$	$[Hex_4CQA-H]^-$
1163.3479	1163.3508	-2.51	13.5	$C_{46}H_{67}O_{34}$	$[Hex_5CQA-H]^-$
1325.3998	1325.4036	-2.92	14.5	$C_{52}H_{77}O_{39}$	$[Hex_6CQA-H]^-$

Table S3. Accurate masses found by LTQ-Orbitrap for the ions identified after roasting of the mixture Man_3-YL. The theoretical mass and the difference between the theoretical and experimental masses for each predicted formula were obtained from Xcalibur software.

Experimental mass (*m/z*)	Theoretical mass (*m/z*)	Mass error (ppm)	RDB equiv.	Composition	Proposed assignment
275.1393	275.139	0.91	7.5	$C_{15}H_{19}O_3N_2$	$[YL-H_2O-H]^-$
292.1184	292.1179	1.58	7.5	$C_{15}H_{18}O_5N$	$[YL-NH_3+O-H]^-$
293.15	293.1496	1.32	6.5	$C_{15}H_{21}O_4N_2$	$[YL-H]^-$
341.1082	341.1078	1.18	2.5	$C_{12}H_{21}O_{11}$	$[Hex_2-H]^-$
383.1606	383.1601	1.15	11.5	$C_{21}H_{23}O_5N_2$	$[HexYL-4H_2O-H]^-$
401.171	401.1707	0.69	10.5	$C_{21}H_{25}O_6N_2$	$[HexYL-3H_2O-H]^-$
419.1815	419.1813	0.6	9.5	$C_{21}H_{27}O_7N_2$	$[HexYL-2H_2O-H]^-$
437.1921	437.1918	0.59	8.5	$C_{21}H_{29}O_8N_2$	$[HexYL-H_2O-H]^-$
455.2027	455.2024	0.67	7.5	$C_{21}H_{31}O_9N_2$	$[HexYL-H]^-$
503.1609	503.1607	0.55	3.5	$C_{18}H_{31}O_{16}$	$[Hex_3-H]^-$
545.2133	545.213	0.53	12.5	$C_{27}H_{33}O_{10}N_2$	$[Hex_2YL-4H_2O-H]^-$
563.2237	563.2235	0.34	11.5	$C_{27}H_{35}O_{11}N_2$	$[Hex_2YL-3H_2O-H]^-$
581.2342	581.2341	0.12	10.5	$C_{27}H_{37}O_{12}N_2$	$[Hex_2YL-2H_2O-H]^-$
599.2448	599.2447	0.24	9.5	$C_{27}H_{39}O_{13}N_2$	$[Hex_2YL-H_2O-H]^-$

Table S3 (Continued). Accurate masses found by LTQ-Orbitrap for the ions identified after roasting of the mixture Man_3-YL. The theoretical mass and the difference between the theoretical and experimental masses for each predicted formula were obtained from Xcalibur software.

Experimental mass (*m/z*)	Theoretical mass (*m/z*)	Mass error (ppm)	RDB equiv.	Composition	Proposed assignment
617.255	617.2552	-0.31	8.5	$C_{27}H_{41}O_{14}N_2$	$[Hex_2YL\text{-}H]^-$
707.266	707.2658	0.25	13.5	$C_{33}H_{43}O_{15}N_2$	$[Hex_3YL\text{-}4H_2O\text{-}H]^-$
725.2764	725.2764	0.11	12.5	$C_{33}H_{45}O_{16}N_2$	$[Hex_3YL\text{-}3H_2O\text{-}H]^-$
731.3499	731.3498	0.16	13.5	$C_{36}H_{51}O_{12}N_4$	$[Hex(YL)_2\text{-}H]^-$
743.287	743.2869	0.13	11.5	$C_{33}H_{47}O_{17}N_2$	$[Hex_3YL\text{-}2H_2O\text{-}H]^-$
761.2977	761.2975	0.25	10.5	$C_{33}H_{49}O_{18}N_2$	$[Hex_3YL\text{-}H_2O\text{-}H]^-$
779.3081	779.3081	0.03	9.5	$C_{33}H_{51}O_{19}N_2$	$[Hex_3YL\text{-}H]^-$
869.3182	869.3186	-0.49	14.5	$C_{39}H_{53}O_{20}N_2$	$[Hex_4YL\text{-}4H_2O\text{-}H]^-$
887.3288	887.3292	-0.41	13.5	$C_{39}H_{55}O_{21}N_2$	$[Hex_4YL\text{-}3H_2O\text{-}H]^-$
905.3395	905.3397	-0.31	12.5	$C_{39}H_{57}O_{22}N_2$	$[Hex_4YL\text{-}2H_2O\text{-}H]^-$
923.3493	923.3503	-1.05	11.5	$C_{39}H_{59}O_{23}N_2$	$[Hex_4YL\text{-}H_2O\text{-}H]^-$
941.3579	941.3609	-3.15	10.5	$C_{39}H_{61}O_{24}N2$	$[Hex_4YL\text{-}H]^-$
1085.4015	1085.4031	-1.49	12.5	$C_{45}H_{69}O_{28}N_2$	$[Hex_5YL\text{-}H_2O\text{-}H]^-$
1247.454	1247.456	-1.54	13.5	$C_{51}H_{79}O_{33}N_2$	$[Hex_6YL\text{-}H_2O\text{-}H]^-$

Table S4. Compounds identified after roasting of the mixture Man_3-YL: the *m/z* values of the ions identified, the proposed assignments, the retention time (RT), and the most abundant product ions observed in the respective LC-MS^2 spectrum, with the indication of the *m/z* values, mass differences relative to the precursor ion, and the identification of the most informative product ions.

m/z	Assignment	RT	LC-MS^2
203	$[Hex+Na]^+$	4.0	[a]No LC-MS^2 spectrum
277	$[YL\text{-}H_2O+H]^+$	10.4-15.6	249 (-28), 136 (-141, a_1), 171 (-106)
294	$[YL\text{-}NH_3+O+H]^+$	17.7-29.9	248 (-46), 276 (-18), 132 (-162, -(Y-NH_3+O)$_{res}$, $[L+H]^+$), 266 (-28), 220 (-74)
295	$[YL+H]^+$	7.2-16.6	136 (-159, a_1), 278 (-17, -NH_3), 249 (-46, -HCO_2H), 119 (-176, a_1-NH_3)
311	$[Hex_2\text{-}3H_2O+Na]^+$	4.3	185 (-126, -(Hex-$3H_2O$)), 149 (-162, -Hex_{res})
329	$[Hex_2\text{-}2H_2O+Na]^+$	3.9	167 (-162, -Hex_{res}), 185 (-144, -(Hex-$2H_2O$)), 203 (-126, -(Hex-$2H_2O$)$_{res}$)
347	$[Hex_2\text{-}H_2O+Na]^+$	4.1	329 (-18), 287 (-60), 185 (-162, -Hex_{res}), 203 (-144, -(Hex-H_2O)$_{res}$)
365	$[Hex_2+Na]^+$	4.2	347 (-18), 305 (-60), 203 (-162, -Hex_{res}), 185 (-180, -Hex)
385	$[HexYL\text{-}4H_2O+H]^+$	35.6	339 (-46), 367 (-18), 357 (-28), 226 (-159, (a_1+(Hex-$4H_2O$)$_{res}$)), 311 (-74), 254 (-131, -L)
385	$[HexYL\text{-}4H_2O+H]^+$	41.1	339 (-46), 367 (-18), 357 (-28), 226 (-159, (a_1+(Hex-$4H_2O$)$_{res}$)), 311 (-74), 254 (-131, -L)
403	$[HexYL\text{-}3H_2O+H]^+$	17.6	385 (-18), 244 (-159, (a_1+(Hex-$3H_2O$)$_{res}$)), 126 (-277), 357 (-46), 278 (-125)
403	$[HexYL\text{-}3H_2O+H]^+$	25.1	272 (-131, -L), 244 (-159, (a_1+(Hex-$3H_2O$)$_{res}$)), 385 (-18), 357 (-46), 279 (-124)
421	$[HexYL\text{-}2H_2O+H]^+$	18.1	290 (-131, -L), 262 (-159, (a_1+(Hex-$2H_2O$)$_{res}$)), 403 (-18), 393 (-28), 375 (-46)
421	$[HexYL\text{-}2H_2O+H]^+$	20.2	290 (-131, -L), 262 (-159, (a_1+(Hex-$2H_2O$)$_{res}$)), 403 (-18), 375 (-46), 391 (-30)
421	$[HexYL\text{-}2H_2O+H]^+$	29.9	262 (-159, (a_1+(Hex-$2H_2O$)$_{res}$)), 403 (-18), 290 (-131, -L), 244 (-177)
439	$[HexYL\text{-}H_2O+H]^+$	18.0	280 (-159, (a_1+(Hex-H_2O)$_{res}$)), 393 (-46), 295 (-144, -(Hex-H_2O)$_{res}$, $[YL+H]^+$)
457	$[HexYL+H]^+$	5.7-16.9	439 (-18), 373 (-84), 421 (-36), 307 (-150), 295 (-162, -Hex_{res}, $[YL+H]^+$), 403 (-54), 136 (-321, a_1), 298 (-159, (a_1+Hex_{res}))
473	$[Hex_3\text{-}3H_2O+Na]^+$	4.2	347 (-126, -(Hex-$3H_2O$)), 311 (-162, -Hex_{res})

Table S4 (Continued). Compounds identified after roasting of the mixture Man_3-YL: the *m/z* values of the ions identified, the proposed assignments, the retention time (RT), and the most abundant product ions observed in the respective LC-MS^2 spectrum, with the indication of the *m/z* values, mass differences relative to the precursor ion, and the identification of the most informative product ions.

m/z	Assignment	RT	LC-MS^2
491	$[Hex_3\text{-}2H_2O+Na]^+$	4.1	329 (-162, -Hex_{res}), 347 (-144, -(Hex-$2H_2O$)), 365 (-126, -(Hex-$2H_2O$)$_{res}$)
509	$[Hex_3\text{-}H_2O+Na]^+$	4.2	347 (-162, -Hex_{res}), 491 (-18), 449 (-60), 365 (-144, -(Hex-H_2O)$_{res}$), 185 (-324, -2xHex_{res})
527	$[Hex_3+Na]^+$	4.2	365 (-162, -Hex_{res}), 347 (-180, -Hex), 509 (-18), 467 (-60), 185 (-342, $[Hex_{res}+Na]^+$)
547	$[Hex_2YL\text{-}4H_2O+H]^+$	18.4-24.9	529 (-18), 388 (-159, (a_1+(Hex_2-$4H_2O$)$_{res}$)), 501 (-46), 385 (-162, -(Hex-H_2O)), 416 (-131, -L), 511 (-36), 421 (-126, -(Hex-$3H_2O$)), 403 (-144, -(Hex-$3H_2O$)$_{res}$)
565	$[Hex_2YL\text{-}3H_2O+H]^+$	16.5	288 (-277), 547 (-18), 403 (-162, -Hex_{res}), 406 (-159, a_1+(Hex_2-$3H_2O$)$_{res}$), 529 (-36)
565	$[Hex_2YL\text{-}3H_2O+H]^+$	19.6-21.2	547 (-18), 406 (-159, a_1+(Hex_2-$3H_2O$)$_{res}$), 403 (-162, -Hex_{res}), 529 (-36), 439 (-126, -(Hex-$3H_2O$))
583	$[Hex_2YL\text{-}2H_2O+H]^+$	24.1	421 (-162, -Hex_{res}), 565 (-18), 262 (-321, a_1+(Hex-$2H_2O$)$_{res}$)), 290 (-293, -(Hex_{res}+L), 403 (-180, -Hex)
601	$[Hex_2YL\text{-}H_2O+H]^+$	16.9	439 (-162, -Hex_{res}), 280 (-321, a_1+(Hex-H_2O)$_{res}$)), 295 (-306, $[YL+H]^+$), 393 (-208, -(Hex_{res}+HCO_2H))
619	$[Hex_2YL+H]^+$	5.8-16.4	457 (-162, -Hex_{res}), 601 (-18), 307 (-312, -(Hex_{res}+150)), 373 (-246, -(Hex_{res}+84))
689	$[Hex_4+Na]^+$	3.9	527 (-162, -Hex_{res}), 365 (-324, -2xHex_{res}), 203 (-486, -3xHex_{res})
709	$[Hex_3YL\text{-}4H_2O+H]^+$	15.7-21.1	547 (-162, -Hex_{res}), 691 (-18), 635 (-74), 529 (-180, -Hex), 583 (-126, -(Hex-$3H_2O$)), 550 (-159, (a_1+(Hex_3-$4H_2O$)$_{res}$))
727	$[Hex_3YL\text{-}3H_2O+H]^+$	16.6-20.6	565 (-162, -Hex_{res}), 709 (-18), 403 (-324, -2xHex_{res}), 547 (-180, -Hex), 295 (-270, $[YL+H]^+$)
733	$[Hex(YL)_2+H]^+$	26.3	439 (-294, -YL), 280 (-453), 602 (-131, -L), 574 (-159), 393 (-340), 715 (-18), 295 (-438, $[YL+H]^+$), 571 (-162, -Hex_{res})
733	$[Hex(YL)_2+H]^+$	31.6	602 (-131, -L), 439 (-294, -YL), 280 (-453), 393 (-340), 715 (-18), 574 (-159), 295 (-438, $[YL+H]^+$), 571 (-162, -Hex_{res})
745	$[Hex_3YL\text{-}2H_2O+H]^+$	22.3	421 (-324, -2xHex_{res}), 583 (-162, -Hex_{res}), 262 (-483, a_1+(Hex-$2H_2O$)$_{res}$)), 403 (-342, -(Hex_{res}+Hex)), 290 (-455, -((2xHex_{res})+L))
745	$[Hex_3YL\text{-}2H_2O+H]^+$	23.2	421 (-324, -2xHex_{res}), 583 (-162, -Hex_{res}), 403 (-342, -(Hex_{res}+Hex)), 262 (-483, a_1+(Hex-$2H_2O$)$_{res}$)), 290 (-455, -((2xHex_{res})+L))
763	$[Hex_3YL\text{-}H_2O+H]^+$	16.7	601 (-162, -Hex_{res}), 439 (-324, -2xHex_{res}), 280 (-483, (a_1+(Hex-H_2O)$_{res}$)), 745 (-18), 295 (-468, $[YL+H]^+$)
763	$[Hex_3YL\text{-}H_2O+H]^+$	18.7	439 (-324, -2xHex_{res}), 601 (-162, -Hex_{res}), 280 (-483, (a_1+(Hex-H_2O)$_{res}$)), 745 (-18), 393 (-370, -((2xHex_{res})+HCO_2H))
781	$[Hex_3YL+H]^+$	5.3-16.2	457 (-324, -2xHex_{res}), 619 (-162, -Hex_{res}), 373 (-408, -((2xHex_{res})+84), 298 (-483, (a_1+Hex_{res}))
851	$[Hex_5+Na]^+$	4.2	689 (-162, -Hex_{res}), 527 (-324, -2xHex_{res}), 671 (-180, -Hex), 365 (-486, -3xHex_{res})
871	$[Hex_4YL\text{-}4H_2O+H]^+$	15.6-21.3	709 (-162, -Hex_{res}), 547 (-324, -2xHex_{res}), 745 (-126, -(Hex-$3H_2O$))
889	$[Hex_4YL\text{-}3H_2O+H]^+$	18.6	727 (-162, -Hex_{res}), 565 (-324, -2xHex_{res}), 871 (-18), 709 (-180, -Hex), 373 (-516)
889	$[Hex_4YL\text{-}3H_2O+H]^+$	19.8	565 (-324, -2xHex_{res}), 727 (-162, -Hex_{res}), 871 (-18), 272 (-455, -((2xHex_{res})+L))
889	$[Hex_4YL\text{-}3H_2O+H]^+$	22.1	565 (-324, -2xHex_{res}), 403 (-486, -3xHex_{res}), 727 (-162, -Hex_{res}), 871 (-18), 709 (-180, -Hex)
895	$[Hex_2(YL)_2+H]^+$	22.8	733 (-162, -Hex_{res}), 439 (-456, -(Hex_{res}+YL)), 602 (-293, -(Hex_{res}+L)), 764 (-131, -L), 280 (-615), 571 (-324, -2xHex_{res})
895	$[Hex_2(YL)_2+H]^+$	27.3	733 (-162, -Hex_{res}), 602 (-293, -(Hex_{res}+L)), 439 (-456, -(Hex_{res}+YL)), 764 (-131, -L), 280 (-615), 571 (-324, -2xHex_{res})
907	$[Hex_4YL\text{-}2H_2O+H]^+$	16.8	745 (-162, -Hex_{res}), 583 (-324, -2xHex_{res})
907	$[Hex_4YL\text{-}2H_2O+H]^+$	27.4	745 (-162, -Hex_{res}), 583 (-324, -2Hex_{res})
925	$[Hex_4YL\text{-}H_2O+H]^+$	16.4	907 (-18), 601 (-324, -2xHex_{res}), 763 (-162, -Hex_{res}), 889 (-36), 583 (-342, -(Hex_{res}+Hex))
943	$[Hex_4YL+H]^+$	4.3-15.3	619 (-324, -2xHex_{res}), 925 (-18), 781 (-162, -Hex_{res}), 457 (-486, -3xHex_{res}), 373 (-570, -((3xHex_{res})+84)), 295 (-648, -4xHex_{res}, $[YL+H]^+$)
1013	$[Hex_6+Na]^+$	4.4	851 (-162, -Hex_{res}), 527 (-486, -3xHex_{res}), 689 (-324, -2xHex_{res}), 953 (-60), 833 (-180, Hex)

Table S4 (Continued). Compounds identified after roasting of the mixture Man_3-YL: the *m/z* values of the ions identified, the proposed assignments, the retention time (RT), and the most abundant product ions observed in the respective LC-MS^2 spectrum, with the indication of the *m/z* values, mass differences relative to the precursor ion, and the identification of the most informative product ions.

m/z	Assignment	RT	LC-MS^2
1057	$[Hex_3(YL)_2+H]^+$	21.7	733 (-324, -2xHex_{res}), 439 (-618, -((2xHex_{res})+YL)), 895 (-162, -Hex_{res}), 602 (-455, -((2xHex_{res})+L))
1057	$[Hex_3(YL)_2+H]^+$	25.3	733 (-324, -2xHex_{res}), 602 (-455, -((2xHex_{res})+L)), 439 (-618, -((2xHex_{res})+YL)), 895 (-162, -Hex_{res}), 926 (-131, -L)
1087	$[Hex_5YL-H_2O+H]^+$	16.2	925 (-162, -Hex_{res}), 1069 (-18), 763 (-324, -2xHex_{res})
1105	$[Hex_5YL+H]^+$	4.2-16.7	943 (-162, -Hex_{res}), 781 (-324, -2xHex_{res})
1249	$[Hex_6YL-H_2O+H]^+$	15.9	[a]No LC-MS^2 spectrum
1267	$[Hex_6YL+H]^+$	5.3-16.7	[a]No LC-MS^2 spectrum

[a]No LC-MS^2 spectrum, but the ion assignment is corroborated by the observation of other ions of the same series, eluting at a similar retention time (RT). Abbreviations: a_1 – peptide fragment (Figure S1A); Hex_{res} – Hexose residue; L – Leucine; L_{res} – Leucine residue; Y – Tyrosine; Y_{res} – Tyrosine residue.

SUPPLEMENTARY FIGURES

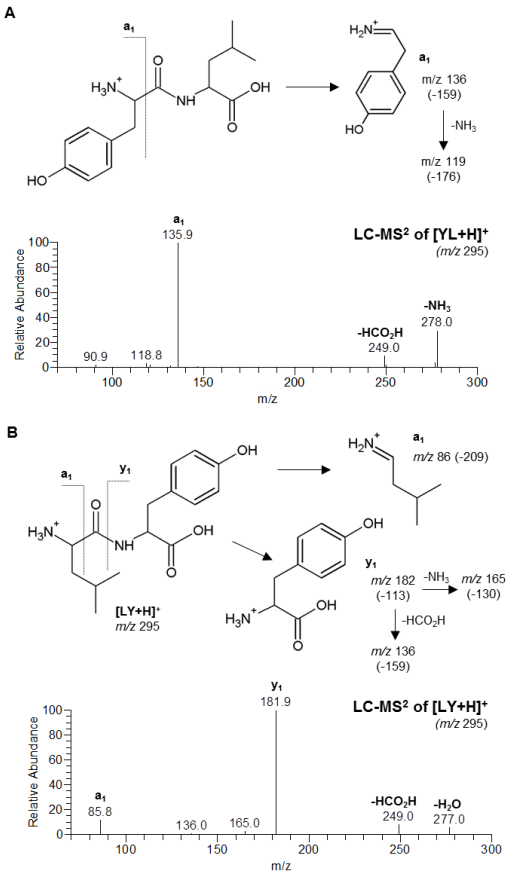

Figure S1. LC-MS^2 spectra of $[M+H]^+$ ions of (A) YL and (B) LY (*m/z* 295) acquired from the unroasted mixtures Man_3-YL and Man_3-LY, respectively. Product ions are assigned as a_1 and y_1 according to the nomenclature originally proposed by Roepstorff & Fohlman (1984), and later modified by Biemann (1988).