Carbohydrate Polymers 80 (2010) 180–185

Contents lists available at ScienceDirect

Carbohydrate Polymers

journal homepage: www.elsevier.com/locate/carbpol

# Structural features of an arabinogalactan-protein isolated from instant coffee powder of *Coffea arabica* beans

Peter Capek [a,*], Mária Matulová [a], Luciano Navarini [b], Furio Suggi-Liverani [b]

[a] Institute of Chemistry, Center for Glycomics, Slovak Academy of Sciences, Dúbravská cesta 9, 845 38 Bratislava, Slovakia
[b] Illycaffè s.p.a., Research & Innovation, via Flavia, 110, I-34147 Trieste, Italy

ARTICLE INFO

*Article history:*
Received 18 September 2009
Received in revised form 4 November 2009
Accepted 9 November 2009
Available online 13 November 2009

*Keywords:*
*Coffea arabica*
Roasted beans
Instant coffee powder
Arabinogalactan-protein
NMR spectroscopy

ABSTRACT

An arabinogalactan-protein (AGP) composed of D-galactose and L-arabinose in 9:1 mol proportion containing about 1.6% of protein has been isolated from the instant coffee powder of roasted *Coffea arabica* beans by water extraction followed by barium hydroxide and ethanol precipitations. AGP was recovered in 3.2% yield of instant coffee powder and its average molecular mass was significantly decreased due to severe manufacturing conditions to 5400. Chemical and spectroscopic studies on the instant coffee AGP revealed a β-(1,3)-linked Gal*p* backbone branched at C6 by side 6-linked Gal*p* residues terminated by Ara*f* and/or Gal*p* and confirming thus an α-arabino-3,6-β-galactan in the instant coffee powder composite. It has been found that about 74% of branched 3,6-linked Gal*p* residues have their origin in the backbone while 26% is located in side chains (6-linked Gal units substituted at C3 by Ara*f* or terminal Gal*p* units). Moreover, about 11% of Gal*p* residues involved in linear 1,6-linkages of AGP side chains survived a severe industrial process.



## 1. Introduction

Coffee beans are important plant products utilizing by human society for preparation of coffee brews for many centuries due to their stimulating effect on organism and delicious flavour typical for this kind of seasoning only. Moreover, coffee became an integral part of our food chain. It has been found that polysaccharide components cellulose, galactomannan and arabinogalactan-proteins (AGPs) represent about 50% of dry weight of coffee beans, from which AGPs constitute about 15%. They are highly glycosylated macromolecules with a large scale of functions in biological processes. Besides, plant AGPs become important due to their prebiotic, cholesterol-lowering, emulsifying, immunomodulatory, and other properties which makes from these glycoconjugates commercially important article for food and pharmaceutical industries (Pettolino, Liao, Zhu, Mau, & Bacic, 2006). On a coffee market, AGPs are very important as the roasted coffee aroma precursors and they play an important role in the determination of some coffee brews functional properties. Due to a highly branched structure, AGPs can be easily extracted from coffee beans. Drastic industrial processing conditions during the instant coffee powder preparation causes AGPs structural modifications including depolymerization of the main as well as side chains. Their extent depends on industrial conditions used, i.e. roasting, extraction and spray drying processes. Up to now little attention has been dedicated to the study of instant coffee powder AGP prepared from roasted *Coffea arabica* beans (Capek, Matulová, Navarini, & Suggi-Liverani, 2009; Wolfrom & Anderson, 1967).

In our previous work (Capek et al., 2009), we have confirmed the reproducibility of two arabinogalactan-protein isolations (AG1 and AG2) recovered from freeze-dried coffee powder of roasted *C. arabica* beans according to Wolfrom and Anderson (1967) procedure. The aim of our further study was to investigate the degree of depolymerization and/or destruction during industrial processing and main structural features of arabinogalactan-protein present in the final commercial coffee product.

## 2. Materials and methods

### 2.1. Coffee material

*Coffea arabica* (100%) blend for espresso coffee was dark roasted and industrially processed, i.e. grinded to particles size of 2.3–2.4 mm with 10% particles less than 1 mm, extracted at 150–180 °C, concentration under vacuum at 50 °C and freeze-dried to give dark brown freeze-dried instant coffee powder.

### 2.2. General methods

Solutions of carbohydrate samples were concentrated under diminished pressure at a bath temperature 45 °C. Polysaccharides

* Corresponding author. Tel.: +421 2 59410209; fax: +421 2 59410222.
*E-mail address:* chemcape@savba.sk (P. Capek).

0144-8617/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carbpol.2009.11.016

were hydrolyzed with 2 M trifluoroacetic acid at 120 °C for 1 h. The quantitative determination of the neutral monosaccharides was carried out in the form of their trifluoroacetates (Shapira, 1969) by gas chromatography on a Hewlett–Packard 5890 Series II chromatograph (Hewlett–Packard, Ramsey, USA) using a PAS-1701 column (0.32 mm × 25) (Hewlett–Packard, Ramsey, USA) at a temperature program of 110–125 °C (2 °C $min^{-1}$) 165 °C (20 °C $min^{-1}$) and a flow rate of hydrogen 20 ml $min^{-1}$. The absolute configuration of monosaccharides was established by the method of Gerwig, Kamerling, & Vilegenthart (1979). Gas chromatography–mass spectrometry of partially methylated alditol acetates (Jansson, Kenne, Liedgren, Lindberg, & Lonngren, 1976) was performed on a FINNIGAN MAT SSQ 710 spectrometer equipped with a SP 2330 column. The uronic acid content was determined with the 3-hydroxybiphenyl reagent (Blumenkrantz & Asboe-Hansen, 1973). The carbohydrate content was determined by the phenol–sulphuric acid assay (Dubois, Gilles, Hamilton, Rebers, & Smith, 1956). Elemental analysis was performed with EA 1108 apparatus (FISONS Instruments, East Grinstead, UK). Protein was calculated from the nitrogen content (%$N$ × 6.25). Optical rotations were measured with an automatic polarimeter Perkin-Elmer Model 241 (Waltham, MA, USA) for 0.5% aqueous solutions at 20 °C.

### 2.3. Isolation of an arabinogalactan-protein

The arabinogalactan-protein (AGP) has been isolated from dark brown freeze-dried instant coffee powder of *C. arabica* blend according already described procedure (Wolfrom & Anderson, 1967) with minor modifications (Capek et al., 2009). Shortly, instant coffee powder was dissolved in distilled water, cooled to 4 °C and treated with formic acid overnight in a refrigerator. The insoluble portion was discarded after centrifugation and the supernatant was treated with five volumes of methanol. The brown precipitated material was removed by centrifugation, dissolved in distilled water and freeze-dried to yield a crude polysaccharide fraction A. The crude material A was dissolved in 0.05 M sodium hydroxide solution and precipitated by saturated aqueous barium hydroxide solution. The dark brown precipitate formed (a crude galactomannan) was removed by centrifugation and the supernatant was treated dropwise with 1.0 N sulphuric acid until a pH of 4 was obtained. The slight precipitate formed was removed by centrifugation (discarded) and the supernatant was poured into 96% ethanol (five volumes). The precipitate was centrifuged, dissolved in distilled water and applied on the column of Amberlite MB 150 ($H^+/OH^-$), and eluted with distilled water. The water eluent containing carbohydrates was concentrated and freeze-dried to give the arabinogalactan-protein (AGP).

### 2.4. Methylation analysis

The dry samples of oligo- and polysaccharides (≈1 mg) were solubilized in dry $Me_2SO$ (1 mL) and methylated by the Hakomori method (Hakomori, 1964). Permethylated carbohydrates were purified using the Sep-Pak $C_{18}$ cartridges (Waters Assoc.), hydrolyzed with 2 M trifluoroacetic acid at 120 °C for 1 h, reduced with $NaBD_4$, acetylated and analyzed by GLCMS (Jansson et al., 1976).

### 2.5. HPLC chromatography

Molecular size determination of polysaccharide samples were performed with Shimadzu apparatus (Vienna, Austria) using a tandem of two columns HEMA-BIO 100 followed HEMA-BIO 40 column (Tessek, Prague, Czech Republic) of dimensions 8 mm × 250 mm. As a mobile phase 0.02 M phosphate buffer pH 7.2 containing 0.1 M NaCl was used at a flow rate 0.8 mL/min. A set of pullulan standards was used for calibration of the column (Gearing Scientific, Polymer Lab., Hertfordshire, UK).

### 2.6. FT-IR and NMR spectroscopies

Fourier transform infrared spectra of carbohydrates were recorded with a Nicolet Magna 750 spectrometer (Nicolet, Madison, USA) with DTGS detector (Nicolet, Madison, USA) and OMNIC 3.2 software (Nicolet, Madison, USA). The carbohydrates were measured in the form of KBr pellets with a sample/KBr ratio 1/200 mg.

Spectra were measured in $D_2O$ at 25 °C on Bruker 300 MHz Avance DPX and Varian 600 MHz UNITY INOVA 600 NB spectrometers, equipped with 5 mm multinuclear probe with inverse detection and 5 mm $^1H\{^{13}C,^{15}N\}$PFG Triple Res IDTG600-5, respectively (both with *z*-gradients). Carbon spectra (100.5 MHz) were measured in 5 mm broadband DBG probe on Varian spectrometer. The acetone was used as the internal standard ($\delta$ 2.225 for $^1H$ and 31.07 for $^{13}C$). Samples were freeze-dried from 95% $D_2O$ and after they were dissolved in 99.98% $D_2O$. For both, $^1H$ and $^{13}C$ NMR spectra, chemical shifts are referenced to internal acetone ($\delta$ 2.217 and 31.07, respectively). $^1H$–$^{13}C$ hetero-correlated HSQC spectra with optimisation on one bond coupling constant $^1J_{CH}$ = 143 Hz and $^1H$–$^1H$ homocorrelated COSY spectra with gradient selection in absolute intensity mode were measured.

## 3. Results and discussion

### 3.1. Isolation and characterization of instant coffee powder AGP

For the instant coffee powder isolation from a blend of *C. arabica* common industrial processes as roasting, grounding, extraction, concentration and freeze or spray drying are used. A crude polysaccharide complex A, isolated according the Wolfrom procedure (Wolfrom & Anderson, 1967) from the instant coffee powder was obtained in 22% yield (w/w) of starting coffee material. In comparison with results (∼19%) for isolation of polysaccharide complex A described by Wolfrom this value was slightly higher. However, in further steps of our isolation the yields of a crude arabinogalactan (5.2% (w/w) of starting coffee powder) and that of the final product (as a colourless fluffy material) purified by ion-exchange chromatography (3.2% of starting coffee powder) were significantly lower (∼45%) in comparison with that of arabinogalactan (AGW) isolated according old procedure (Wolfrom & Anderson, 1967). This discrepancy could be caused by many reasons such as due to different coffee bean species used, geographic region of cultivated coffee plants or different conditions of industrial processes as roasting, grinding, extraction and drying used for preparation of instant coffee powder (Fischer, Reimann, Trovato, & Redgwell, 2001; Nunes & Coimbra (2002); Oosterveld, Harmsen, Voragen, & Schols, 2003; Wragg, Mendl, Adeoye, Hague, & Jauch, 2004).

The nitrogen content ∼0.3% (equivalent to about 1.6% of protein) and signals observed in the region $\delta$ 3.1–0.8 of the $^1H$ NMR spectrum of AG (not shown) suggested that the polymer is present in instant coffee powder in the form of arabinogalactan-protein complex (AGP). Its average molecular mass, determined by HPLC, was 5400 and optical rotation +14°. Although its molecular mass was higher in comparison with AGW (∼1800) isolated from not specified plant material, their optical rotation values were comparable (+14° versus +9.5°). It is evident that drastic industrial processing conditions during the instant coffee powder preparation significantly depolymerize the molecular mass (∼500–1500 × $10^3$) of the native AGP (Redgwell, Curti, Fischer, Nicolas, & Fay, 2002).

Compositional analysis of AGP revealed the presence of galactose (85%) and arabinose residues (8.2%), and smaller amounts of

accompanying sugars as mannose (2.7%), heptose (2.7%), glucose (1.4%), and xylose (traces) residues as contaminants. No rhamnose and uronic acid residues, which are common constituents of coffee AGPs, were determined. The characteristic galactose/arabinose ratio in AGP was found to be 9:1, while in AGW its ratio was about 16:1 (Wolfrom & Anderson, 1967). Higher content of arabinose in AGP in comparison with AGW could be explained by different processing conditions of coffee beans during the instant coffee powder preparation. In particular, under stronger processing conditions the cleavage of labile arabinofuranosyl residues is a lot faster. Both facts (lower molecular weight and different galactose/arabinose ratio) indicate that for the AGW isolation an instant coffee powder prepared under much more severe processing conditions was used. In comparison with the native green coffee AGPs, where the galactose/arabinose ratio varies from 0.9 to 3.1, it is evident that labile linkages of arabinose in side chains are preferentially pyrolyzed (Redgwell et al., 2002). However, during drastic industrial processes the 1,3-linked backbone of AGP has been significantly depolymerized as well, as it can be seen from its molecular size (5400).

### 3.2. Sugar linkage analysis of instant coffee powder AGP

Methylation analysis data of AGP (Table 1) revealed the presence of linkages characteristic for type II arabino-3,6-galactan. Galactose was found in four main methylated products arising from terminal, 3-, 6- and 3,6-linked galactopyranose residues. They constitute over 92% of all sugar derivatives of which 3- and 3,6-linked galactopyranose residues represent about 60%. As it can be seen from the Table 1, the ratio 3- and 3,6-linked galactopyranose residues (1.6:1) indicating thus relatively high branching of the polymer. The backbone of the native green coffee AGP is supposed to be branched at each third or fourth units by side chains composed of highly branched 6-linked Gal units mainly at C3 contributing thus also to the quantity of 3,6-linked Gal units (Redgwell et al., 2002). However, the splitting of all linkages is random in the native green coffee AGP during the industrial processing resulting in fragments with lower molecular weight. Further, the data of the methylation analysis show that over 60% of the backbone Gal units (internal 3-linked Gal) are non-substituted. The amount of terminal Gal unit (≈22%) is relatively higher and is very close to that one of 3,6-linked Gal residues which can have their origin in the backbone (3-linked Gal substituted at O6 by Gal of the side chains) or in side chains (6-linked Gal units substituted at O3 by Ara*f* or terminal Gal). Moreover, some of terminal Gal residues occupy non-reducing terminal ends of the backbone and their ratio to 3,6-linked Gal residues depends on length of polysaccharide chains due to high polydispersity of AGP. The presence of the terminal Gal residue linked to O3 of the neighbouring Gal unit was confirmed by NMR. In the structure of coffee AGP proposed by Redgwell et al. (2002) Ara*f* is located in side chains. Thus accordingly over 6% of terminal Ara*f* residues found by methylation analysis should be located in side chains at C3 of 6-linked Gal units indicating that 6% of 3,6-linked Gal have their origin in side chains. Consequently, this fact indicates that about 17% of total 23% 3,6-linked Gal residues should come from the backbone. These conclusions are in accordance with NMR data (Table 2), which correlate very roughly with the results of sugar linkage analysis. According the abovementioned data and because of the fact that the length of side chains was not proven, we could hypotheses that the backbone could be substituted by single terminal Gal or short chains of two–three Gal units carrying low content of Ara*f* residues.

**Table 1**
Sugar linkage analysis of instant coffee powder arabinogalactan-protein (AGP).

| Sugar derivative | AGP (mol%) | Mode of linkage |
|---|---|---|
| 2,3,5-$Me_3$-Ara[a] | 6.4 | Ara*f*-(1→ |
| 2,3-$Me_2$-Ara | 1.3 | →5)-Ara*f*-(1→ or →4)-Ara*p*-(1→ |
| 2,3,4,6-$Me_4$-Gal | 22.0 | Gal*p*-(1→ |
| 2,4,6-$Me_3$-Gal | 36.4 | →3)-Gal*p*-(1→ |
| 2,3,4-$Me_3$-Ga | 11.2 | →6)-Gal*p*-(1→ |
| 2,4-$Me_2$-Gal | 22.7 | →3,6)-Gal*p*-(1→ |
| 2,3,6-$Me_3$-Gal/Glc | tr | 4)-Galp/Glc*p*-(1→ |

[a] 2,3,5-$Me_3$-Ara = 1,4-di-*O*-acetyl-2,3,5-tri-*O*-methyl-arabinitol, etc., tr = traces.

### 3.3. NMR spectroscopy of AGP

In the $^1$H NMR spectrum of AGP, besides signals due to carbohydrates, also those due to a protein part of the molecule could be observed in the region $\delta$ 3.1–0.8. Signals due to aromatic rings were of very low intensity, however, no signals characteristic of chlorogenic acid were between them. Low molecular weight of AGP suggested by HPLC (≈5400) was confirmed also by NMR because H1/C1 anomeric signals due to reducing α and βGal could be detected by $^{13}$CNMR as well as by 2D $^1$H–$^{13}$C HSQC NMR spectra (Fig. 1). In general, literature data concerning the assignment of arabinogalactan signals are not coherent and show some discrepancies. Assignment of our data was mainly based on the 2D NMR and selective 1D experiments. Signals due to individual spin systems were attributed on the basis of 2D $^1$H–$^1$H homo-nuclear correlated COSY and TOCSY experiments as well as 1D TOCSY spectra with selective excitation for chosen proton signals. Complete set of data, obtained after $^1$H data correlation with $^{13}$C in HSQC and HMBC spectra (Figs. 1 and 2) is presented in the Table 2. Sets of H1/H2/C1 chemical shifts could be used as a finger print for an identification of individual types of sugars in a given linkage type.

**Table 2**
$^1$H and $^{13}$C NMR data of arabinogalactan-protein isolated from *Arabica* instant coffee measured in $D_2O$ at 25 °C and acetone as an internal standard.

| Sample | Unit | Chemical shift $\delta$ (ppm) | | | | | | Type of linkage |
|---|---|---|---|---|---|---|---|---|
| | | H1/C1 | H2/C2 | H3/C3 | H4/C4 | H5/C5 | H6,H6′/C6 | |
| AGP | Araf(1→3) | 5.24/110.2 | 4.21/82.1 | 3.94/77.3 | 4.12/84.7 | 3.82, 3.71/62.0 | | Term 1→3 |
| | Araf(1→5) | 5.09/108.7 | 4.12/82.1 | 3.94/77.3 | 4.07/84.7 | 3.90, 3.77/68.3 | | Term/int 1→5 |
| | Galβ(1→3) | 4.62/105.0 | 3.61/71.7 | 3.66/73.6 | 3.92/69.4 | 3.70/75.9 | 3.82, 3.70/61.8 | Term 1→3 |
| | →3)Galβ(1→ | 4.69/104.7 | 3.77/75.6 | 3.85/82.9 | 4.24/69.3 | 3.70/76.0 | 3.77/61.8 | Int 1→3 |
| Backbone | →3,6)Galβ(1→ | 4.52–4.48/103.9 | 3.68/75.8 | 3.82/82.9 | 4.20, 4.10/69.3 | 3.91/74.4 | 4.04, 3.91/70.3 | Int 1→3,6 |
| | →6)Galβ(1→ | 4.45/104.6 | 3.54/ 71.6 | 3.64/73.6 | 3.92/69.4 | 3.91/74.4 | 4.04, 3.91/70.1 | Int 1→6 |
| | Galβ(1→6) | 4.43/104.4 | 3.51/71.6 | 3.64/73.6 | 3.92/69.4 | 3.69/76.0 | 3.77/61.8 | Term 1→6 |
| | →6)Galβ(1→[a] | 4.58/103.9 | 3.38/73.6 | 3.55/76.0 | 3.70/75.6 | 3.99/75.3[b] | 70.4 | Unknown |
| | →3)Galβ | 4.63/97.1 | 3.66/73.4 | 3.81/83.3 | 4.20/69.4 | 3.73/75.6 | 3.76/61.8 | Red 1,3 |
| | →3)Galα | 5.28/93.1 | 3.98/68.3 | 3.99/80.3 | 4.26/70.2 | 4.12/71.3 | 3.72/62.0 | Red 1,3 |

[a] Aglycon part of the molecule not identified yet.
[b] Tentative assignment from HSQC spectrum.

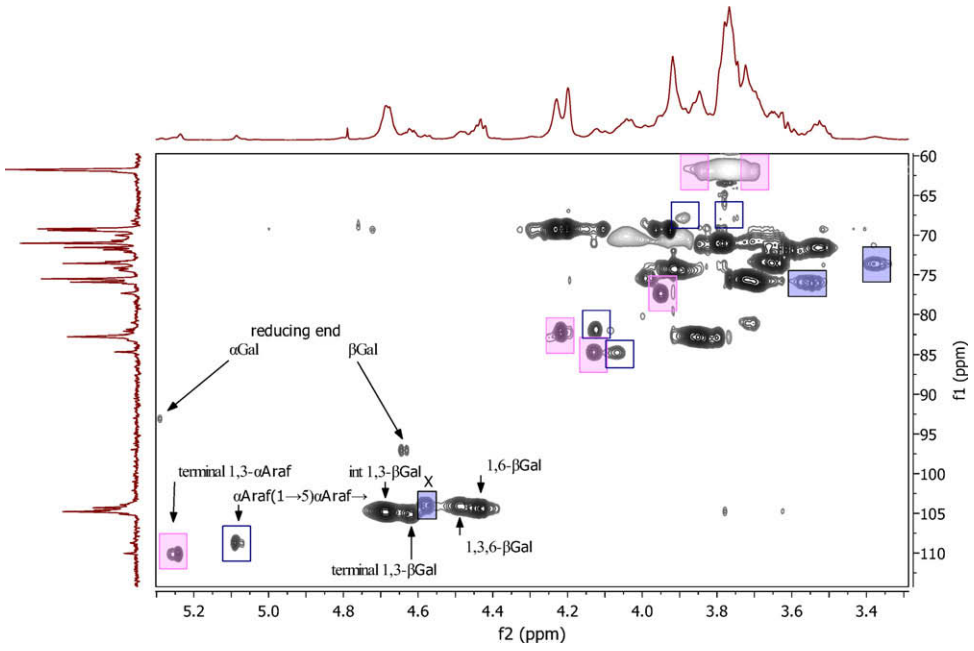


**Fig. 1.** $^{1}H–^{13}C$ hetero-correlated HSQC spectrum of AGP isolated from instant *Arabica* coffee measured in $D_2O$ at 25 °C with acetone as an internal standard ($\delta$ 2.225, 31.07). Terminal 1,3: αAra*f*: terminal arabinofuranose linked to O3 of the neighbouring Gal unit; αAraf(1→5)αAraf: terminal/internal arabinofuranose linked to O5 of the neighbouring Ara*f* unit; int 1,3 Gal: internal 1,3-linked Gal unit; terminal 1,3-βGal: terminal Gal unit linked to O3 of the neighbouring βGal unit; 1,3,6 βGal: 1,3,6-linked galactose unit; 1,6 βGal: 1,6-linked galactose units (internal and terminal); X: signal due to 1,6-linked galactose of unknown origin.

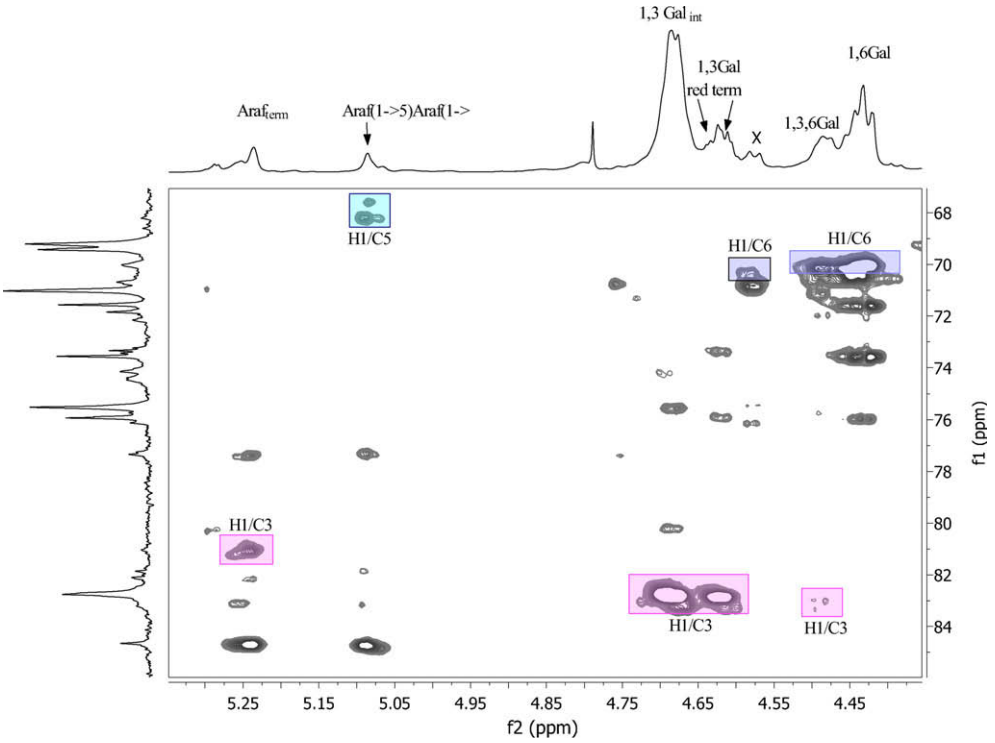


**Fig. 2.** Part of the $^{1}H–^{13}C$ long range HMBC spectrum of AGP. Ara$f_{term}$: terminal arabinofuranose linked to O3 of the neighbouring Gal unit; Araf(1→5)Araf: terminal/internal arabinofuranose linked to O5 of the neighbouring Araf unit; 1,3 $Gal_{int}$: internal 1,3-linked Gal unit; 1,3 Gal red term: 3-linked Gal at the reducing end and terminal Gal unit linked to O3 of the neighbouring Gal unit (terminal 1,3-linked Gal); 1,3,6 Gal: 1,3,6-linked galactose unit; 1,6 Gal: 1,6-linked galactose units (internal and terminal); X: signal due to 1,6-linked galactose of unknown origin. H1/C3: cross peak between H1 and C3; H1/C5: cross peak between H1 and C5; H1/C6: cross peak between H1 and C6.

Signals due to H1 at $\delta$ 4.43 and 4.45 showed in long range $^{1}H–^{13}C$ hetero-correlated HMBC spectrum (Fig. 2) H1/C6 cross peaks at $\delta$ 4.43/70.1 and 4.45/70.1 indicating its long range interaction with C6 of βGal involved in the 1,6-linkage. A comparison of HSQC spectra of AGP isolated from instant coffee with another ones isolated from green coffee highly substituted by Ara*f* (unpublished results) suggested that the signal at $\delta$ 4.43 is due to terminal βGal residues linked to O6 of the neighbouring Gal one, while that one at $\delta$ 4.45 due to internal 1,6-linked βGal units of side chains showing a long range interaction to C6 at ~70.2 in HMBC spectrum

(Table 2). Group of H1 signals at $\delta$ 4.52–4.45 showed also in HMBC a cross peak to C6 ($\delta \sim 70.2$) suggesting also their origin in 1,6-linked βGal of side chains. However, for some of these signals, particularly for the signal at $\delta$ 4.49 and 4.52, a long range interaction to C3 at $\delta$ 83.2 could be observed, giving thus an evidence that these signals belong to 3,6-linked Gal. In the case of H1 signals at $\delta$ 4.69 and 4.62–4.63 no long range interaction to C6 was detected. Instead, the H1/C3 cross peaks to C3 at $\delta$ 82.9 was present in both cases indicating their origin in 3-linked Gal of the backbone. This attribution is in accordance with coffee arabinogalactan literature data (Nunes, Reis, Silva, Domingues, & Coimbra, 2008).

The origin of signals (H1, H2, C1) at $\delta$ 4.58/3.39/103.9 (Table 2), denoted as X in the spectra, remained unclear. $^1$H NMR data show that it is galactose unit, but its H3/C3 chemical shifts did not indicate the presence of 1,3-glycosidic linkage. Results of the methylation analysis showed the presence of 1,4-linked hexose (Gal or Glc). Downfield shift was occurred for H4/C4 atoms of this galactose residue indicating the substitution at this position. Small $^3J_{H4,H5}$ coupling constants in galactose spin system did not allow the identification of H5, H6, H6′ by TOCSY and thus their exact values were not determined. H2/C2 and H3/C3 chemical shifts of this Gal unit show an unusual chemical shifts in the HSQC spectrum (Fig. 1). In the HMBC spectrum (Fig. 2) a cross peak at $\delta$ 4.58/70.4 due to H1/C6 indicates that this galactose unit is 6-linked as the signal at $\delta$ 70.4 was identified by DEPT as $CH_2$ signal. Further cross peaks at $\delta$ 4.58/70.7 observed in HMBC was very intense and it was identified as CH signal. These facts indicate that this Gal unit could be linked to some peptide by O1 and an unusual upfield shift of H2 supports this hypothesis. In the case on hydroxyproline (Hyp)-arabinogalactan from transgenic *Nicotiana tabactum* (Tan, Qui, Lamport, & Kieliszewski, 2004) very same H1 downfield shift of 1,3-linked Gal at the reducing end of the backbone due to the linkage to O4 of hydroxyproline (Hyp) was observed. In spite of the similarity of H1 chemical shifts all other NMR data (H2–H6/C2–C6) in our case were very different. Moreover, any of reported chemical shifts of Hyp was similar to those observed for peptidic part of the molecule in our spectra (Bolling et al., 2007; Tan et al., 2004). Further studies will be necessary to clarify the origin of this Gal unit in AGP.

Characteristic sets of H1/H2/C1 chemical shifts values for β-galactose residues in the AGP, which could be use a finger print, were following: Gal in the backbone: $\delta$ 4.69/3.77/104.7 for internal 1,3-linked βGal, $\delta$ 4.62/3.61/105.0 for terminal βGal linked to O3 of the neighbouring Gal unit, $\delta$ 4.63/3.66/97.1, for the reducing end 3-linked βGal; in side chains: H1/H2/C1 signals at $\delta$ 4.45/3.54/104.6 for internal 1,6-linked βGal, $\delta$ 4.43/3.51/104.4 for terminal βGal residue.

Obtained $^1$H NMR data for signals in the region $\delta$ 5.5–4.8 suggested the presence of Ara*f* involved in more types of linkages then estimated on the basis of methylation analysis. Here, besides H1/H2/C1 signal at $\delta$ 5.29/3.99/93.1 due to the reducing 3-linked αGal, two intense signals due to arabinofuranose (Ara*f*) at $\delta$ 5.24/4.21 and 5.09/4.12 were dominant. The H1/H2/C1 signals at $\delta$ 5.24/4.21/110.2 ($^1J_{C1,H1}$ 174.3 Hz) were attributed to the terminal Ara*f* unit linked to O3 of βGal (e.g. terminal 1,3-linked Gal) on the basis of the H1/C3 cross peak at $\delta$ 5.24/81.2 in HMBC indicating the substitution of some Gal unit at C3. This observation is in full agreement with literature data (Tan et al., 2004). On the basis the H1/C5 cross peak at $\delta$ 5.09/68.3 observed in HMBC spectrum the H1/H2/C1 signals at $\delta$ 5.09/4.12/108.7 ($^1J_{C1,H1}$ 175.6 Hz) were assigned to Ara*f* residue linked to C5 of neighbouring Ara*f*. Chemical shifts for the whole spin system of this Ara*f* unit are in full agreement with literature data (Polle, Ovodova, Chizhov, Shashkov, & Ovodov, 2002) and that of the methylation analysis. Chemical shifts of C1 anomeric carbon indicated α-configuration for both arabinose units (Joseleau, Chambat, Vignon, & Barnoud, 1977). In addition, COSY spectrum showed the presence of further cross peaks due to Ara*f* at $\delta$ 5.30/4.29, 5.27/4.22, 5.07/4.09 ($^1J_{C1,H1}$ 174.3 Hz) and 5.03/4.15 which suggest that Ara*f* could be involved into another types of linkages which were not determined by methylation analysis due to their low content. Complete assignment of signals observed in AGP spectra is presented in the Table 2.

Preliminary structural studies showed relatively complex structure of *Arabica* AGP isolated from commercial instant coffee powder. As it was concluded above, about 17% of 3,6-linked Gal unit could have an origin in the galactopyranose backbone bearing side chains. Thus from the total amount (23%) of 3,6-linked Gal units over 6% should have their origin in side chains. This conclusion was based on the amount of Ara*f*, which is supposed to occupy O3 position of 6-linked Gal residues (Redgwell & Fischer, 2006). However, 22% of terminal Gal units found indicate that in addition to those having their origin in side chains (≈17%) still about 5% of their location remain unclear. In the NMR spectra both terminal Gal residues linked to O3 or O6 of neighbouring Gal units have been identified at $\delta$ 4.62 and 4.43, respectively. The presence of signal at $\delta$ 4.62 in the spectra, together with those of reducing end α,β Gal, confirm low molecular weight of the instant coffee AGP.

## 4. Conclusions

It can be concluded that the arabinogalactan-protein composed of D-galactose and L-arabinose (9:1) and protein (∼1.6%) from instant coffee powder of *C. arabica* in about 3.2% yield has been isolated. Its high molecular mass (∼several hundred thousand in the native form) was significantly depolymerized to the value 5400 due to severe industrial process during which highly branched side chains of AGP were rapidly pyrolyzed to mostly linear short ones. Linkage sugar analysis as well as NMR studies of AGP revealed 3-linked galactopyranosyl backbone branched at C6 by side 6-linked galactosyl side chains terminated by non-substituted galactosyl residues. Some of side chain internal Gal units remained substituted by arabinosyl residues indicating that ≈6% branched 3,6-linked galactopyranosyl residues have their origin in side chain, while 17% of them are in the backbone.

## Acknowledgements

This project has been supported by the Illycaffe grant, Trieste, Italy, the Slovak Grant Agency VEGA No. 2/0155/08, APVV Grant No. 0030-08 and Slovak State Program 2003SP200280203.

## References

Blumenkrantz, N., & Asboe-Hansen, O. (1973). A new method for quantitative determination of uronic acid. *Analytical Biochemistry, 54*, 484–489.

Bolling, K., Lamshöft, M., Schweimer, K., Marner, F.-J., Budzikiewicz, H., & Waffenschmidt, S. (2007). Structural analysis of linear hydroxyproline-bound O-glycans of *Chlamydomonas reinhardtii*-conservation of the inner core in *Chlamydomonas* and land plants. *Carbohydrate Research, 342*, 2557–2566.

Capek, P., Matulová, M., Navarini, L., & Suggi-Liverani, F. (2009). A comparative study of arabinogalactan-protein isolates from instant coffee powder of *Coffea arabica* beans. *Journal Food and Nutrition Research, 48*(2), 80–86.

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*(3), 350–356.

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

Gerwig, G. J., Kamerling, J. P., & Vliegenthart, J. F. G. (1979). Determination of the absolute configuration of monosaccharides in complex carbohydrates by capilary g.l.c. *Carbohydrate Research, 77*, 1–8.

Hakomori, S. J. (1964). A rapid permethylation of glycolipid and polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide. *Biochemistry (Tokyo), 55*, 205–208.

Jansson, P. E., Kenne, L., Liedgren, H., Lindberg, B., & Lonngren, J. (1976). A practical guide to the methylation analysis of carbohydrates. *Chemical Communications University of Stockholm, 8*, 1–75.

Joseleau, J.-P., Chambat, G., Vignon, M., & Barnoud, F. (1977). Chemical and $^{13}C$ NMR studies on two arabinans from the inner bark of young stems of *Rosa glauca*. *Carbohydrate Research, 58*, 165–175.

Nunes, F. M., & Coimbra, M. A. (2002). 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., 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., Harmsen, J. S., Voragen, A. G. J., & Schols, H. A. (2003). Extraction and characterization of polysaccharides from green roasted *Coffea arabica* beans. *Carbohydrate Polymer, 52*, 285–296.

Pettolino, F., Liao, M., Zhu, Y., Mau, S., & Bacic, A. (2006). Structure, function and cloning of arabinogalactan-proteins (AGPs): An overview. *Food and Food Ingredients Journal of Japan, 211*, 12–25.

Polle, A. Ya., Ovodova, R. G., Chizhov, A. O., Shashkov, A. S., & Ovodov, Yu. S. (2002). Structure of tanacetan, a pectic polysaccharide from Tansy *Tanacetum vulgare* L.. *Biochemistry (Moscow), 67*, 1371–1376.

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

Redgwell, R. J., & Fischer, M. (2006). Coffee carbohydrates. *Brazilian Journal of Plant Physiology, 18*(1), 165–174.

Shapira, J. (1969). Identification of sugars as their trifluoroacetylpolyol derivatives. *Nature, 222*, 792–793.

Tan, L., Qui, F., Lamport, D. T. A., & Kieliszewski, M. J. (2004). Structure of a hydroxyproline (Hyp)-arabinogalactan polysaccharide from repetitive Ala-Hyp expressed in transgenic *Nicotiana tabacum*. *Journal of Biological Chemistry, 279*, 13156–13165.

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

Wragg, A., Mendl, C., Adeoye, I., Hague, R., & Jauch, K. (2004). Analytical method and model for measurements of mannan oligomer size distribution in high yield soluble coffees. In *Proceedings of 20th international colloquium on the chemistry of coffee* (pp. 233–241). Paris: Association for Science and Information on Coffee.