



**HAL**  
open science

## Gum Arabic in solution: Composition and multi-scale structures

Marina Atgié, Jean-Christophe Garrigues, Alexis Chennevière, Olivier Masbernat, Kevin Roger

► **To cite this version:**

Marina Atgié, Jean-Christophe Garrigues, Alexis Chennevière, Olivier Masbernat, Kevin Roger. Gum Arabic in solution: Composition and multi-scale structures. *Food Hydrocolloids*, 2019, 91, pp.319-330. 10.1016/j.foodhyd.2019.01.033 . hal-02194435

**HAL Id: hal-02194435**

**<https://hal.science/hal-02194435v1>**

Submitted on 12 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.






## Open Archive Toulouse Archive Ouverte

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible

This is an author's version published in: <http://oatao.univ-toulouse.fr/24244>

**Official URL:** <https://doi.org/10.1016/j.foodhyd.2019.01.033>

**To cite this version:**

Atgié, Marina  and Garrigues, Jean-Christophe and Chennevière, Alexis and Masbernat, Olivier  and Roger, Kevin  *Gum Arabic in solution: Composition and multi-scale structures*. (2019) *Food Hydrocolloids*, 91. 319-330. ISSN 0268-005X

Any correspondence concerning this service should be sent to the repository administrator: [tech-oatao@listes-diff.inp-toulouse.fr](mailto:tech-oatao@listes-diff.inp-toulouse.fr)

# Gum Arabic in solution: Composition and multi-scale structures

M. Atgié<sup>a</sup>, J.C. Garrigues<sup>b</sup>, Alexis Chennevière<sup>c</sup>, O. Masbarnat<sup>a</sup>, K. Roger<sup>a,\*</sup>

<sup>a</sup>Laboratoire de Génie Chimique, Université de Toulouse, CNRS, INPT, UPS, Toulouse, France

<sup>b</sup>Laboratoire des IMRCP, Université de Toulouse, CNRS UMR 5623, Université Paul Sabatier, 31062 Toulouse Cedex 09, France

<sup>c</sup>Laboratory Léon Brillouin, CEA-CNRS, Gif sur Yvette, France

---

Gum Arabic is a natural acacia tree exudate containing hyperbranched polysaccharides and proteins. Here, we perform a dual chromatographic separation together with small-angle X-ray and neutron scattering structural characterizations. We show that the different species present in Gum Arabic can not be easily classified in distinct families. They are rather build from various combinations of two building blocks that are evidenced by a mismatch between small-angle X-ray and neutron scattering. One block corresponds to hyperbranched polysaccharides, which we describe as three-dimensional multi-scale porous colloids possessing three length scales of 7, 2 and 0.7 nm. The other block corresponds to protein chains that organize as Gaussian chains in solution and are prone to aggregation. A large array of polysaccharide/protein conjugates was identified, which differs in size, hydrophobicity and amino-acid content. Still, their structure is always the juxtaposition of the two building blocks structures. Additionally, small-angle neutron scattering reveals that large-scale structures are ubiquitous in Gum Arabic solutions and originate from the self-association of both free and conjugated polypeptide chains. Despite its compositional complexity, Gum Arabic solutions thus possess a robust multi-scale structure that is mainly impacted by concentration and ionic repulsions.

---

## 1. Introduction

Gum Arabic, or acacia gum, is an exudate produced by acacia trees in sub-Saharan countries, which operates as a natural wound plaster, thus shielding trees against insects, molds and droughts. Highly water soluble, gum Arabic has been under extensive use since the dawn of civilization with an extensive array of applications (Sanchez et al., 2018; Zipkin, Wagner, McGrath, Brooks, & Lucas, 2014), notably as a binder and emulsifier. This natural hydrocolloid is composed of a complex mixture of biopolymers. In order to understand gum Arabic interfacial properties, its composition and associated structures have been extensively studied. The gum was first identified as a heterogeneous mixture of protein and polysaccharides differing in their molar mass and hydrophobicity (Churms, Merrifield, & Stephen, 1983; Jermny, 1962; Lewis & Smith, 1957; Vandeveld & Fenyo, 1985). Chromatographic analysis has been the most employed technique in the literature to characterize gum Arabic structural composition. Gum species were first separated through their molar mass using size exclusion chromatography (Randall, Phillips, & Williams, 1988; Vandeveld & Fenyo, 1985). It was shown that the gum presents a continuum of molar mass ranging from 5 kg mol<sup>-1</sup> to 1000 kg mol<sup>-1</sup> with a mean molar mass around 500 kg mol<sup>-1</sup>. Gum Arabic polypeptide-rich species were associated with the higher molar mass fraction. Hydrophobic interaction chromatography separation was later performed on gum Arabic, enabling the sorting of species according to

their hydrophobicity. From these separations the gum was described as a mixture of three fractions: an arabinogalactan rich polysaccharide fraction constituting the bulk of the gum (AG) (88%), a polysaccharide-protein conjugates (AGP) (10%) and a fraction of glycoproteins (GP) (2%) (Al-Assaf, Phillips, Aoki, & Sasaki, 2007; Islam, Phillips, Sljivo, Snowden, & Williams, 1997; Osman, Menzies, Williams, & Phillips, 1994; Osman, Menzies, Williams, Phillips, & Baldwin, 1993; Randall, Phillips, & Williams, 1989; Renard, Lavenant-Gourgeon, Ralet, & Sanchez, 2006; Williams, Phillips, & Stephen, 1990). Yariv reagent was used on each of these recovered fractions and it was observed that protein/polysaccharide conjugates were present in each (Osman et al., 1993). While the exact composition in these different species was shown to depend on the geography, soil and age of the trees, the overall chemical and physico-chemical characteristics are nevertheless fairly robust (Al-Assaf, Phillips, & Williams, 2005; Idris, Williams, & Phillips, 1998; Osman et al., 1994; Osman, Williams, Menzies, & Phillips, 1993). A simple two-dimensional separation of gum species was performed using size exclusion chromatography as the first dimension, followed with a hydrophobic interaction separation. Renard et al. thus showed that the gum contains a hydrophilic fraction (associated with the lowest protein content) with a small polydispersity in size, whereas the other fractions (with higher protein content) displayed a broad range of sizes (Renard et al., 2006). The authors concluded that gum Arabic was composed of a continuum of species.

The structure of gum Arabic conjugates was extensively studied and

---

\* Corresponding author.

E-mail address: [kevin.roger@ensiacet.fr](mailto:kevin.roger@ensiacet.fr) (K. Roger).

several models were proposed. The prominent model describes gum conjugates as wattle blossoms consisting of a polypeptide chain bearing approximately five spheroids composed of highly branched polysaccharides (Randall et al., 1988). The gum was alternatively described as a twisted hairy rope by Qi et al. while Goodrum et al. proposed a symmetric representation of the gum conjugates based on a palindromic polypeptide sequence (Goodrum, Patel, Leykam, & Kieliszewski, 2000; Qi, Fong, & Lampion, 1991). Finally, Mahendran and co-workers proposed a slightly different representation with the main difference arising from the molar mass of the polysaccharide spheroids but in accordance with the wattle blossom model (Mahendran, Williams, Phillips, Al-Assaf, & Baldwin, 2008).

However, these models were elaborated mainly from rheological analysis, chemical analysis and degradation studies, which are rather indirect means to characterize structures. Direct methods such as Small Angle Neutron and X-ray Scattering measurements (SANS and SAXS) were only performed by two separate research groups. Dror, Cohen and coworkers (Dror, Cohen, & Yerushalmi-Rozen, 2006) used both SANS and SAXS to study gum Arabic aqueous solutions at various ionic strengths and concentrations. However, they pointed out the necessity to characterize gum Arabic fractions for a better structural understanding.

Renard, Sanchez and coworkers used either SANS or SAXS in a series of four papers dedicated to the characterization of the structure of three gum Arabic fractions in solution obtained through a hydrophobic interaction chromatographic separation (Renard, Garnier, Lapp, Schmitt, & Sanchez, 2012; Renard, Lavenant-Gourgeon, Lapp, Nigen, & Sanchez, 2014; Renard, Lepvrier, et al., 2014; Sanchez et al., 2008). Conjugates fraction (AGP) was described as a thin triaxial ellipsoid and the polysaccharide-rich fraction (AG) as a thin disk. However no spectra of the whole gum Arabic in solution was reported, which precluded any comparison with the fractions they separated. Also, since either SAXS or SANS was used, it was not possible to directly compare the fractions between themselves. Furthermore, large length-scales were not probed and the solutions parameters such as ionic strength and concentrations were not varied.

At present, the literature consensus (Sanchez et al., 2018) has been to integrate these two-dimensional units into a wattle blossom model, using also the palindromic sequence for polypeptide chains, thus assuming their homogeneity in all gum species. Additionally, a light scattering study (Wang, Burchard, Cui, Huang, & Phillips, 2008), rheology studies (Li et al., 2009, 2011) as well as a few spectra presented by Dror et al. (Dror et al., 2006) suggest some self-association in solution.

In this study, we examine in details the heterogeneity of gum Arabic, the colloidal structure of its constituting species and the resulting structure of gum Arabic aqueous solutions. Notably we performed a two dimensional chromatographic separation using three detections and very narrow fractions of the first separation to gain in resolution. Small-angle X-ray and neutron scattering experiments were performed in order to characterize the different structures present in gum Arabic solutions and its fractions over a large range of length scales.

## 2. Materials and methods

### 2.1. Materials

Spray-dried powder of gum arabic (*Acacia senegal*) was a gift from Caragum International<sup>®</sup> (Marseille, France) and the composition was 2.5 wt% of proteins (Nx7), 11 wt% moisture, 3.16 wt% ash and nearly no lipid (CG103). This results in a protein content of 2.9wt% of proteins (Nx7) for the dry gum Arabic without ash. Sodium chloride (99.5 > % BioXtra), hexadecane (99% ReagentPlus), pentane and hydrochloric acid were purchased from Sigma Aldrich. Distilled water was used for all experiments.

### 2.2. Nitrogen content analysis

Nitrogen content analyses were performed at the Laboratoire de Coordination Chimique (Toulouse, France). A Perkin-Elmer 2400 CHN series II was operated at 1050 °C with oxygen as carrier gas. Each sample was analyzed in duplicate. Protein fraction content was deduced from the nitrogen content through the protein conversion number calculated from the gum Arabic proteins amino acid distribution (Renard et al., 2006) as shown in (1) where  $\rho_{\text{protein}}$  and  $\rho_{\text{nitrogen}}$  are the fractions of protein and nitrogen, respectively, in wt%:

$$\rho_{\text{protein}} = \rho_{\text{nitrogen}} \times 7 \quad (1)$$

It is worth noting that this value slightly differs from the common value of 6.6 taken in the majority of gum Arabic publications. However, we must stress that the 6.6 value does not correspond to the measured amino-acid distribution of gum Arabic and is therefore incorrect.

### 2.3. Size exclusion chromatography

Size exclusion chromatography was used to separate Arabic gum species as a function of their relative hydrodynamic volume. A 7.8mmx300 mm BioSuite 450 Å SEC column (Waters<sup>®</sup>) packed with 8 µm porous silica beads was used. The average pore size of the silica beads is 450 Å.

The separation was performed on an Alliance HPLC unit (Waters 2695 separations module), a 0.5mol.L<sup>-1</sup> NaCl aqueous solution at 25 °C was used as the eluent phase at a flow rate of 0.8 mL.min<sup>-1</sup>. Each sample was filtrated with a nylon 0.2 µm membrane and a volume of 50 µL was injected. UV detection was performed at 210 nm and 280 nm using a Waters<sup>®</sup> 2487 UV detector. A refractive index detection was used to measure the mass percentage of each eluted moiety (Waters<sup>®</sup> 410 differential refractometer). The refractive index detection was only available until 10.5 mL of elution due to a negative peak appearing from the difference in refractive index between the sodium chloride of the eluent phase and the water from the injected sample. The column was calibrated using branched dextran standards (Waters<sup>®</sup>) (Kato, Tokuya, & Takahashi, 1983) (see Fig. 7 in supporting information).

Variation coefficients for each detection were calculated from three injections of the same sample. For UV detection at 210 nm a 0.3% variation coefficient was measured. It was of 0.5% for UV detection at 280 nm and of 13% for the refractive index detection.

UV detection at 280 nm is more sensitive to the amount of aromatic amino acid (tyrosine and phenylalanine), which are minor constituents of gum Arabic polypeptides, as already observed in the literature (Goodrum et al., 2000; Randall et al., 1989; Ray, Bird, Iacobucci, & Clark Jr, 1995). UV detection at 210 nm is mostly sensitive to the peptide bond ( $\pi \rightarrow \pi^*$  transitions from C=O bond of the amides linkages). A much lower, but non zero, absorbance is expected from polysaccharides at this wavelength, notably due to the presence of carboxylic groups (uronic acid moieties).

### 2.4. Hydrophobic interaction chromatography

Hydrophobic interaction chromatography was used to separate Arabic gum fractions as a function of their hydrophobic properties. We used a 7.5mmx75 mm 10 µm Biosuite Phenyl column (Waters<sup>®</sup>), which consists of phenyl groups grafted to a methacrylic ester based polymeric resin. The average pore size of the column is 1000 Å to accommodate macromolecules with high molar masses. A salt concentration gradient is required to create a "salting-out" effect and progressively desorb hydrophobic species. Less hydrophobic species are eluted first with highest salt concentration, while more hydrophobic species are eluted last.

The separation was performed on an Alliance HPLC unit (Waters 2695<sup>®</sup> separations module). The mobile phase was composed of a solution of NaCl at a constant flowrate of 0.5 mL.min<sup>-1</sup>. A continuous

time gradient of salt concentration from 4 mol.L<sup>-1</sup> to 0 mol.L<sup>-1</sup> was applied during 22 min, followed by pure water during 12 more minutes (see Fig. 8 in supporting information). Each sample was filtered with a nylon 0.2 μm membrane and a volume of 30 μL was injected. UV absorbance was measured at 210 nm and 280 nm (Waters® 2487 UV detector). We observed that UV absorbance (especially at 210 nm) was very sensitive to NaCl concentration. The baseline was not constant along the separation time. In order to correct this deviation, each chromatogram was subtracted by the chromatogram of distilled water injected in same elution conditions.

Variation coefficients for each detection were estimated from comparison of three injections of the same sample. A 6% (respectively 11%) variation coefficient was observed at 210 nm (respectively 280 nm).

## 2.5. Two dimensional chromatographic separation

Five fractions were collected using size exclusion chromatography as described before. A volume of 80 μL of a gum solution at 5w/w% was injected in the SEC column. Two separations were carried out and fractions of 120 μL were collected and freeze-dried. These fractions were then dissolved in 50 μL of distilled water and 20 μL of each resulting solution was injected in the HIC column for the second dimension separation. Since the resulting sample concentrations were low, we could only use a 210 nm UV detection.

## 2.6. Small angle neutron scattering (SANS) experiments and data treatment

SANS measurements were mostly performed at the Laboratoire Léon Brillouin neutron facility on both PACE and PAXY instruments, with some measurements performed at ISIS on the SANS2D instrument. The majority of spectra was acquired on the PAXY small angle diffractometer located at the Orphée reactor in Saclay, France. Samples were dissolved in 20 mM NaCl in D<sub>2</sub>O solution. Spectra were recorded using four different spectrometer configurations:  $\lambda = 6 \text{ \AA}$  (incident wavelength),  $d = 1 \text{ m}$  (sample to the detector distance),  $\lambda = 6 \text{ \AA}$  and  $d = 3 \text{ m}$ ,  $\lambda = 8.5 \text{ \AA}$  and  $d = 5 \text{ m}$  and  $\lambda = 15 \text{ \AA}$  and  $d = 6.7 \text{ m}$ . The range of wave vectors  $q$  covered extends between  $6.5 \cdot 10^{-3}$  and  $4.5 \cdot 10^{-2} \text{ \AA}^{-1}$ . Contributions from the Quartz Helma cells and ambient noise were subtracted. The data were normalized using H<sub>2</sub>O spectra and set to absolute scale thanks to normalization factor measurements (Brület, Lairez, Lapp, & Cotton, 2007). The subtraction of the incoherent background was performed using the sample transmission at high  $q$  values and a calibration of D<sub>2</sub>O/H<sub>2</sub>O solutions at different ratios (transmission as a function of scattered intensity). Several power laws were observed and the typical absolute uncertainty on each power law exponent was evaluated to  $\pm 0.1$ .

## 2.7. Small angle X-ray scattering (SAXS) experiments and data treatment

SAXS experiments were performed on the ID02 instrument at the ESRF synchrotron facility in Grenoble, France and in-house on a Xeuss 2.0 instrument from XENOCs. For a few samples, spectra were collected on both instruments and were identical for the measurements durations chose, which precludes any radiation damaging artefacts. Samples were injected in disposable quartz capillaries (1.5 mm inner diameter). Measurements were taken at 1.5 and 7 m. Azimuthal averaging was performed after mask subtraction, yielding one-dimensional spectra after normalization by transmissions. One-dimensional spectra were subtracted by the adequate reference spectrum from either a water-filled or buffer-filled capillary. Several power laws were observed and the typical absolute uncertainty on each power law exponent was evaluated to  $\pm 0.1$ .

## 2.8. Circular dichroism

Circular dichroism measurements were performed using a native gum Arabic sample and fractions separated by SEC on a Jasco J-815 CD spectrometer at IMRCP laboratory in Toulouse, France. A quartz cell with a 10 mm path length was used. Measurements were performed between 185 and 250 at 25 °C. A concentration of 0.025% was used for gum Arabic and of 0.00625% for FA, FB and FC. FD was not measured due to its low solubility. Results were normalized by the concentration and the path length to be compared.

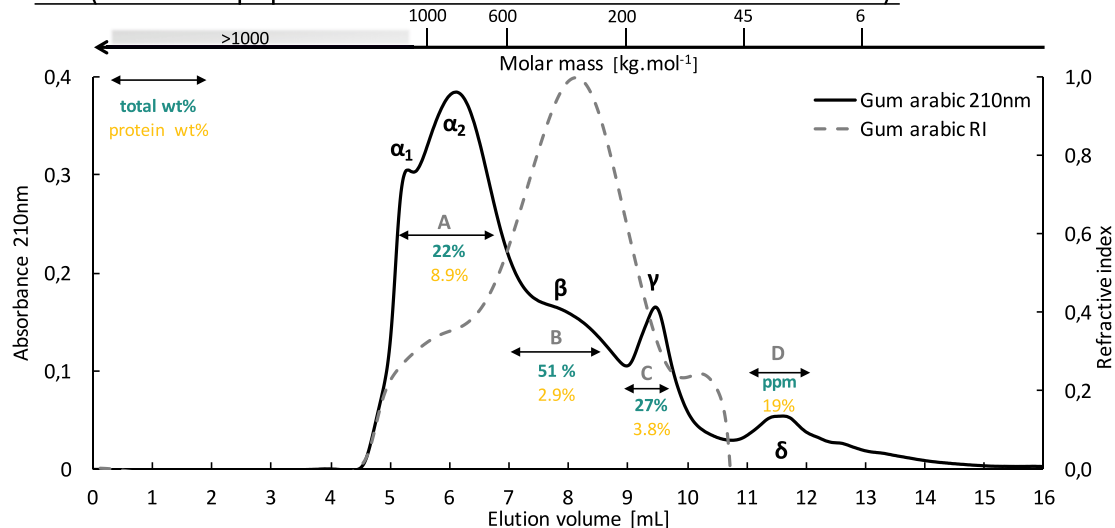
## 3. Results

### 3.1. Chromatographic separation

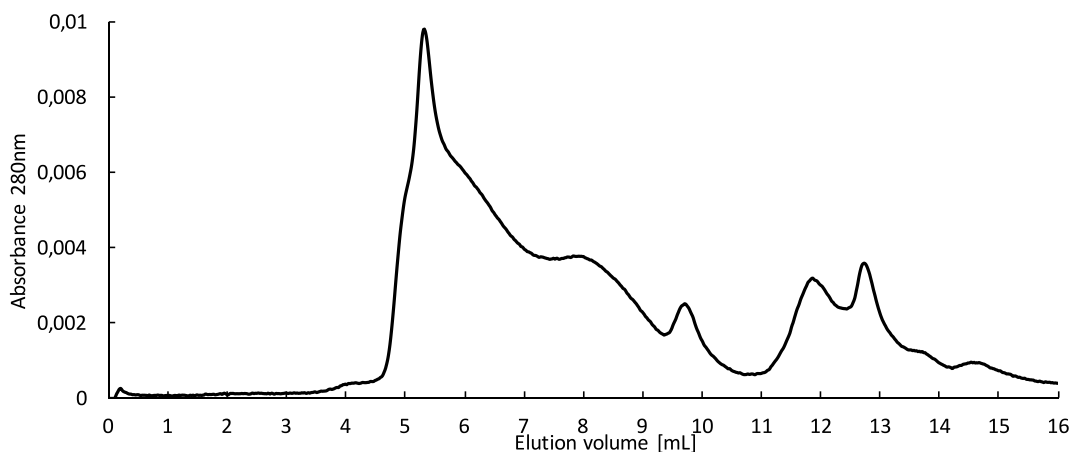
Since gum Arabic is a hydrocolloid mixture, its proper description would require a separation into distinct populations. Chromatographic separation was thus performed to sort species with respect to their size, using size exclusion chromatography, and to their hydrophobicity, using hydrophobic interaction chromatography. In this study, we use a gum Arabic possessing a 2.9 wt% protein content (see materials and methods).

Fig. 1 A displays the size exclusion chromatogram of our batch of gum Arabic using both a UV detection at 210 nm, which is mostly sensitive to the peptide bond (Kobayashi, Utsugi, & Matsuda, 1986; Kuipers & Gruppen, 2007), and a refractive index detection, which is proportional to concentration. This chromatogram is similar to previously published gum Arabic chromatograms (Idris et al., 1998; Padala, Williams, & Phillips, 2009; Renard et al., 2006), which ensures the representability of our gum Arabic batch. We observed that gum Arabic displays a broad range of molar mass ranging from 6 kg mol<sup>-1</sup> to more than 1000 kg mol<sup>-1</sup>. From the 210 nm UV detection, we observe five peaks (Peak  $\alpha_1$ : > 1000 kg mol<sup>-1</sup>, peak  $\alpha_2$ : 800 kg mol<sup>-1</sup>, peak  $\beta$ : 420 kg mol<sup>-1</sup>, peak  $\gamma$ : 135 kg mol<sup>-1</sup>, peak  $\delta$ : 23 kg mol<sup>-1</sup>). Using a preparative size exclusion chromatographic separation, we collected broad fractions centered around these peaks. Fraction A corresponds to the largest species, represents 22 wt% of gum Arabic mass and displays a protein content of 8.9 wt%. Fraction B represents 51 wt% of gum Arabic mass and has a protein content of 2.9 wt%. Fraction C represents 27 wt% of gum Arabic mass and has a protein content of 3.8 wt%. Fraction D, which corresponds to the smallest species, is a trace fraction of gum Arabic with a high protein content of 19 wt%. This mass partition of gum Arabic in these different species is quantitatively consistent with the refractive index detection, which displays a broad peak centered on fraction B and two adjacent shoulders on fractions A and C. Interestingly, peak magnitudes thus differ between refractive index detection and 210 nm UV detection, which is quantitatively consistent with the differences in protein contents between the different fractions. A 280 nm UV detection was also performed and is displayed in Fig. 1B. This detection is mostly sensitive to the aromatic amino acids (tryptophan and tyrosine in gum Arabic). The relative signal of Fraction D largely increases compared to the 210 nm detection and four peaks can be distinguished. The signal of fraction C decreases and within fraction A, peak  $\alpha_1$  increases while peak  $\alpha_2$  decreases. This difference between the two UV detections indicates differences in protein composition between the different fractions. Fraction D, B and part of A centered on peak  $\alpha_1$  are enriched in aromatic amino-acids, which are hydrophobic, compared to fractions C and part of fraction A centered on peak  $\alpha_2$ . All fractions obtained through size exclusion chromatography contain proteins. It is thus interesting to quantify the amphiphilic behavior of the different species contained in Gum Arabic. To this end, we performed a hydrophobic interaction chromatography, which is based on the interaction of the eluted species with a phenyl grafted column. Unlike other studies (Randall et al., 1989; Ray, Bird, Iacobucci, & Clark, 1995; Renard et al., 2006), our experiments have been performed with a continuous gradient of salt concentration in order to avoid artefacts

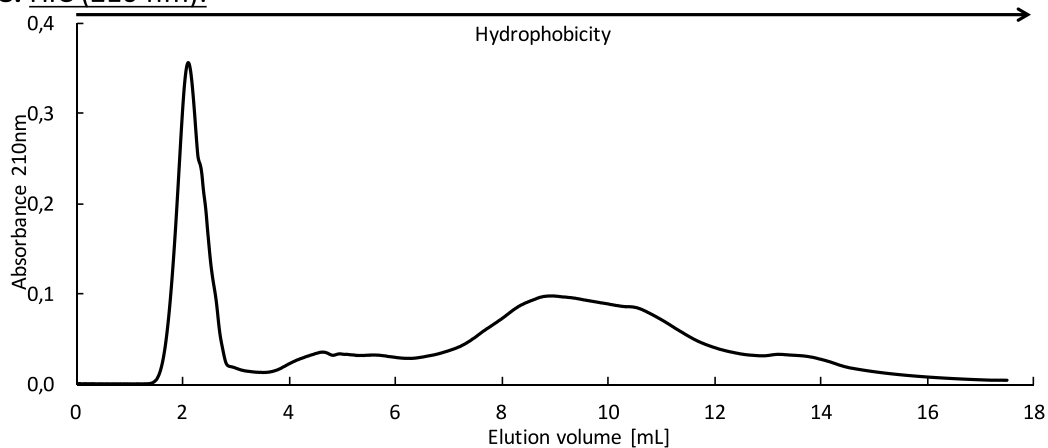
**A: SEC (210nm  $\propto$  peptide bond + Refractive index  $\propto$  concentration):**



**B: SEC (280 nm  $\propto$  aromatic amino acid):**



**C: HIC (210 nm):**

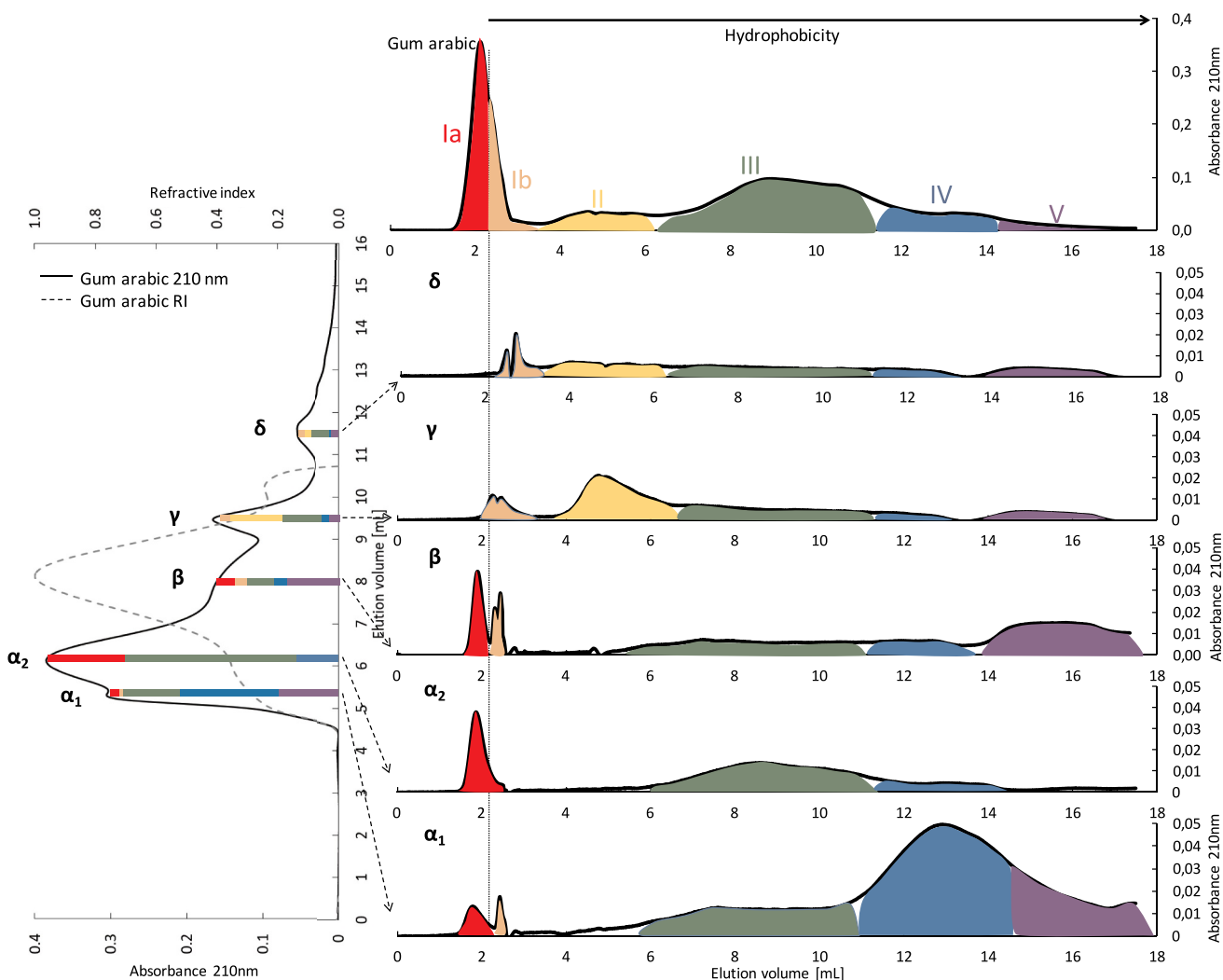


**Fig. 1.** (A) Size exclusion chromatogram at 210 nm, (B) at 280 nm and (C) Hydrophobic interaction chromatogram of a 30 g/L aqueous solution of gum Arabic. Dotted line: Refractive index detection. (A) Weight, or mass, percentages of recovered fraction using a preparative size exclusion column (same column filling, larger dimensions) are indicated as well as the protein content for these fractions. Scale for molar mass as a function of elution volumes is displayed on the top of the chromatograms. (C) Less hydrophobic species of the gum are not retained by the column and last eluted species are the more hydrophobic ones as indicated by the scale on top of the chromatogram.

during abrupt salinity changes. The resulting chromatogram is displayed in Fig. 1C. The first peak is actually the convolution of several peaks as shown in Fig. 2, with the first convoluted peak corresponding to species that are not retained by the phenyl grafted column and the

following convoluted peaks to species that are only weakly retained. This observation suggests that there is a rather homogeneous family of hydrophilic species in gum Arabic. In contrast, the rest of the chromatogram displays a broad elution profile, which emphasizes the





**Fig. 2.** Two-dimensional separation of gum Arabic with size exclusion chromatography as the first dimension column and hydrophobic interaction as the second dimension. Five fractions were collected by size exclusion chromatography with a narrow volume collection (120  $\mu$ L). ( $\alpha_1$ : > 1000 kg.mol<sup>-1</sup>,  $\alpha_2$ : 800 kg.mol<sup>-1</sup>,  $\beta$ : 420 kg.mol<sup>-1</sup>,  $\gamma$ : 135 kg.mol<sup>-1</sup>,  $\delta$ : 23 kg.mol<sup>-1</sup>).

heterogeneity of the gum as a continuum of species of increasing hydrophobicity.

### 3.2. Two-dimensional separation

Both size exclusion and hydrophobic interaction chromatographic separations yield broad elution profiles, which indicates a continuum of sizes and amphiphilicities. However, the relationship between size and amphiphilicity cannot be deduced from a separate analysis and calls for a coupled, or two dimensional, chromatographic separation. We thus performed a size exclusion chromatographic separation followed by a hydrophobic interaction chromatographic separation. This order is justified by the much lower salt concentrations used in the size exclusion chromatographic separation and the higher resolution of this latter technique. Five narrow fractions were first collected by size exclusion separation using a fraction collector. These fractions are centered on the five peaks observed in the 210 nm UV detection (see Fig. 1,  $\alpha_1$ : 5.2 mL,  $\alpha_2$ : 6.04 mL,  $\beta$ : 8 mL,  $\gamma$ : 9.5 mL,  $\delta$ : 11.5 mL). Once collected, these fractions were concentrated prior to injection in the second dimension column by freeze-drying. The dissolved fractions were injected in the size exclusion column (first dimension) for a control. The resulting chromatograms shown in Fig. 1 in supporting information are single-peak shaped and centered around an elution volume. The five fractions

differing in their molar mass ( $F_{\alpha_1}$ : > 1000 kg mol<sup>-1</sup>,  $F_{\alpha_2}$ : 800 kg mol<sup>-1</sup>,  $F_{\beta}$ : 420 kg mol<sup>-1</sup>,  $F_{\gamma}$ : 135 kg mol<sup>-1</sup>,  $F_{\delta}$ : 23 kg mol<sup>-1</sup>) were then separated in the hydrophobic interaction column. The two-dimensional separation resulting from this procedure is reported in Fig. 2. A UV at 210 nm detection was used for the two-dimensional separation.

This two-dimensional separation immediately reveals that it is not possible to associate a given molar mass to a given hydrophobicity for all species composing gum Arabic, as shown in Fig. 2. This means that a chromatographic fraction either obtained through SEC or HIC separation, does not correspond to a single family. Fractions and families should thus be distinguished when discussing gum Arabic constituents.

Fractions Ia and Ib include species of low hydrophobicity that are not or barely retained by the hydrophobic interaction column. Their size distribution spreads over the whole range of gum Arabic but mostly covers molar mass ranging from 200 to 800 kg mol<sup>-1</sup>, comprised between peaks  $\alpha_2$  and  $\gamma$  in Fig. 2. Fraction II is composed of species of intermediate molar mass (100 kg mol<sup>-1</sup>) and hydrophobicity (4 mL < elution volume < 6 mL), which is centered on peak  $\gamma$ . Fraction III comprises species of high molar mass (500–1000 kg mol<sup>-1</sup>) and intermediate hydrophobicity (6 mL < elution volume < 11 mL) lying between peaks  $\alpha_1$  and  $\alpha_2$ . Fraction IV corresponds to the highest molar mass (> 1000 kg mol<sup>-1</sup>) and high hydrophobicity (11 mL < elution

volume < 14 mL) and is centered on peak  $\alpha_1$ . Fraction V is a group of several families comprising species of highest hydrophobicity (elution volume > 14 mL), and covering the whole size spectrum of gum Arabic, with a minimum around peak  $\alpha_2$ .

Ideally, this two-dimensional separation could be used to produce thin fractions of given molar mass and hydrophobicity, which could be further analyzed by other techniques to assess their colloidal structure. However, this was not judged feasible due to the very small amounts collected at each separation step.

In the rest of the results section, we will thus present a structural characterization of gum Arabic as well as broad fractions collected through size exclusion chromatography, in solution. The discussion will link this structural insight with the different fractions separated through the two-dimensional chromatographic analysis.

### 3.3. Structure of gum Arabic probed by SANS and SAXS

Aqueous solutions of gum Arabic are colloidal systems composed of several types of hydrocolloids. For such systems, a non-intrusive and representative structural characterization can be performed using scattering methods. Wang et al. described a self-association behavior for gum Arabic in solution using light scattering (Wang et al., 2008) but small-angle scattering neutron (SANS) or X-ray (SAXS) scattering are more adequate regarding the characteristic length scale of hydrocolloids.

Whilst highly similar, SAXS and SANS largely differ in respect to the origin of contrast. In SAXS, the scattering contrast originates from differences in electronic density, while in SANS it originates from differences in nucleus types and is particularly sensitive to isotopic H/D substitution. A dual characterization using both techniques is thus a powerful mean to investigate the heterogeneous mesostructure of hydrocolloids.

Here we report an extensive structural characterization of both gum Arabic and its fractions separated through size exclusion chromatography, using both SAXS and SANS methods over an extended range, which notably includes large length scales.

First of all, we compared our gum Arabic batch (Caragum international<sup>®</sup>) to another batch from a different supplier, which was used by Renard, Sanchez et al. (Nexira<sup>®</sup>). Both gums were dissolved at 5 g/L and 50 g/L in D<sub>2</sub>O, 0.02M NaCl and the resulting solutions were measured using SANS. Since both scattering curves are nearly identical (see Fig. 2 in supporting information), we can conclude that the two gums are similar from a structural standpoint.

Strikingly, we observe an abrupt increase of the scattering intensity with decreasing  $q$ , which corresponds to large correlation length scales (Fig. 3). This intensity upturn scaled at  $q^{-3}$  at the 5 g/L and  $q^{-2}$  at 50 g/L. Interestingly, the signature of large correlation lengths was not observed by Renard and coworkers since it occurs at a lower  $q$ -range than what they investigated. In the work of Dror and coworkers, this low- $q$  range was only investigated for a 5 w/w% gum Arabic solution without salt. They also observed a strong intensity upturn at low  $q$  values with a  $q^{-2}$  power law. Importantly, filtrating our aqueous solution with a 200 nm filter did not result in any visible changes on its SANS spectra (see Fig. 3 in supporting information, SAXS spectra are also unchanged), which suggests that these large scale structures are highly flexible and possibly result from a self-assembly behavior as previously inferred from light scattering measurements (Al-Assaf, Sakata, McKenna, Aoki, & Phillips, 2009). Another noticeable feature of the SANS spectra is the presence of a structure correlation peak in the intermediate- $q$  range (red and dark blue curves in Fig. 3), which is screened upon salt addition at low gum concentration (light blue curve) and still present at higher gum concentration (50 g/L orange curve). Such a behavior has been previously reported and is consistent with the weakly polyelectrolyte nature of gum Arabic species (Dror et al., 2006). In the case of the higher gum concentration studied (50 g/L), a 20 mM NaCl concentration was not sufficient to fully screen the charges.

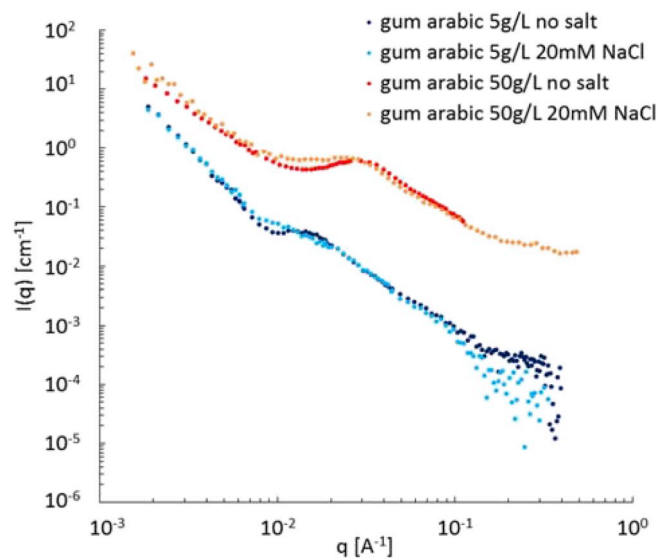


Fig. 3. Small angle neutron scattering spectrum of gum sample at 5 g/L or 50 g/L with (20 mM NaCl) or without salt. Samples were measured in D<sub>2</sub>O. Data from gum spectra at 50 g/L and no salt from Dror et al., 2006 (Dror et al., 2006).

Finally, in the high  $q$ -range and thus small length scales, the scattered intensity does not scale with  $q^{-4}$ , which would be expected for smooth interfaces according to the Porod law. A  $q^{-2}$  law is observed, which may correspond to several possible structures such as Gaussian chains, a porous network with a fractal dimension of 2 or two-dimensional objects such as disks (Glatter, 2017). However, the signal-to-noise ratio is too high in this  $q$ -range to consider a finer analysis.

In order to gain a better resolution at high  $q$  values, SAXS measurements were also performed. Similarly to the work of Dror et al., the effect of concentration and ionic strength was also investigated but over a larger  $q$ -range and using a point collimation method, which unambiguously yields the scattered intensity versus  $q$ . Fig. 4 displays SAXS spectra collected for various gum Arabic concentrations in water at three different ionic strengths. In the high- $q$  range ( $5 \cdot 10^{-2}$  -  $3 \cdot 10^{-1} \text{ \AA}^{-1}$ ), all curves overlap when the intensity is normalized by the concentration (see Fig. 4 in supporting information), which indicates that the structure at small length scales is not impacted by either concentration or ionic strength. This corresponds to form factors of one or several colloidal populations. In this range, a  $q^{-2.1}$  power law is followed by a  $q^{-2.5}$  power law, which provides a more resolved characterization than what could be achieved using SANS, but with another contrast sensitive mostly to counter ions clouds. Such a signature is more consistent with form factors of mass fractals rather than two-dimensional objects, which should scale exactly as  $q^{-2}$ . Furthermore, the presence of oscillations suggests that well-defined correlation lengths are present over these few nanometers length scales. Displaying the SAXS spectra in a Kratky plot ( $I(q)q^2$  vs.  $q$ ), which is commonly used to investigate intrinsically disordered proteins (Kikhney & Svergun, 2015; Receveur-Bréchet & Durand, 2012), allows to conveniently extract the gyration radius associated with each oscillation through  $q_{\text{peak}} \cdot R_g = 3^{1/2}$  (see Fig. 5). We obtain three gyration radius of respectively 0.7, 2 and 7 nm from peak positions in the Kratky plot at 0.246, 0.854 and  $2.54 \text{ nm}^{-1}$ .

In the intermediate- $q$  range ( $10^{-2}$  -  $5 \cdot 10^{-2} \text{ \AA}^{-1}$ ), a correlation peak is observed, at the same abscissa as in the corresponding SANS spectrum. Its magnitude decreases with the ionic strength, which is expected for colloidal objects interacting through long-range ionic repulsions. Consistently with previous reports (Dror et al., 2006; Renard et al., 2012; Renard, Lepvrier, et al., 2014; Sanchez et al., 2008), 20 mM and 50 mM NaCl ionic strengths were sufficient to suppress this correlation peak at low gum concentrations. Without addition of salt



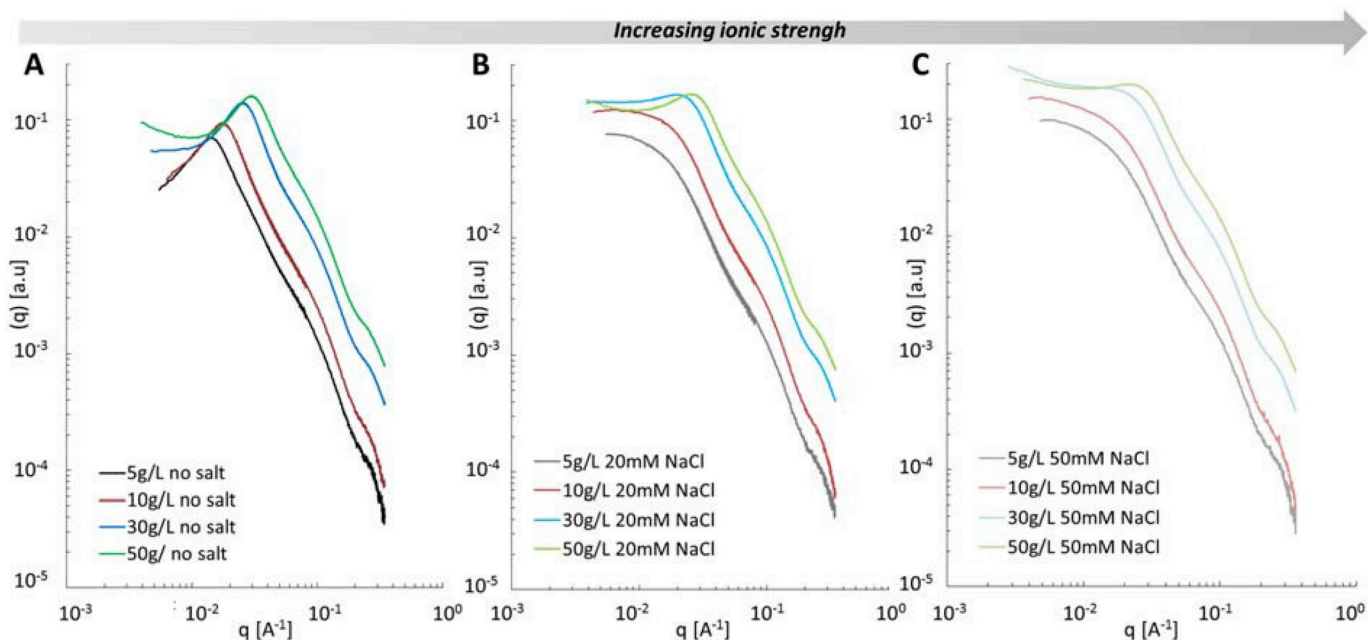


Fig. 4. Small angle x-ray scattering spectra of gum arabic at different gum and NaCl concentrations (A): no salt, (B): 20 mM NaCl and (C): 50 mM NaCl.

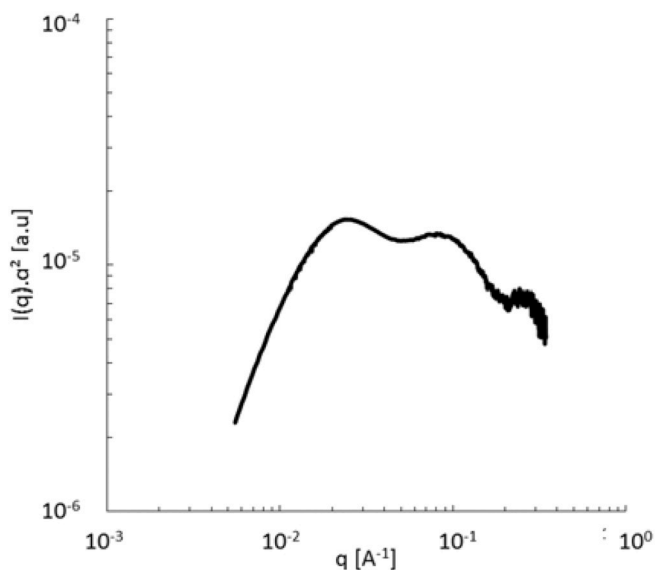


Fig. 5. Kratky-plot type representation of a gum Arabic solution at 5 g/L in 20 mM NaCl by small angle x-ray scattering.

(Fig. 4A), the  $q$ -abscissa of this structure peak varied as the gum concentration to the power  $1/3$ , as also observed by Dror et al. (Dror et al., 2006). This scaling law suggests a repulsion-controlled interaction between colloidal objects rather than within a concentrated polyelectrolyte matrix (Dror et al., 2006). For the 5 g/L, the correlation distance  $2\pi/q_{\text{peak}}$  is 44 nm, while for the 50 g/L solution it is 21 nm. Since the correlation peak significantly broadens at a 100 g/L gum concentration (Dror et al., 2006), we can deduce that the onset of the overlap occurs at a gum concentration ranging between 50 and 100 g/L, corresponding to 15–20 nm correlation distances.

At low gum concentrations and 20–50 mM NaCl concentrations, the correlation peak disappears and the intensity tends towards a constant value in the limit of low  $q$  values, indicating a finite size of the scattering objects. This value is similar to the spacing value deduced from the correlation peak  $q$ -abscissa at the overlap concentration, which suggests that these finite objects are the charged objects. In the

corresponding SANS spectra, the intensity never reaches a constant value due to its strong upturn in the low- $q$  range, which is not observed in SAXS spectra. Only a slight increase of the scattered intensity at low  $q$  could be detected for larger gum concentrations in SAXS spectra. This SANS/SAXS mismatch is uncommon and calls for an explanation.

We verified that SAXS spectra of both  $\text{H}_2\text{O}$  gum solutions and  $\text{D}_2\text{O}$  gum solutions were identical. The SANS/SAXS mismatch is thus unrelated to any slight interaction differences in the two systems, which may occur in the vicinity of a phase transition. This leaves only another explanation related to the differences in contrast distribution within a heterogeneous scattering system, the parts of which display important contrast differences with the solvent, between SAXS and SANS experiments. SANS experiments were performed by dissolving gum Arabic in deuterated water ( $\text{D}_2\text{O}$ ) so the contrast difference is mainly due to the hydrogen/deuterium ratio, while in SAXS experiments the contrast difference originates from differences in electron density. SANS/SAXS mismatch likely results from two superimposing phenomena, which both highlight the polypeptide backbone compared to the polysaccharides. Firstly, hydrogen-deuterium exchange is expected to take place on labile hydrogen (Banc et al., 2016), which are present on both hydroxyl (sugars) and amine groups (proteins). This exchange will tend to decrease the contrast between polypeptide/polysaccharides and deuterated water. This effect is more pronounced for polysaccharides than for proteins since they possess more hydroxyl groups than proteins possess amine groups. Secondly, polysaccharides are likely to be more hydrated than the polypeptide chains, which possess hydrophobic amino-acids and secondary structures. This hydration level will drastically reduce the neutronic contrast of polysaccharides out of the range of molecular length scales (where hydration is directly probed). We can thus conclude that the intensity upturn in the low- $q$  range in the SANS spectra is the signature of polypeptide chains aggregate-like structures, as a result of a strong self-association mechanism of polypeptide chains in solution.

#### 3.4. Structure of gum Arabic fractions probed by SANS and SAXS

Scattering spectra contain the contributions from all colloidal objects in solution. These objects may display different structures, concentrations and contrasts. It would thus seem preferable to analyze separately the different components of the gum. In practice, this is

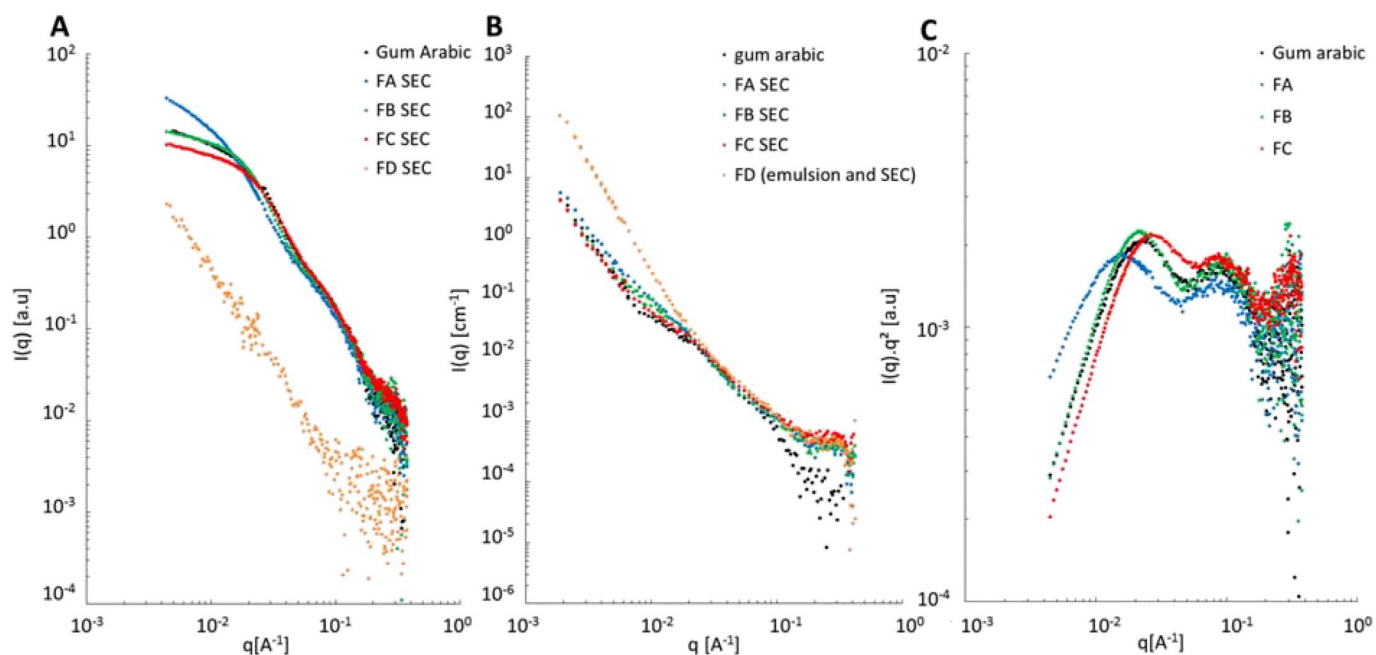


Fig. 6. (A) Small angle x-ray scattering spectra and (B) Small angle neutron scattering spectra of gum Arabic fractions separated by preparative size exclusion chromatography (FA represents the fraction with the higher molar mass and FD the smaller) measured at 5 g/L in 20 mM NaCl D<sub>2</sub>O solution (C) Kratky representation of SAXS spectra.

challenging since the gum is a continuum of sizes and hydrophobicity so that ideally a dual chromatographic separation should be implemented to collect fractions, which is hardly feasible in practice. Still, it is interesting to analyze relatively broad fractions separated from size exclusion chromatography. Fig. 6 displays SANS and SAXS spectra of fractions A, B, C and D collected through a size exclusion chromatographic separation (chromatograms in Fig. 1A and B). Strikingly, fractions A, B and C display SANS and SAXS spectra very similar to those of the whole gum Arabic in solution. The behavior of the three fractions SAXS spectra slightly differs in the low  $q$ -range: the scattered intensity of fraction A is higher than B, itself higher than C, following their size exclusion separation rank ( $M_{WA} > M_{WB} > M_{WC}$ ). This overall similarity between the fractions spectra is not trivial since it could be expected that species of different sizes would display different structures and also that the smallest species would display a finite size, which would be visible at a constant scattered intensity in the low- $q$  range. Such a finite size is observed in the SAXS spectra but not in the SANS spectra.

From the  $I \cdot q^2$  representation, we find a gyration radius of 11.6 nm for fraction A, 8.3 nm for fraction B and 6.4 nm for fraction C (Fig. 6C). Over the rest of the  $q$ -range, spectra are identical, notably the two other oscillations visible on SAXS spectra in the  $I \cdot q^2$  representation, which correspond to gyration radius of 2 and 0.7 nm. The signal-to-noise ratio and the resolution are too low in the SANS spectra to assess whether these oscillations are present or not. These oscillations could arise from several structural features: (i) several scattering populations (ii) a single scattering population displaying a multi-scale structure (iii) a single scattering population possessing a complex shape.

This similarity between gum Arabic and its largest fractions does not extend to its smallest fraction, fraction D, which exhibits a completely different SAXS spectrum. The scattered intensity scales with  $q^2$  over the  $q$ -range investigated. It is also much lower than for the other fractions spectra, suggesting a probable partial aggregation of fraction D species. Such an aggregation was visually observed in solutions of fraction D at higher concentrations. Interestingly, the corresponding SANS spectrum displays an opposite behavior with a much higher scattering intensity in the low  $q$ -range. Moreover, the SANS intensity for fraction D at low  $q$  scales as  $q^{-3.7}$  differing from the other fractions. This shows that the

fractal structure of aggregation can be influenced by the composition of gum Arabic. Fraction D corresponds to the smallest and most proteinaceous species of gum Arabic. The contrast between SANS and SAXS spectra of fraction D confirms the previously exposed hypothesis underlying the SANS/SAXS mismatch: SANS is mostly sensitive to polypeptide backbones that self-associate in solution, while SAXS is mostly sensitive to the hyperbranched polysaccharides groups that rather behave as porous colloids.

These recovered fractions were also characterized using circular dichroism measurements and compared with a native gum Arabic sample. Results are not presented here (see Fig. 6 in supporting information). It was observed that fractions B and C did not present an optical activity contrary to native gum Arabic and fraction A. The optical activity of fraction D could not be measured due to solubilization issues regarding this sample.

These observations are in good concordance with the protein content measured for each recovered fraction through a preparative size exclusion chromatography. Indeed, fraction A being richer in protein display a non negligible optical activity unlike the other fractions.

#### 4. Discussion

A dual chromatographic separation of gum Arabic unveiled several populations differing in size and amphiphilicity. Neutron scattering reveals an unresolved sub-micrometric polymeric network, while X-ray scattering highlights multi-scale nanometric porous colloids. We will now discuss the composition and colloidal structure of the different populations of gum Arabic, before addressing the overall structure of gum Arabic in solution.

##### 4.1. Gum Arabic populations: multiple combinations of hyperbranched polysaccharides with polypeptides

Two-dimensional chromatography shows that neither size nor hydrophobicity can be considered as truly efficient separation criteria. Each chromatographic fraction thus contains several hydrocolloid families. Still, in the literature, gum Arabic is commonly described as a mixture of three populations, based on a single chromatographic

separation, size exclusion or more frequently hydrophobic interaction chromatography. Furthermore, the terminology chosen is somehow ambiguous since it consists of three synonyms for covalent assemblies of polysaccharides and polypeptides, Arabinogalactan-peptide (AGp), Arabinogalactan-protein (AGP) and glycoproteins (GP). Nonetheless, this terminology points out a crucial parameter to sort out gum Arabic constituents, which is the polypeptide/polysaccharide ratio. Interestingly, recent literature studies emphasize the absence of free polysaccharides (Osman et al., 1993; Padala et al., 2009), which would not be bound to polypeptide chains, while older studies displayed the hypothesis of free polysaccharides units with the proteins part conjugated to the higher molecular weight fraction (Connolly, Fenyo, & Vandeveld, 1987; Randall et al., 1988). This consensus was established on the basis of a substantial protein content in each fraction separated either through SEC or HIC. However, the two dimensional analysis performed in this study reveals that in a protein-poor fraction such as fraction B, there is a coexistence of non-hydrophobic species with species exhibiting a broad spectrum of hydrophobicity. We thus argue that it is actually more likely that in such low-protein fractions, there is a coexistence between free polysaccharides and a small amount of polysaccharide/protein conjugates. Similarly, the definition of a glycoprotein fraction is actually impossible, since different amino-acid compositions in polypeptides are revealed by comparing the 210 nm absorption with the 280 nm absorption for instance. Also hydrophobic species are spread all over the size exclusion chromatogram as observed from the two-dimensional separation. This leads to the likely interpretation that the diverse populations in gum Arabic are built from polysaccharides and various polypeptide chains. We will thus now turn towards a more precise description of the main populations in terms of the polypeptide/polysaccharide ratio.

#### 4.2. Polysaccharide populations are made of multi-scale porous colloids

Both our SEC/HIC chromatogram and the HIC/SEC chromatogram of Renard and coworkers (Renard et al., 2006) indicate that two rather homogeneous fractions of intermediate size species exists, corresponding to the lowest hydrophobicity ((Ia) and (Ib) in Fig. 2). Furthermore, these fractions were shown to remain intact during enzymatic attacks of polypeptide chains (Goodrum et al., 2000). We know from SEC separation that these populations are found in protein-poor fractions and as explained above we argue that the protein content in such fractions actually corresponds to traces of other families. Fraction (Ia) corresponds to species of the lowest hydrophobicity, while (Ib) mainly includes smaller species displaying a slightly larger hydrophobicity. We lack the resolution to really distinguish (Ia) and (Ib) in terms of structural analysis and will thus from then discuss it as an average fraction (I) spreading all over the size spectrum. We observed that SAXS spectra of the whole gum and its fractions A, B and C were highly similar and interpreted it as the signature of hyperbranched polysaccharides, which we describe as a multi-scale porous colloid. SANS spectra evidenced on their part the role of proteins in establishing longer-range networks. We thus propose that population (I) would correspond to the hyperbranched polysaccharides, with no or little polypeptide chains. The average gyration radius of the whole gum estimated as 7 nm yields a spherical radius of 9 nm, matching previous measurements of hydrodynamic radius (Sanchez et al., 2008).

Since the ionic repulsion correlation peak falls in the same  $q$ -range close to overlap as the contact distance between two 7 nm units, we argue that these colloids bear the charge, consistently with the well-established presence of uronic acids in the periphery of hyperbranched polysaccharides (Islam et al., 1997).

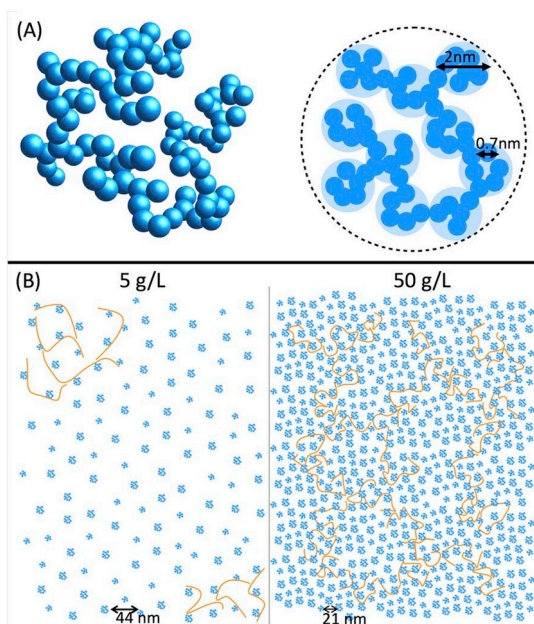
The SAXS intensity at small length-scales follows multiple power laws with exponents comprised between 2 and 2.5, which are separated by oscillations. Renard, Sanchez and co-workers argued in favor of two-dimensional objects, such as flat disks or tri-axial ellipsoids to describe these objects. However, their scattering curves are too noisy at large  $q$

values to distinguish oscillations and the two-dimensional shape was mainly supported by additional structural characterizations using techniques that bring the sample in contact with a surface (TEM, Cryo TEM, AFM). However, as acknowledged by the authors, these techniques may alter the structures of gum Arabic species. Flat disks or tri-axial ellipsoid models failed to fit our SAXS curves, which are more resolved than in Renard and coworkers studies (Renard et al., 2012; Sanchez et al., 2008) (see Fig. 5 in Supporting Information). In our opinion, the deviation from the Porod law, which states that for smooth interfaces the scattered intensity should decay as  $q^{-4}$ , is rather indicative of three-dimensional porous networks. Scattering from porous networks is commonly predicted using fractal models (Teixeira, 1988). For instance, the intensity scattered by a mass fractal of dimension 2.5 will follow a  $q^{-2.5}$  power law. The scattering spectra consisting of several oscillations separated by power law decays can thus describe a three dimensional porous network displaying multiple porosities. Such a three-dimensional porous structure is perfectly compatible with chemical characterizations of polysaccharides that have proposed a hyperbranched structure. For instance, the most recent chemical linkage model was proposed by Nie et al. using methylation and 2D NMR (Nie et al., 2013). Furthermore, a few additional arguments support the three-dimensional porous structure. First, hydrodynamic radii match exactly SAXS and SANS radii estimated from the radius of gyration of a spherical shape, while a 50% discrepancy was reported by Sanchez et al. with their flat disk model. Secondly, the gum concentration-dependent signature of the structure peak of gum Arabic, which stems from ionic repulsions, is consistent with a population of repulsive objects up to a 50 g/L concentration, with a shift of the peak position to higher  $q$  values and an increase of the peak height as the gum concentration is increased. However, further increase of the concentration leads to a broadening and decrease in peak height (Dror et al., 2006). This observation suggests that objects interpenetrate and reorganize, which seems difficult to understand with a two-dimensional structure. Furthermore, the 15–20 nm distance at overlap is more consistent with a spherical object.

#### 4.3. Several protein building blocks

Once we have isolated population (I) that contains no or little polypeptide chains, we are left with many populations or groups of populations that contain various but substantial amounts of polypeptides. Enzymatic degradation of both gum Arabic and gum Arabic intermediate HIC fraction, shows that the largest family of gum Arabic degrades into the two polysaccharide-rich families (Ia and Ib) (Connolly et al., 1987; Mahendran et al., 2008; Randall et al., 1989). Interestingly, a mild reducing attack as performed by Mahendran et al., leads to an increase of the polysaccharide-rich population, contrarily to what was stated in their paper, with a peak around 4–5.10<sup>4</sup> g mol<sup>-1</sup> (Mahendran et al., 2008). This peak corresponds in size to the highly proteinaceous fraction D of gum Arabic, as can be seen on our chromatograms. This result is consistent with the absence of this population after enzymatic treatment, since such a treatment is expected to cut accessible polypeptide chains, while these chains are expected to remain in the mild reducing treatment that mainly attacks o-linkages between peptides and sugars. We can thus conclude that polysaccharide/protein conjugates are made from distinct polysaccharide and protein blocks. We have described above the polysaccharide building block and now turn towards describing the protein building blocks. Fraction D contains the highest protein amount, more than 50 wt%. Despite its extremely small amount in gum Arabic, it is easily visible in the 210 nm absorption chromatogram, even more so in the 280 nm absorption chromatogram, consistently with a high protein content. The large differences between the 210 and 280 nm chromatograms in terms of peak intensities suggest that these polypeptide rich populations are highly heterogeneous in their amino-acid content. Rather than referring to a glycoprotein population, these observations indicate that several glycoprotein





**Fig. 7.** (A): Representation of gum Arabic polysaccharides structure in solution as a result of the SAXS measurements, (B): True scale schematic representation of gum Arabic macromolecules in solution as a function of concentration. Orange segments represent the glycoproteins moieties while the blue globular elements represent the branched arabinogalactan polysaccharides. Segments linked with polysaccharides represent the arabinogalactan-protein conjugates. The scheme respects the multiscale structure of gum Arabic species in solution for two concentrations: dilute at 5 g/L and close to overlap of AGp at 50 g/L. The scale is given by the correlation distances extracted from structure peak positions at low ionic strength. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

populations exist with different backbones. This casts some doubts about the generality of both palindromic pattern for gum Arabic proposed by Goodrum et al. and the two bands obtained after HF deglycosylation by Mahendran et al. (Goodrum et al., 2000; Mahendran et al., 2008).

The scattering spectrum of fraction D largely differs from the other fractions and rather correspond to polymeric networks. Notably, the signature of hyperbranched polysaccharides, the three oscillations, is not observed.

Importantly, this fraction (FD) and its narrower counterpart (F  $\delta$ ) behaved in an original fashion after being collected through SEC chromatography and freeze-drying. While the other fractions yielded a cloud-like powder after freeze-drying, this fraction was collected as a thin film, which could not be properly solubilized in water. This suggests a strong aggregation in concentrated conditions, which was also observed in SANS scattering curves of this fraction. This aggregation behavior suggests that glycoproteins backbones are at the core of amphiphilic properties of gum Arabic, through their hydrophobic amino-acids content.

Interestingly, the two-dimensional separation does not provide evidence of a specific population that would be abundant in fraction  $\delta$ . Indeed, it rather suggests a broad mixture of many other populations, which doesn't seem consistent with the high protein content that should yield high hydrophobicity. However, the aggregation behavior of fraction D or the narrowest fraction  $\delta$  may likely impact the chromatographic analysis, which always involves a 200 nm filtration prior to injection. We thus believe that some very hydrophobic aggregated species are missing in fraction  $\delta$  chromatogram.

#### 4.4. Polysaccharide/protein conjugates

Apart from fraction I, most fractions display a substantial hydrophobicity, which suggests they correspond to polysaccharide/protein conjugates of different sizes and compositions.

Fraction II is rather homogeneous and corresponds to intermediate sizes. It is not distinguishable from the main peak in the refractive index detection but corresponds to a clear peak in UV absorption at 210 and 280 nm ( $F\gamma$ : 135 kg mol<sup>-1</sup>, and FC). Since it is much more intense with respect to other fractions at 210 nm than at 280 nm, we can deduce it contains less aromatic amino-acids (tryptophan and tyrosine), which are hydrophobic, than average.

Fraction IV is also well-defined and corresponds to the largest species displaying a large hydrophobicity. Compared to population III that is broader and associated to smaller species, fraction IV also contains more aromatic amino-acids, as shown by comparing absorption at 210 and 280 nm. Fraction V seems bimodal and corresponds likely to a large diversity of different species.

#### 4.5. Polypeptide chains link the building blocks

SAXS and SANS spectra are very contrasted at large length scales due to a probable highlighting of polypeptide chains in SANS experiments. Since length scales up to 300 nm were probed without any signs of reaching a finite size, it is likely that the polypeptide network is composed of several chains. This would be consistent with the rheological behavior of gum Arabic and light scattering observations of self-association (Al-Assaf et al., 2009; Li et al., 2009, 2011; Wang et al., 2008). Two power law behaviors are observed depending on gum concentration. At low gum concentration (5 g/L with salt), for decreasing (resp. increasing)  $q$  (resp. length scales), we first observe an increase in  $q^{-1}$  in the intermediate- $q$  range, followed by an increase in  $q^{-3}$  (Fig. 3). At higher gum concentration (50 g/L no salt), we observe a more extended increase in  $q^{-2}$  (Fig. 3). Therefore, increasing the gum concentration leads to an increase of the compactness in the range 10–100 nm but a decrease of compactness in the range 100–300 nm. The low concentration behavior would be then consistent with collapsed networks build from rigid blocks, which would swell upon increasing their concentrations. The driving force must originate rather from polypeptide/solvent interactions rather than ionic repulsions since the slopes were unchanged upon screening the structure correlation peak. We can thus conclude that polypeptide chains structure solutions of gum Arabic at large length scales. These polypeptide chains are ubiquitous in gum Arabic, and only the less hydrophobic fraction collected in HIC chromatography does not seem to contain it. Fig. 7 displays a true scale scheme of the multi-scale structure of the different gum Arabic species in aqueous solutions under two concentration regimes.

## 5. Conclusion

Gum Arabic is a heterogeneous polysaccharide/protein natural blend, which cannot be separated into distinct families through a single chromatographic separation. Using a dual size/hydrophobicity chromatographic separation, it was demonstrated in this work that even narrow size fractions contained species with different hydrophobicity. The three categories classically described in the literature, arabino-galactan (AG), glycoprotein (GP) and arabino-galactan protein conjugates (AGP) should thus rather be thought of as a representation of gum Arabic as a mixture of species differing in their polysaccharide/protein ratio.

We performed an extensive neutron and X ray scattering characterization to present a structure of gum Arabic in solution, focusing respectively on the polypeptide backbone or hyper branched polysaccharide. The presented results are in agreement with a three-dimensional multi-scale porous structure for the hyperbranched

polysaccharides, rather than two-dimensional structures. The largest length scale is around 7 nm and contains two smaller length scales of 2 and 0.7 nm. These length scales separate two regimes of distinct porosity, associated with fractal dimensions of 2.1 and 2.5. Several protein populations are present and differ in their amino-acid content. In solution they organize themselves as Gaussian chains, which corresponds to a polymer in theta-solvent. Several populations of polysaccharide/protein conjugates were identified, largely differing in size and hydrophobicity. As a result, we can expect different functional properties for the different types of conjugates, rather than a single AGP behavior. Their structure is overall consistent with the wattle-blossom picture, with polysaccharides porous structures linked to polypeptide chains.

We will show in separate studies the importance of this more extensive characterization of gum Arabic heterogeneity and colloidal structure in problems involving interfaces, such as emulsion stabilization. However, an immediate conclusion of the present work is that the description of gum Arabic in solution requires only the distinction between polysaccharide-rich and protein-rich species. Indeed, its structure in solution is the juxtaposition of two other structures. One stems from the abundance of hyperbranched polysaccharides, which adopt the structure of a multi-scale porous spheroidal particles. Due to this structure, the resulting colloidal dispersion displays a much weaker viscosity than the corresponding linear polymer. This explains why gum Arabic can be solubilized in water up to very large concentrations and with a lower impact on the solution viscosity. Another structure results from the presence of polypeptide chains that adopt a Gaussian chain conformation in water. A dilute network is thus observed in solution, either in clusters at low gum concentrations or in the whole solution at higher gum concentrations. Interestingly, this dual structure is similar to the one of covalently-linked polysaccharide-protein conjugates but largely relies on weaker intermolecular interactions that encompass all species present.

## Acknowledgements

Authors would like to thank Agence Nationale de la Recherche for financial support of Laboratoire Commun SOPHY (Projet-ANR-14-LAB3-0011), a partnership between Laboratoire de Génie Chimique and CARAGUM Int.™ company.

Authors are grateful to Sabine HEINISCH for her help on the two dimensional chromatography experiments at the Institut des Sciences Analytiques in Lyon and to Isabelle Borget from Laboratoire de Coordination Chimique de Toulouse for the measurements of nitrogen content.

The Léon Brillouin laboratory (CEA Saclay) is acknowledged for providing SANS beamtime. The European Synchrotron Radiation Facility (ESRF, Grenoble, France) is acknowledged for providing SAXS beamtime on the ID02 instrument. The research federation FERMAT (Université de Toulouse, France) is acknowledged for providing access to the Xeuss 2.0 SAXS instrument (Xenocs) through the CPER IMATECBIO grant and Pierre Roblin from the Laboratoire de Génie Chimique de Toulouse for technical support.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodhyd.2019.01.033>.

## References

- Al-Assaf, S., Phillips, G. O., Aoki, H., & Sasaki, Y. (2007). Characterization and properties of Acacia senegal (L.) Willd. var. senegal with enhanced properties (Acacia (sen) SUPER GUM™): Part 1—Controlled maturation of Acacia senegal var. senegal to increase viscoelasticity, produce a hydrogel form and convert a poor into a good emulsifier. *Food Hydrocolloids*, 21(3), 319–328. <https://doi.org/10.1016/j.foodhyd.2006.04.011>.
- Al-Assaf, S. A., Phillips, G. O., & Williams, P. A. (2005). Studies on acacia exudate gums. Part I: the molecular weight of Acacia senegal gum exudate. *Food Hydrocolloids*, 19(4), 647–660. <https://doi.org/10.1016/j.foodhyd.2004.09.002>.
- Al-Assaf, S., Sakata, M., McKenna, C., Aoki, H., & Phillips, G. O. (2009). Molecular associations in acacia gums. *Structural Chemistry*, 20(2), 325. <https://doi.org/10.1007/s11224-009-9430-3>.
- Banc, A., Charbonneau, C., Dahesh, M., Appavou, M.-S., Fu, Z., Morel, M.-H., et al. (2016). Small angle neutron scattering contrast variation reveals heterogeneities of interactions in protein gels. *Soft Matter*, 12(24), 5340–5352. <https://doi.org/10.1039/C6SM00710D>.
- Brület, A., Lairez, D., Lapp, A., & Cotton, J.-P. (2007). Improvement of data treatment in small-angle neutron scattering. *Journal of Applied Crystallography*, 40(1), 165–177. <https://doi.org/10.1107/S0021889806051442>.
- Churms, S. C., Merrifield, E. H., & Stephen, A. M. (1983). Some new aspects of the molecular structure of Acacia senegal gum (gum arabic). *Carbohydrate Research*, 123(2), 267–279. [https://doi.org/10.1016/0008-6215\(83\)88483-3](https://doi.org/10.1016/0008-6215(83)88483-3).
- Connolly, S., Fenyo, J.-C., & Vandeveld, M.-C. (1987). Heterogeneity and homogeneity of an arabinogalactan-protein: Acacia senegal gum. *Food Hydrocolloids*, 1(5–6), 477–480. [https://doi.org/10.1016/S0268-005X\(87\)80045-0](https://doi.org/10.1016/S0268-005X(87)80045-0).
- Dror, Y., Cohen, Y., & Yerushalmi-Rozen, R. (2006). Structure of gum arabic in aqueous solution. *Journal of Polymer Science Part B: Polymer Physics*, 44(22), 3265–3271. <https://doi.org/10.1002/polb.20970>.
- Glatter, O. (2017). *Scattering Methods and their Application in Colloid and Interface Science*. Elsevier.
- Goodrum, L. J., Patel, A., Leykam, J. F., & Kieliszewski, M. J. (2000). Gum arabic glycoprotein contains glycomodules of both extensin and arabinogalactan-glycoproteins. *Phytochemistry*, 54(1), 99–106. [https://doi.org/10.1016/S0031-9422\(00\)00043-1](https://doi.org/10.1016/S0031-9422(00)00043-1).
- Idris, O. H. M., Williams, P. A., & Phillips, G. O. (1998). Characterisation of gum from Acacia senegal trees of different age and location using multidetection gel permeation chromatography. *Food Hydrocolloids*, 12(4), 379–388. [https://doi.org/10.1016/S0268-005X\(98\)00058-7](https://doi.org/10.1016/S0268-005X(98)00058-7).
- Islam, A. M., Phillips, G. O., Slijvo, A., Snowden, M. J., & Williams, P. A. (1997). A review of recent developments on the regulatory, structural and functional aspects of gum arabic. *Food Hydrocolloids*, 11(4), 493–505. [https://doi.org/10.1016/S0268-005X\(97\)80048-3](https://doi.org/10.1016/S0268-005X(97)80048-3).
- Jermny, M. (1962). Chromatography of acidic polysaccharides on Deae-cellulose. *Australian Journal of Biological Sciences*, 15(4), 787–792.
- Kato, T., Tokuya, T., & Takahashi, A. (1983). Comparison of poly(ethylene oxide), pullulan and dextran as polymer standards in aqueous gel chromatography. *Journal of Chromatography A*, 256, 61–69. [https://doi.org/10.1016/S0021-9673\(01\)88212-1](https://doi.org/10.1016/S0021-9673(01)88212-1).
- Kikhney, A. G., & Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters*, 589(19PartA), 2570–2577. <https://doi.org/10.1016/j.febslet.2015.08.027>.
- Kobayashi, M., Utsugi, H., & Matsuda, K. (1986). Intensive UV absorption of dextrans and its application to enzyme reactions. *Agricultural & Biological Chemistry*, 50(4), 1051–1053. <https://doi.org/10.1080/00021369.1986.10867514>.
- Kuipers, B. J. H., & Gruppen, H. (2007). Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography – mass spectrometry analysis. *Journal of Agricultural and Food Chemistry*, 55(14), 5445–5451. <https://doi.org/10.1021/jf070337f>.
- Lewis, B. A., & Smith, F. (1957). The heterogeneity of polysaccharides as revealed by electrophoresis on glass-fiber paper. *Journal of the American Chemical Society*, 79, 3929–3931.
- Li, X., Fang, Y., Al-Assaf, S., Phillips, G. O., Nishinari, K., & Zhang, H. (2009). Rheological study of gum arabic solutions: Interpretation based on molecular self-association. *Food Hydrocolloids*, 23(8), 2394–2402. <https://doi.org/10.1016/j.foodhyd.2009.06.018>.
- Li, X., Fang, Y., Zhang, H., Nishinari, K., Al-Assaf, S., & Phillips, G. O. (2011). Rheological properties of gum arabic solution: From Newtonianism to thixotropy. *Food Hydrocolloids*, 25(3), 293–298. <https://doi.org/10.1016/j.foodhyd.2010.06.006>.
- Mahendran, T., Williams, P. A., Phillips, G. O., Al-Assaf, S., & Baldwin, T. C. (2008). New Insights into the structural characteristics of the Arabinogalactan – Protein (AGP) fraction of gum arabic. *Journal of Agricultural and Food Chemistry*, 56(19), 9269–9276. <https://doi.org/10.1021/jf800849a>.
- Nie, S.-P., Wang, C., Cui, S. W., Wang, Q., Xie, M.-Y., & Phillips, G. O. (2013). A further amendment to the classical core structure of gum arabic (Acacia senegal). *Food Hydrocolloids*, 31(1), 42–48. <https://doi.org/10.1016/j.foodhyd.2012.09.014>.
- Osman, M. E., Menzies, A. R., Williams, P. A., & Phillips, G. O. (1994). Fractionation and characterization of gum arabic samples from various African countries. *Food Hydrocolloids*, 8(3–4), 233–242. [https://doi.org/10.1016/S0268-005X\(09\)80335-4](https://doi.org/10.1016/S0268-005X(09)80335-4).
- Osman, M. E., Menzies, A. R., Williams, P. A., Phillips, G. O., & Baldwin, T. C. (1993a). The molecular characterisation of the polysaccharide gum from Acacia senegal. *Carbohydrate Research*, 246(1), 303–318. [https://doi.org/10.1016/0008-6215\(93\)84042-5](https://doi.org/10.1016/0008-6215(93)84042-5).
- Osman, M. E., Williams, P. A., Menzies, A. R., & Phillips, G. O. (1993b). Characterization of commercial samples of gum arabic. *Journal of Agricultural and Food Chemistry*, 41(1), 71–77. <https://doi.org/10.1021/jf00025a016>.
- Padala, S. R., Williams, P. A., & Phillips, G. O. (2009). Adsorption of gum Arabic, egg white protein, and their mixtures at the oil – water interface in limonene oil-in-water emulsions. *Journal of Agricultural and Food Chemistry*, 57(11), 4964–4973. <https://doi.org/10.1021/jf803794n>.
- Qi, W., Fong, C., & Lampport, D. T. A. (1991). Gum Arabic glycoprotein is a twisted hairy rope a new model based on o-galactosylhydroxyproline as the polysaccharide attachment site. *Plant Physiology*, 96(3), 848–855. <https://doi.org/10.1104/pp.96.3.848>.
- Randall, R. C., Phillips, G. O., & Williams, P. A. (1988). The role of the proteinaceous



- component on the emulsifying properties of gum arabic. *Food Hydrocolloids*, 2(2), 131–140. [https://doi.org/10.1016/S0268-005X\(88\)80011-0](https://doi.org/10.1016/S0268-005X(88)80011-0).
- Randall, R. C., Phillips, G. O., & Williams, P. A. (1989). Fractionation and characterization of gum from Acacia senegal. *Food Hydrocolloids*, 3(1), 65–75. [https://doi.org/10.1016/S0268-005X\(89\)80034-7](https://doi.org/10.1016/S0268-005X(89)80034-7).
- Ray, A. K., Bird, P. B., Iacobucci, G. A., & Clark, B. C., Jr. (1995). Functionality of gum arabic. Fractionation, characterization and evaluation of gum fractions in citrus oil emulsions and model beverages. *Food Hydrocolloids*, 9(2), 123–131. [https://doi.org/10.1016/S0268-005X\(09\)80274-9](https://doi.org/10.1016/S0268-005X(09)80274-9).
- Receveur-Bréchet, V., & Durand, D. (2012). How random are intrinsically disordered proteins? a small angle S. *Ingenta Connect*, 13, 55–75.
- Renard, D., Garnier, C., Lapp, A., Schmitt, C., & Sanchez, C. (2012). Structure of arabinogalactan-protein from Acacia gum: From porous ellipsoids to supramolecular architectures. *Carbohydrate Polymers*, 90(1), 322–332. <https://doi.org/10.1016/j.carbpol.2012.05.046>.
- Renard, D., Lavenant-Gourgeon, L., Lapp, A., Nigen, M., & Sanchez, C. (2014). Enzymatic hydrolysis studies of arabinogalactan-protein structure from Acacia gum: The self-similarity hypothesis of assembly from a common building block. *Carbohydrate Polymers*, 112, 648–661. <https://doi.org/10.1016/j.carbpol.2014.06.041>.
- Renard, D., Lavenant-Gourgeon, L., Ralet, M.-C., & Sanchez, C. (2006). Acacia senegal Gum: continuum of molecular species differing by their protein to sugar ratio, molecular weight, and charges. *Biomacromolecules*, 7(9), 2637–2649. <https://doi.org/10.1021/bm060145j>.
- Renard, D., Lepvrier, E., Garnier, C., Roblin, P., Nigen, M., & Sanchez, C. (2014). Structure of glycoproteins from Acacia gum: An assembly of ring-like glycoproteins modules. *Carbohydrate Polymers*, 99, 736–747. <https://doi.org/10.1016/j.carbpol.2013.08.090>.
- Sanchez, C., Nigen, M., Mejia Tamayo, V., Doco, T., Williams, P., Amine, C., et al. (May 2018). Acacia gum: History of the future. *Food Hydrocolloids*, 78, 140–160. <https://doi.org/10.1016/j.foodhyd.2017.04.008>.
- Sanchez, C., Schmitt, C., Kolodziejczyk, E., Lapp, A., Gaillard, C., & Renard, D. (2008). The Acacia Gum Arabinogalactan Fraction is a thin oblate ellipsoid: a new model based on small-angle neutron scattering and Ab initio calculation. *Biophysical Journal*, 94(2), 629–639. <https://doi.org/10.1529/biophysj.107.109124>.
- Teixeira, J. (1988). Small-angle scattering by fractal systems. *Journal of Applied Crystallography*, 21(6), 781–785. <https://doi.org/10.1107/S0021889888000263>.
- Vandeveldel, M.-C., & Fenyo, J.-C. (1985). Macromolecular distribution of Acacia senegal gum (gum arabic) by size-exclusion chromatography. *Carbohydrate Polymers*, 5(4), 251–273. [https://doi.org/10.1016/0144-8617\(85\)90034-7](https://doi.org/10.1016/0144-8617(85)90034-7).
- Wang, Q., Burchard, W., Cui, S. W., Huang, X., & Phillips, G. O. (2008). Solution properties of conventional gum Arabic and a matured gum Arabic (Acacia (sen) SUPER GUM). *Biomacromolecules*, 9(4), 1163–1169. <https://doi.org/10.1021/bm7011696>.
- Williams, P. A., Phillips, G. O., & Stephen, A. M. (1990). Spectroscopic and molecular comparisons of three fractions from Acacia senegal gum. *Food Hydrocolloids*, 4(4), 305–311. [https://doi.org/10.1016/S0268-005X\(09\)80207-5](https://doi.org/10.1016/S0268-005X(09)80207-5).
- Zipkin, A. M., Wagner, M., McGrath, K., Brooks, A. S., & Lucas, P. W. (2014). An experimental study of hafting adhesives and the implications for compound tool technology. *PLoS One*, 9(11), e112560. <https://doi.org/10.1371/journal.pone.0112560>.