Carbohydrate Polymers 99 (2014) 736–747

Contents lists available at ScienceDirect

Carbohydrate Polymers

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

Structure of glycoproteins from Acacia gum: An assembly of ring-like glycoproteins modules

D. Renard [a,*], E. Lepvrier [b], C. Garnier [b], P. Roblin [c], M. Nigen [d], C. Sanchez [d]

[a] INRA, UR1268 Biopolymères Interactions Assemblages, F-44300 Nantes, France
[b] UMR CNRS 6290 Equipe TAF—Translation and Folding, Université de Rennes 1, Campus Beaulieu, F-35042 Rennes cedex, France
[c] Synchrotron Soleil, Saint-Aubin, F-91192 Paris, Gif-sur-Yvette, France
[d] UMR1208 Ingénierie des Agropolymères et Technologies Emergentes INRA-Montpellier SupAgro-CIRAD-Université Montpellier 2 Place Pierre Viala, F-34060 Montpellier, France

ARTICLE INFO

Article history:
Received 24 July 2013
Received in revised form 26 August 2013
Accepted 28 August 2013
Available online 4 September 2013

Keywords:
HPSEC–MALLS
Small-angle X-ray scattering
Synchrotron radiation circular dichroism
Transmission electron microscopy
Glycoprotein
Acacia gum

ABSTRACT

The glycoprotein (GP) molecular fraction structure of the gum exudate of *Acacia senegal* (gum Arabic) isolated from hydrophobic interaction chromatography was investigated using high-performance size exclusion chromatography–multi angle laser light scattering (HPSEC–MALLS), small angle X-ray scattering (SAXS), synchrotron radiation circular dichroism (SRCD) and transmission electron microscopy (TEM) observations. In solution, GP would be a mixture of spheroidal monomers and more anisotropic oligomers as suggested by the two exponent values found in the R_g vs. M_w relationship and TEM observations. The GP conformation probed by SAXS was ascribed to a thin object with a triaxial ellipsoid morphology, certainly attributed to GP oligomers. A 9 nm diameter particle was also identified by SAXS in agreement with the dimensions identified by TEM on single isolated ring-like structures. The GP oligomerization process, as probed by TEM, would be the result of ring-like subunits self-association. This self-association would lead to more linear or, sometimes, cyclised assembly. At the molecular level, GP fraction was found to have secondary structures mainly made of β-sheets and turns (64%) but also, to a lesser extent, made of polyproline II (PPII) and α-helices (19%). These features were characteristic of hydroxyprolin-rich glycoproteins with arabinosylated and arabinogalactan polysaccharide side chains grafted to the polypeptide backbone. The GP molecular fraction structure from Acacia gum would be an assembly of ring-like glycoproteins modules. These ring-like structures were certainly due to hydroxyproline (Hyp)–arabinogalactan (AG) subunits.

1. Introduction

Exudate gums are among the oldest natural gums being used as thickening and stabilizing agents. Exudate gums are secreted by trees during gummosis, this wound response results in a plastic sealant that plugs cracks in the bark. The gum drying, in contact with air and sunlight, leads to the formation of hard, glasslike lumps which can be easily collected (Verbecken, Dierckx, & Dewettinck, 2003). Acacia gum and one of its molecular fractions, arabinogalactan-protein (AGP) fraction, have remarkable emulsifying and flavour stabilizing properties for the wide industrial use of Arabic gum, especially for the food and soft drink industries (Chikamai, Banks, Anderson, & Weiping, 1996; Phillips, Takigami, & Takigami, 1996; Prakash & Mangino, 1990; Randall, Phillips, & Williams, 1988; Ray, Bird, Iacobucci, & Clark, 1995; Thevenet, 1995). It is a branched, neutral or slightly acidic, polysaccharidic complex obtained as a mixed calcium, magnesium and potassium salt. It is composed of a continuum of heteropolysaccharides rich in arabinose and galactose (PRAGs) (Idris, Williams, & Phillips, 1998). The main chain consists of 1,3-linked β-D-galactopyranosyl units. Side chains are composed of two to five 1,3-linked β-D-galactopyranosyl units, joined to the main chain by 1,6-linkages. Both the main and the side chains contain units of α-L-arabinofuranosyl, α-L-rhamnopyranosyl, β-D-glucuronopyranosyl, and 4-*O*-methyl-β-D-glucuronopyranosyl, the two latter being mostly end-units (Anderson & Stoddart, 1966). Acacia gum would be composed by 64 ramified 1,3-linked homogalactan symmetrically arranged sub-units, each of molecular mass 8000 g mol^{-1} (Churms, Merrifield, & Stephen, 1983).

Acacia gum contains about 2% of a polypeptide (Anderson, Bridgeman, Farquhar, & McNab, 1983). It is a highly heterogeneous material, because it has, on the one hand, a variation in monomer composition and/or in the linking and branching of the monomer units; and on the other hand, a variation in molecular masses

* Corresponding author.
E-mail address: drenard@nantes.inra.fr (D. Renard).

0144-8617/$ – see front matter.
http://dx.doi.org/10.1016/j.carbpol.2013.08.090

distribution. The consequence of this heterogeneity is reflected through the molecular species collected after the Acacia gum fractionation and the mode of both separation and detection used. In general, three main fractions can be isolated by hydrophobic interaction chromatography (Randall, Phillips, & Williams, 1989). Analysis of the fractions showed that each contained similar proportions of the various sugars and differed essentially in their molecular masses and protein contents. The bulk of the gum (88.3 wt% of the total) was shown to be comprised by an arabinogalactan-peptide (AG) fraction with a weight-average molecular weight of 2.86×10^5 g mol^{-1} and a low protein content (1.1 wt%). The second major fraction (10.4 wt% of the total) was identified as an arabinogalactan-protein (AGP) with a 1.86×10^6 g mol^{-1} molecular weight and contained a greater proportion of protein (9 wt%). The third minor fraction (1.3 wt%) will consist of one, or possibly two, glycoproteins (GP). One of the GP had a molecular weight of 2.95×10^5 g mol^{-1} and the highest protein content (24.6 wt%) (Renard, Lavenant-Gourgeon, Ralet, & Sanchez, 2006). The proteinaceous component of the first two fractions had similar amino acid distributions, with hydroxyproline and serine being the most abundant (Randall et al., 1989). The amino acid composition of the third fraction was significantly different with aspartic acid being the most abundant (Renard et al., 2006). As the three main fractions are precipitated by the Yariv reagent, they can all be classified as AGP type macromolecule.

From HPSEC–MALLS measurements and rheology, it was suggested that Acacia gum molecules displayed a random coil shape with some unusual rheological behaviour ascribed to interplay between surface and bulk rheological properties of Acacia gum dispersions (Sanchez, Renard, Robert, Schmitt, & Lefebvre, 2002). We showed that AG from Acacia gum was an oblate ellipsoid with a ~20 nm diameter and a ~1.5 nm thickness with an inner interspersed chains network (Sanchez et al., 2008).

The tertiary structure of the arabinogalactan-protein (AGP) molecular fraction has been described in terms of a wattle-blossom macromolecular assembly by virtue of which few (~5) discrete polysaccharide domains of $M_w \sim 2 \times 10^5$ g mol^{-1} are held together by a short peptide backbone chain (Fincher, Stone, & Clarke, 1983). More recently, Mahendran, Williams, Phillips, Al-Assaf, and Baldwin (2008) identified smaller carbohydrate blocks of $\sim 4.5 \times 10^4$ g mol^{-1} linked by *O*-serine and *O*-hydroxyproline residues to the polypeptide chain of approximately 250 amino acids in length. Therefore, the authors proposed a new model where the shape of the macromolecule would be a kind of spheroidal random coil consistent with the wattle-blossom structure. This spheroidal random coil would be composed of a folded polypeptide chain carrying large sugar blocks with a possible thin oblate ellipsoid morphology (Mahendran et al., 2008). An alternative model was proposed which described the Acacia gum AGP, with $M_w \sim 1.5 \times 10^6$ g mol^{-1}, as a statistical model with a fundamental 7×10^3 g mol^{-1} subunit, where polysaccharide side chains attach to a polypeptide backbone of probably more than 400 residues in a highly regular and ordered fashion (every 10 to 12 residues repetitive peptide units), forming a 'twisted hairy rope' of ~150 nm long, 5 nm diameter and an axial ratio of ~30:1 (Qi, Fong, & Lamport, 1991). This model would be consistent with the fact that large macromolecules, like AGPs, could migrate through a primary cell wall with a 4 to 5 nm porosity by reptation (Carpita, 1982).

We recently identified a triaxial ellipsoid conformation for AGP in solution corresponding to a distribution of spheroidal (10–40 nm) and more anisotropic (20–60 nm) flat particles with thicknesses below 3–5 nm (Renard, Garnier, Lapp, Schmitt, & Sanchez, 2012, 2013). All particles were porous supramolecular assemblies of smaller structural subunits with dimensions of about 2–10 nm. In addition, it was demonstrated that these building structural subunits were mainly branched chains and ring-like structures with diameters of about 1–5 nm. For the first time, we identified a linear chain with branched building blocks closely looking like the famous wattle-blossom model so many times hypothetically described in the literature as the structure of AGP from Acacia gum (Renard et al., 2013).

The objectives of this present study were to probe the structure of the GP molecular fraction using HPSEC–MALLS, small angle X-ray scattering (SAXS), circular dichroism and microscopic methods. To our best knowledge, no literature data exist on these glycoproteins. A better understanding of the glycoprotein fraction structure from Acacia gum should improve the knowledge of both biological and technological function/properties of this important class of macromolecules.

2. Materials and methods

2.1. Materials

Spray-dried Acacia gum (lot 97J716) from *Acacia senegal* trees was a gift from CNI company (Rouen, France). Before purification, Acacia gum was dialyzed against deionized water and then freeze-dried. All chemicals were of analytical grade and were obtained from Merck, Sigma and Aldrich.

2.2. Purification of GP molecular fraction

The Acacia gum glycoprotein (GP) molecular fraction was purified by Hydrophobic Interaction Chromatography (HIC), as previously described (Renard et al., 2006). Purified GP was dialysed and then freeze-dried before further characterization.

2.3. Preparation of GP dispersions

Known amounts of previously dialyzed and freeze-dried GP powder were dispersed in the appropriate solvent conditions (10 mM acetate buffer containing 100 mM NaCl or 50 mM NaCl or 50 mM $NaNO_3$) one night at room temperature under gentle stirring conditions. The dispersions were centrifuged 40 min at 16,000 g to remove air bubbles and undissolved material. After centrifugation, contrary to AG and AGP, a high quantity (~30 wt%) of undissolved material remained in GP dispersions. We presumed that it could be linked to the hydrophobic nature of the GP fraction and/or to its high propensity towards aggregation. The GP dispersions supernatants were then filtered through 0.2 μm Anotop membranes (Anotop, Alltech, France) for light scattering experiments.

2.4. HPSEC–MALLS

The determination of molecular weight and size distributions of GP were performed by coupling on line a high-performance size exclusion chromatography to a multi-angle laser light scattering detector, a differential refractometer and a differential viscosimeter. GP dispersions at 0.04 wt% into 50 mM $NaNO_3$ buffer containing 0.02% NaN_3 as preservative were filtered through 0.2 μm Anotop membranes (Anotop, Alltech, France), and injected on a HPSEC system constituted of a Shodex OH SB-G pre-column followed by two Shodex OH-pack 804 HQ and 805 HQ columns used in series. The samples were eluted at 0.7 mL min^{-1} with a 50 mM $NaNO_3$ solution. On-line molecular weight and intrinsic viscosity determinations were performed at room temperature using a multi-angle laser light scattering (MALLS) detector (mini-Dawn, Wyatt, Santa Barbara, CA, operating at three angles: 41°, 90° and 138°), a differential refractometer (ERC 7547 A) (dn/dc = 0.153 mL g^{-1}) and a differential viscosimeter (T-50A, Viscotek). Weight (M_w), number (M_n), average molecular weights (g mol^{-1}) and radius of gyration (R_g, nm) were

calculated using Astra 1.4 software (Wyatt). Intrinsic viscosity [η] was calculated using TriSEC software, version 3.0 (Viscotek).

2.5. Dynamic light scattering (DLS)

DLS experiments on filtrated GP (C = 0.3 wt%) dissolved in 50 mM $NaNO_3$ solution were performed at T = 25 °C (±0.1 °C) using an ALV-5000 multi-bit, multi-tau correlator in combination with a Malvern goniometer and a Spectra-Physics laser emitting vertically polarised light at 514.5 nm. The correlation functions obtained by dynamic light scattering (DLS) at θ = 30°, 90° and 150° scattering angle were analysed using the inverse Laplace transform routine REPES (Štepánek, 1993) in order to obtain the corresponding relaxation time distribution (and hydrodynamic radii distribution).

Only one population with a hydrodynamic radius (R_h) of 16.1 nm was determined for GP in solution whatever the scattering angle. However, small aggregates with a 300 nm R_h value were also identified at low scattering angle (less than 2% of the total GP sample).

2.6. SAXS experiments and data treatment

SAXS experiments on GP (C = 1.7 wt%) dissolved in 10 mM acetate buffer pH 5 containing 100 mM NaCl were performed on the SWING beamline at Synchrotron SOLEIL, Gif-sur-Yvette, France. The wavelength was set to λ = 1.033 Å. The 17 × 17 cm^2 low-noise Aviex CCD detector was positioned at 1000, 1800 and 5000 mm from the sample. The resulting exploitable q-range was 0.002–0.8 $Å^{-1}$, where $q = 4\pi \sin\theta/\lambda$, considering 2θ as the scattering angle. Samples passed through a thermostatted quartz capillary (T = 25 °C) with a 1.5 mm diameter and a 10 μm wall thickness positioned in a vacuum chamber. The sample volume was 20 μL and the flow through the capillary was 40 μL min^{-1}. Samples were separated from the pushing liquid (water) by two air volumes of 6 μL each, as described previously (David & Perez, 2009). A total of 15 frames of 2 s each were recorded. In all cases, the transmitted intensity was permanently measured with an accuracy of 0.1%, using a diode embedded in the beam-stop. The recorded curves were normalized to transmitted intensity and subsequently averaged using "Foxtrot", a dedicated homemade application. The same protocol was applied to buffer scattering.

The pair distance distribution function $P(r)$ together with the structural parameter derived from $P(r)$, i.e. the maximum dimension of the particle (D_{max}) and the radius of gyration R_g, were computed from the form factor by the indirect Fourier transform program GNOM (Svergun, Semenyuk, & Feigin, 1988; Svergun, 1992). Fit of experimental GP form factor by different theoretical form factors was assayed using the SansView software v2.0.1 (http://danse.chem.utk.edu/sansview.html).

2.7. SRCD

Previous circular dichroism spectrum on GP molecular fraction (Renard et al., 2006) performed on a CD 6 dichrograph instrument (Jobin-Yvon, Longjumeau, France) in the 185–260 nm far UV range, identified PPII, β-sheet and unordered secondary structures, using the self-consistent method (Sreerama & Woody, 1993). Synchrotron radiation circular dichroism was performed in the 170–280 nm wavelength range in order to get deeper insight into the secondary structures of the GP molecular fraction. The very low wavelength vacuum UV (VUV) data obtainable using synchrotron radiation as a bright light source considerably enhances the accuracy of secondary structures prediction by taking advantage of good signals below 180 nm (Whitmore & Wallace, 2008). SRCD spectra on GP (C = 0.325 wt%) were thus performed on the DISCO beamline at Soleil synchrotron (Gif-sur-Yvette, France). Spectra were recorded in far-UV region (170–280 nm) with 1 nm steps. Each spectrum was the average of four acquisitions and was corrected for the contribution of the buffer solution. The 263–270 nm region was set to zero, and resulting spectra were calibrated with d-10-camphorsulfonic acid, using the CDtool software (Lees, Smith, Wien, Miles, & Wallace, 2004). Spectra were normalized to protein concentrations. Secondary structures estimates were obtained after far-UV CD spectra were analysed using ContinLL program (Van Stokkum, Spoelder, Bloemendal, Van Grondelle, & Groen, 1990) in Dichroweb (Lobley, Whitmore, & Wallace, 2002) and using the protein data set SP175 as the reference set (Whitmore & Wallace, 2004, 2008). To date, SP175 database contains the largest set of spectra (72 soluble proteins) (Lees, Miles, Wien, &, Wallace, 2006). For all measurements, spectra were analyzed down to 175 nm, corresponding to a photomultiplier high tension remaining below half of its total variation.

2.8. TEM and image analyses

GP samples were prepared in the same conditions as those reported in Renard et al. (2012, 2013) for microscopy observations. 4 μL of GP samples (0.01–0.1 wt%) were applied on carbon-coated copper grids (300 mesh), during 30 to 120 s. Then, the grids were rinsed with filtered water and negatively stained for 60 s with 2% (w/v) uranyl acetate, and wicked dry prior to analysis using a 100 CXII transmission electron microscope from JEOL operating at accelerating voltages of 120 kV. Pictures obtained from TEM were treated using the ImageJ freeware v1.45o. Generally, image analyses consisted in capturing one object, changing the image contrast and filtering using a bandpass FFT filter. Sometimes an additional filtering step was made using the CLAE (Contrast Limited Adaptive Histogram Equalization) ImageJ plugin that enhances the local contrast of images.

3. Results

3.1. HPSEC–MALLS measurements

Previous data obtained using HPSEC–MALLS identified two main populations in the GP distribution from Acacia gum (peaks 1 and 2 in Fig. 1a) with weight average molecular weights (M_w) of 2.67×10^6 and 2.95×10^5 g mol^{-1}, radii of gyration R_g of 41.3 and 19.5 nm, polydispersity indices (M_w/M_n) of 1.13 and 1.01 and intrinsic viscosity [η] values of 41.3 and 19.5 mL g^{-1} for peaks 1 and 2, respectively (Renard et al., 2006). Another population with a broad size distribution located between peaks 1 and 2 was also previously described (Renard et al., 2006).

Dimensions extrapolated to the infinite dilution depend on the molecular weight (at a constant temperature) as a simple power law, for monodisperse polymers and for populations with a sufficiently high molecular weight, according to:

$$R_g = K_G M_w^{\nu_g} \quad (1)$$

$$R_h = K_h M_w^{\nu_h} \quad (2)$$

$$[\eta] = K_\alpha M_w^{\alpha} \quad (3)$$

where R_g, R_h, [η] and M_w are the z-average radius of gyration, the z-average hydrodynamic radius, the z-average intrinsic viscosity and the weight average molecular weight, respectively; ν_g, ν_h and α are the corresponding hydrodynamic coefficients; and K_G, K_h and K_α are the corresponding constants. The exponents remain universal in the context of these laws. The values of ν_g, ν_h and α are dependent upon polymer shape, temperature, and polymer–solvent interactions: $\nu_g = \nu_h = 0.33$, $\alpha = 0$ for a sphere, $\nu_g = \nu_h = 0.5$–0.6, $\alpha = 0.5$–0.8

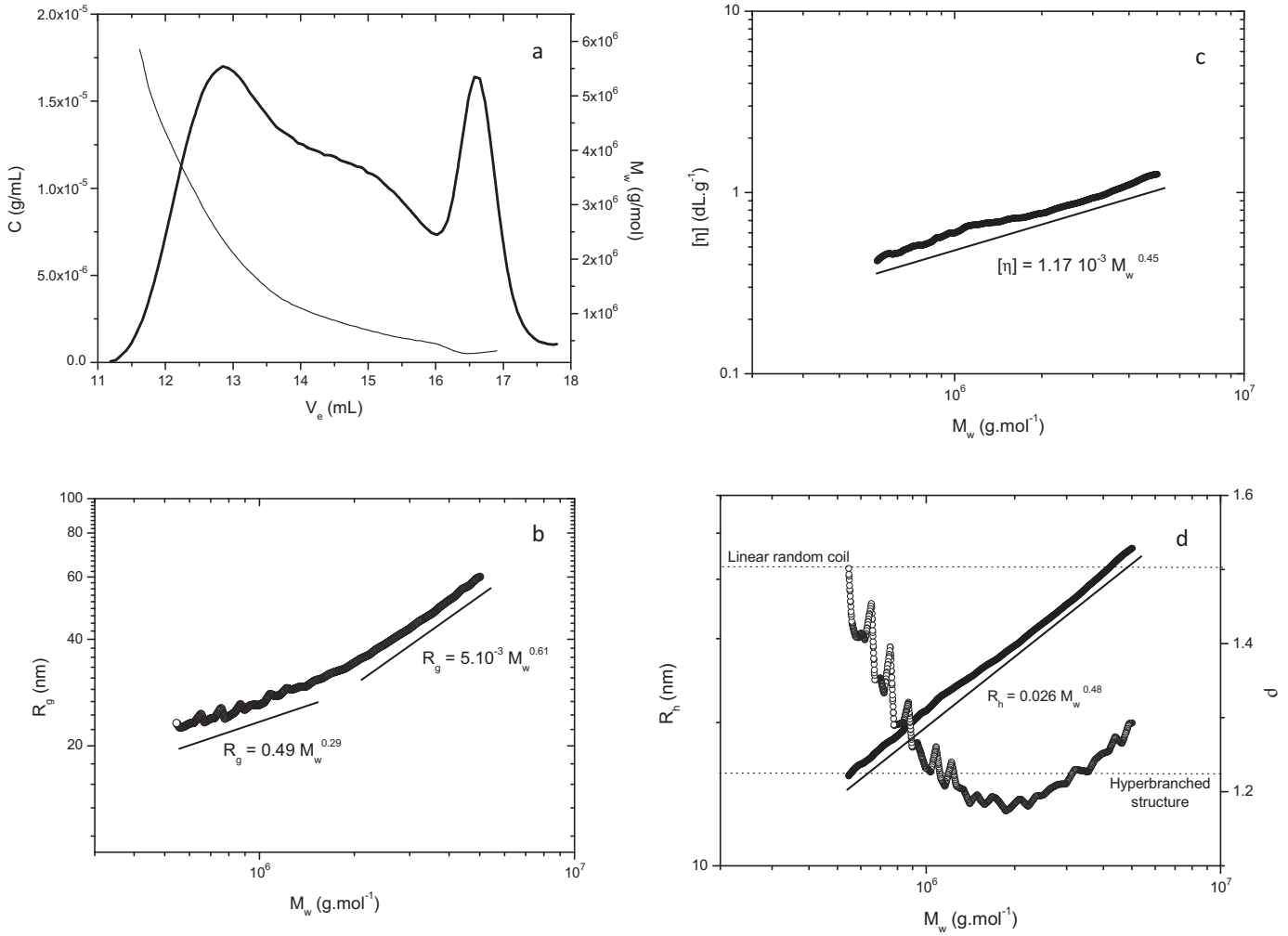

Fig. 1. (a) HPSEC chromatogram showing the elution profile monitored by refractive index (C, g/mL) and light scattering (M_w, g mol^{-1}) for Acacia gum glycoprotein (GP), (b) Radius of gyration (R_g, nm) (c) intrinsic viscosity ([η], dL/g) and (d) hydrodynamic radius (R_h, nm) and the ρ structure factor as a function of molecular weight (M_w, g mol^{-1}) obtained from HPSEC coupled to on-line viscosimeter and multi-angle light scattering (MALLS) detector; R_h was calculated from [η], results are displayed in (c).

for a linear random coil and $\nu_g = \nu_h = 1$, $\alpha = 1.8$ for a rod (Table 1) (Burchard, 1999). The pre-factors are constant and non-universal: they depend on both the detailed monomer structure and the solvent chosen. Relationships between R_g, [η], R_h vs. M_w were thus analysed taking into account the best signal to noise ratio of light scattering detector, i.e. size distribution ranging from $M_w \sim 5 \times 10^5$ g mol^{-1} to $M_w \sim 5 \times 10^6$ g mol^{-1} and corresponding to 73% of the total weight distribution (Fig. 1b–d). It is important to note that the second population (peak 2 in Fig. 1a), which may be ascribed to isolated GP macromolecules considering the very low M_w/M_n value. This population was excluded from these analyses due to the lack of sensitivity of the light detector. The log–log plot of R_g vs. M_w (Fig. 1b) exhibited two slopes with exponent values ν_g of 0.29 in the $\sim 5 \times 10^5$–1.6×10^6 g mol^{-1} M_w range and 0.61 in the ~ 1.6–5×10^6 g mol^{-1} M_w range. The existence of two experimental exponent values would mean that GP would adopt two conformations depending on the M_w distribution: a spheroidal shape for the intermediate M_w range (the lowest M_w range being excluded from the analysis) and a more linear conformation for the high M_w. These two possible conformations could be due to different morphologies

Table 1
Exponent (ν_g, α and ν_h) values were determined by the $R_G = f(M_w)$, $[\eta] = f(M_w)$ and $R_H = f(M_w)$ relationships, respectively, established from HPSEC–MALLS measurements for the glycoprotein (GP) molecular fraction from Acacia gum. The ρ ($\rho = R_g/R_h$) structure-sensitive ρ-parameter, determined by HPSEC–MALLS (1) and SAXS, HPSEC and DLS (2), were also reported. Theoretical values for different conformations were also given in comparison.

Theory					Experiment
Exponent	Sphere	Hyperbranched polymer	Linear random coil	Rod	
ν_g	0.3		0.5–0.6	1	0.29 0.61
α	0	<0.5	0.5–0.8	1.7–1.8	0.45
ν_h	0.3		0.5–0.6	1	0.48
ρ	0.778	1.225–1.534	1.504–1.780		1.5–1.3(1) 1.05–1.2(2)

(2) R_g (SAXS) = 17 nm; R_g (HPSEC) = 19.5 nm (third population, peak 2 in Fig. 1); R_h (DLS) = 16.1 nm.

or similar morphologies with different affinities for the solvent. An alternative hypothesis would be that isolated and self-associated GP macromolecules would be present in solution giving rise to two different conformations.

Further analyses using the Mark–Houwink–Sakurada (MHS) relationship vs. M_w should help to better define the GP conformation with increasing M_w as exponent values in this latter case would also reflect quality of the solvent. The $[\eta]$ vs. M_w log–log plot (Fig. 1c) showed one slope with an α exponent value of 0.45 with increasing M_w. This experimental α exponent value would be partly in agreement with a random coil conformation as the theoretical value of 0.5 for flexible chains was reported in theta solvent (Table 1) (Ross-Murphy, 1994). However, a better agreement of the experimental α exponent value would be with a hyperbranched polymer conformation for which theoretical $\alpha < 0.5$. The pre-factor in the $[\eta]$ vs. M_w relationship is in part related to the solvent polarity (Shen, Mu, Yu, & Chen, 2004). Moreover, for flexible chains, it typically ranges from 10^{-5} to 10^{-3} dL g^{-1} (Kawahigashi, Sumida, & Yamamoto, 2005). K value of 1.17 10^{-3} dL g^{-1} was obtained from linear fitting. The pre-factor and exponent value obtained from the $[\eta]$ vs. M_w relationship indicated that GP adopted a rather flexible chain conformation and that water was a moderately good solvent. A 0.47 exponent value was reported for total Acacia gum, the authors indicating that such a value would be indicative of a compact globular structure (Idris et al., 1998).

Using M_w and $[\eta]$, we calculated the GP hydrodynamic radius (R_h), assuming a spherical shape, using the Tanford equation $R_h = \left(3M_w[\eta]/10\pi N_a\right)^{1/3}$ (Tanford, 1961), with N_a the Avogadro's number. A linear relationship between R_h and M_w was found on a log–log plot representation (Fig. 1d) with an exponent value ν_h of 0.48 with increasing M_w. This value is in agreement with a random coil in θ solvent condition (Table 1) (Burchard, 1994). Again, the aqueous solvent chosen would not be a good solvent for GP macromolecules. A 0.42 value was previously found for total Acacia gum, in agreement with a moderately solvent affinity (Idris et al., 1998).

Molecular parameters and diluted solution properties can be gained from their R_g and R_h values for linear, hyperbranched, dendrimers and dendrimer-like synthetic polymers (Burchard, 1994, 1999). In a given polymer solution, the ratio of geometric to hydrodynamic radius ($\rho = R_g/R_h$) depends on its chain architecture and conformation. The ρ ratio is known to be affected by the macromolecular flexibility and polydispersity (Adolphi & Kulicke, 1997). Flexible linear polymers in a good solvent have ρ value ranging from 1.5 to 1.7 to up to >2, whereas ρ values are 0.778 for a homogeneous sphere and 1.225 for a hyperbranched polymer (Table 1). Fig. 1d showed the plot of the ρ value vs. M_w. The ρ values were not independent of M_w and decreased from ~1.5 to 1.3 with M_w. The decrease of ρ values with M_w was previously described for hyperbranched dextran (Rolland-Sabaté, Mendez-Montealvo, Colonna, & Planchot, 2008). For hyperbranched polymers, the theory states that the branching effect becomes noticeable at molar masses lower than 10^5 g mol^{-1} (branching degree increases the ρ-parameter in the low M_w region). Modifications of ρ values, in the low M_w region for GP, did not correspond to the branching effect, as predicted by the hyperbranched theory.

As observed in the chromatogram (Fig. 1a), GP was not completely fractionated by the HPSEC columns used in these experiments. Polydispersity is known to increase the z-average of the mean-square radius of gyration value and, thereby, ρ values (Burchard, 1999). Indeed, an increase of ρ values at low GP M_w was in agreement with this explanation. Considering the high M_w region, ρ values were in accordance with the ρ parameter of hyperbranched structures theoretically predicted to be 1.22 (in the Kirkwood approximation for the hydrodynamic interaction) (Burchard, 1999). A ρ value around 1 is also compatible with a two-dimensional object such as an oblate ellipsoid (He & Niemeyer, 2003; Matsuoka et al., 2006). A relevant example is amylopectin, a branched polysaccharide from starch, that displays an oblate ellipsoid shape (Lelievre, Lewis, & Marsden, 1986) and a ρ parameter around 1 (Bello-Pérez, Roger, Colonna, & Paredes-López, 1998). For GP, ρ values upper than 1, found in the low M_w range (5×10^5–1.6×10^6 g mol^{-1}), would be much more ascribed to polydispersity than to a linear random coil conformation. However, the increase of ρ values from 1.18 to 1.3 in the high M_w range (1.6×10^6–5×10^6 g mol^{-1}) would be in agreement with an increasing branching degree, as stated by the hyperbranched polymers theory.

From the 1.59×10^6 g mol^{-1} average GP M_w and the 75.8 mL g^{-1} $[\eta]$ value, previously determined (Renard et al., 2006), a 26.8 nm R_h value was calculated. This value was slightly higher than the 16.1 nm experimental R_h determined by DLS (Renard et al., 2006). The discrepancy observed between the calculated and experimental R_h values could be due to the presence of aggregated species eluted in the first peak of the HPSEC chromatogram (Fig. 1a) giving rise to higher M_w and $[\eta]$ values. Another hypothesis is that, in solution, GP would adopt conformations different from globular conformation (and this could be particularly true if aggregates are also present). A rough estimation from the average M_w of GP ($M_w = 1.6 \times 10^6$ g mol^{-1}), using a branched polysaccharide model (Protein Utilities software, Malvern Instruments) gave R_h = 19.7 nm, which is much more in agreement with the DLS result. However, globular and close-packed shape of Acacia gum molecules was suggested previously based on the low viscosity of gum solutions (Swenson, Kaustinen, Kaustinen, & Thompson, 1968; Anderson and Dea, 1971).

3.2. SAXS measurements

GP scattering form factor in 10 mM acetate buffer pH 5, 100 mM NaCl was obtained using SAXS. The absence of any correlation peak in the scattering profile demonstrated that 110 mM ionic strength was sufficient to suppress intermolecular repulsive interactions

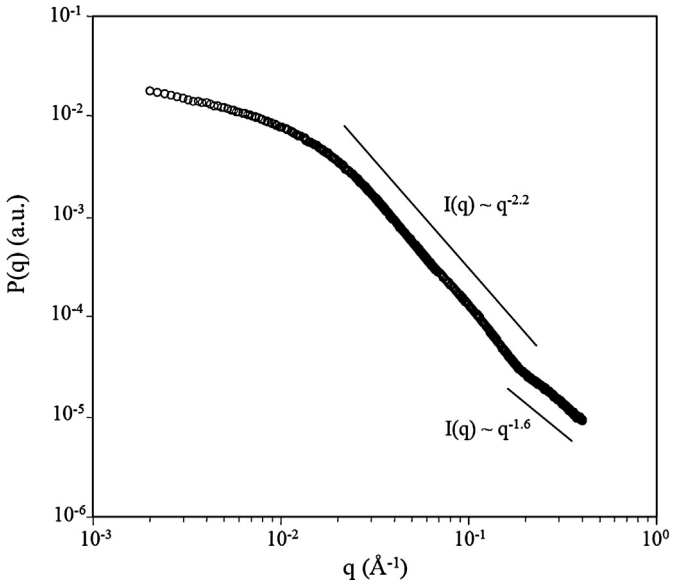

Fig. 2. Glycoprotein (GP) form factor from Acacia gum (1.7 wt%) obtained by SAXS at 25 °C in 10 mM acetate buffer pH 5, 100 mM NaCl. The GP radius of gyration (R_g) was 17 nm, as calculated from the Guinier region. From the power law exponent, in the intermediate q range, a 0.45 (1/2.2) GP affinity for the solvent was calculated (theta solvent). The exponent, in the high q range (1.6), indicated a local conformation intermediate between a rod and a freely jointed chain.

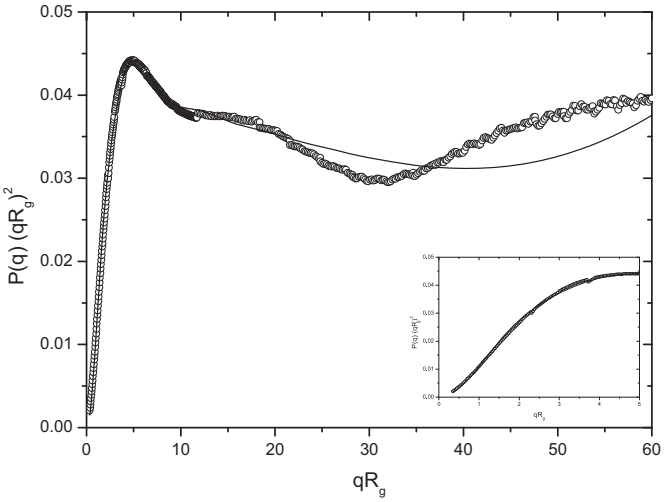

Fig. 3. Kratky-type plot using a 0.45 affinity for the solvent of the GP scattering curve (1.7 wt%) in 10 mM acetate buffer pH 5, 100 mM NaCl and theoretical triaxial ellipsoid form factor with 0.8, 8.2 and 26.9 nm (–) semi axes. Inset: Enlargement of the small qR_g range.

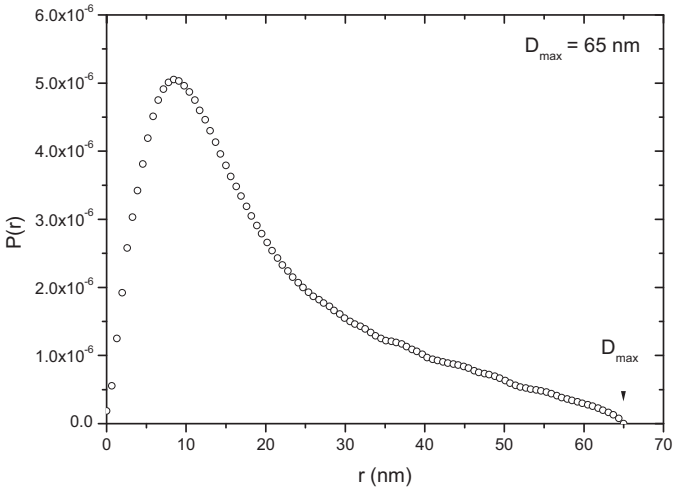

Fig. 4. Pair distance distribution function $P(r)$ calculated by indirect Fourier transform for the glycoprotein (GP) from Acacia gum form factor using GNOM software. A 65 nm maximum dimension for GP (D_{max}) was determined. The corresponding radius of gyration R_g was 17.2 nm.

(Fig. 2) and confirmed that GP was a weakly charged polyelectrolyte (Renard et al., 2006). Similar result was obtained by SANS on total Acacia gum (Dror, Cohen, & Yerushalmi-Rozen, 2006).

At low scattering wave vectors q, the GP radius of gyration (R_g) was calculated from the Guinier region as the condition $qR_g < 1$ was fulfilled in the q range explored. A R_g value of 17 nm was calculated. It was interesting to note that the R_g calculated from the third population in the HPSEC chromatogram (peak 2, Fig. 1a) was 19.5 nm. This value was very close to those found using SAXS. At intermediate scattering wave vectors q, the scattering function also followed a power law with an exponent value α of 2.2. A value of 2 is indicative of the presence of an isolated population of disk-like particles, and more generally of a 2D particles morphology (Svergun & Koch, 2003), or of a fractal structure. From the α exponent, the GP affinity ν for the solvent was estimated to be 0.45 ($\alpha = 1/\nu$), value corresponding to a theta solvent, intermediate between a poor affinity ($\nu = 0.33$) and a very good affinity ($\nu = 0.6$) for the solvent. This result was in total agreement with the solvent affinity found from the ν_h and α exponents determined by HPSEC coupled to on-line viscosimeter. In the high q range ($q > 0.1$ Å^{-1}), where the local structure of scattering objects is probed, scattering intensity followed a power law with an exponent of 1.6. This exponent value would be characteristic of an intermediate conformational state between a rod (q^{-1} power law) and a freely jointed chain (q^{-2} power law). The extended or flexible conformation probed at high q was confirmed by the Debye–Porod representation ($I(q) \times q^4$ vs. q) where no defined plateau was noticed (data not shown).

We previously demonstrated that the arabinogalactan-peptide fraction (AG) from Acacia gum could be described as a thin oblate ellipsoid (Sanchez et al., 2008) and that of the arabinogalactan-protein fraction (AGP) by a tri-axial ellipsoid form factor or an elliptical cylinder (Renard et al., 2012). Even if GP molecular fraction displayed at least two populations in the HPSEC chromatogram, a triaxial ellipsoid model was used to fit experimental SAXS data using semi-axes values of 0.8, 8.2 and 26.9 nm (Fig. 3). However, the fit was not completely satisfying, only the overall characteristic of the form factor was approached. It is important to note that the fit was better at small q with longer particle, emphasizing the possible existence of two particle populations as suggested by the R_g vs. M_w relationship (see Fig. 1b). It can be also noted that the theoretical form factor suggested a rather "two-dimensional" conformation concerning GP, as opposed to a more globular conformation. The GP possible conformation can be estimated as a first approximation from the shape of the pair distance distribution function $P(r)$, calculated from the form factor $P(q)$ by Indirect Fourier transform. The GP $P(r)$ function was displayed in Fig. 4. From this function, the maximum distance (D_{max}) found for GP was 65 nm. Note that this value was used to fit the experimental form factor by the tri-axial ellipsoid theoretical form factor. The $P(r)$ function shape suggested the presence of rather globular particles (i.e. maximum of the peak centred at ~10 nm) with flexible polymeric extension (i.e. long tail extending to distances r up to 65 nm) for a homogeneous solution of macromolecules (Glatter, 1982). However, HPSEC–MALLS chromatogram and R_g vs. M_w conformation plot (Fig. 1a and b) evidenced the coexistence of at least two different macromolecular GP populations in solution. Hence, the shape of $P(r)$ function could result from the scattering signals of both globular macromolecules and more elongated particles as associated GP macromolecules. In solution, the scattering of GP particles could thus come from two different electron density contrasts ascribed to two particle populations.

3.3. Far-UV circular dichroism

GP molecular fraction was identified to be the richest proteinaceous component of Acacia gum with 24.6% of protein (Renard et al., 2006). In addition, far-UV circular dichroism identified 9% PPII type conformation but also a significant amount of β-sheet and unordered (random coil) structure for GP in solution. We refined this secondary structure analyses performing far-UV circular dichroism in the VUV region using SRCD at Soleil (Gif-sur-Yvette, France). Far-UV circular dichroism spectra of the glycoprotein (GP) molecular fraction from Acacia gum were displayed in Fig. 5. Similar curves were obtained using classical dichrograph and synchrotron source available at Soleil. A strong minimum in the mean residue ellipticity at 205–208 nm and a weak maximum at 226 nm were partial signature of a polyproline II (PPII) conformation: a left-handed helix with three residues/turns and a pitch of 0.94 nm. A second minimum at 183 nm clearly appeared in the SRCD spectrum of GP molecular fraction, which is a signature of arabinogalactan polysaccharide substitution to the polypeptide backbone and characteristic of hydroxyproline-rich glycoproteins (HRGP) (Shpak, Barbar, Leykam, & Kieliszewski, 2001). The advantage of SRCD compared to classical dichrograph comes from the opportunity to enlarge the wavelength range and more particularly in the

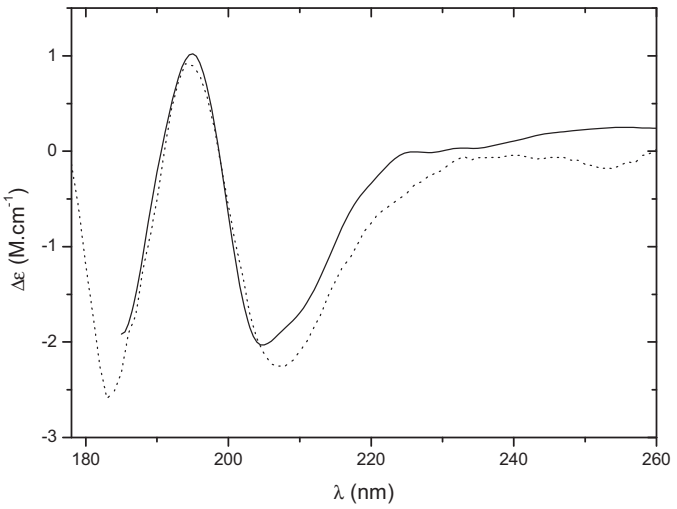

Fig. 5. Far-UV circular dichroism spectra of glycoprotein (GP) from Acacia gum. (·····) SRCD spectrum of GP (C = 0.325 wt%) in 10 mM acetate buffer pH 5, 100 mM NaCl; (–) CD spectrum of GP (C = 0.3 wt%) in water.

UV region considerably enhancing the reliability of the secondary structures prediction. Table 2 compared the secondary structures prediction obtained from the self-consistent method using classical CD (Sreerama & Woody, 1993; Deléage & Geourjon, 1993) and the ContinLL method using SRCD (Provencher & Glockner, 1981; Lobley, Whitmore & Wallace, 2002). PPII, turns and unordered structures were found to be equivalent whatever the method and algorithm used while a significant higher proportion of α-helix (9%) was identified by SRCD to the detriment of β-sheet structures. These original SRCD data obtained on GP in solution highlighted the importance of secondary structures in the conformational constraints necessary for the folding of the glycoproteins domains in the whole macromolecule.

3.4. TEM

Typical micrographs at three magnifications were shown in Fig. 6. These micrographs displayed moderately contrasted isolated particles with some ring-like morphology (white arrows, Fig. 6). The ring-like structures were often self-associated giving rise to linear or circular "string-of-rings" (black arrows, Fig. 6). Self-association of GP was in accordance with all the microscopy observations we previously performed on AG (Sanchez et al., 2008) and AGP (Renard et al., 2012, 2013). This confirmed again the high propensity of Acacia gum molecular fractions towards adhesion and therefore self-association on solid surfaces (i.e. carbon-coated copper grids). The GP remarkable surface properties again rendered the single particles observation difficult. However, from micrographs in Fig. 6, it was possible to highlight structural details of ring-like structures and self-associated ring-like structures enhancing the contrast of images (Figs. 7 and 8). All the identified isolated particles had a spheroidal shape while slight anisotropy appeared when ring-like structures self-associated. Contrary to what was

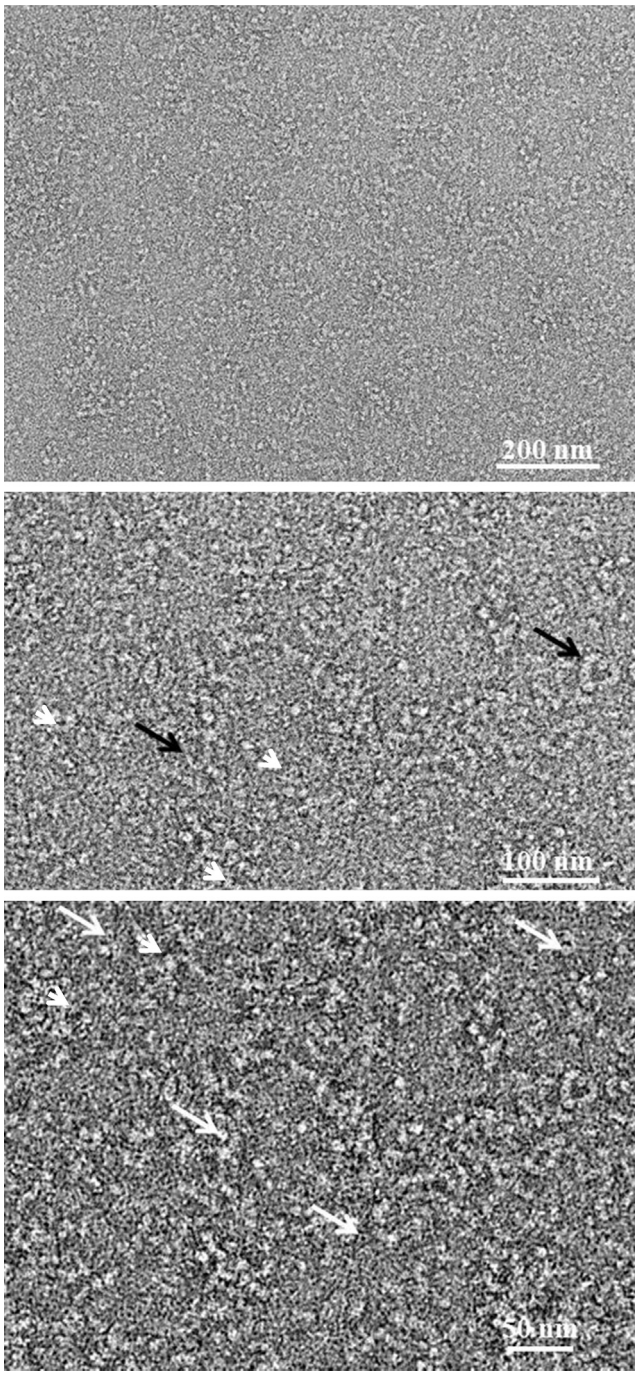

Fig. 6. Glycoprotein (GP) molecular fraction micrographs obtained by TEM at three magnifications dispersed at 1 wt% in 50 mM NaCl. White arrows: ring-like structures; black arrows: self-associated ring-like structures.

Table 2
Secondary structure percentages predicted from CD or SRCD. The SRCD glycoprotein (GP) molecular fraction spectra are displayed in Fig. 5.

	α-Helix	β-sheet	Turn	PPII	Unordered
SRCD (175–260 nm) ContinLL method[a,b]	9	38	26	10	17
CD (185–260 nm)Self-consistent method[c,d]	1	49	23	9	18

[a] Provencher and Glockner (1981).
[b] http://dichro web.cryst.bbk.ac.uk/); (Lobley et al., 2002; Whitmore & Wallace, 2008).
[c] Sreerama and Woody (1993).
[d] http://dic roprot-pbil.ibcp.fr/); (Deléage & Geourjon, 1993).

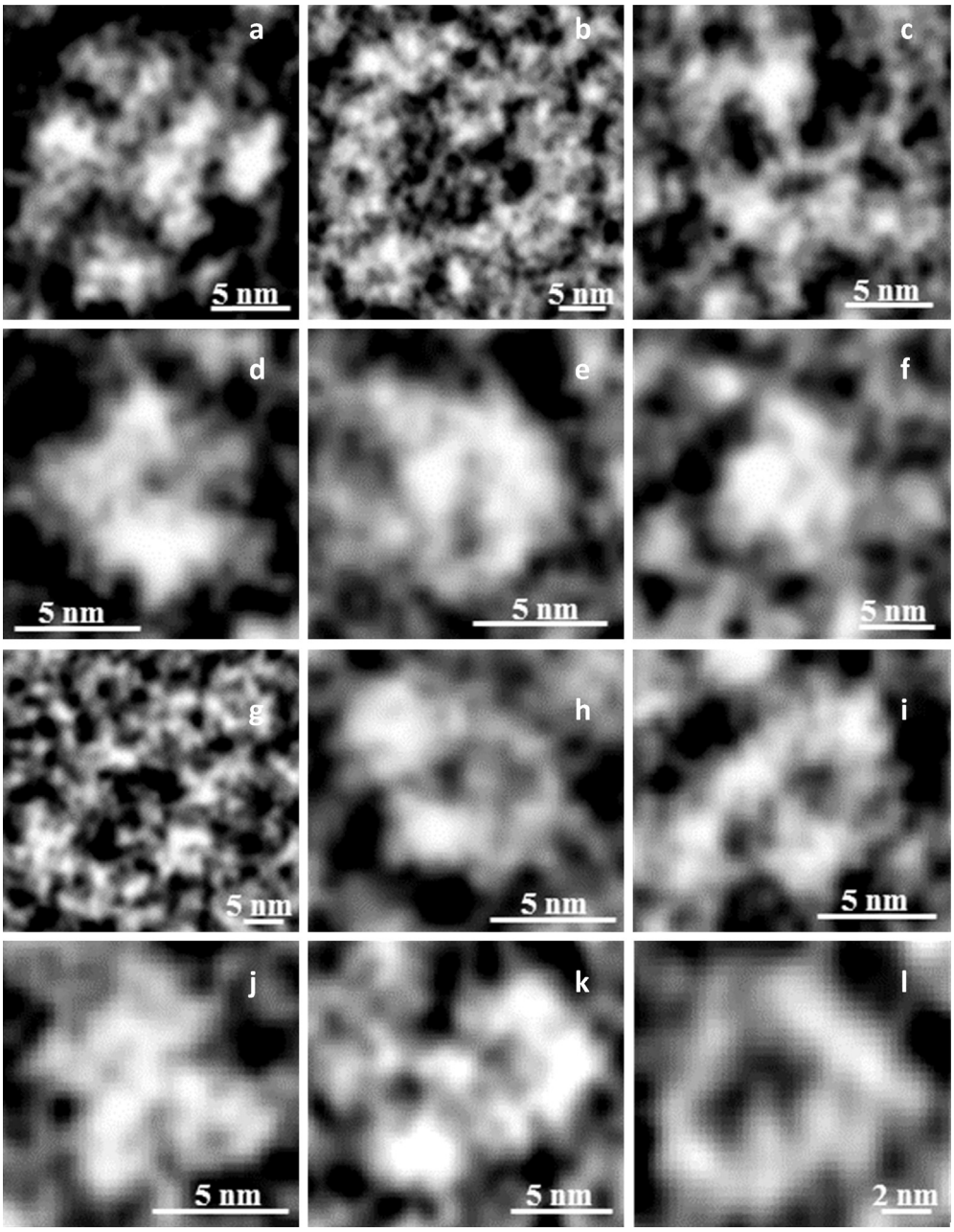

Fig. 7. Structural details of glycoprotein (GP) particles obtained from TEM micrographs dispersed at 1 wt% in 50 mM NaCl. After capturing one particle from raw micrographs (Fig. 6), image's contrast was modified when necessary, and images were filtered using a bandpass FFT filter. An additional filter was then made using the CLAE (Contrast Limited Adaptive Histogram Equalization) ImageJ plugin.

previously observed on AG and AGP, no outer structure combined to an inner porous network of interspersed chains was observed in the spheroidal particles morphology. These spheroidal particles were structurally made of an inhomogeneous outer thick shell (shell width ranging from 2 to 5 nm) and a central hole giving rise to the particles a typical ring-like morphology with a 8 to 11 nm diameters (Fig. 8a–c, g). The higher heterogeneous thickness observed on some ring-like structures could come from the condensation of smaller rings. It was interesting to note that, by SAXS, the maximum (i.e. peak) in the $P(r)$ function was located at 9 nm, a distance in agreement with the diameter values estimated using microscopy for single ring-like structures. Smaller (Fig. 8f) or higher (Fig. 8h) ring-like structures diameters were also observed. Self-associated ring-like structures also appeared forming a kind of interconnected network (Figs. 7b and g and 8i and k) highlighting again the active surface properties of GP molecular fraction from Acacia gum in particular when adsorbed onto solid surfaces.

4. Discussion

The glycoprotein (GP) molecular fraction from Acacia gum was found to be highly polydisperse after hydrophobic interaction chromatography (Fig. 1a) (Renard et al., 2006). Indeed, two main populations (denoted peaks 1 and 2, Fig. 1a) were identified separated by an ill-defined third population located between peaks 1 and 2. The integration of the whole chromatogram obtained by HPSEC–MALLS gave a 2.4 (M_w/M_n) size polydispersity value. This high size polydispersity value could explain the increase of the ρ-parameter values with a decreasing of M_w (Fig. 1d). To be much more relevant towards conformation analyses, it should

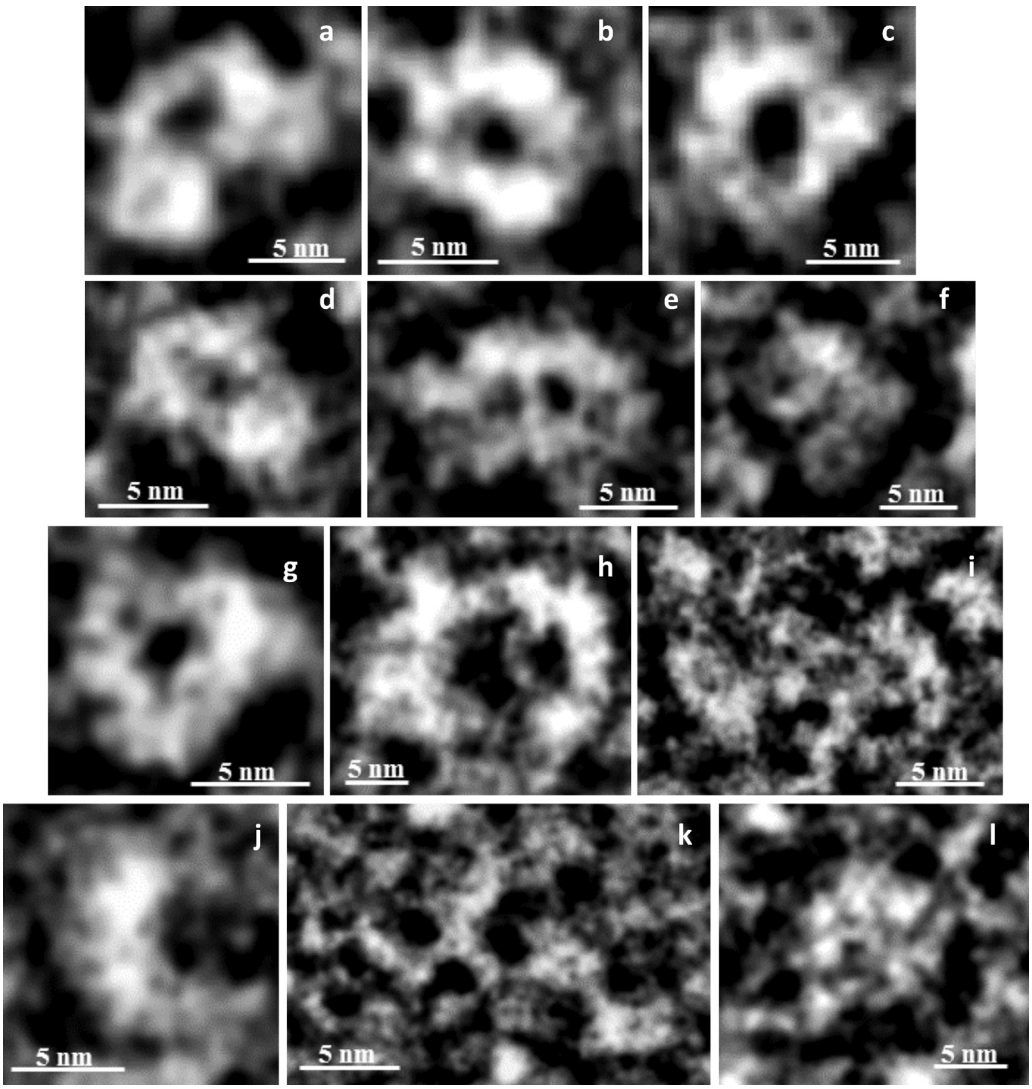

Fig. 8. Structural details, at high resolution, of GP particles captured from TEM micrographs dispersed at 1 wt% in 50 mM NaCl. Image analyses were performed as in Fig. 7.

be better to analyse each population in the chromatogram separately to get more reliable data. Therefore, each population had a very narrow size polydispersity with M_w/M_n values of 1.13 and 1.01 for peaks 1 and 2, respectively (Renard et al., 2006). From HPSEC–MALLS data analyses, it was uneasy to know whether these populations were isolated glycoproteins or a self-association phenomenon leading to aggregated glycoproteins. If we referred to literature data on hydroxyprolin-rich glycoproteins (HRGP), previous studies identified HRGP oligomerization, suggesting a highly adhesive property. These HRGPs were transmitting tissue-specific (TTS) protein, a stylar transmitting tissue AGP (Cheung, Wang & Wu, 1995; Wu, Wang & Cheung, 1995) and carrot AGP (Baldwin, McCann & Roberts, 1993). The TTS oligomerization was supposed to be driven by charge–charge interactions in a head-to-tail fashion between the *N* and *C* termini leading *in vitro* to cyclic "string-of-ellipsoids", ellipsoid being the shape of the TTS monomer (Cheung et al., 1995). In our case, M_w obtained for each HPSEC–MALLS population, and assuming that the second population would be the monomer (peak 2, Fig. 1a; $M_w = 2.95 \times 10^5$ g mol^{-1}), the oligomerization degree would be $N = 9$ for the first population (peak 1, Fig. 1a; $M_w = 2.67 \times 10^6$ g mol^{-1}) and $N \sim 3$ for the intermediate population ($M_w = 7.76 \times 10^5$ g mol^{-1}). This hypothesis of a self-association process taking place in the GP molecular fraction would come from the fact that the polydispersity indices M_w/M_n calculated from HPSEC–MALLS data for each population were very low. This calculation highlighted the presumably absence of heterogeneity coming from random self-association and/or molecular size polydispersity due to heterogeneous glycomodules branching on polypeptide backbone. In the particular case of the second population (peak 2, Fig. 1A), a 1.01 (M_w/M_n) value clearly indicated the absence of heterogeneity in the branching degree and/or its length. The presence of monomeric and oligomeric GP macromolecules could also explain the two ν_G exponent values found from the R_g vs. M_w relationship (Fig. 1b): a spheroidal shape for the smallest oligomers and a more linear conformation or compact structures with some linear branches for the biggest oligomers (Table 1). Deviation of the M_w dependence of R_g and therefore ν_G exponent for high M_w would be clearly indicative of short-chain branching for GP oligomers (Rolland-Sabaté et al., 2008). The short-chain branching hypothesis for the highest M_w population was confirmed by the calculation of the apparent density (d_{app}) according to the relation $d_{app} = M_w/N_a(4\pi/3)R_g{}^3$, where the increase of M_w was related to a decrease of d_{app} (data not shown). In addition, the short-chain branching was accompanied by an increasing branching degree in the high M_w region, as suggested by the increase of the ρ-parameter values with M_w (Fig. 1d). Therefore, it is important to remind that

the R_g vs. M_w relationship did not identify the dependence of the GP monomer subunit, the sensitivity of the light detector being too low to measure R_g values lower than ~10–15 nm (i.e. when there is no or little angular dependence of the scattered intensity) and M_w values lower than 2×10^5 g mol^{-1}. However, the GP oligomers hydrodynamic properties were found to be close to a linear random coil conformation as suggested by the unique α and ν_H exponent value found with 0.45 and 0.48 values, respectively, whatever M_w (Table 1). Further structural analyses using SAXS confirmed the more elongated shape for the whole GP macromolecule (i.e. corresponding to GP oligomers) with a tri-axial ellipsoid shape and a low thickness (1.6 nm). These hypotheses were confirmed by the fractal dimension values determined from the exponents of the R_g vs. M_w relationship with d_f values of 3.4 and 1.6 for the intermediate and high M_w population, respectively. These values were in agreement with globular (i.e. denser) and rather elongated shapes for the two different mass populations, respectively. However, the fractal dimension measured in the intermediate q range (10^{-2}–10^{-1} Å^{-1}) of the GP form factor, $d_f = 2.2$, was in accordance with those calculated from the α exponent in the $[\eta]$ vs. M_w relationship, $d_f = 2.07$ (with $d_f = 3/1 + \alpha$), highlighting again the more globular or denser shape of the GP molecular fraction at higher q range (i.e. corresponding to an observation scale close to the GP monomer). In addition, the GP monomer would have a local conformation between a rod and a linear random coil, as suggested by the 1.6 exponent value found from SAXS measurement at high q range. This value revealed the potential long-chain branched structures at the molecular scale. In addition, the presence of long-chain branches could be reflected in the form factor through the discrepancy observed for qR_g values >20 nm between the experimental and fitted data (Fig. 3). Whether these findings, and in particular the GP self-association mechanism, could be closely related to the moderate affinity of GP for an aqueous environment or to specific interactions between GP monomers remains unclear. The pair distance distribution function $P(r)$ (Fig. 4) gave also a useful information looking to the size distribution of GP oligomerization in solution, with r ~ 9 nm, corresponding certainly to GP monomer, and $D_{max} = 65$ nm, corresponding certainly to the size of the biggest oligomers. This distance of 9 nm, that could be attributed to GP monomer, was in well accordance with the diameter (8 to 11 nm) of particles measured from TEM micrographs (Figs. 7 and 8).

To summarize, GP monomer, with a rather globular shape and homogeneous long-chain branches, would be able to self-associate giving rise to small oligomers with a rather compact conformation and bigger oligomers with a more extended conformation closely related to the self-association mode.

The homogeneity or, at least, microheterogeneity of GP monomer's branches could be explained by the GP amino acid sequence, and the previously demonstrated hydroxyproline (Hyp) contiguity statement (Goodrum, Patel, Leykam, & Kieliszewski, 2000; Kieliszewski & Shpak, 2001). The Hyp contiguity hypothesis predicts that Hyp arabinosylation increases with Hyp contiguity and that clustered non-contiguous Hyp residues are sites of arabinogalactan polysaccharide addition in the AGPs and gums. In the case of the GP monomer from Acacia gum, with a very low molecular weight polydispersity, assuming that the Hyp contiguity hypothesis is valid for this glycoprotein, it implies that more non-contiguous Hyp residues would be present in the sequence giving rise to long and uniform arabinogalactosylated polysaccharide branches. The GP molecular fraction secondary structures probed by synchrotron radiation circular dichroism (SRCD) (Table 2, Fig. 5) identified a 19% helical fraction (α-helix and PPII) and a 17% random coil (i.e. unordered) fraction. The nearly equivalent proportion of helices and unordered structures could not unambiguously conclude to a higher proportion of contiguous or non-contiguous Hyp residues in the GP molecular fraction amino acid sequence. However, our hypothesis would be that the conformational constraints induced by the helices and turns secondary structures (accounted for 45% of the total of GP secondary structures), could be at the origin of the more homogeneous branching giving rise to a very low molecular weight polydispersity for each peak identified by HPSEC–MALLS. This hypothesis would mean that a steric hindrance would occur for the enzymes involved in the post-translational modifications of the polypeptide backbone leading to a more uniform grafting of oligoarabinosides and polysaccharides side chains onto the protein.

TEM observations revealed smaller objects appearing more spherical than larger ones that were somewhat more elongated or sometimes cyclised (Fig. 6). The smallest objects appeared as ring-like structures while the biggest ones appeared as a ring-like subunits assembly. These conclusions came from the structural details observed in Figs. 7 and 8 where individual ring-like structures with a 10 nm average diameter coexisted with larger objects coming from the ring-like subunits self-association. However, in this case, it was very difficult to estimate the oligomeric species diameters as these structures were often connected to each other, giving rise to a sort of interspersed network. The biggest diameter of 65 nm probed by SAXS for GP in solution were thus very difficult to confirm by TEM probably due to adsorption and as a consequence of self-association phenomena occurring during samples preparation. However, the two types of structures observed by TEM were in agreement with the two ν_G exponents found by HPSEC–MALLS and confirmed the existence of two morphologies ascribed to more compact and spheroidal structures (i.e. rings) together with more anisotropic and elongated structures (i.e. assembly of rings). Ring-like structures with smaller diameters (1–5 nm) were previously observed in AGP molecular fraction (Renard et al., 2012, 2013) and were ascribed to the ring-like structures of the galactan backbone calculated by molecular modelling in a Hyp-AG subunit consisting of 15-sugar residues attached to one Hyp (Lamport & Várnai, 2013). As suggested by the authors, these structural features obtained on a small arabinogalactan-peptide could be conserved on larger Hyp-AGs.

5. Conclusion

The structure of the glycoprotein (GP) molecular fraction of the gum exudate from *A. senegal* (gum Arabic) isolated using HIC was identified to be a mixture of spheroidal ring-like monomers with presumably homogeneous long branches and more anisotropic oligomers resulting from monomers self-association as suggested by HPSEC–MALLS and SAXS analyses together with TEM observations. The conformation of GP in solution probed by SAXS was ascribed to a thin particle with a triaxial ellipsoid morphology certainly ascribed to GP oligomers. Therefore, a particle with a 9 nm diameter was identified by SAXS in agreement with the dimensions identified by TEM on single isolated ring-like structures. GP oligomerization, as probed using TEM, would be the result of ring-like subunits self-association forming more linear or sometimes cyclised assemblies. At the molecular level, GP fraction was found to have secondary structures mainly made of β-sheets and turns but also, to a lesser extent, PPII and α-helices. These features were characteristic of HRGP with arabinosylated and arabinogalactan polysaccharide side chains grafted to the polypeptide backbone. The structure of GP would thus be an assembly of ring-like glycoproteins modules, the ring-like structures being certainly ascribed to hydroxyproline (Hyp)-arabinogalactan (AG) subunits. Table 3 summarized the main structural and hydrodynamic parameters determined in this study for the GP molecular fraction from Acacia gum.

Table 3
Summary of the glycoprotein (GP) molecular fraction from Acacia gum structural parameters determined by SAXS, DLS, HPSEC–MALLS, or calculated.

Structural parameters	Values
M_w (g mol^{-1})	1.59×10^6
M_w/M_n	2.4
D_{max} (nm)	65
R_g (nm)	19.1[a], 17[b]
R_h (nm)	16.1[c]
Translational diffusion coefficient D (10^7 cm^2 s^{-1})	1.36
$[\eta]$ (mL g^{-1})	75.8
Hydration parameter δ (g g^{-1})	4[d]
Hydrated volume (mL)	15.9×10^{-18}
Flory–Fox parameter Φ	15.5×10^{23}[e]
Solvent affinity	0.45[f]
Fractal dimension D_f	2.2[g], 2.07[h]

[a] From HPSEC–MALLS (third population).
[b] From SAXS.
[c] Determined by dynamic light scattering at three scattering angles (30°, 90° and 150°).
[d] Taken from Phillips et al. (1996).
[e] From the Flory–Fox equation $\Phi = [\eta]M_w/6^{3/2}(R_g)^3$.
[f] (1/2.2) power law in the intermediate q range of $P(q)$.
[g] Determined in the intermediate q range from the power law $I(q) \sim q_f^{-D}$.
[h] Determined from $[\eta] = KM_w{}^{\alpha}$ in the low M_w range with $d_f = 3/1 + \alpha$.

Acknowledgements

We would like to acknowledge L. Lavenant-Gourgeon for glycoprotein (GP) purification and characterization, M.-J. Crepeau and M.-C. Ralet for HPSEC–MALLS measurements and data treatments. Special thanks to Pierre Briozzo and Y. Gohon for their technical assistance during SRCD measurements.

References

Adolphi, U., & Kulicke, W.-M. (1997). Coil dimensions and conformation of macromolecules in aqueous media from flow field-flow fractionation/multi-angle laser light scattering illustrated by studies on pullulan. *Polymer*, *38*, 1513–1519.
Anderson, D. M. W., & Stoddart, J. F. (1966). Studies on uronic acid materials. *Carbohydrate Research*, *2*, 104–114.
Anderson, D. M. W., & Dea, I. C. M. (1971). Recent advances in the chemistry of Acacia gums. *Journal of Society of Cosmetic Chemistry*, *22*, 61–76.
Anderson, D. M. W., Bridgeman, M. M. E., Farquhar, J. G. K., & McNab, C. G. A. (1983). The chemical characterization of the test article used in toxicological studies of gum Arabic (*Acacia senegal* (L.) Willd). *International Tree Crops Journal*, *2*, 245–254.
Baldwin, T. C., McCann, M. C., & Roberts, K. (1993). A novel hydroxyproline-deficient arabinogalactan protein secretd by suspension-cultured cells of *Daucus carota* (purification and partial characterization). *Plant Physiology*, *103*, 115–123.
Bello-Pérez, L. A., Roger, P., Colonna, P., & Paredes-López, O. (1998). Laser light scattering of high amylose and high amylopectin materials stability in water after microwave dispersion. *Carbohydrate Polymers*, *37*, 383–394.
Burchard, W. (1994). Light scattering. In S. B. Ross-Murphy (Ed.), *Physical techniques for the study of food biopolymers* (pp. 343–392). London: Blackie Academic & Professional.
Burchard, W. (1999). Solution properties of branched macromolecules. *Advances in Polymer Science*, *143*, 113–194.
Carpita, N. (1982). Limiting diameters of pores and the surface structure of plant cell walls. *Science*, *218*, 813–814.
Cheung, A. Y., Wang, H., & Wu, H.-M. (1995). A floral transmission tissue-specific glycoprotein attracts pollen tubes and stimulates their growth. *Cell*, *82*, 383–393.
Chikamai, B. N., Banks, W. B., Anderson, D. M. W., & Weiping, W. (1996). Processing of gum Arabic and some new opportunities. *Food Hydrocolloids*, *10*, 309–316.
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*, 267–279.
David, G., & Perez, J. (2009). Combined sampler robot and high-performance liquid chromatography: A fully automated system for biological small-angle X-ray scattering experiments at the synchrotron SOLEIL SWING beamline. *Journal of Applied Crystallography*, *42*, 892–900.
Deléage, G., & Geourjon, C. (1993). An interactive graphic program for calculating the secondary structure content of proteins from circular dichroism spectrum. *Computer Applied Biosciences*, *2*, 197–199.
Dror, Y., Cohen, Y., & Yerushalmi-Rozen, R. (2006). Structure of gum Arabic in aqueous solutions. *Journal of Polymer Science Part B: Polymer Physics*, *44*, 2365–3271.
Fincher, G. B., Stone, B. A., & Clarke, A. E. (1983). Arabinogalactan-proteins: Structure biosynthesis and function. *Annual Review of Plant Physiology*, *34*, 47–70.
Glatter, O. (1982). Interpretation. In O. Glatter, & O. Kratky (Eds.), *Small angle X-ray scattering* (pp. 167–196). London: Academic Press Inc.
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*, 99–106.
He, L., & Niemeyer, B. (2003). A novel correlation for protein diffusion coefficients based on molecular weight and radius of gyration. *Biotechnological Progress*, *19*, 544–548.
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*, 379–388.
Kawahigashi, M., Sumida, H., & Yamamoto, K. (2005). Size and shape of soil humic acids estimated by viscosity and molecular weight. *Journal of Colloid and Interface Science*, *284*, 463–469.
Kieliszewski, M. J., & Shpak, E. (2001). Synthetic genes for the elucidation of glycosylation codes for arabinogalactan-proteins and other hydroxyproline-rich glycoproteins. *Cellular and Molecular Life Sciences*, *58*, 1386–1398.
Lamport, D. T. A., & Várnai, P. (2013). Periplasmic arabinogalactan glycoproteins act as a calcium regulator that regulates plant growth and development. *New Phytologist*, *197*, 58–64.
Lees, J. G., Smith, B. R., Wien, F., Miles, A. J., & Wallace, B. A. (2004). CDtool-an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Analytical Biochemistry*, *332*, 285–289.
Lees, J. G., Miles, A. J., Wien, F., & Wallace, B. A. (2006). A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics*, *22*, 1955–1962.
Lelievre, J., Lewis, J. A., & Marsden, K. (1986). The size and shape of amylopectin: A study using analytical ultracentrifugation. *Carbohydrate Research*, *153*, 195–203.
Lobley, A., Whitmore, L., & Wallace, B. A. (2002). DICHROWEB: An interactive website for the analysis of protein secondary structure from circular dichroism data. *Bioinformatics*, *18*, 211–212.
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*, 9269–9276.
Matsuoka, K., Yonekawa, A., Ishii, M., Honda, C., Endo, K., Moroi, Y., Abe, Y., & Tamura, T. (2006). Micellar size, shape and counterion binding of *N*-(1,1-dihydroperfluoroalkyl)-*N*,*N*,*N*-trimethylammonium chloride in aqueous solutions. *Colloid and Polymer Science*, *285*, 323–330.
Phillips, G. O., Takigami, S., & Takigami, M. (1996). Hydration characteristics of the gum exudate from *Acacia senegal*. *Food Hydrocolloids*, *10*, 11–19.
Prakash, A., & Mangino, M. E. (1990). The effects of added proteins on the functionality of gum Arabic in soft drink emulsion systems. *Food Hydrocolloids*, *4*, 177–184.
Provencher, S. W., & Glockner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. *Biochemistry*, *20*, 33–37.
Qi, W., Fong, C., & Lamport, 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*, 848–855.
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*, 131–140.
Randall, R. C., Phillips, G. O., & Williams, P. A. (1989). Fractionation and characterization of gum from *Acacia senegal*. *Food Hydrocolloids*, *3*, 65–75.
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*, 123–131.
Renard, D., Lavenant-Gourgeon, L., Ralet, M.-C., & Sanchez, C. (2006). *Acacia senegal* gum: Continuum of molecular species differing by their protein to sugar ratio, molecular weight and charges. *Biomacromolecules*, *7*, 2637–2649.
Renard, D., Garnier, C., Lapp, A., Schmitt, C., & Sanchez, C. (2012). Structure of arabinogalactan-protein from Acacia gum: From porous ellipsoids to supramolecular architectures. *Carbohydrate Polymers*, *90*, 322–332.
Renard, D., Garnier, C., Lapp, A., Schmitt, C., & Sanchez, C. (2013). Corrigendum to structure of arabinogalactan-protein from Acacia gum: From porous ellipsoids to supramolecular architectures. *Carbohydrate Polymers*, *97*, 864–867.
Rolland-Sabaté, A., Mendez-Montealvo, M. G., Colonna, P., & Planchot, V. (2008). Online determination of structural properties and observation of deviations from power law behavior. *Biomacromolecules*, *9*, 1719–1730.
Ross-Murphy, S. B. (1994). Rheological methods. In S. B. Ross-Murphy (Ed.), *Physical techniques for the study of food biopolymers* (pp. 343–392). London: Blackie Academic & Professional.
Sanchez, C., Renard, D., Robert, P., Schmitt, C., & Lefebvre, J. (2002). Structure and rheological properties of acacia gum dispersions. *Food Hydrocolloids*, *16*, 257–267.
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*, 629–639.
Shen, Q., Mu, D., Yu, L.-W., & Chen, L. (2004). A simplified approach for evaluation of the polarity parameters for polymer using the *K* coefficient of the Mark–Houwink–Sakurada equation. *Journal of Colloid and Interface Science*, *275*, 30–34.
Shpak, E., Barbar, E., Leykam, J. F., & Kieliszewski, M. J. (2001). Contiguous hydroxyproline residues direct hydroxyproline arabinosylation in *Nicotiana tabacum*. *Journal of Biological Chemistry*, 11272–11278.

Sreerama, N., & Woody, R. W. (1993). A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Analytical Biochemistry*, *209*, 32–44.

Štepánek, P. (1993). Data analysis in dynamic light scattering. In W. Brown (Ed.), *Dynamic light scattering: The methods and some applications* (pp. 177–241). Oxford: Clarendon Press.

Svergun, D. I., Semenyuk, A. V., & Feigin, L. A. (1988). Small-angle-scattering-data treatment by the regularization method. *Acta Crystallographica*, *A44*, 244–250.

Svergun, D. I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography*, *25*, 495–503.

Svergun, D. I., & Koch, M. H. J. (2003). Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics*, *66*, 1735–1782.

Swenson, H. A., Kaustinen, H. M., Kaustinen, O. A., & Thompson, N. S. (1968). Structure of gum arabic and its configuration in solution. *Journal of Polymer Science Part A-2*, *6*, 1593–1606.

Tanford, C. (1961). Transport processes—viscosity. In C. Tanford (Ed.), *Physical chemistry of macromolecules* (pp. 317–456). New York: John Wiley & Sons Inc.

Thevenet, F. (1995). Acacia gums—Natural encapsulation agent for food ingredients. In S. J. Risch, & G. A. Reineccius (Eds.), *Encapsulation and controlled release of food ingredients. American Chemical Society Symposium series* (vol. 590)

Van Stokkum, I., Spoelder, H., Bloemendal, M., Van Grondelle, R., & Groen, F. (1990). Estimation of protein secondary structure and error analysis from CD spectra. *Analytical Biochemistry*, *191*, 110–118.

Verbecken, D., Dierckx, S., & Dewettinck, K. (2003). Exudate gums: Occurrence production and applications. *Applied Microbiology Biotechnology*, *63*, 10–21.

Whitmore, L., & Wallace, B. A. (2004). DICHROWEB, an on-line server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, *32*, W668–W673.

Whitmore, L., & Wallace, B. A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*, *89*, 392–400.

Wu, H.-M., Wang, H., & Cheung, A. Y. (1995). A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower. *Cell*, *82*, 393–403.