



Arabinogalactan-like glycoproteins from *Ulva lactuca* (Chlorophyta) show unique features compared to land plants AGPs

Tereza Přerovská, Svatopluk Henke, Roman Bleha, Vojtěch Spiwok, Simona Gillarová, Jean-Claude Yvin, Vincent Ferrières, Eric Nguema-Ona, Petra Lipovová

► To cite this version:

Tereza Přerovská, Svatopluk Henke, Roman Bleha, Vojtěch Spiwok, Simona Gillarová, et al.. Arabinogalactan-like glycoproteins from *Ulva lactuca* (Chlorophyta) show unique features compared to land plants AGPs. *Journal of Phycology*, 2021, 57 (2), pp.619-635. 10.1111/jpy.13121 . hal-03102151

HAL Id: hal-03102151

<https://hal.science/hal-03102151>

Submitted on 26 Jan 2021

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.

MISS TEREZA PŘEROVSKÁ (Orcid ID : 0000-0002-2918-9383)

Article type : Regular Article

Arabinogalactan-like glycoproteins from *Ulva lactuca* (Chlorophyta) show unique features compared to land plants AGPs

Tereza Přerovská²

Department of Biochemistry and Microbiology, University of Chemistry and Technology Prague,
Technická 3, Prague, 16625, Czech Republic

Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR-UMR 6226, 35000
Rennes, France

Svatopluk Henke, Roman Bleha

Department of Carbohydrates and Cerials, University of Chemistry and Technology Prague,
Technická 3, Prague, 16625, Czech Republic

Vojtěch Spiwok

Department of Biochemistry and Microbiology, University of Chemistry and Technology Prague,
Technická 3, Prague, 16625, Czech Republic

Simona Gillarová

Department of Carbohydrates and Cerials, University of Chemistry and Technology Prague,
Technická 3, Prague, 16625, Czech Republic

Jean-Claude Yvin

Centre Mondial de l'Innovation Roullier, Laboratoire de Nutrition Végétal, 18 Avenue Franklin
Roosevelt, Saint-Malo, 35400, France

Vincent Ferrières

Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR-UMR 6226, 35000
Rennes, France

Eric Nguema-Ona

Centre Mondial de l'Innovation Roullier, Laboratoire de Nutrition Végétal, 18 Avenue Franklin
Roosevelt, Saint-Malo, 35400, France

Petra Lipovová

Department of Biochemistry and Microbiology, University of Chemistry and Technology Prague,
Technická 3, Prague, 16625, Czech Republic

²Corresponding author: Phone, +420 220443028; Email: prerovst@vscht.cz

Condensed running title: AGP-like proteins from *Ulva lactuca*

Editorial Responsibility: M. Herron (Associate Editor)

ABSTRACT

Arabinogalactan-proteins (AGPs) encompass a diverse group of plant cell wall proteoglycans, which play an essential role in plant development, signaling, plant-microbe interactions, and many others. Although they are widely distributed throughout the plant kingdom and extensively studied, they remain largely unexplored in the lower plants, especially in seaweeds. *Ulva* species have high economic potential since various applications were previously described including bioremediation, biofuel production, and as a source of bioactive compounds. This article presents the first experimental confirmation of AGP-like glycoproteins in *Ulva* species and provides a simple extraction protocol of *Ulva lactuca* AGP-like glycoproteins, their partial characterization and unique comparison to scarcely described *Solanum lycopersicum* AGPs. The reactivity with primary anti-AGP antibodies as well as Yariv reagent showed a great variety between *Ulva lactuca* and *Solanum lycopersicum* AGP-like glycoproteins. While the amino acid analysis of the AGP-like glycoproteins purified by the β -D-glucosyl Yariv reagent showed a similarity between algal and land plant AGP-like glycoproteins, neutral saccharide analysis revealed unique glycosylation of the *Ulva lactuca* AGP-like glycoproteins. Surprisingly, arabinose and galactose were not the most prevalent monosaccharides and the most outstanding was the presence of 3-O-methyl-hexose, which has never been described in the AGPs. The exceptional structure of the *Ulva lactuca* AGP-like glycoproteins implies a specialized adaptation to the marine environment and might bring new insight into the evolution of the plant cell wall.

Key index words: AGP-like glycoproteins; cell wall; extraction; green algae; methylated hexose; *Ulva lactuca*.

Abbreviations: AGPs, arabinogalactan-proteins; ESI, electrospray ionization; ESTs, expressed sequence tags; FMOC, 9-fluorenylmethyl chloroformate; GABA, 4-aminobutanoic acid; GPI, glycosylphosphatidylinositol; HPAEC, high-performance anion-exchange chromatography; HRGPs, hydroxyproline-rich glycoproteins; Hyp, hydroxyproline; MES, 4-morpholineethanesulfonic acid; MS, mass spectrometry; OPA, o-phthalaldehyde; PAD, pulsed amperometric detection; PMSF, phenylmethylsulfonyl fluoride; PRPs, proline-rich proteins; PVDF, polyvinylidene difluoride; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; TBS, Tris-buffered saline.

INTRODUCTION

Arabinogalactan-proteins (AGPs) encompass an extremely diverse class of cell wall glycoproteins/proteoglycans ubiquitous in land plants, from bryophytes to angiosperms (Popper et al. 2011, Classen et al. 2019). Recently, they have also been detected in a marine angiosperm (Pfeifer et al. 2020). AGPs have been and remain the subject of several studies, particularly in higher plants, because they were reported to play an essential role in many aspects of plant growth and development including embryogenesis, cell growth, and differentiation, in plant responses to biotic and abiotic stress, signalling, as well as in sexual reproduction (Seifert and Roberts 2007, Nguema-Ona et al. 2012, Ma et al. 2018, Su and Higashiyama 2018, Mareri et al. 2019). AGPs are found either attached to the plasma membrane via glycosylphosphatidylinositol (GPI) anchor, to the components of the cell wall matrix (pectins, hemicelluloses) or free in apoplastic space and secretions (Borderies et al. 2004, Hervé et al. 2015). Along with extensins and proline-rich proteins, AGPs belong to the hydroxyproline-rich superfamily of plant cell wall proteins (HRGPs; Showalter 1993, Kieliszewski et al. 2010). A key feature of HRGPs is the occurrence of proline and/or hydroxylated proline residues (Hyp). The hydroxylation of the proline residues by prolyl-4-hydroxylases is a prerequisite for further glycosylation of HRGP members (Gorres and Raines 2010, Nguema-Ona et al. 2014). Within the AGP family, repetitive dipeptide motifs Ser-Pro, Ala-Pro and Thr-Pro, were reported to template further AGP glycosylation based on the Hyp contiguity hypothesis (Kieliszewski and Shpak 2001). Moreover, an abundance of these four amino acids (Ala, Ser, Thr, Pro) within AGP proteic backbone (known as PAST) was reported to account for more than 50% of the total amino acids in classical AGPs (Showalter et al. 2010). Typically, the carbohydrate moiety of known land plant AGPs accounts for 90% [w/w] of the mass of the glycoprotein. They often harbor a larger β -D-(1,3)-galactan backbone with β -D-(1,6)-galactan side chains, which can be further substituted by arabinose, rhamnose, galactose, fucose or glucuronic acid (Tan et al. 2003, Ellis et al. 2010, Tryfona et al. 2012, Ma et al. 2018). The exact sugar composition varies within phylogenetically distant groups, species as well as between tissues and organs (Nothnagel 1997, Showalter 2001, Bartels et al. 2017). Moreover, developmentally-regulated changes in terms of sugar composition, within a given population of AGP, have also been reported and associated with various biological processes (Knox 1995, Van Hengel et al. 2002). For instance, the presence of epiphytic bacteria plays crucial role in *Ulva* spp. morphogenesis and can affect amino acid and saccharide composition of *Ulva* macromolecules, possibly also of AGPs (Segev et al. 2016, Alsufyani et al.

2020, Polikovskiy et al. 2020). Another key feature of AGPs is their ability to bind to synthetic phenylazo dyes, namely β -D-glucosyl and β -D-galactosyl Yariv reagents (Yariv et al. 1962, Kitazawa et al. 2013, Paulsen et al. 2014). This feature was already exploited for AGP localization, purification as well as for functional studies (Guan and Nothnagel 2004, Tang et al. 2006, Nguema-Ona et al. 2007). Monoclonal antibodies directed against AGPs represent another group of essential tools to study AGP localization and function (Knox 2008, Rydahl et al. 2018). Besides, the bioinformatic approach can be used for AGP identification, classification, and evolution analysis (Showalter et al. 2010, Johnson et al. 2017, Ma et al. 2017). Unfortunately, so far only few algal genomes were published disabling the use of bioinformatics in most cases. Still, the knowledge of the algal genome sequence does not guarantee AGP annotation.

Nevertheless, AGPs/HRGPs and AGP/HRGP-like structures were reported to occur across green and brown algae lineages (Sørensen et al. 2011, Hervé et al. 2015). However, contrasting with the wealth of information available on land plant AGPs, much less is known about AGP occurrence, structure, and function in algae. Although, the composition of the algal cell wall is extremely miscellaneous even in phylogenetically close species, sometimes the occurrence of common, structural polysaccharide or glycoprotein components can be found across the different lineages (Popper et al. 2011, Domozych et al. 2012). Using immunocytochemistry, the presence of AGPs was described in several green microalgae of the freshwater-originated Charophyta division, specifically in Desmidiaceae, Coleochaetaceae, Mesotaeniaceae, Zygnemataceae, Chlorokybaceae and Peniaceae families (Domozych et al. 2007, Eder et al. 2008, Sørensen et al. 2011, Palacio-López et al. 2019). Furthermore, AGPs were detected as well in the Charale order, representing the multicellular algae with stem-like and leaf-like structures (Domozych et al. 2010). Within the Chlorophyta division, AGPs were also reported in Oedogoniaceae, and Codiaceae families (Estevez et al. 2008, 2009, Fernández et al. 2010, 2015). Recent bioinformatic studies taking advantage of various sequenced plant genomes selected across the plant kingdom, or 1000 transcriptomes, has exposed the presence of AGPs throughout the entire plant kingdom except for red algae (Johnson et al. 2017, Ma et al. 2017, He et al. 2019). These studies have highlighted the predominant presence of various hybrid and chimeric AGPs within algal species and showed significant structural differences between algal and land plant AGPs with regard to protein sequences.

Ulva lactuca is a green marine benthic alga belonging to the class Ulvophyceae, which dominates shallow marine environments and display outstanding diversity regarding cytological

and morphological characteristics (Wichard et al. 2015). In addition to fibrillar cellulose, mannan or xylan polysaccharides, the green algae of this class produce sulfated polysaccharides, which are either i) uronic acid-limited in *Bryopsis*, *Caulerpa* and *Codium* genera or ii) uronic acid-rich in *Gayralia*, *Acetabularia*, *Monostroma* and *Ulva* genera (Domozych et al. 2012). *Ulva* spp. has tremendous economic potential in various sectors of the industry when renewable biofuel production belongs to the most extensively assessed applications due to the diminishing petroleum reserves (Suganya et al. 2016, Sudhakar et al. 2018). The biggest obstacle to the biofuel production is the recalcitrant nature of algal cell walls, which is caused by the presence of matrix and microfibrillar polysaccharides as well as the proteoglycans (Mishra et al. 2017). Thus, understanding the structure of algal cell walls is crucial for the right choice of the pretreatment method resulting in the improvement of biofuel extraction yield (Maneein et al. 2018, Kostas et al. 2020). Interestingly, the cell wall formation and composition can be modulated by bacteria-seaweed symbioses (Spoerner et al. 2012, Alsufyani et al. 2020, Polikovsky et al. 2020). AGPs were already successfully identified in the cell walls of *Codium* species, and typical AGP glycan structures were found in *Codium fragile* (Estevez et al. 2009, Fernández et al. 2010, 2015). Although the whole-genome sequence of *Ulva mutabilis* was recently published, no AGPs were annotated in the genome (Clerck et al. 2018). Mining a collection of EST from *Ulva linza*, Stanley et al. (2005) found 39 sequences with similarities to adhesive and cell wall proteins. Most of them resembled PRPs or HRGPs, precisely pherophorins, extensins, arabinogalactan-proteins, collagen-like protein 3, and others. Prolyl-4-hydroxylase was also present among ESTs (Stanley et al. 2005). However, so far there was no experimental proof of their presence as well as information about localization and structure of AGPs in the *Ulva* species.

Here, using a panel of biochemical techniques commonly used to isolate and analyze land plant-derived AGPs, including the use of the anti-AGP directed antibodies or the Yariv reagent coupled to analytic techniques, we confirm the occurrence of AGP-like components in *Ulva lactuca*. Structural information of the *Ulva lactuca* AGP-enriched fraction is presented. Parallel isolation of the AGP population from tomato (*Solanum lycopersicum*) leaves also allowed to easily compare the features of AGP-like population extracted from this marine green macroalgae with a typical land plant.

MATERIALS AND METHODS:

Algal (Ulva lactuca) and tomato (Solanum lycopersicum) plant material

Ulva lactuca materials collected in Brittany (France) were purchased from the European Marine Biological Resource Center (EMBRC, Station Biologique de Roscoff; <https://embrc-france.obs-banyuls.fr>) in 2017. *Ulva lactuca* was identified based on the sequence and phylogenetic analysis of *rbcL*, ITS, and *tufA* (plastid elongation factor) genes according to Vieira et al. (2016) and Lin et al. (2012). The material used for further extractions and analyses was freeze-dried and ground to a fine powder in CryoMill.

Tomato (*Solanum lycopersicum* cv Brenda F1) plants were grown in the greenhouse under a 16:8 h light:dark photoperiod. Tomato leaves were used for protein and AGP extraction when they reached five or six fully expanded leaves (about 5 weeks after sowing).

Extractions of AGPs

Extraction of AGPs according to Schultz et al. (2000)

1 g of ground freeze-dried *Ulva lactuca* was mixed with 4 mL of extraction buffer (50 mM Tris-HCl, pH 8, 10 mM EDTA, 4 mM sodium dithionite, 1% Triton X-100, 1 mM phenylmethylsulfonyl fluoride (PMSF)) and incubated at 4°C for 3 h using the rotary mixer.

Samples were centrifuged at 14 000g at 4°C for 10 min and the supernatant was precipitated with five volumes of ethanol at 4°C overnight without mixing. Following the centrifugation at the same conditions, the pellet was resuspended in 4 mL of 50 mM Tris-HCl buffer, pH 8. The suspension was centrifuged, the supernatant was retained, the pellet was resuspended and centrifuged as described before and the supernatant was retained.

Salt extraction

The 4 mL of 50 mM 4-morpholineethanesulfonic acid (MES) buffer, pH 6 with 0.2 M CaCl₂ and 1 mM PMSF were added to 1 g of ground freeze-dried *Ulva lactuca*. Extraction with buffer without salt was used for comparison. The extractions mixtures were incubated for 24 h at 4°C using a rotary mixer. Samples were centrifuged at 22,000g for 20 min at 4°C.

Immunodot-blot assay

25 µL of crude extracts were applied by 5 × 5 µL on the nitrocellulose membrane, always drying the sample at room temperature before another application. After the final drying of the membrane, blots were blocked with 5% low-fat milk in Tris-buffered saline (TBS) with 0.05% Tween 20 [v/v] (TBST) overnight at 4°C on a rocking platform. Primary antibodies against

arabinogalactan-proteins (JIM16, JIM13, and MAC207), all provided by PlantProbes, were used in 1:500 dilution in 5% low-fat milk in TBST for 1.5 h at room temperature on rocking platform 100 rpm. After washing with TBST three times for 20 min at room temperature on a rocking platform, blots were incubated with an anti-rat IgG secondary antibody coupled to horseradish peroxidase in dilution 1:10 000 in 5% low-fat milk in TBST for 1.5 h at room temperature on rocking platform 100 rpm. After washing as described before, the membranes were developed in SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific, Waltham, MA, USA) for 5 min at room temperature and the chemiluminescence was detected by ChemiDoc Imaging System (Bio-Rad, Hercules, CA, USA).

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot
Samples were mixed with 4X Laemmli buffer (Bio-Rad, USA), boiled for 10 min and 4-25 μL were loaded on 4-15% Mini-PROTEAN® TGX Stain-Free™ precast polyacrylamide gels (Bio-Rad). Gels were run at constant current 200 V for approximately 35 min, then they were imaged by stain-free technology in ChemiDoc Imaging System (Bio-Rad) or by Pierce Silver Stain Kit (Thermo Fisher Scientific). Separated proteins were transferred to nitrocellulose membrane via the Trans-Blot Turbo system (Bio-Rad), using the 10-min program for high molecular proteins and were checked for the efficiency of transfer. The membranes were blocked and incubated with antibodies (see the description in the immunodot-blot essay section).

Yariv gel-diffusion assay

Based on van Holst and Clarke (1985), 1% agarose gels containing the β -D-galactosyl, β -D-glucosyl or α -D-mannosyl Yariv reagent ($25 \mu\text{g} \cdot \text{mL}^{-1}$), 1% NaCl and 0.02% sodium azide were poured into Petri dishes. Extracts from *Ulva lactuca* and *Solanum lycopersicum* were loaded by 60 μL into the wells. While the *Solanum lycopersicum* samples were used at concentration $5 \text{ mg} \cdot \text{mL}^{-1}$, the *Ulva lactuca* samples at concentration $20 \text{ mg} \cdot \text{mL}^{-1}$. 60 μL of 1% NaCl was used as a negative control and as a positive control the gum arabic in concentrations 0.25, 0.5, and $1 \text{ mg} \cdot \text{mL}^{-1}$ was used. Petri dishes were kept in a wet chamber for 24 hours at room temperature to allow the halo to develop. The gels were then destained by 1% NaCl at 4°C . The formation of red precipitate was observed.

AGP purification using Yariv reagent according to Schultz et al. (2000)

The β -D-glucosyl Yariv reagent dissolved in 1% (w/v) NaCl was added to *Ulva lactuca* extract to the final concentration of $1 \text{ mg} \cdot \text{mL}^{-1}$. The mixture was incubated overnight at 4°C to precipitate. The insoluble complex was collected by centrifugation at $20,000g$ for 1 h at 4°C . The remaining Yariv reagent was removed by three washes with 1% (w/v) NaCl and then by two with methanol. Afterward, the washed pellet was air-dried (to remove the excess of methanol, the pellet should remain wet) and dissolved in a minimal volume of dimethyl sulfoxide. Few crystals of sodium dithionite were added, and the solution was mixed. Finally, water was slowly added with constant mixing until the mixture changed color from red to pale yellow or colorless. The solution was then desalted on a PD10 column and freeze-dried.

Identification of putative classical AGPs in Ulva mutabilis genome

Based on Showalter et al. (2010), putative classical AGPs were identified in genome of *Ulva mutabilis*. Briefly, classical AGPs were defined as protein containing more than 50% PAST, together with the presence of dipeptide motifs (AP, PA, SP and TP) distributed throughout the protein sequence. Identified sequences were analyzed for presence of signal peptide by SignalP-5.0 (<http://www.cbs.dtu.dk/services/SignalP/>) and for GPI-anchor by big-PI Predictor (http://mendel.imp.ac.at/gpi/cgi-bin/gpi_pred.cgi). The search of most similar sequences was done by BLAST analysis.

Amino acid composition analysis

Yariv purified AGP-like glycoproteins were separated by SDS-PAGE and blotted to polyvinylidene difluoride (PVDF) membrane. Major bands were cut out, the membrane was washed several times with distilled water and air-dried. Proteins were hydrolyzed with 6 M HCl for 24 h at 110°C in vacuo airtight vials and acid was evaporated in a desiccator. The hydrolysate was resuspended in $100 \mu\text{L}$ of 0.1 M HCl, containing internal standard 4-aminobutanoic acid (GABA). Amino acids were pre-column derivatized by autosampler with o-phthalaldehyde (OPA) for primary amino acids and 9-fluorenylmethyl chloroformate (Fmoc) for secondary amino acids.

The separation was performed on a Waters AccQ.Tag Ultra C18 ($1.7 \mu\text{m}$ particle size, $2.1 \times 100 \text{ mm}$, Waters, USA) column. The column temperature was maintained at 40°C and the volume injected on the column was $20 \mu\text{L}$. Gradient elution was performed using 10 mM borate buffer pH 8.2 as eluent A and acetonitrile/MeOH/ H_2O in ratio 45:45:10 as eluent B. The flow rate was kept constant at $0.4 \text{ mL} \cdot \text{min}^{-1}$ with the following gradient: 2-57% B during 0.35-13.3 min.

Fluorometer was used for the detection of derivatized amino acids, with an excitation wavelength of 340 nm and an emission wavelength of 450 nm.

Determination of saccharide content based on Ludwig and Goldberg (1956)

Anthrone reagent was prepared by dissolving 0.2 g of anthrone in mixture of 5 mL of ethanol and 95 mL of 75% sulfuric acid on ice. For measurement of calibration line was used glucose in concentrations 0, 0.01, 0.1, 1, 10, 100 $\mu\text{g} \cdot \text{mL}^{-1}$. To 100 μL of samples (0.1 $\text{mg} \cdot \text{mL}^{-1}$) and calibration solutions was added 500 μL of anthrone reagent on ice and solutions were vortexed. Afterward the mixture was incubated for 10 min at 100°C, chilled on ice and the content of hexoses was determined spectrophotometrically at 625 nm.

Determination of uronic acid content based on Blumenkrantz and Asboe-Hansen (1973)

For measurement of calibration line was used galacturonic acid in concentrations 0, 40, 80, 120, 160, 200 and 240 $\mu\text{g} \cdot \text{mL}^{-1}$. Samples (1 $\text{mg} \cdot \text{mL}^{-1}$) and calibration solutions were diluted by distilled water 1:4 to final volume 500 μL and 3 mL of 12.5 mM sodium tetraborate decahydrate (0.478 g dissolved in 100 mL of 96% sulfuric acid) were added and the mixture was vortexed. The tubes were kept at 100°C for 5 min, chilled on ice and 50 μL of 0.15% 3-hydroxybiphenyl in 0.5% NaOH was added. Into the sample blanks was added only 0.5% NaOH. The solutions were vortexed and kept at room temperature for 30 minutes. The absorbance was measured at 520 nm. From the samples was also subtracted control solution of β -glucan (1 $\text{mg} \cdot \text{mL}^{-1}$), as correction of neutral saccharide interference.

Saccharide composition analysis

1 mg of Yariv purified AGP-like glycoproteins or desalted extracts were dissolved in 0.5 mL of 1 M H_2SO_4 and were hydrolyzed for 8 h at 90°C. To neutralize the samples, 150 mg of BaCO_3 were added and incubated overnight on vortex. Samples were centrifuged at 10,000g for 15 min, the supernatant was taken and diluted 1:1. Samples were filtrated and pH was checked. If needed, samples were further diluted to get within the calibration range of the following analysis.

The samples were analyzed using high-performance anion-exchange chromatography (HPAEC) with pulsed amperometric detection (PAD) system Dionex DX-600 (Dionex, Sunnyvale, CA, USA) with anion-exchange column CarboPac PA1, 2 mm \times 250 mm (Thermo Fisher Scientific) for the possible presence of about 20 saccharides and sugar alcohols. The

Dionex ECD-50 detector (Dow, USA) was switched to the PAD mode. The injection volume was 10 μL . The mobile phase flow rate was $0.25 \text{ mL} \cdot \text{min}^{-1}$ with isocratic elution for 30 min in 20 mM NaOH followed by 10 min of regeneration with 200 mM NaOH and 15 min of re-equilibration again with 20 mM NaOH. The column temperature was maintained at 25°C . The same system (DX-600 with PA1 column) was used for isolation of a specific fraction containing the unknown saccharide, with the detector in the off state to avoid the oxidation of saccharide.

The composition of this unknown saccharide was verified using system HPAEC/PAD system DX-500 (Dow, USA) with column Dionex CarboPac MA1, $4 \text{ mm} \times 250 \text{ mm}$ (Dow, Midland, MI, USA) and ECD-40 detector. The injection volume was 20 μL . The column temperature was maintained at 30°C . The method was run with 480 mM NaOH as the mobile phase in isocratic mode $0.4 \text{ mL} \cdot \text{min}^{-1}$ with total time 95 min for one sample.

Once the purity was verified, the 20 mM NaOH present in the collected fraction was immediately removed by Pierce™ Strong Cation Exchange Spin Column (Thermo Fisher Scientific) with H^{+} as the counterion.

Analysis of unknown monosaccharide

Electrospray ionization mass spectrometry

Mass spectra were measured using LTQ Orbitrap XL hybrid mass spectrometer (Thermo Fisher Scientific) equipped with an electrospray ion source. The mobile phase was methanol/water (4:1, by vol.) at the flow rate of $100 \mu\text{L} \cdot \text{min}^{-1}$. The desalted samples were dissolved in methanol and injected using a 5- μL loop into the mobile phase flow. For the negative ion mode, spray voltage, capillary voltage, tube lens voltage, and capillary temperature were 4.6 kV, -25 V, -125 V, and 275°C , respectively. For the positive ion mode, spray voltage, capillary voltage, tube lens voltage, and capillary temperature were 4.8 kV, 9 V, 150 V, and 275°C , respectively. The spectra were recorded at the resolution of 100,000. The analysis was performed by Mass Spectrometry Group of Josef Cvačka at the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences.

Distinction between 3-O-methyl- and 4-O-methyl-hexoses by gas chromatography – mass spectrometry according to Pfeifer and Classen (2020)

Collected, desalted and lyophilized samples were reduced by the 1 M sodium borodeuteride solution in DMSO and were incubated at 40°C for 1.5 h. The remaining sodium borodeuteride was

inactivated by the addition of 50 μL of glacial acetic acid for two times. The volatile alditol acetates were obtained by the addition of 200 μL of 1-methylimidazole, 2 mL of acetic anhydride and the reaction run at laboratory temperature for 20 min. The remaining acetic anhydride was destroyed by the addition of 10 mL of distilled water. The alditol acetates were extracted with 1 mL of dichloromethane after the acidification with 1 mL of 0.1 M sulfuric acid.

Obtained derivatives were analyzed by Shimadzu single quadrupole GCMS-QP2010 SE gas chromatograph-mass spectrometer (Japan) equipped with a capillary column HP-5 (30 m length, 0.25 mm internal diameter, 0.15 μm film thickness). Helium was used as the carrier gas (flow 35 $\text{cm}^3 \cdot \text{s}^{-1}$). The temperature of injector was 220°C. The oven temperature program was following: 45°C rose to 140°C (10°C $\cdot \text{min}^{-1}$), 140°C for 5 min, then rose to 170°C (0.5°C $\cdot \text{min}^{-1}$), 170°C for 1 min, then rose to 280°C, final °C for 5 min, total time 90 min.

Tandem mass spectrometry

Tandem mass spectrometry (MS/MS) with collision induced dissociation (CID) spectra were measured in a linear ion trap using LTQ Orbitrap XL hybrid mass spectrometer (Thermo Fisher Scientific) equipped with an electrospray ion source. The sample (or standard) was dissolved in methanol/water (4:1, v/v) and directly infused at 3 $\mu\text{L} \cdot \text{min}^{-1}$ into the ion source. The spray voltages were 5.5 kV (for the positive ion mode) and 4.6 kV (for the negative ion mode); the capillary temperature was 275°C. The precursors were isolated using a 1 Da window and the normalized collision energy was optimized in the 15 – 24 range. The analysis was performed by Mass Spectrometry Group of Josef Cvačka at the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences.

RESULTS

Very little is known about AGP-like glycoproteins from algae. In this article, we dealt with their identification, partial purification and characterization of the carbohydrate and peptide part. For comparison, tomato AGPs were also characterized as representative of higher plants.

Validation of the anti-AGP antibodies as valuable probes to track AGP-like glycoproteins in *Ulva lactuca* and use of salt in extraction protocol to enhance their yield

To establish a suitable antibody for the AGP probing in *Ulva lactuca*, the extraction using the standard protocol of AGPs extraction according to Schultz et al. (2000) was performed. Extract from *Solanum lycopersicum* was prepared in the same way to compare the abundance of AGP-

epitopes recognized by different primary antibodies between green macroalga and higher plant. The occurrence of AGP-associated epitopes was evaluated by immunodot-blot assay with different anti-AGP monoclonal antibodies, which were chosen based on literature research (Fig. 1). While JIM13 is known to have a strong response with tomato AGPs (Fragkostefanakis et al. 2012), JIM16 and MAC207 were previously successfully used for AGP detection in *Codium* sp., a member of Ulvophyceae (Estevez et al. 2009, Fernández et al. 2010). The best result for *Ulva lactuca* was obtained with the anti-AGP JIM16 antibody (Fig. 1), which is in agreement with Fernandez et al. (2010). Interestingly, epitopes recognized by the anti-AGP JIM13 antibody, often abundant in higher plants, were not detected in our *Ulva lactuca* extract, suggesting that the composition and the structures of these AGPs might be different from higher plants to green macroalgae. While for *Ulva lactuca* the MAC207 showed a response with intermediate intensity, there was almost none in *Solanum lycopersicum*. JIM13 performed slightly better than JIM16 for *Solanum lycopersicum* according to the immunodot-blot assay. Thus, JIM16 was the best probe for *Ulva lactuca* and JIM13 for *Solanum lycopersicum*. However, this extraction protocol yielded a too low amount of AGPs based on the applied long exposition time.

To improve the yield and simplify the extraction protocol of AGPs, we took advantage of the existing literature on cell wall protein extractions, where inorganic salts were shown to be used frequently with calcium chloride being the most efficient (Jamet et al. 2008, Printz et al. 2015).

Extracts were analyzed for total protein content (Fig. 2A), and the presence of glycoproteins bearing anti-AGP JIM16 (Fig. 2B) epitopes. The tomato extract was analyzed for comparison of macroalgal to the land plant AGPs. The best staining of the generated SDS-PAGE gels was obtained with either silver staining or stain-free technique, while the Coomassie Brilliant Blue and the periodic acid-Schiff (PAS) method failed to stain them (data not shown).

All extracts consisted of several proteins ranging from low to very high molecular weight (Mw) above 250 kDa (Fig. 2A). The results also showed that overall protein concentration was much higher in *Solanum lycopersicum* extract than in *Ulva lactuca* extracts. The JIM16 antibody exposed at least three distinct bands in *Ulva lactuca* extract. Mobility of the first band corresponds to the AGP-like glycoprotein with a molecular weight above 250 kDa, the second to the AGP-like glycoprotein of approximately 200 kDa, and the third of 100 kDa (Fig. 2B). Interestingly the third AGP-like glycoprotein is mainly extracted by 0.2 M CaCl₂ since a very weak signal is present in extracts without the presence of CaCl₂ (Fig. 2B). From the results, it is obvious that salt extraction was better for *Ulva lactuca* than the well-known AGP extraction protocol according to Schultz et

al. (2000). The JIM16 response of *Solanum lycopersicum* extract was very weak (Fig. 2B, line 4), revealing two faint bands corresponding to AGPs with approximate molecular weights of 150 and 100 kDa. The explanation of difference in JIM16 response with *Solanum lycopersicum* extract during immunodot-blot and western blot assays lies within the techniques themselves and also tangled specificity of primary antibodies. The immunodot-blot assay is more prone to false positive results, because no separation step like in case of western blot is included and whole extract is loaded on the membrane, which can result in cross-reactivity issues and non-specific interactions. Consequently, the *Solanum lycopersicum* extract was also probed with JIM13 primary antibody (Fig. 2C) and as was expected JIM13 had a much stronger response and exposed three bands, the first one corresponds to AGP with a high molecular weight above 250 kDa, which is present also in extracts from *Ulva lactuca*. The second most significant one with Mw around 150 kDa and the third one with 75 kDa. Expectedly the SDS-PAGE bands corresponding to the AGPs and AGP-like glycoproteins are blurry, which is typical for heterogeneously and heavily glycosylated proteins (Brooks 2009).

The confirmation of AGP-like glycoprotein presence by Yariv reagent

For further confirmation of the AGP-like glycoproteins presence in *Ulva lactuca* and *Solanum lycopersicum* extracts the Yariv gel-diffusion assay was performed. Extract prepared with the help of 0.2 M CaCl_2 formed a halo with both β -D-galactosyl and β -D-glucosyl Yariv reagent (Fig. 3). Also, there was no precipitation with α -D-mannosyl Yariv reagent (data not shown), which was in good agreement with what is known from higher plants (Guan and Nothnagel 2004, Kitazawa et al. 2013). However, compared to *Solanum lycopersicum* extract, *Ulva* extract shows much lower reactivity with Yariv reagent. Since no difference between β -D-galactosyl and β -D-glucosyl Yariv reagent was observed, only β -D-glucosyl Yariv reagent, the most commonly used, was chosen for the following purification.

Yariv reagent was also used for purification of AGPs according to Schultz et al. (2000). Samples after Yariv purification were analyzed for total protein content (Fig. 4A) and the presence of glycoproteins bearing anti-AGP JIM16 resp. JIM13 (Fig. 4B resp. 4C) epitopes. The protein content of both extracts rapidly decreased after Yariv purification and only faint smear was detected in all cases, which was expected due to the nature of AGPs (Fig. 4A). Interestingly immunodetection with JIM16 showed differences in composition of AGP-like glycoproteins before and after Yariv purification (Fig. 4B) in the case of *Ulva lactuca*. While in the extract at

least three AGP-like glycoproteins were detected after Yariv purification mainly high molecular AGP-like glycoprotein was prevalent. Since JIM16 is not the best probe for tomato, JIM13 was used analysis of tomato extract before and after Yariv purification and, in this case, the composition of purified AGPs remained identical and was corresponding to the AGPs in the raw extract. These results indicate different structural properties of extracted *Ulva lactuca* AGP-like glycoproteins and explain the low reactivity with Yariv reagent during radial gel-diffusion assay compared to *Solanum lycopersicum*. When silver staining protocol was used for analyzing the total protein content of *Ulva lactuca* and *Solanum lycopersicum* β -D-glucosyl Yariv precipitates, more proteins were revealed, pointing out higher sensitivity compared to the stain-free technique (Fig. 4D). The most interesting was a low molecular protein/glycoprotein after Yariv purification with approximate molecular weight 20 kDa, which is almost invisible with stain-free visualization (Fig. 4A). This protein did not interact with the JIM16 antibody, therefore other anti-AGP antibodies were tested, but in all cases, no interaction was detected (data not shown). Highlighted areas were chosen for further amino acid composition analysis.

The amino acid composition of Yariv purified AGP-like glycoproteins from Ulva lactuca and Solanum lycopersicum

The amino acid composition of four different Yariv purified AGP-like glycoproteins after electrophoretic separation was analyzed, two from *Ulva lactuca* and two from *Solanum lycopersicum* (Fig. 4). The experimental values were compared with the genomic information (Fig. 5). Although the genome of closely related *Ulva mutabilis* was published recently, no AGPs were annotated (Clerck et al. 2018). Therefore, the analysis of *Ulva mutabilis* genome was performed and 6 putative classical AGPs were identified, which had PAST content above 50%, contained dipeptide motifs and signal sequence. None of them contained GPI-anchor and BLAST search either failed to identify any similar proteins or just found uncharacterized proteins with low cover query and identity. Typical extensin motif (SP₃₋₄) was present in one putative AGP. All of these putative AGPs were used for boxplot comparison with experimental data of amino acid content (Table S1 in the Supporting Information). In the case of *Solanum lycopersicum* were experimental values compared to 44 AGPs annotated in its genome (GenBank assembly accession: GCA_000188115.3). *Solanum lycopersicum* AGPs were separated to classical AGPs, fasciclin-like AGPs and other non-classical AGPs (Table S2 in the Supporting Information). From the genomic data themselves were obvious big differences and variability between the individual

types of AGPs as well as between *Ulva* and tomato AGPs (Fig. S1 in the Supporting Information). If we focus on the most prevalent amino acids in land plant classical AGPs (Thr, Ala, Ser and Pro), their contents in *Ulva* and tomato classical AGPs were very similar (Fig. 5). On the other hand, fasciclin-like and other non-classical AGPs had lower amount of alanine and in case of fasciclin-like AGPs also of proline resulting in lower PAST content compared to classical AGPs.

Hydroxyproline was present in all tested proteins. Amino acid composition of two AGP-like glycoproteins from *Ulva lactuca* was in the most cases similar or was differing just by few percents (Fig. 5), even though U1 represents high molecular weight AGP-like glycoprotein, which gives strong response with JIM16, while U2 is 20 kDa protein/glycoprotein, which has no interaction with tested antibodies. In the case of *Solanum lycopersicum*, both AGPs had the similarity even higher with minimal differences (Fig. 5). The biggest variance was in proline content, which was greater in AGP named T2 with approximate molecular weight 160 kDa. Interestingly, the amino acid compositions of *Solanum lycopersicum* and *Ulva lactuca* AGP-like glycoproteins were very much alike with PAST content in range between 27-34%. Surprisingly, even AGPs annotated in *Solanum lycopersicum* genome as classical have lower PAST content, exactly 8 out of 16 AGPs did not fulfill the criteria of 50%. Based on the results, it seems that extracted and analyzed AGP-like glycoproteins belong to non-classical AGPs.

Glycosylation of Yariv purified AGP-like glycoproteins isolated from Ulva lactuca and Solanum lycopersicum

Neutral monosaccharide composition of the Yariv precipitated fraction as well as crude extracts was examined. In this case, a huge difference between monosaccharide composition between *Ulva lactuca* and *Solanum lycopersicum* AGP-like glycoproteins was observed (Table 1). Although fucose, rhamnose, arabinose, galactose, glucose, mannose, and unknown saccharide (Unknown2) with retention time 9.9 min were detected in all tested samples, their content varied greatly. Also, the overall saccharide content was much higher in *Solanum lycopersicum* Yariv precipitate. The major monosaccharide in *Ulva lactuca* extract was rhamnose, the main component of ulvan polysaccharide. Interestingly, whereas the content of arabinose was small, significant unknown saccharide (Unknown1) with retention time 6.12 min appeared only in *Ulva lactuca* chromatograms. Yariv purification of *Ulva lactuca* extract led to a major decrease in rhamnose and xylose content, while in the case of the remaining monosaccharides increased. Unexpectedly, the most prevalent monosaccharide in *Ulva lactuca* Yariv precipitate was saccharide Unknown1,

which was further analyzed. On the contrary, galactose and arabinose, the main components of the glycan structures in AGPs, were the most prevalent monosaccharides even in crude *Solanum lycopersicum* extract. While the quantity of galactose, arabinose, rhamnose, and fucose was further increased in the sample after Yariv purification, the content of the remaining monosaccharides decreased. The contents of unknown monosaccharides are just estimations, because the coefficients used for their calculations were either of the most similar monosaccharide or average of all used standards, if the nature remains completely elusive. Since the presence of glucuronic acids in glycan structure of arabinogalactan-proteins is well established, content of uronic acids was determined in both extracts and Yariv precipitated fractions. In the case of *Ulva lactuca* the ratio between uronic acid and neutral sugar content increased from 1:1.2 to 1:5.2 after Yariv purification of AGP-like glycoproteins. On the contrary, the ratio between uronic acid and neutral sugar content in *Solanum lycopersicum* AGPs decreased from 1:11.5 to 1:4 (data not shown). Thus, *Ulva lactuca* Yariv precipitated fraction contained slightly smaller amount of uronic acids than *Solanum lycopersicum* Yariv precipitated AGPs.

The structural analysis of the Unknown1 monosaccharide

The Unknown1 saccharide was chosen for further analysis, due to its unique presence and high content only in *Ulva lactuca* samples. When the Unknown1 saccharide was collected in sufficient purity and analyzed using a different column operating in higher pH and interestingly, the retention time has shifted closer to rhamnose (Table 2, Fig. S2 in the Supporting Information). Considering the extremely low amount of unknown saccharide, the only option was MS analysis. The MS analysis revealed a $C_7H_{14}O_6Na$ molecule with a molecular mass of $217.06835 \text{ g} \cdot \text{mol}^{-1}$ (Fig. S3 in the Supporting Information). After the subtraction of sodium adduct, the molecular mass corresponded exactly to a methylated hexose.

In order to confirm this hypothesis and also gain extra information, we did GC-MS analysis of unknown monosaccharide according to Pfeifer and Classen (2020), which allow us to distinguish between 3-*O*-methyl- and 4-*O*-methyl-hexoses. Following their protocol, we obtained exactly the same spectra of our unknown saccharide as of 3-*O*-methyl-glucose standard and also corresponding to the results in the article (Fig. 6). However, the retention time of our unknown saccharide and 3-*O*-methyl-glucose differed, so it has to be different hexose. This conclusion is also supported by results of tandem mass spectrometry since 3-*O*-methyl-glucose standard and our unknown saccharide had same fragmentation pattern, but differing in intensity of individual

fragments (Fig. S4 in the Supporting Information). Unfortunately, no standards of 3-*O*-methyl-hexoses except for glucose are available and extremely low amount of unknown saccharide hampers further precise identification.

DISCUSSION

Plant cell walls play crucial roles in many aspects of plant development, cellular functions, and interactions with the environment. They have been extensively studied and many different probes, which have been extensively described by Rydahl et al. (2018), were developed to elucidate their structure. Among these probes, antibodies represent the most important and largest group. Since the 1980's hundreds of cell wall-related antibodies have been reported covering all the main polymeric components, including arabinogalactan-proteins (Pattathil et al. 2010, Nguema-Ona et al. 2012). The reactivity of *Ulva lactuca* extract with selected primary anti-AGP antibodies was in agreement with results from *Codium fragile* (Estevez et al. 2009) and *Codium vermilara* (Fernández et al. 2010), both being members of Ulvophyceae. The results suggest some similarities between *Ulva* and tomato AGP-like glycoproteins considering their interaction with the JIM16 antibody. On the other hand, also some structural differences can be expected since *Ulva* AGP-like glycoproteins reacted positively with MAC207 while tomato AGPs with JIM13 antibody. Interestingly, MAC207 and JIM13 should according to the literature recognize identical epitope β -GlcA-(1 \rightarrow 3)- α -GalA-(1 \rightarrow 2)- α -Rha (Yates et al. 1996), but experimental data show big discrepancies in their responses to the same AGPs, pointing out to additional factors affecting the interaction (Sørensen et al. 2011, Hervé et al. 2015). Recently epitope of JIM16 was elucidated and it was found out that JIM16 interacts with at least three β -(1,3)-galactose residues substituted with β -(1,6)-galactose (Ruprecht et al. 2017). Thus, both extracts had a response with two antibodies with different specificity, suggesting the presence of different AGP epitopes within the extracts (Pattathil et al. 2010, Ruprecht et al. 2017, Rydahl et al. 2018). Nevertheless, although these probes can give us extremely valuable insight in to cell wall composition, organization and development, their usage has also its drawbacks, mainly the cross-reactivity issues and non-specific interactions (Pattathil et al. 2010). For instance, discrepancies in JIM16 response with tomato extract on immunodot-blot and western blot can be explained by JIM16 non-specific interactions. Since weak non-specific interactions of JIM16 with tomato pectic polysaccharide, xyloglucan, glucomannan and RG I were reported by manufacturer. Despite the fact, that according to manufacturer screening for cross-reactivity and non-specific interaction, JIM16 has

much lower cross-reactivity and non-specific interactions than JIM13. Generally, immunodot-blot assay is more prone to non-specific interactions and can lead to false positive results, therefore the results should be verified by another technique. In the case of *Ulva lactuca*, the JIM16 cross-reactivity is unlikely since based on literature *Ulva* do not contain any polymers with similar epitope (Lahaye and Robic 2007, Domozych et al. 2012). Still possibility of non-specific interactions cannot be excluded.

Over the last years, seaweed research experienced an extreme boom due to a chase for new bioactive compounds, valuable metabolites, or simply new sources of proteins for constantly growing human population. To obtain these molecules, various extraction methods were tested, established, and reviewed (Kadam et al. 2015, Michalak and Chojnacka 2015). Although extensive literature background is available on protein extraction from seaweed, these extractions have never focused directly on the cell wall proteins, especially arabinogalactan-proteins. Extraction of land plant cell wall proteins usually contains two steps: cell wall enrichment and extraction, which is mostly done with the help of inorganic salts (Jamet et al. 2008, Printz et al. 2015). Our results support this well-established consensus since inorganic salt helped extraction of AGP-like glycoproteins from the *Ulva* cell wall (Table S3 in the Supporting Information). Surprisingly, the most well-known and used extraction protocol of AGPs in land plants (Schultz et al. 2000), is not suitable for *Ulva lactuca*. The same phenomenon was previously described also for green microalga *Micrasterias denticulata* (Eder et al. 2008). However the problem is probably not in the composition of extraction solution itself, but with the ethanol precipitation step, which is highly influenced by the structure and composition of polysaccharides (Xu et al. 2014). So far Wijesekara et al. (2017) is the only paper focusing on the extraction of glycoproteins from *Ulva* species, mentioning the possibility of AGP-like glycoprotein presence within these extracts. Compared to the extraction procedures in this paper, which are complicated and hard to scale up, our AGP extraction protocol is very simple and effective, which also ensures high reproducibility of results. Expectedly, the differences between *Ulva* and tomato extracts on western blot suggest diverse AGP-like glycoproteins in the extracts, not only varying in the size but also epitopes. On the contrary, the protein with an approximate molecular weight above 250 kDa is present in both extracts, but the intensity with used antibodies is different. Without further analysis, it is impossible to say, whether we are dealing with conserved AGP-like glycoprotein, in which glycosylation pattern was modified throughout the evolution or two completely different biopolymers. Despite the phylogenetic closeness of *Codium vermilara* and *Codium fragile*,

different bands were revealed by JIM16 antibody on western blots in their water extracts (Estevez et al. 2009, Fernández et al. 2010). These big differences in the AGP structure and distribution within the same genus are implying a great variability of these glycoproteins in algae. It should be also noted, that the molecular weights determined from western blots are most likely overestimated due to the poor mobility of highly glycosylated AGPs in SDS-PAGE gels.

Yariv reagents (mainly β -D-glucosyl and β -D-galactosyl) are widely used in studies focused on AGPs, thanks to their ability to specifically bind AGPs (Yariv et al. 1962). Many protocols including *in situ* localization, Yariv diffusion assay, purification, and functional studies of AGPs were developed in land plants (Schultz et al. 2000, Guan and Nothnagel 2004, Tang et al. 2006, Kitazawa et al. 2013). However, utilization of Yariv reagent among algal species is very rare, up to now only six papers were published, one concerning brown algae and the remaining ones green algae (Eder et al. 2008, Estevez et al. 2009, Fernández et al. 2010, 2015, Hervé et al. 2015, Palacio-López et al. 2019). Yariv radial gel-diffusion assay showed the presence of AGP-like glycoproteins capable of binding both β -D-glucosyl and β -D-galactosyl Yariv reagent within the *Ulva lactuca* extract. Compared to the identically prepared *Solanum lycopersicum* extract, it seems that they are present in very small amounts. Further, yields of Yariv precipitation of *Ulva lactuca* extract were much lower compared to *Solanum lycopersicum*, even though they were prepared identically (data not shown). The question remains whether the poor reactivity is caused by a small amount of AGP-like glycoproteins in *Ulva lactuca* extract or due to structural differences of some AGP-like glycoproteins, which hamper proper binding to Yariv reagent. Although the Yariv reagents are well accepted in AGP research since the 1960s, their target structure, as well as the mechanism, remains elusive. Nevertheless, it is generally accepted that they interact with AGPs containing type II arabinogalactan, made of β -(1,3)-galactan the main chain, and β -(1,6)-galactan side chains. Recently this hypothesis was confirmed for β -D-galactosyl Yariv reagent (Kitazawa et al. 2013, Sato et al. 2018). They also showed that while the presence of protein moiety is not required, the extent of β -(1,6)-galactan substitution affects the Yariv reagent reactivity. The difference in Yariv reagent and JIM16 specificity can explain the discrepancies on western blot before and after Yariv purification. While JIM16 recognizes at least three β -(1,3)-galactose residues substituted with β -(1,6)-galactose (Ruprecht et al. 2017), Yariv reagent binds to β -(1,3)-galactan chains longer than five residues (Kitazawa et al. 2013, Sato et al. 2018). The higher β -(1,6)-galactan substitution could explain the loss of other AGP-like glycoproteins after β -D-glucosyl Yariv purification. The inability of binding Yariv reagent and at the same time positive

response with multiple anti-AGPs monoclonal antibodies was already described in several green algae, namely *Micrasterias*, *Chlorokybus*, *Coleochaete* and *Penium* (Eder et al. 2008, Palacio-López et al. 2019). Noteworthy, even lower yields of Yariv purification were reported for *Zostera marina*, a seagrass evolved from monocotyledonous land plants. The structure of Yariv purified AGPs showed unique features compared to the land plant AGPs, i.e. having a high degree of branching and high amounts of terminal 4-*O*-methyl-glucuronic acid (Pfeifer et al. 2020). Thus, the interaction of substituted β -(1,3)-galactan main chain and Yariv reagent is possible, but the interaction is very weak, resulting in a low yield of purification.

Various methods were established for protein detection after electrophoretic separation. While the benefits of a stain-free technique are the fast detection and possibility of immediate visualization of proteins, the main benefit of silver staining is its high sensitivity (Westermeyer 2016). When silver staining was used instead of stain-free technique, more proteins interacting with Yariv reagent were exposed making the difference in the specificity of Yariv reagent and JIM16 antibody even more profound. The stain-free technique is limited by the presence of tryptophan residues in protein backbone and analysis of tryptophan content in *Solanum lycopersicum* found that 28 out of 44 AGPs did not contain any and the remaining ones in average only 0.63% (data not shown). Thus, it seems that the stain-free technique is not the best choice of AGP detection. The lacking interaction of 20 kDa protein/glycoprotein with tested antibodies is not surprising since the antibodies interact with specific epitopes and any modification will disturb its recognition, in this case, it might be missing β -(1,6)-galactan substitution (Ruprecht et al. 2017, Ma et al. 2018).

For further characterization of the AGP-like protein portion, amino acid composition analysis was performed. It is a classical widely used method for protein characterization (Rutherford and Gilani 2009). So far, there is no publication regarding the characterization of purified AGP-like glycoproteins from algae. On the other hand, in the case of *Solanum lycopersicum* few AGPs were previously characterized, mainly the most well-known LeAGP-1 (Gao et al. 1999, Zhao et al. 2002). All these AGPs had hydroxyproline content around 30% and PAST content around 70%, fulfilling all the parameters of classical AGPs. Nonetheless, similarly analyzed AGPs from *Echinacea purpurea* and *Triticum aestivum* had a lower content of hydroxyproline and PAST did not exceed 45% (Bossy et al. 2009, Göllner et al. 2010). A large variation in AGP composition between species as well as in the same species was previously published (Bobalek and Johnson 1983, Putoczki et al. 2007). The analysis of all 44 annotated

AGPs in the *Solanum lycopersicum* genome itself proved great variability within their composition. Interestingly, although the putative classical AGPs identified in *Ulva mutabilis* genome had similar content of proline, alanine, serine and threonine as *Solanum lycopersicum* classical AGPs, there were significant differences in content of remaining amino acids (Fig. S1). However, the presence of signal and GPI-anchor sequence was not taken into the consideration and therefore the real compositions might be slightly altered. Although the hydroxyproline content of examined AGP-like glycoproteins might seem quite low, AGPs with much lower content or even completely lacking hydroxyproline were already described (Bobalek and Johnson 1983, Baldwin et al. 1993, Happ and Classen 2019, Pfeifer et al. 2020). Besides, the concentrations of various amino acids can be altered during this complex experimental procedure, for example, due to hydrolysis or oxidation. All these shortcomings are nicely reviewed and should be taken into account during result interpretation (Golaz et al. 1996, Fountoulakis and Lahm 1998, Rutherford and Gilani 2009). Nevertheless, altogether the data based on PAST a Hyp content are suggesting that purified glycoproteins belong to non-classical AGP-like glycoproteins.

The unique monosaccharide composition of *Ulva lactuca* β -D-glucosyl Yariv precipitate further supports this hypothesis. Unexpectedly, the amount of saccharides in *Ulva lactuca* AGP-like glycoproteins was much lower than of *Solanum lycopersicum*. The possible explanations might be less extensive glycosylation of *Ulva lactuca* AGP-like glycoproteins or incomplete hydrolysis due to their unique glycan structure. The degradation of labile sugars during acid hydrolysis as well as variation in sugar composition among different species based on different types of depolymerization and analytical method has been previously described (Willför et al. 2009). For example, mild acid hydrolysis of *Zostera marina* AGPs led to the loss of arabinofuranose (Pfeifer et al. 2020). The unusual saccharide composition of *Ulva* Yariv precipitate might be also explained by interaction of AGP-like glycoproteins with the cell wall polysaccharides as previously described (Immerzeel et al. 2006, Tan et al. 2013, Hijazi et al. 2014). However, when the results are compared to previously published composition of water-soluble polysaccharides of *Ulva* spp., the contents of neutral saccharides differ greatly (Siddhanta et al. 2001, Yaich et al. 2013, Tabarsa et al. 2018, Gao et al. 2020, Wahlström et al. 2020). Moreover, since AGPs are also known to contain uronic acids, and in some cases, their content was substantial, the content of uronic acids was assessed (Serpe and Nothnagel 1995, Zhao et al. 2002). The ratio between neutral saccharides and uronic acid changed significantly after Yariv purification. In the case of *Ulva lactuca*, the decrease in uronic acid content results from removal

of uronic acid-rich polysaccharide, ulvan. The uronic acid content of well-described *Solanum lycopersicum* LeAGP-1 is quite high, which explains the increase in uronic acid content after Yariv purification (Zhao et al. 2002). Uronic acid content varies greatly in between different species, developmental stages and even individual AGPs (Pellerin et al. 1995, Leszczuk et al. 2020). For instance, *Zostera marina* AGPs contained high amounts of glucuronic acid and terminal methylated glucuronic acid, rare to land plant AGPs (Pfeifer et al. 2020). Uronic acids are responsible for calcium binding capacity of AGPs, creating pH dependant periplasmic AGP–Ca²⁺ capacitor, which can be fine-tuned according to environmental factors for example by amount of arabinose and rhamnose within AGP glycan (Lamport and Várnai 2012). Unfortunately, the heterogeneity and complexity of AGP glycans make their structural and functional analysis extremely demanding (Ma et al. 2018, Su and Higashiyama 2018). Therefore, information regarding glycosylation of AGPs in algae is very scarce and so far, only Estevéz et al. (2009) found characteristic type-II AGP-glycan structures within *Codium fragile* AGPs, nevertheless, the exact glycan structure remains unknown. Unusual glycan structures and monosaccharide composition were previously described in the cell wall fraction precipitated by β -D-glucosyl Yariv reagent of *Physcomitrella patens* or *Zostera marina* (Fu et al. 2007, Pfeifer et al. 2020). Our data suggest that *Ulva lactuca* AGP-like glycoprotein glycosylation differs from *Codium fragile*, but this hypothesis must be further examined.

Outstanding was the presence of 3-*O*-methyl-hexose within the glycan structure of Yariv purified *Ulva lactuca* AGP-like glycoproteins, which has never been found in any AGPs. This discovery is important for further understanding of the evolution of AGPs. Methylation of AGP glycans was previously described in higher plants, where β -(1,6)-galactan side chains are often terminated by 4-*O*-methyl glucuronic acid (Temple et al. 2019, Pfeifer et al. 2020). Unique 3-*O*-methyl-rhamnose was found in AGPs of various mosses, ferns, and liverwort *Marchantia polymorpha* (Bartels and Classen 2017, Bartels et al. 2017, Happ and Classen 2019). Although no information regarding the presence of methylated hexose in AGPs can be found, exceptional 6-methyl-galactofuranose was identified in *Chlamydomonas reinhardtii* extensins (Bollig et al. 2007). Methylated hexoses are also widely distributed within *N*-glycans of algal glycoproteins (Staudacher 2012). For instance, 6-*O*-methyl mannose can be found in *Porphyridium* sp. and *Chlamydomonas reinhardtii*, whereas 3-*O*-methyl and even 3,6-di-*O*-methyl mannoses were described in *Chlorella vulgaris* (Levy-Ontman et al. 2011, Mathieu-Rivet et al. 2013, Mócsai et al. 2019). Methylated xylose and rhamnose were identified in extracellular proteoglycan of *Rhodella*

grisea (Capek et al. 2008). Nevertheless, if we do not focus strictly on glycoproteins, the most commonly methylated hexose in plants and algae is galactose (Staudacher 2012, Pfeifer and Classen 2020). Further studies will be needed to precisely identify the 3-*O*-methyl-hexose and the structure of *Ulva lactuca* AGP-like glycoproteins, but preliminary results suggest a unique glycosylation pattern.

In summary, this study is the first direct experimental evidence of the presence of AGP-like glycoproteins in *Ulva* sp. The suitable extraction protocol was established and an extraordinary comparison of the green algal and land plant AGPs composition was presented. All the data are suggesting a unique structure of *Ulva lactuca* AGP-like glycoproteins, especially regarding their glycosylation. 3-*O*-methyl-hexose was for the first time detected in Yariv precipitated AGP-like glycoproteins, however further analyses are needed for exact identification. Revealing the structure of *Ulva lactuca* AGP-like glycoproteins might help to understand the evolution of these indispensable plant glycoproteins.

ACKNOWLEDGMENTS

This work was supported by the grant of Specific university research (A1_FPBT_2020_001). We would like to thank the French Ministry of Foreign Affairs and the Centre Mondial de l'Innovation – Roullier for financial support. Thanks to Josef Cvačka from Mass Spectrometry Group at the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences for his kind help with MS analysis of unknown saccharide.

REFERENCES

- Alsufyani, T., Califano, G., Deicke, M., Grueneberg, J., Weiss, A., Engelen, A.H., Kwantes, M., Mohr, J.F., Ulrich, J.F. & Wichard T. 2020. Macroalgal–bacterial interactions: identification and role of thallusin in morphogenesis of the seaweed *Ulva* (Chlorophyta). *J. Exp. Bot.* 71:3340-9.
- Avci, U., Pattathil, S. & Hahn, M.G. 2012. Immunological approaches to plant cell wall and biomass characterization: immunolocalization of glycan epitopes. In Himmel, M. E. [Ed.] *Biomass Conversion: Methods and Protocols*. Humana Press, Totowa, NJ, pp. 73–82.
- Baldwin, T.C., McCann, M.C. & Roberts, K. 1993. A novel hydroxyproline-deficient arabinogalactan protein secreted by suspension-cultured cells of *Daucus carota* (purification and partial characterization). *Plant Physiol.* 103:115–23.

- Bartels, D., Baumann, A., Maeder, M., Geske, T., Heise, E.M., von Schwartzberg, K. & Classen, B. 2017. Evolution of plant cell wall: arabinogalactan-proteins from three moss genera show structural differences compared to seed plants. *Carbohydr. Polym.* 163:227–35.
- Bartels, D. & Classen, B. 2017. Structural investigations on arabinogalactan-proteins from a lycophyte and different monilophytes (ferns) in the evolutionary context. *Carbohydr. Polym.* 172:342–51.
- Blumenkrantz, N. & Aboe-Hansen, G. 1973. New method for quantitative determination of uronic acids. *Anal. Biochem.* 54:484–89.
- Bobalek, J.F. & Johnson, M.A. 1983. Arabinogalactan-proteins from douglas fir and loblolly pine. *Phytochemistry*. 22:1500–3.
- Bollig, K., Lamshöft, M., Schweimer, K., Marner, F.-J., Budzikiewicz, H. & Waffenschmidt, S. 2007. Structural analysis of linear hydroxyproline-bound O-glycans of *Chlamydomonas reinhardtii*—conservation of the inner core in *Chlamydomonas* and land plants. *Carbohydr. Res.* 342:2557–66.
- Borderies, G., le Béhec, M., Rossignol, M., Lafitte, C., Le Deunff, E., Beckert, M., Dumas, C. & Matthys-Rochon E. 2004. Characterization of proteins secreted during maize microspore culture: arabinogalactan proteins (AGPs) stimulate embryo development. *Eur. J. Cell Biol.* 83:205–12.
- Bossy, A., Blaschek, W. & Classen, B. 2009. Characterization and immunolocalization of arabinogalactan-proteins in roots of *Echinacea purpurea*. *Planta Med.* 75:1526–33.
- Brooks, S.A. 2009. Strategies for analysis of the glycosylation of proteins: current status and future perspectives. *Mol. Biotechnol.* 43:76–88.
- Capek, P., Matulová, M. & Combourieu, B. 2008. The extracellular proteoglycan produced by *Rhodella grisea*. *Int. J. Biol. Macromol.* 43:390–393.
- Classen, B., Baumann, A. & Utermohlen, J. 2019. Arabinogalactan-proteins in spore-producing land plants. *Carbohydr. Polym.* 210:215–24.
- DeClerck, O., Kao, S.-M., Bogaert, K.A., Blomme, J., Foflonker, F., Kwantes, M., Vancaester, E., Vanderstraeten, L., Aydogdu, E., Boesger, J., Califano, G., Charrier, B., Clewes, R., Del Cortona, A., D'Hondt, S., Fernandez-Pozo, N., Gachon, C.M., Hanikenne, M., Lattermann, L., Leliaert, F., Liu, X., Maggs, C.A., Popper, Z.A., Raven, J.A., Van Bel, M., Wilhelmsson, P.K.I., Bhattacharya, D., Coates, J.C., Rensing, S.A., Van Der Straeten, D.,

- Vardi, A., Sterck, L., Vandepoele, K., Van de Peer, Y., Wichard, T. & Bothwell, J. 2018. Insights into the evolution of multicellularity from the sea lettuce genome. *Curr. Biol.* 28:2921-33.
- Domozych, D., Ciancia, M., Fangel, J.U., Mikkelsen, M.D., Ulvskov, P. & Willats, W.G.T. 2012. The cell walls of green algae: a journey through evolution and diversity. *Front. Plant Sci.* 3:82.
- Domozych, D.S., Elliott, L., Kiemle, S.N. & Gretz, M.R. 2007. *Pleurotaenium trabecula*, a desmid of wetland biofilms: the extracellular matrix and adhesion mechanisms. *J. Phycol.* 43:1022-38.
- Domozych, D.S., Sørensen, I., Pettolino, F.A., Bacic, A. & Willats, W.G.T. 2010. The cell wall polymers of the charophycean green alga *Chara corallina*: immunobinding and biochemical screening. *Int. J. Plant Sci.* 171:345-61.
- Eder, M., Tenhaken, R., Driouich, A. & Lütz-Meindl, U. 2008. Occurrence and characterization of arabinogalactan-like proteins and hemicelluloses in *Micrasterias* (Streptophyta). *J. Phycol.* 44:1221-34.
- Ellis, M., Egelund, J., Schultz, C.J. & Bacic, A. 2010. Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol.* 153:403-19.
- Estevez, J.M., Fernández, P.V., Kasulin, L., Dupree, P. & Ciancia, M. 2009. Chemical and in situ characterization of macromolecular components of the cell walls from the green seaweed *Codium fragile*. *Glycobiology.* 19:212-28.
- Estevez, J.M., Leonardi, P.I. & Alberghina, J.S. 2008. Cell wall carbohydrate epitopes in the green alga *Oedogonium bharuchae* f. *minor* (Oedogoniales, Chlorophyta). *J. Phycol.* 44:1257-68.
- Fernández, P.V., Ciancia, M., Miravalles, A.B. & Estevez, J.M. 2010. Cell-wall polymer mapping in the coenocytic macroalga *Codium vermilara* (Bryopsidales, Chlorophyta). *J. Phycol.* 46:456-65.
- Fernández, P.V., Raffo, M.P., Alberghina, J. & Ciancia, M. 2015. Polysaccharides from the green seaweed *Codium decorticans*. Structure and cell wall distribution. *Carbohydr. Polym.* 117:836-44.
- Fountoulakis, M. & Lahm, H.-W. 1998. Hydrolysis and amino acid composition analysis of proteins. *J. Chromatogr. A* 826:109-34.

- Fragkostefanakis, S., Dandachi, F. & Kalaitzis, P. 2012. Expression of arabinogalactan proteins during tomato fruit ripening and in response to mechanical wounding, hypoxia and anoxia. *Plant Physiol. Biochem.* 52:112–8.
- Fu, H., Yadav, M.P. & Nothnagel, E.A. 2007. *Physcomitrella patens* arabinogalactan proteins contain abundant terminal 3-O-methyl-L-rhamnosyl residues not found in angiosperms. *Planta*. 226:1511-24.
- Gao, M., Kieliszewski, M.J., Lamport, D.T.A. & Showalter, A.M. 1999. Isolation, characterization and immunolocalization of a novel, modular tomato arabinogalactan-protein corresponding to the LeAGP-1 gene. *Plant J.* 18:43–55.
- Gao, X., Qu, H., Shan, S., Song, C., Baranenko, D., Li, Y. & Lu, W. 2020. A novel polysaccharide isolated from *Ulva Pertusa*: structure and physicochemical property. *Carbohydr. Polym.* 233:115849.
- Golaz, O., Wilkins, M.R., Sanchez, J.-C., Appel, R.D., Hochstrasser, D.F. & Williams, K.L. 1996. Identification of proteins by their amino acid composition: an evaluation of the method. *Electrophoresis* 17:573–9.
- Göllner, E.M., Blaschek, W. & Classen, B. 2010. Structural investigations on arabinogalactan-protein from wheat, isolated with Yariv reagent. *J. Agric. Food Chem.* 58:3621–6.
- Gorres, K.L. & Raines, R.T. 2010. Prolyl 4-hydroxylase. *Crit. Rev. Biochem. Mol. Biol.* 45:106–24.
- Guan, Y. & Nothnagel, E.A. 2004. Binding of arabinogalactan proteins by yariv phenylglycoside triggers wound-like responses in Arabidopsis cell cultures. *Plant Physiol.* 135:1346–66.
- Happ, K. & Classen, B. 2019. Arabinogalactan-proteins from the liverwort *Marchantia polymorpha* L., a member of a basal land plant lineage, are structurally different to those of angiosperms. *Plants* 8:460.
- He, J., Zhao, H., Cheng, Z., Ke, Y., Liu, J. & Ma, H. 2019. Evolution analysis of the fasciclin-like arabinogalactan proteins in plants shows variable fasciclin-AGP domain constitutions. *Int. J. Mol. Sci.* 20:1945.
- Hervé, C., Siméon, A., Jam, M., Cassin, A., Johnson, K.L., Salmeán, A.A., Willats, W.G.T., Doblin, M.S., Bacic, A. & Kloareg, B. 2015. Arabinogalactan proteins have deep roots in eukaryotes: identification of genes and epitopes in brown algae and their role in *Fucus serratus* embryo development. *New Phytol.* 209:1428–41.

- Hijazi, M., Roujol, D., Nguyen-Kim, H., del Rocio Cisneros Castillo, L., Saland, E., Jamet, E. & Albenne, C. 2014. Arabinogalactan protein 31 (AGP31), a putative network-forming protein in *Arabidopsis thaliana* cell walls? *Ann. Bot.* 114:1087-97.
- Immerzeel, P., Eppink, M.M., de Vries, S.C., Schols, H.A. & Voragen, A.G.J. 2006. Carrot arabinogalactan proteins are interlinked with pectins. *Physiol. Plant.* 128:18-28.
- Jamet, E., Boudart, G., Borderies, G., Charmont, S., Lafitte, C., Rossignol, M., Canut, H. & Pont-Lezica, R. 2008. Isolation of plant cell wall proteins. In Posch, A. [Ed.] *2D PAGE: Sample Preparation and Fractionation*. Humana Press, Totowa, NJ, pp. 187–201.
- Johnson, K.L., Cassin, A.M., Lonsdale, A., Wong, G.K.-S., Soltis, D.E., Miles, N.W., Melkonian, M., Melkonian, B., Deyholos, M.K., Leebens-Mack, J., Rothfels, C.J., Stevenson, D.W., Graham, S.W., Wang, X., Wu, S., Pires, J.C., Edger, P.P., Carpenter, E.J., Bacic, A., Doblin, M.S. & Schultz, C.J. 2017. Insights into the evolution of hydroxyproline-rich glycoproteins from 1000 plant transcriptomes. *Plant Physiol.* 174:904–21.
- Kadam, S.U., Álvarez, C., Tiwari, B.K. & O'Donnell, C.P. 2015. Chapter 9 - Extraction of biomolecules from seaweeds. In Tiwari, B. K. & Troy, D. J. [Eds.] *Seaweed Sustainability*. Academic Press, San Diego, USA, pp. 243–69.
- Kieliszewski, M.J., Lampion, D.T.A., Tan, L. & Cannon, M.C. 2010. Hydroxyproline-rich glycoproteins: form and function. In Ulvskov, P. [Ed.] *Annual Plant Reviews Volume 41*. John Wiley & Sons, Ltd., Chichester, UK, pp. 321–42.
- Kieliszewski, M.J. & Shpak, E. 2001. Synthetic genes for the elucidation of glycosylation codes for arabinogalactan-proteins and other hydroxyproline-rich glycoproteins. *Cell. Mol. Life Sci. CMLS*. 58:1386–98.
- Kitazawa, K., Tryfona, T., Yoshimi, Y., Hayashi, Y., Kawauchi, S., Antonov, L., Tanaka, H., Takahashi, T., Kaneko, S., Dupree, P., Tsumuraya, Y. & Kotake, T. 2013. β -Galactosyl Yariv reagent binds to the β -1,3-galactan of arabinogalactan proteins. *Plant Physiol.* 161:1117–26.
- Knox, J.P. 1995. Developmentally regulated proteoglycans and glycoproteins of the plant cell surface. *FASEB J.* 9:1004–12.
- Knox, J.P. 2008. Revealing the structural and functional diversity of plant cell walls. *Curr. Opin. Plant Biol.* 11:308–13.

- Kostas, E.T., White, D.A. & Cook, D.J. 2020. Bioethanol production from UK seaweeds: investigating variable pre-treatment and enzyme hydrolysis parameters. *BioEnergy Res.* 13:271–85.
- Lahaye, M. & Robic, A. 2007. Structure and functional properties of ulvan, a polysaccharide from green seaweeds. *Biomacromolecules.* 8:1766-74.
- Lamport, D.T.A. & Várnai, P. 2013. Periplasmic arabinogalactan glycoproteins act as a calcium capacitor that regulates plant growth and development. *New Phytol.* 197:58-64.
- Leszczuk, A., Cybulska, J., Skrzypek, T. & Zdunek, A. 2020. Properties of arabinogalactan proteins (AGPs) in apple (*Malus × Domestica*) fruit at different stages of ripening. *Biology* 9:225.
- Levy-Ontman, O., Arad, S. (Malis), Harvey, D.J., Parsons, T.B., Fairbanks, A. & Tekoah, Y. 2011. Unique N-glycan moieties of the 66-kDa cell wall glycoprotein from the red microalga *Porphyridium* sp. *J. Biol. Chem.* 286:21340–52.
- Lin, Z., Lin, Z., Li, H. & Shen, S. 2012. Sequences analysis of ITS region and 18S rDNA of *Ulva*. *Int. Scholarly Res. Not.* 2012:468193.
- Ludwig, T.G. & Goldberg, H.J.V. 1956. The anthrone method for the determination of carbohydrates in foods and in oral rinsing. *J. Dent. Res.* 35:90-94.
- Ma, Y., Yan, C., Li, H., Wu, W., Liu, Y., Wang, Y., Chen, Q. & Ma, H. 2017. Bioinformatics prediction and evolution analysis of arabinogalactan proteins in the plant kingdom. *Front. Plant Sci.* 8:66.
- Ma, Y., Zeng, W., Bacic, A. & Johnson, K. 2018. AGPs through time and space. In Roberts, J.A. [Ed.] *Annual Plant Reviews Online*. John Wiley & Sons, Ltd., Chichester, UK, pp. 767–804.
- Maneein, S., Milledge, J.J., Nielsen, B.V. & Harvey, P.J. 2018. A review of seaweed pre-treatment methods for enhanced biofuel production by anaerobic digestion or fermentation. *Fermentation* 4:100.
- Mareri, L., Romi, M. & Cai, G. 2019. Arabinogalactan proteins: actors or spectators during abiotic and biotic stress in plants? *Plant Biosyst.* 153:173–85.
- Mathieu-Rivet, E., Scholz, M., Arias, C., Dardelle, F., Schulze, S., Mauff, F.L., Teo, G., Hochmal, A.K., Blanco-Rivero, A., Loutelier-Bourhis, C., Kiefer-Meyer, M.C., Fufezan, C., Burel, C., Lerouge, P., Martinez, F., Bardor, M. & Hippler, M. 2013. Exploring the N-

- glycosylation pathway in *Chlamydomonas reinhardtii* unravels novel complex structures. *Mol. Cell. Proteomics* 12:3160–83.
- Michalak, I. & Chojnacka, K. 2015. Algae as production systems of bioactive compounds. *Eng. Life Sci.* 15:160–76.
- Mishra, V., Dubey, A. & Prajapati, S.K. 2017. Algal biomass pretreatment for improved biofuel production. In Gupta, S. K., Malik, A. & Bux, F. [Eds.] *Algal Biofuels: Recent Advances and Future Prospects*. Springer International Publishing, Cham, pp. 259–80.
- Mócsai, R., Figl, R., Troschl, C., Strasser, R., Svehla, E., Windwarder, M., Thader, A. & Altmann, F. 2019. N-glycans of the microalga *Chlorella vulgaris* are of the oligomannosidic type but highly methylated. *Sci. Rep.* 9:1–8.
- Nguema-Ona, E., Bannigan, A., Chevalier, L., Baskin, T.I. & Driouich, A. 2007. Disruption of arabinogalactan proteins disorganizes cortical microtubules in the root of *Arabidopsis thaliana*. *Plant J.* 52:240–51.
- Nguema-Ona, E., Coimbra, S., Vitré-Gibouin, M., Mollet, J.-C. & Driouich, A. 2012. Arabinogalactan proteins in root and pollen-tube cells: distribution and functional aspects. *Ann. Bot.* 110:383–404.
- Nguema-Ona, E., Vitré-Gibouin, M., Gotté, M., Plancot, B., Lerouge, P., Bardor, M. & Driouich, A. 2014. Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. *Front. Plant Sci.* 5:499.
- Nothnagel, E.A. 1997. Proteoglycans and related components in plant cells. In Jeon, K. W. [Ed.] *International Review of Cytology*. Academic Press, pp. 195–291.
- Palacio-López, K., Tinaz, B., Holzinger, A. & Domozych, D.S. 2019. Arabinogalactan proteins and the extracellular matrix of charophytes: a sticky business. *Front. Plant Sci.* 10:447.
- Pattathil, S., Avci, U., Baldwin, D., Swennes, A.G., McGill, J.A., Popper, Z., Bootten, T., Albert, A., Davis, R.H., Chennareddy, C., Dong, R., O'Shea, B., Rossi, R., Leoff, C., Freshour, G., Narra, R., O'Neil, M., York, W.S. & Hahn, M.G. 2010. A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiol.* 153:514–25.
- Paulsen, B.S., Craik, D.J., Dunstan, D.E., Stone, B.A. & Bacic, A. 2014. The Yariv reagent: behaviour in different solvents and interaction with a gum arabic arabinogalactanprotein. *Carbohydr. Polym.* 106:460–8.

- Pellerin, P., Vidal, S., Williams, P. & Brillouet, J.-M. 1995. Characterization of five type II arabinogalactan-protein fractions from red wine of increasing uronic acid content. *Carbohydr. Res.* 277:135-43.
- Pfeifer, L. & Classen, B. 2020. Validation of a rapid GC-MS Procedure for quantitative distinction between 3-*O*-methyl- and 4-*O*-methyl-hexoses and its application to a complex carbohydrate sample. *Separations* 7:42.
- Pfeifer, L., Shafee, T., Johnson, K.L., Bacic, A. & Classen, B. 2020. Arabinogalactan-proteins of *Zostera marina* L. contain unique glycan structures and provide insight into adaption processes to saline environments. *Sci. Rep.* 10:8232.
- Polikovskiy, M., Califano, G., Dunger, N., Wichard, T. & Golberg, A. 2020. Engineering bacteria-seaweed symbioses for modulating the photosynthate content of *Ulva* (Chlorophyta): significant for the feedstock of bioethanol production. *Algal Res.* 49 :101945.
- Popper, Z.A., Michel, G., Hervé, C., Domozych, D.S., Willats, W.G.T., Tuohy, M.G., Kloareg, B. & Stengel, D.B. 2011. Evolution and diversity of plant cell walls: from algae to flowering plants. *Annu. Rev. Plant Biol.* 62:567–90.
- Printz, B., Dos Santos Morais, R., Wienkoop, S., Sergeant, K., Lutts, S., Hausman, J. F. & Renaut, J. 2015. An improved protocol to study the plant cell wall proteome. *Front. Plant Sci.* 6:237.
- Putoczki, T.L., Pettolino, F., Griffin, M.D.W., Möller, R., Gerrard, J.A., Bacic, A. & Jackson, S.L. 2007. Characterization of the structure, expression and function of *Pinus radiata* D. Don arabinogalactan-proteins. *Planta*. 226:1131-42.
- Raven, J.A. 2003. Long-distance transport in non-vascular plants. *Plant Cell Environ.* 26:73–85.
- Ruprecht, C., Bartetzko, M.P., Senf, D., Dallabernadina, P., Boos, I., Andersen, M.C.F., Kotake, T., Knox, J.P., Hahn, M.G., Clausen, M.H. & Pfrengle, F. 2017. A synthetic glycan microarray enables epitope mapping of plant cell wall glycan-directed antibodies. *Plant Physiol.* 175:1094–104.
- Rutherford, S.M. & Gilani, G.S. 2009. Amino acid analysis. *Curr. Protoc. Protein Sci.* 58:11.9.1-11.9.37.
- Rydahl, M.G., Hansen, A.R., Kračun, S.K. & Mravec, J. 2018. Report on the current inventory of the toolbox for plant cell wall analysis: proteinaceous and small molecular probes. *Front. Plant Sci.* 9:581.

- Sato, K., Hara, K., Yoshimi, Y., Kitazawa, K., Ito, H., Tsumuraya, Y. & Kotake, T. 2018. Yariv reactivity of type II arabinogalactan from larch wood. *Carbohydr. Res.* 467:8–13.
- Schultz, C.J., Johnson, K.L., Currie, G. & Bacic, A. 2000. The classical arabinogalactan protein gene family of Arabidopsis. *Plant Cell.* 12:1751–67.
- Seifert, G.J. & Roberts, K. 2007. The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* 58:137–61.
- Segev, E., Wyche, T.P., Kim, K.H., Petersen, J., Ellebrandt, C., Vlamakis, H., Barteneva, N., Paulson, J.N., Chai, L., Clardy, J. & Kolter, R. 2016. Dynamic metabolic exchange governs a marine algal-bacterial interaction. *eLife.* 5:e17473.
- Serpe, M.D. & Nothnagel, E.A. 1995. Fractionation and structural characterization of arabinogalactan-proteins from the cell wall of rose cells. *Plant Physiol.* 109:1007–16.
- Showalter, A.M. 1993. Structure and function of plant cell wall proteins. *Plant Cell* 5:9–23.
- Showalter, A.M. 2001. Arabinogalactan-proteins: structure, expression and function. *Cell. Mol. Life Sci. CMLS* 58:1399–417.
- Showalter, A.M., Keppler, B., Lichtenberg, J., Gu, D. & Welch, L.R. 2010. A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. *Plant Physiol.* 153:485–513.
- Siddhanta, A.K., Goswami, A.M., Ramavat, B.K., Mody, K.H. & Mairh, O.P. 2001. Water soluble polysaccharides of marine algal species of *Ulva* (Ulvales, Chlorophyta) of Indian waters. *Indian J. Mar. Sci.* 30:166-72.
- Sørensen, I., Pettolino, F.A., Bacic, A., Ralph, J., Lu, F., O'Neill, M.A., Fei, Z., Rose, J.K.C., Domozych, D.S. & Willats, W.G.T. 2011. The charophycean green algae provide insights into the early origins of plant cell walls. *Plant J.* 68:201–11.
- Spoerner, M., Wichard, T., Bachhuber, T., Stratmann, J. & Oertel, W. 2012. Growth and thallus morphogenesis of *Ulva mutabilis* (Chlorophyta) depends on a combination of two bacterial species excreting regulatory factors. *J. Phycol.* 48:1433–47.
- Stanley, M.S., Perry, R.M. & Callow, J.A. 2005. Analysis of expressed sequence tags from the green alga *Ulva linza* (Chlorophyta). *J. Phycol.* 41:1219–26.
- Staudacher, E. 2012. Methylation – an uncommon modification of glycans. *Biol. Chem.* 393:675–85.
- Su, S. & Higashiyama, T. 2018. Arabinogalactan proteins and their sugar chains: functions in plant reproduction, research methods, and biosynthesis. *Plant Reprod.* 31:67–75.

- Sudhakar, K., Mamat, R., Samykano, M., Azmi, W.H., Ishak, W.F.W. & Yusaf, T. 2018. An overview of marine macroalgae as bioresource. *Renew. Sustain. Energy Rev.* 91:165–79.
- Suganya, T., Varman, M., Masjuki, H.H. & Renganathan, S. 2016. Macroalgae and microalgae as a potential source for commercial applications along with biofuels production: abiorefinery approach. *Renew. Sustain. Energy Rev.* 55:909–41.
- Tabarsa, M., You, S., Dabaghian, E.H. & Surayot, U. 2018. Water-soluble polysaccharides from *Ulva intestinalis*: molecular properties, structural elucidation and immunomodulatory activities. *J. Food Drug Anal.* 26:599-608.
- Tan, L., Leykam, J.F. & Kieliszewski, M.J. 2003. Glycosylation motifs that direct arabinogalactan addition to arabinogalactan-proteins. *Plant Physiol.* 132:1362–9.
- Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan C., Hao, Z., Zhu, X., Avci, U., Miller, J.S., Baldwin, D., Pham, C., Orlando, R., Darvill, A., Hahn, M.G., Kieliszewski, M.J. & Mohnen, D. 2013. An Arabidopsis cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* 25:270-87.
- Tang, X. C., He, Y. Q., Wang, Y. & Sun, M. X. 2006. The role of arabinogalactan proteins binding to Yariv reagents in the initiation, cell developmental fate, and maintenance of microspore embryogenesis in *Brassica napus* L. cv. Topas. *J. Exp. Bot.* 57:2639–50.
- Temple, H., Mortimer, J.C., Tryfona, T., Yu, X., Lopez-Hernandez, F., Sorieul, M., Anders, N. & Dupree, P. 2019. Two members of the DUF579 family are responsible for arabinogalactan methylation in Arabidopsis. *Plant Direct* 3:1-4.
- Tryfona, T., Liang, H. C., Kotake, T., Tsumuraya, Y., Stephens, E. & Dupree, P. 2012. Structural characterization of *Arabidopsis* leaf arabinogalactan polysaccharides. *Plant Physiol.* 160:653–66.
- Van Hengel, A.J., Van Kammen, A. & De Vries, S.C. 2002. A relationship between seed development, arabinogalactan-proteins (AGPs) and the AGP mediated promotion of somatic embryogenesis. *Physiol. Plant.* 114:637–44.
- van Holst, G. J. & Clarke, A.E. 1985. Quantification of arabinogalactan-protein in plant extracts by single radial gel diffusion. *Anal. Biochem.* 148:446–50.
- Vieira, H.H., Bagatini, I.L., Guinart, C.M. & Vieira A.A.H. 2016. *tufA* gene as molecular marker for freshwater Chlorophyceae. *Algae* 31:155-65.
- Wahlström, N., Edlund, U., Pavia, H., Toth, G., Jaworski, A., Pell, A.J., Choong, F.X., Shirani, H., Nilsson, K.P.R. & Richter-Dahlfors, A. 2020. Cellulose from the green macroalgae *Ulva*

lactuca: isolation, characterization, optotracing, and production of cellulose nanofibrils. *Cellulose* 27:3707-25.

Westermeyer, R. 2016. Protein detection. In Westermeyer, R. [Ed.] *Electrophoresis in Practice*. John Wiley & Sons, Ltd., Chichester, UK, pp. 131–64.

Wichard, T., Charrier, B., Mineur, F., Bothwell, J.H., Clerck, O.D. & Coates, J.C. 2015. The green seaweed *Ulva*: a model system to study morphogenesis. *Front. Plant Sci.* 6:72.

Wijesekara, I., Lang, M., Marty, C., Gemin, M.P., Boulho, R., Douzenel, P., Wickramasinghe, I., Bedoux, G. & Bourgougnon, N. 2017. Different extraction procedures and analysis of protein from *Ulva* sp. in Brittany, France. *J. Appl. Phycol.* 29:2503–11.

Willför, S., Pranovich, A., Tamminen, T., Puls, J., Laine, C., Suurnäkki, A., Saake, B., Uotila, K., Simolin, H., Hemming, J. & Holmbom, B. 2009. Carbohydrate analysis of plant materials with uronic acid-containing polysaccharides—A comparison between different hydrolysis and subsequent chromatographic analytical techniques. *Ind. Crops Prod.* 29:571–80.

Xu, J., Yue, R. Q., Liu, J., Ho, H. M., Yi, T., Chen, H. B. & Han, Q. B. 2014. Structural diversity requires individual optimization of ethanol concentration in polysaccharide precipitation. *Int. J. Biol. Macromol.* 67:205–9.

Yaich, H., Garna, H., Besbes, S., Paquot, M., Blecker, C. & Attia, H. 2013. Effect of extraction conditions on the yield and purity of ulvan extracted from *Ulva lactuca*. *Food Hydrocolloids*. 31:375-82.

Yariv, J., Rapport, M.M. & Graf, L. 1962. The interaction of glycosides and saccharides with antibody to the corresponding phenylazo glycosides. *Biochem. J.* 85:383–8.

Yates, E.A., Valdor, J.F., Haslam, S.M., Morris, H.R., Dell, A., Mackie, W. & Knox, J.P. 1996. Characterization of carbohydrate structural features recognized by anti-arabinogalactan-protein monoclonal antibodies. *Glycobiology* 6:131-9.

Zhao, Z.D., Tan, L., Showalter, A.M., Lamport, D.T.A. & Kieliszewski, M.J. 2002. Tomato LeAGP-1 arabinogalactan-protein purified from transgenic tobacco corroborates the Hyp contiguity hypothesis. *Plant J.* 31:431–44.

Figure 1: Immunodot-blot results of extracts after AGP extraction according to Schultz et al. (2000) of *Ulva lactuca* and *Solanum lycopersicum* probed with anti-AGP monoclonal primary antibodies JIM13, JIM16 and MAC207. Exposition time 1 minute. (-) No signal detected; (+/-) almost no signal detected; (+) weak signal detected; (++) medium signal detected; (+++) strong signal detected.

Figure 2: SDS-PAGE and western blot results of different extraction protocols to compare the content of AGP-like glycoproteins. (A) Total protein content visualized by stain-free detection; (B) immunolabeling with anti-AGP JIM16 primary antibody; (C) immunolabeling of *Solanum lycopersicum* extract with anti-AGP JIM13 primary antibody. (1) Classical AGP extraction according to Schultz et al. (2000); (2) 50 mM MES pH 6, 24 h at 4°C; (3) 0.2 M CaCl₂ in 50 mM MES pH 6, 24 h at 4°C; (4) Extraction from tomato: 0.2 M CaCl₂ in 50 mM MES pH 6, 24 h at 4°C. The exposition time of western blots was 10 s.

Figure 3: Yariv gel-diffusion assay of various extracts. 1% NaCl served as negative control, while gum arabic (concentrations 0.25, 0.5, 1 mg · mL⁻¹) as positive control. Schultz: *Ulva lactuca* extract (1.2 mg per well) prepared according to Schultz et al. (2000); MES CaCl₂: *Ulva lactuca* extract (1.2 mg per well), 0.2 M CaCl₂ in MES pH 6 for 24 h at 4°C; Tomato: tomato extract (0.3 mg per well) used as land plant control of extraction, 0.2 M CaCl₂ in 50 mM MES pH 6 for 24 h at 4°C.

Figure 4: SDS-PAGE and western blot analysis of β-D-glucosyl Yariv purification of *Ulva lactuca* and *Solanum lycopersicum* extract. (A) Total protein content visualized by stain-free detection; (B) immunolabeling with anti-AGP JIM16 primary antibody, 1-minute exposition time; (C) immunolabeling of *Solanum lycopersicum* extracts with anti-AGP JIM13 primary antibody, 30 s exposition time; (D) total protein content visualized by silver staining. (1) *Ulva lactuca* extract: 0.2 M CaCl₂ in 50 mM MES pH 6, 24 h at 4°C; (2) β-D-glucosyl Yariv purified *Ulva lactuca* extract; (3) *Solanum lycopersicum* extract: 0.2 M CaCl₂ in 50 mM MES pH 6, 24 h at 4°C; (4) β-D-glucosyl Yariv purified *Solanum lycopersicum* extract. Volume loaded on gel: 4 μL of extracts and 9 μL of Yariv purified samples (2 mg · mL⁻¹). Highlighted areas were chosen for amino acid composition analysis.

Figure 5: The comparison of the most prevalent amino acid (alanine, serine, threonine, proline and their sum known as PAST) contents in land plant AGPs. The amino acid contents of β -D-glucosyl Yariv precipitated AGP-like glycoproteins after electrophoretic separation from *Ulva lactuca* (U1 and U2) and *Solanum lycopersicum* (T1 and T2) were compared with genomic data. AGP-like glycoproteins U1, U2, T1 and T2 were cut out of PVDF membranes after western blotting. U1 and T1 represent glycoproteins with higher molecular weight, while U2 and T2 with lower molecular weight corresponding to the highlighted areas in Figure 4. The experimental data were compared with 44 AGPs annotated in *Solanum lycopersicum* genome, which were further divided to three groups to classical AGPs (T Clas.), fasciclin-like AGPs (T FLA) and other non-classical AGPs (T Non-Clas.). The analysis of *Ulva mutabilis* genome identified 6 putative classical AGPs, which were added to the comparison (U Clas.). The genomic data are visualized as Tukey's boxplots, while experimental data are presented as the mean \pm standard error of the mean (n = 5, collected from three independent experiments).

Figure 6: GC-MS spectra of alditol acetate derivative of Unknown1 saccharide from *Ulva lactuca* β -D-glucosyl Yariv precipitate (**A**) compared to GC-MS spectra of alditol acetate derivative of 3-O-methyl-glucose standard (**B**) based on Pfeifer and Classen (2020).

Table S1: Putative classical AGPs identified in *Ulva mutabilis* genome and used for comparison of amino acid contents.

Table S2: AGPs annotated in *Solanum lycopersicum* genome used for comparison of amino acid contents.

Figure S1: The comparison of the amino acid contents (except Ala, Ser, Thr and Pro) of β -D-glucosyl Yariv precipitated AGP-like glycoproteins from *Ulva lactuca* (U1 and U2) and *Solanum lycopersicum* (T1 and T2) with genomic data. AGP-like glycoproteins U1, U2, T1 and T2 were cut out of PVDF membranes after western blotting. U1 and T1 represent glycoproteins with higher molecular weight, while U2 and T2 with lower molecular weight corresponding to the highlighted areas in Figure 4. The experimental data were compared with 44 AGPs annotated in *Solanum lycopersicum* genome, which were further divided to three groups to classical AGPs (T Clas.),

fasciclin-like AGPs (T FLA) and other non-classical AGPs (T Non-Clas.). The analysis of *Ulva mutabilis* genome identified 6 putative classical AGPs, which were added to the comparison (U Clas.). Distinction between glutamic/aspartic acid and their amides plus determination of tryptophan content is impossible due to the acid hydrolysis. In addition, cysteine cannot be analyzed as a free amino acid because of side reaction during hydrolysis (Fountoulakis and Lahm 1998). The genomic data are visualized as Tukey's boxplots, while experimental data are presented as the mean \pm standard error of the mean ($n = 5$, collected from three independent experiments).

Figure S2: The composition of collected Unknown1 saccharide present in *Ulva lactuca* β -D-glucosyl Yariv precipitate was verified using system HPAEC/PAD with column Dionex CarboPac MA1 run in isocratic mode with 480 mM NaOH.

Figure S3: Positive ESI-MS spectra of unknown saccharide from *Ulva lactuca* β -D-glucosyl Yariv precipitate measured using LTQ Orbitrap XL hybrid mass spectrometer equipped with an electrospray ion source. For the positive ion mode, spray voltage, capillary voltage, tube lens voltage, and capillary temperature were 4.8 kV, 9 V, 150 V, and 275°C, respectively.

Figure S4: Tandem mass spectrometry (MS/MS) with collision induced dissociation (CID) spectra of (A) $[M + Na]^+$ ion with m/z 217; (B) $[M + NH_4]^+$ ion with m/z 212; (C) $[M - H]^-$ ion with m/z 193 of (1) desalted unknown1 saccharide from *Ulva lactuca* β -D-glucosyl Yariv precipitate and (2) 3-O-methyl-glucose standard. Spectra were measured in a linear ion trap using LTQ Orbitrap XL hybrid mass spectrometer equipped with an electrospray ion source.

Table S3: Influence of salts on the extraction of AGPs.

Table 1: Neutral monosaccharide composition (mass % of total neutral saccharides) of *Ulva lactuca* and *Solanum lycopersicum* extract and β -D-glucosyl Yariv precipitate (YP). Values in the table represent mean \pm standard error of the mean (n = 9, collected from three independent experiments).

	<i>Ulva</i> extract	<i>Ulva</i> YP	Tomato extract	Tomato YP	Retention time* [min]
Fucose	traces ^c	2.4 \pm 0.3	1.0 \pm 0.2	0.5 \pm 0.1	4.78
Unknown1 ^a	15.5 \pm 0.8	27.0 \pm 1.5	n.d. ^d	n.d.	6.12
Rhamnose	64.8 \pm 1.3	17.7 \pm 1.1	8.0 \pm 0.5	9.9 \pm 0.4	7.50
Arabinose	0.7 \pm 0.1	1.8 \pm 0.1	22.8 \pm 0.6	26.2 \pm 0.6	8.55
Unknown2 ^b	2.1 \pm 0.3	3.2 \pm 0.2	2.7 \pm 0.3	traces	9.90
Galactose	3.4 \pm 0.1	16.0 \pm 0.6	43.3 \pm 0.4	61.3 \pm 0.6	11.10
Glucose	4.0 \pm 0.5	13.4 \pm 1.3	12.2 \pm 0.5	1.3 \pm 0.1	11.92
Mannose	1.7 \pm 0.2	16.6 \pm 0.1	3.6 \pm 0.3	traces	12.70
Xylose	7.5 \pm 0.3	2.0 \pm 0.1	6.4 \pm 0.2	traces	13.18

^a Approximate content, concentration calculated with coefficient of 3-*O*-methyl-glucose.

^b Approximate content, concentration calculated with average coefficient of all standards.







^c Traces: content < 0.5%.

^d n.d.: not detected

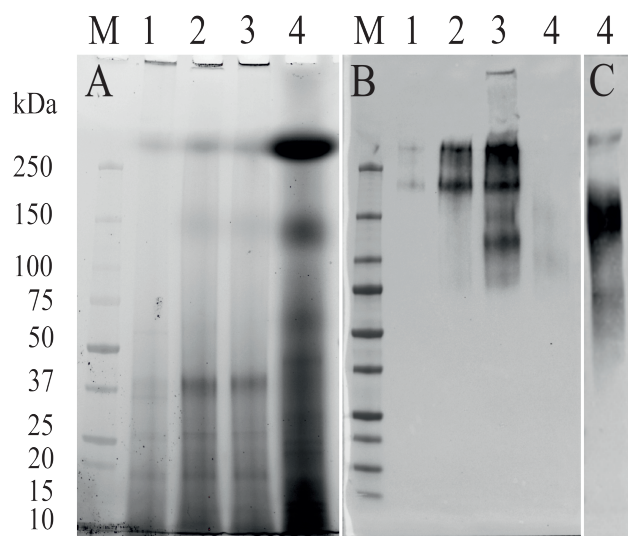
* Retention times for column CarboPac PA1 used for their separation in isocratic mode with 20 mM NaOH.

Table 2: Retention times of standard sugars and unknown1 saccharide on anion-exchange column CarboPac MA1 run in isocratic mode with 480 mM NaOH.

Standard	t_R [min]
Glycerol	8.70
Unknown1	12.28
Rhamnose	12.55
Fucose	12.83
Mannosamine	13.90
Glucosamine	14.10
Trehalose	14.45
Ribitol	15.30
Mannitol	16.30
Mannose	17.70
Arabinose	17.80
Glucose	19.20
Xylose	19.95
Galactose	21.08
Lactulose	21.30
Maltitol	21.58
Isomaltose	22.68
Fructose	22.75
Lactose	22.90
Ribose	24.60
Sucrose	32.60
Raffinose	36.60
Maltose	43.10
Maltotriose	81.20

Antibody	JIM13		JIM16		MAC207	
<i>Ulva lactuca</i>	-		++		+	
Tomato	++		++		+/-	

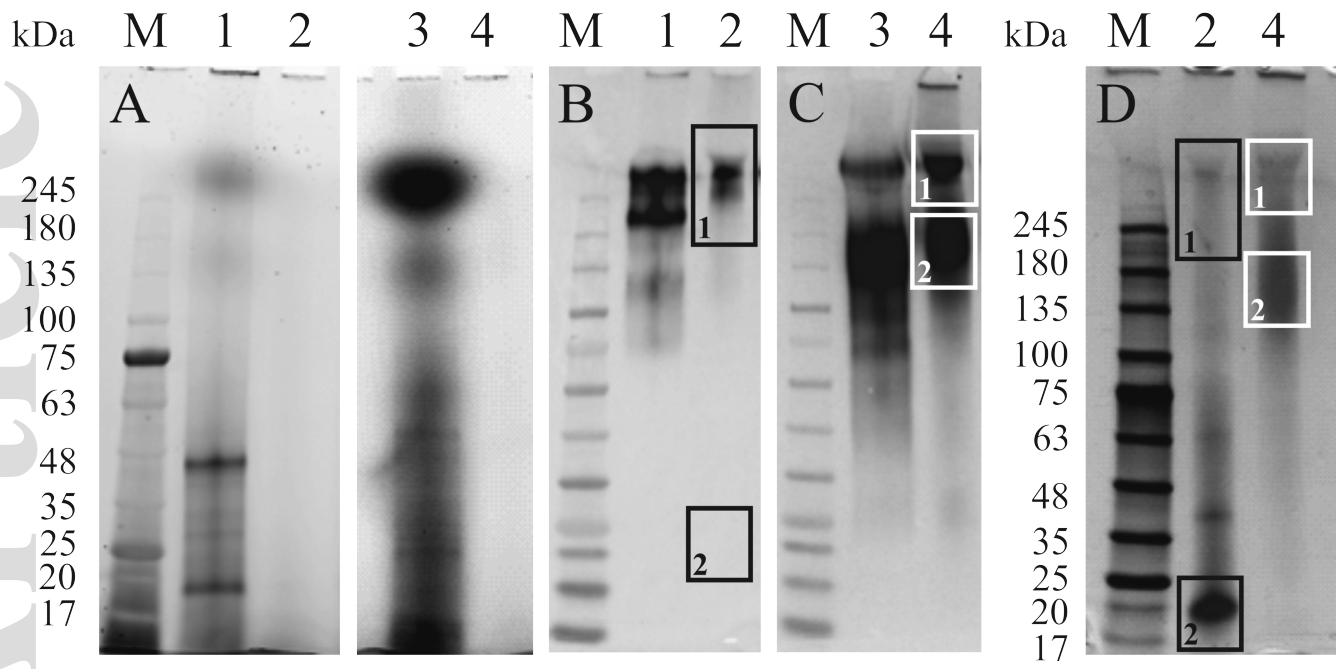
jpy_13121-20-139_f1.tif



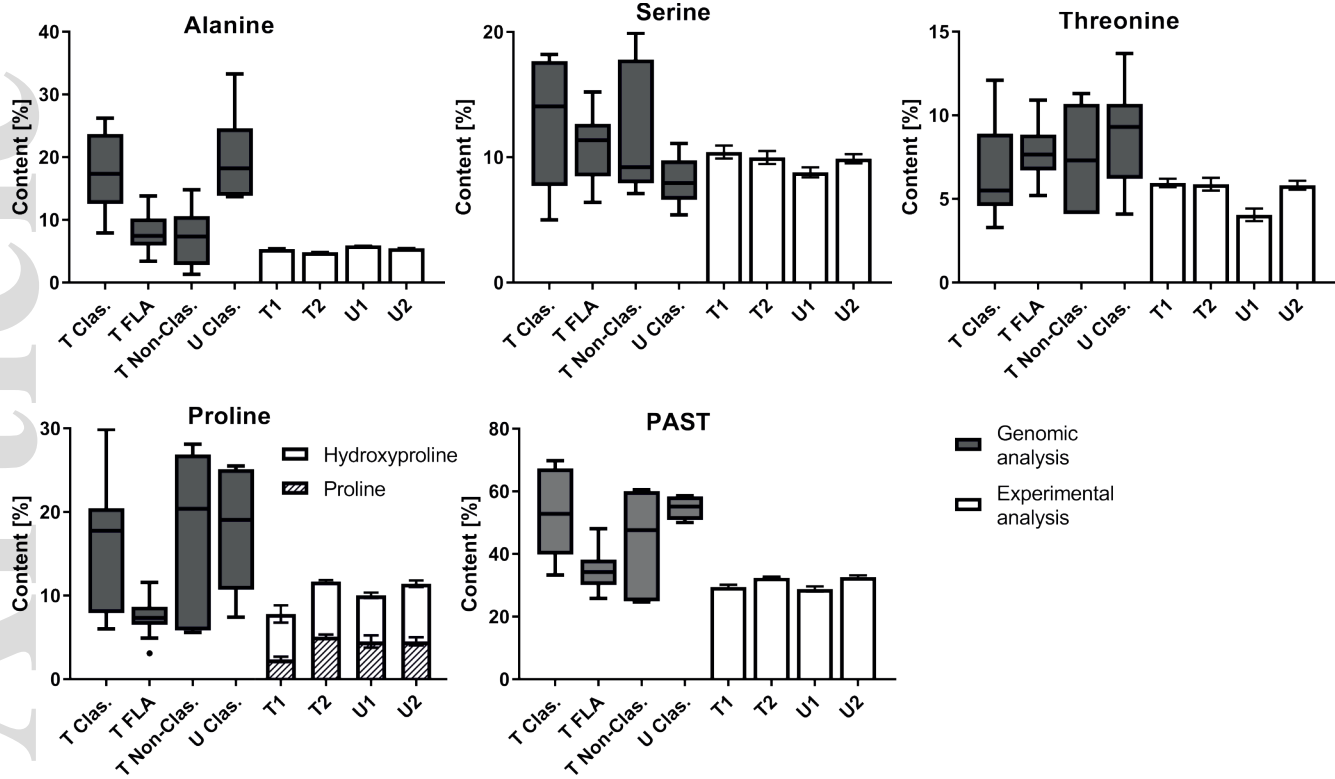
jpy_13121-20-139_f2.tif

	1% NaCl	Gum arabic [mg·mL ⁻¹]			Schultz	MES CaCl ₂	Tomato
		0.25	0.5	1			
β-D-Galactosyl Yariv reagent							
β-D-Glucosyl Yariv reagent							

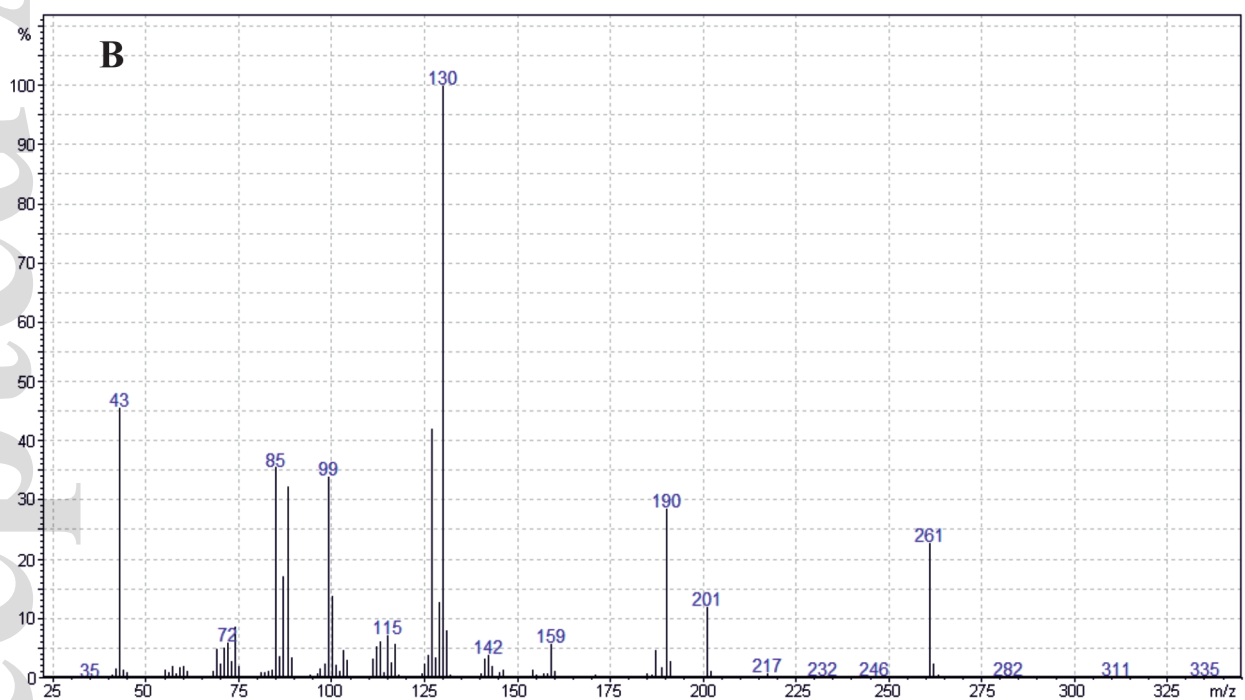
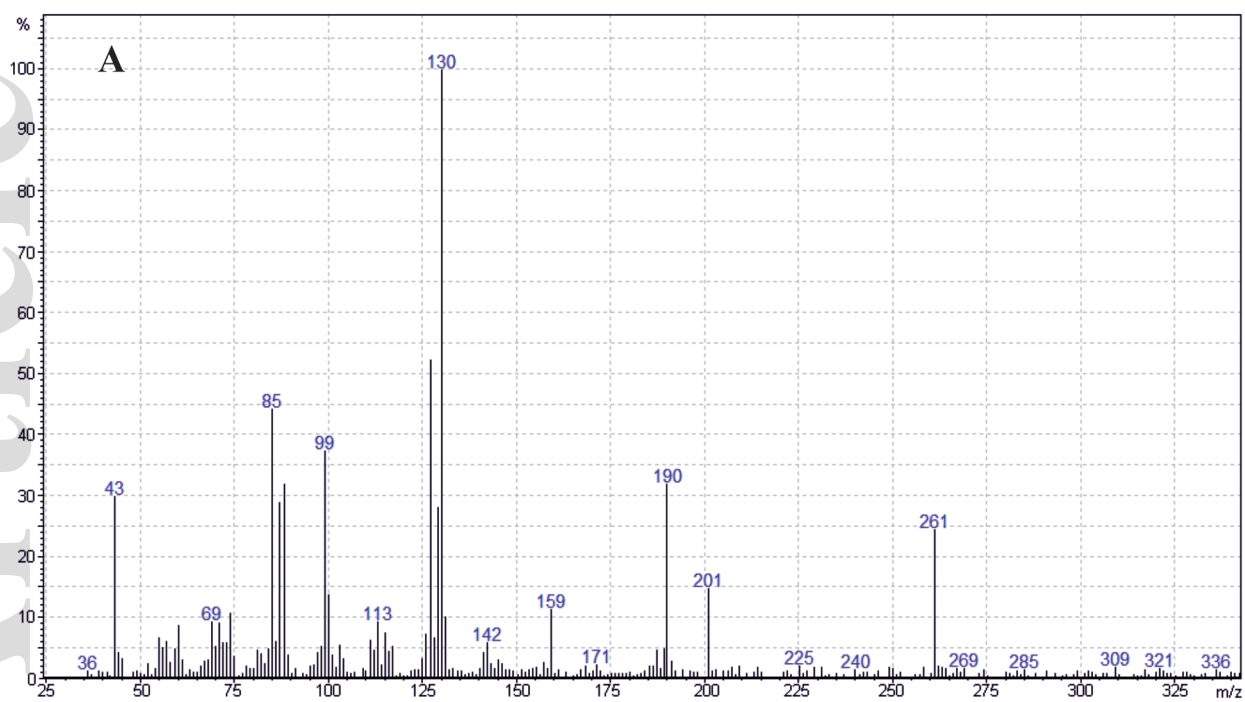
jpy_13121-20-139_f3.tif



jpy_13121-20-139_f4.tif



jpy_13121-20-139_f5.tif



jpy_13121-20-139_f6.tif