



OPEN

# Arabinogalactan-proteins of *Zostera marina* L. contain unique glycan structures and provide insight into adaption processes to saline environments

Lukas Pfeifer<sup>1</sup>, Thomas Shafee<sup>2</sup>, Kim L. Johnson<sup>2,3</sup>, Antony Bacic<sup>2,3</sup> & Birgit Classen<sup>1</sup>✉

Seagrasses evolved from monocotyledonous land plants that returned to the marine habitat. This transition was accomplished by substantial changes in cell wall composition, revealing habitat-driven adaption to the new environment. Whether arabinogalactan-proteins (AGPs), important signalling molecules of land plants, are present in seagrass cell walls is of evolutionary and plant development interest. AGPs of *Zostera marina* L. were isolated and structurally characterised by analytical and bioinformatics methods as well as by ELISA with different anti-AGP antibodies. Calcium-binding capacity of AGPs was studied by isothermal titration calorimetry (ITC) and microscopy. Bioinformatic searches of the *Z. marina* proteome identified 9 classical AGPs and a large number of chimeric AGPs. The glycan structures exhibit unique features, including a high degree of branching and an unusually high content of terminating 4-O-methyl-glucuronic acid (4-OMe GlcA) residues. Although the common backbone structure of land plant AGPs is conserved in *Z. marina*, the terminating residues are distinct with high amounts of uronic acids. These differences likely result from the glycan-active enzymes (glycosyltransferases and methyltransferases) and are essential for calcium-binding properties. The role of this polyanionic surface is discussed with regard to adaption to the marine environment.

Seagrasses evolved from early monocotyledonous land plants, returning to the sea around 140 million years ago. Today, these marine angiosperms are a polyphyletic group containing about 60 species in four families (*Zosteraceae*, *Hydrocharitaceae*, *Posidoniaceae* and *Cymodoceaceae*). Seagrasses form important coastal ecosystems worldwide and support marine species, providing food and habitat. Human activities pose substantial risks for these sea meadows, prompting the need for protection and understanding of these valuable resources.

During adaption to the marine environment, several genes/gene families have been either lost (e.g. stomatal genes) or reduced (e.g. genes involved in the synthesis of terpenoids) whereas others have been regained (e.g. genes involved in sulfation of polysaccharides<sup>1</sup>). In contrast to other marine organisms, the cell wall of seagrasses is poorly understood. Beside ancestral traits of land plants, one would anticipate a habitat-driven adaption process to the new environment, which is characterised by multiple stressors (high amounts of salt, different seagrass grazers and bacterial colonization).

Genome sequencing of two *Zostera* species revealed that return to the marine habitat was accomplished by dramatic changes in cell wall composition<sup>1,2</sup>. Besides polysaccharides known from angiosperm land plants, the cell walls of seagrasses are characterised by sulfated polysaccharides, a common attribute of the macroalgae. For example, a sulfated D-galactan composed of the regular tetrasaccharide repeating unit [3-β-D-Gal-2(OSO<sub>3</sub>)-(1,4)-α-D-Gal-(1,4)-α-D-Gal-(1,3)-β-D-Gal-4(OSO<sub>3</sub>)-1,] was characterised from *Ruppia maritima*<sup>3</sup>. This ability to synthesize sulfated polysaccharides was lost during adaption to terrestrial as well as freshwater habitats and

<sup>1</sup>Pharmaceutical Institute, Department of Pharmaceutical Biology, Christian-Albrechts-University of Kiel, Gutenbergstr. 76, 24118, Kiel, Germany. <sup>2</sup>La Trobe Institute for Agriculture & Food, Department of Animal, Plant and Soil Sciences, La Trobe University, Melbourne, Victoria, 3086, Australia. <sup>3</sup>Sino-Australia Plant Cell Wall Research Centre, School of Forestry and Biotechnology, Zhejiang A&F University, Hangzhou, China. ✉e-mail: [bclassen@pharmazie.uni-kiel.de](mailto:bclassen@pharmazie.uni-kiel.de)

Monosaccharide	AGP_UR whole plant	AGP_UR leaves	AGP_UR rhizome	AGP_UR roots	AGP_AH	AGP_Ox	AGP_UR+Ox
Gal	47.2	50.8	43.1	49.2	47.7	84.0	78.1
Ara	36.0	34.9	34.5	30.8	40.6	3.7	3.4
Rha	4.0	1.4	6.8	7.2	4.5	3.5	3.4
Man	2.5	2.6	3.3	2.0	3.3	5.0	1.3
4-OMe Glc	2.6	1.8	0.8	2.0	n.d.	n.d.	3.0
Glc	6.9	6.9	11.5	8.8	—	2.9	10.8
Xyl	0.8	1.6	—	—	3.9	0.9	—
Ara:Gal	1: 1.3	1: 1.5	1: 1.2	1: 1.6	1: 1.2	1: 22.7	1: 22.6

**Table 1.** Neutral monosaccharide composition (mol %) of the uronic acid reduced AGPs from the different *Z. marina* organs and from partial hydrolyses of whole plant AGP. n.d., not detectable; AH, alkaline hydrolysis; Ox, oxalic acid hydrolysed; UR, uronic acid reduced.

later independently regained by marine angiosperms. Another unique feature of cell walls of seagrasses is the occurrence of unusual polyanionic, low-methyl esterified pectins rich in apiose (Apif). In *Z. marina*, this characteristic polysaccharide is called zosterin and has a backbone of  $\alpha$ -1,4-D-galacturonic acid (GalAp), substituted by 1,2-linked Apif oligosaccharides or single Apif residues<sup>4,5</sup>. Thus, cell walls of seagrasses are characterised by new combinations of structural polysaccharides known from both marine macroalgae and angiosperm land plants. In *Z. marina*, zosterin contributes to the polyanionic character of the cell wall matrix and possibly has crucial roles in osmotic adjustment to salt stress.

In addition to polysaccharides, glycoproteins are particularly important components of primary cell walls of land plants which can exhibit signalling functions during cell expansion and development. The highly glycosylated AGPs have gained particular interest due to their involvement in both wall architecture and cellular regulatory processes. Members of the hydroxyproline-rich glycoprotein family (classified in<sup>6</sup>), AGPs are ubiquitous in seed land plants<sup>7</sup> and have also been found in ferns, lycophytes and mosses<sup>8</sup>. They are structurally characterised by large carbohydrate moieties comprised of arabinogalactans (AGs, normally >90% of the molecule) which are covalently linked via hydroxyproline (Hyp) to relatively small protein/peptide backbones (normally around 1–10% of the molecule). Distinct glycan modifications have been identified in different species and tissues and are suggested to influence both their physical properties and function. AGPs are therefore a good model to study cell wall adaptations to different environments. The AGs of seed plants mainly consist of type II (3,6)-galactans with 3-, 6- and 3,6-linked  $\beta$ -D-galactose (Galp) residues, substituted with  $\alpha$ -L-arabinose (Araf) and often minor amounts of glucuronic acid (GlcAp) residues<sup>7</sup>. AGPs have been implicated in different developmental processes like cell growth, cell proliferation, pattern formation and reproduction<sup>9–12</sup> and are sometimes also important in plant-microbe interactions<sup>13</sup>. Interestingly, massive upregulation of AGPs has been observed in salt-stressed tobacco cells, suggesting a role of AGPs in osmoregulation<sup>14</sup>.

To date, the presence of AGPs has not been established in any seagrass species. Therefore, we collected *Z. marina*, which is one of the ecologically most important macrophytes of the Baltic Sea, to search for AGPs. Investigation of the proteome of *Z. marina* L. identified sequences predicted to encode the highly glycosylated classical AGPs as well as low to moderately glycosylated chimeric AGPs. We utilized the specific interaction of AGPs with the dye  $\beta$ -glucosyl Yariv reagent ( $\beta$ GlcY) to isolate these glycoproteins and detect them by light microscopy. We established their major structural characteristics by different analytical methods as well as by interaction with different anti-AGP monoclonal antibodies. The AGPs from *Z. marina* exhibit special features not known for AGPs from land plants suggestive of a marine environment specialisation, which sheds further light on cell wall evolution, especially with regard to adaption to the marine habitat.

## Results

**Yield and composition of AGPs from *Z. marina*.** High molecular weight fractions (HMFs) were isolated from the whole plant and from the different organs (roots, rhizome and leaves) by ethanol-precipitation of the supernatants prepared after centrifugation of the aqueous extracts. The HMF extracts had essentially similar monosaccharide compositions characterised by Gal- and Ara-rich polysaccharides and glycoproteins, except for in the root that had a lower content of Ara (Supplementary Table S1).

Precipitation of AGPs with  $\beta$ GlcY was undertaken from HMFs of whole plant and organs. The resulting AGP yields differed significantly. While the whole plant contained approximately 0.096% ( $\text{m m}^{-1}$  dry weight) AGPs, the rhizome and root AGP contents were much higher (0.144% ( $\text{m m}^{-1}$ ) and 0.138% ( $\text{m m}^{-1}$ ), respectively, whereas leaves contained very low AGP levels (0.005% ( $\text{m m}^{-1}$ )).

Organic nitrogen quantification of  $\beta$ GlcY-precipitable AGPs from whole plant revealed 2.03% nitrogen, which corresponds to an approximate protein content ( $\times 6.25$ ) of 12.69%<sup>15</sup>. Colorimetric determination of Hyp detected 0.39% ( $\text{m m}^{-1}$ ) in the AGP from the whole plant sample, which equates to a Hyp content in the protein moiety of approximately 3.1% (w/w).

The AGPs from the different organs had Gal and Ara as their major monosaccharides but showed differences in their total Gal + Ara content (78–86%) as well as their Ara: Gal ratio (Table 1). Rhamnose (Rhap) was noticeably lower in the leaf AGP.

Compared with the neutral monosaccharide composition prior to carboxyl reduction (Supplementary Table S2), a noticeable increase in Glc content originating from GlcA was detectable in all AGP samples after

Monosaccharide	Deduced linkage type	<i>Z. marina</i> UR	<i>Z. marina</i> UR + Ox
Galp	1,3,6-	40.0	32.3
	1,6-	1.7	39.1
	1,3-	7.4	10.7
	ter	—	1.9
Araf	1,3-	4.2	1.0
	1,2-	1.3	—
	ter	23.7	1.7
GlcP*	1,4-	5.1	1.6
	ter	10.5	9.1
Manp	1,6-	3.0	1.0
Rhap	ter	3.1	1.6

**Table 2.** Linkage analysis (mol %) of *Z. marina* AGPs before and after partial acid hydrolysis. \*Derived from GlcAp by UR. Ox, oxalic acid hydrolysed; UR, uronic acid reduced; ter, non-reducing terminal residues.

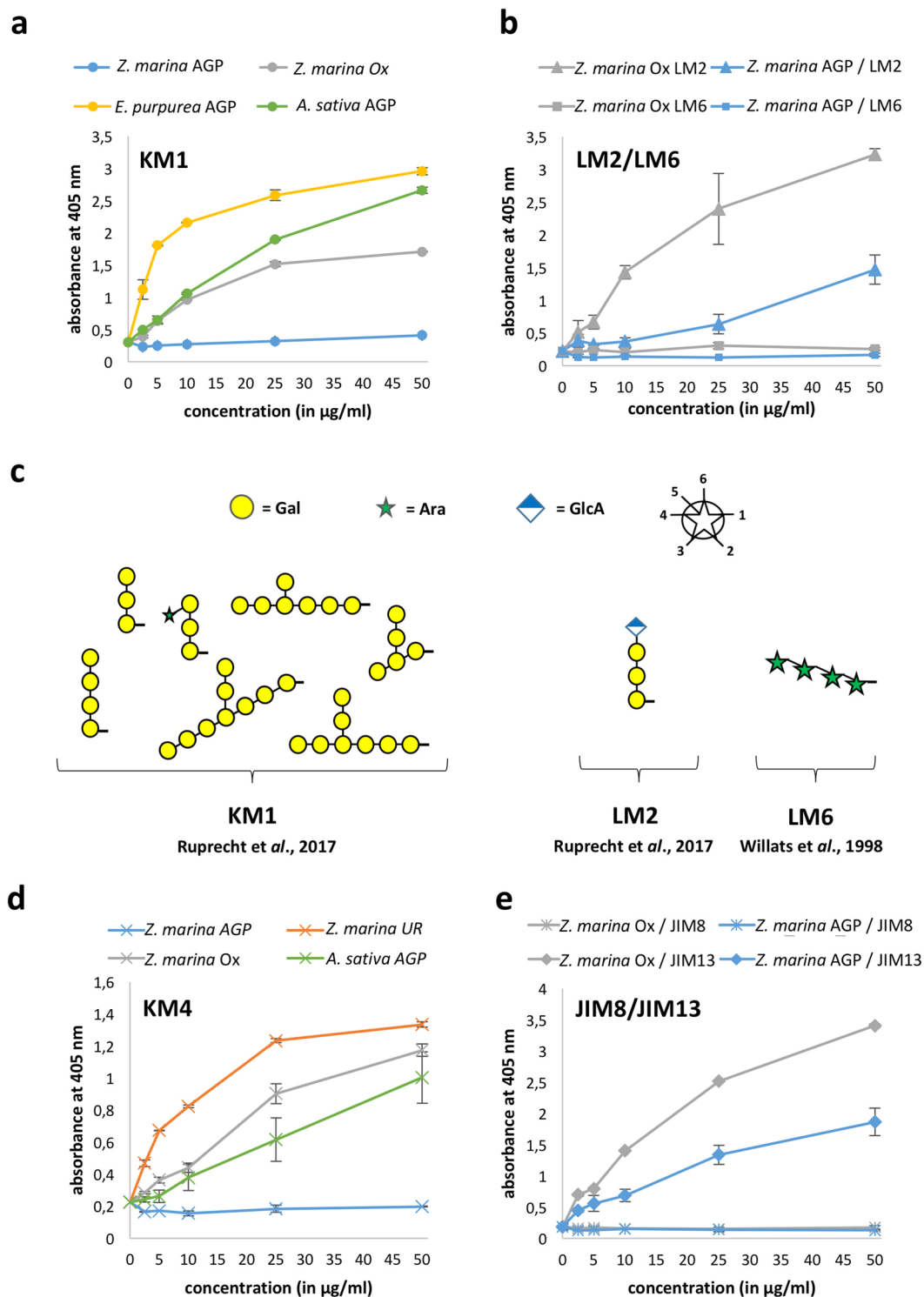
uronic acid reduction (Table 1). The mass spectra of the peaks of Glc clearly revealed the presence of the characteristic fragments of a di-deuterated C6 (reduction of the carboxyl group with NaBD<sub>4</sub>). Furthermore, in all chromatograms of uronic acid reduced (UR) samples a peak appeared, which was not detected in the non-reduced AGPs. This peak with the relative retention time of 0.862 (relative to *myo*-inositol-hexaacetate) was clearly identified by its mass spectrum as C6 di-deuterated 4-O-methyl-GlcP (Supplementary Fig. S1).

Different partial hydrolyses with the whole plant AGP fractions were performed to gain insights into structural details of *Zostera* AGPs (Table 1). Through alkaline hydrolysis (AH) the protein backbone is removed whereas the carbohydrate composition remained largely unchanged. After AH, no GlcP was detected, which could be a consequence of the removal of trace levels of βGlcY under alkaline conditions. Mild acid hydrolysis (Ox) led to loss of most Araf residues while reduction of uronic acids revealed the presence of both di-deuterated 4-OME GlcP and GlcP originating from GlcAp.

**Structure of AGPs from *Z. marina*.** AGP containing samples were subjected to linkage analysis by methylation after reduction of carboxyl groups and also after additional partial acid hydrolysis (Ox). For AGP after reduction of carboxyl groups, the molar recovery of branched residues (1,3,6-linked Galp) was approximately equal to the molar recovery of non-reducing terminal residues (Araf, GlcP and Rhap), indicating complete methylation. The analysis (Table 2) of the uronic acid reduced AGPs from whole plant revealed the different Galp units typical for type II AGs with an extraordinarily high content of 1,3,6-linked Galp, which indicates a highly branched structure, which was confirmed by size exclusion chromatography (SEC) (see below). The Araf is mainly terminal and located in sidechains of the molecule. Interestingly, no 1,5-linked Araf typical for many land plant AGPs was detected. Small amounts of 1,6-linked Manp might be part of N-glycans present on chimeric AGPs<sup>16</sup> and have therefore not been included in the proposed structure (see below, Fig. 2). Mild acid hydrolysis of the sample prior to methylation led to near complete loss of Araf residues and to an increase of 1,6-linked Galp, indicating that Araf is bound to Gal at C-3 of 1,6-linked Galp branches. Deuterium-labelled GlcP was present as both terminal and 1,4-GlcAp. The primary fragments 207 (for deuterated terminal GlcP) and 235 (for deuterated 1,4-linked GlcP) were accompanied by no or only trace levels of the corresponding non-deuterated fragments (205 and 233, respectively), indicating that only GlcAp and not GlcP is present in the native AGP.

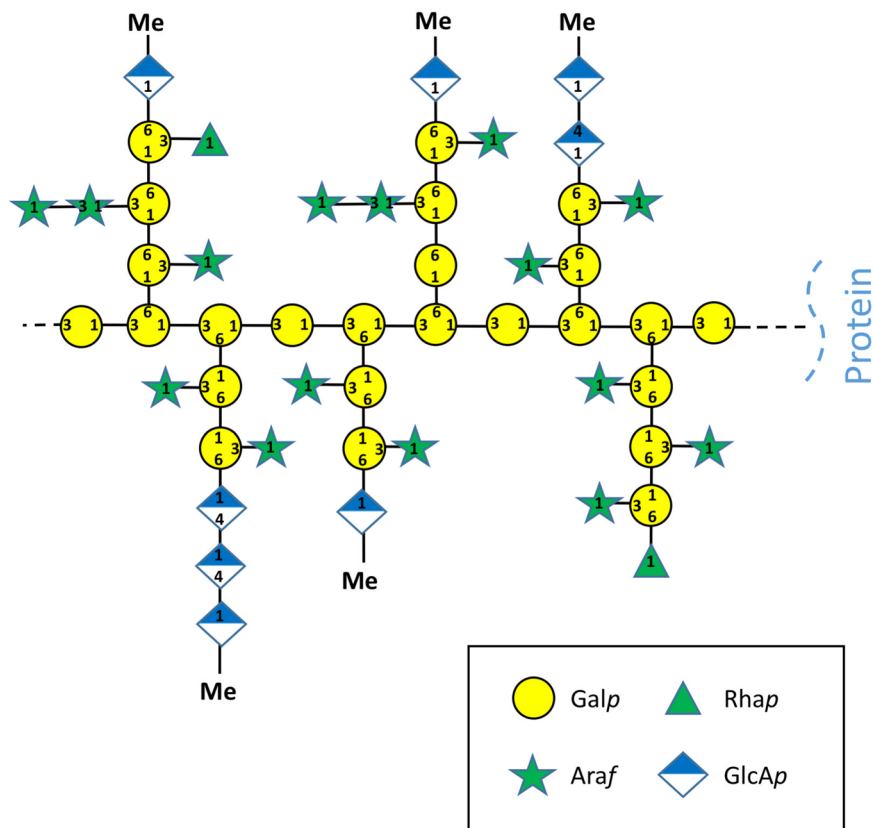
**Determination of molecular weight by size-exclusion chromatography.** Absolute molecular weights of *Z. marina* AGPs and degraded products were determined by SEC separation and multi-angle light scattering (MALS) detection and hydrodynamic volumes were calculated using commercially available pullulan standards (Supplementary Table S3). The absolute molecular masses determined by MALS were in a range typical for AGPs and always much higher compared to their hydrodynamic volumes. This showed a highly branched structure of all the AGP molecules (explaining a large mass in a small volume). Partly, these differences could be caused by the different structures of pullulan standards, which are linear and unbranched polymers<sup>17</sup>, compared with branched AGPs. As expected, chemical modifications by either reduction of uronic acids or oxalic hydrolysis (Ox) decreased the absolute molecular masses and hydrodynamic volumes, due to loss of hydration (UR<sup>18</sup>) and loss of mainly Araf by treatment with Ox. In all chromatograms some higher order aggregates resulting from self association were present; whether or not this reflects an *in vivo* property or is an artefact of the fractionation procedure cannot be distinguished. It is not unusual to see such aggregation unless the chromatography is conducted under strong dissociating conditions, such as chaotropic reagents (eg urea/guanidine hydrochloride) but this was not compatible with the SEC-MALS detection<sup>19</sup>.

**Binding of *Z. marina* AGPs to antibodies raised against land plant AGPs.** The native *Z. marina* AGPs and their partially degraded products were investigated for their ability to bind to the antibodies LM2, LM6, JIM8, JIM13<sup>20–23</sup>, KM1 (raised against *Echinacea* AGP<sup>24</sup>), and KM4 (raised against *Avena* AGP<sup>25</sup>) (Fig. 1). With KM1, the native AGP from *Z. marina* showed no reactivity. After mild acid hydrolysis good binding was evident (Fig. 1a), due to increase in 1,6-linked Galp residues (Table 2). These 1,6-linked Galp residues are present in the original antigen (*Echinacea* AGP<sup>24</sup>) and an AGP from the terrestrial monocotyledonous plant *Avena sativa*<sup>26</sup> and



**Figure 1.** Reactivity of *Zostera marina* AGPs and their partial hydrolyses (with oxalic acid = Ox; by reduction of uronic acids = UR) with antibodies directed against AGP glycans by ELISA. **(a)** KM1 antibody with *E. purpurea* AGP and *A. sativa* AGP as positive controls. **(b)** LM2 and LM6 antibodies. **(c)** oligosaccharides with strong binding affinity to KM1, LM2 and LM6, respectively<sup>22</sup>. **(d)** KM4 antibody with *A. sativa* AGP as positive control. **(e)** JIM8 and JIM 13 antibodies.

known to be part of the binding epitope (Fig. 1c<sup>22</sup>). The epitope of LM2 comprises some 1,6-linked Galp residues with a terminal GlcAp (Fig. 1c<sup>22</sup>). The native *Z. marina* AGPs show moderate binding which increases after mild acid hydrolysis (Fig. 1b), known to increase 1,6-linked Galp residues. LM6, which is directed against a 1,5-linked AraF chain (Fig. 1c), does not bind to *Z. marina* AGPs, thus confirming that this linkage is absent in this seagrass



**Figure 2.** Structural proposal for the polysaccharide moiety of *Zostera marina* AGP based on molar portions detected in linkage analysis (see Table 2).

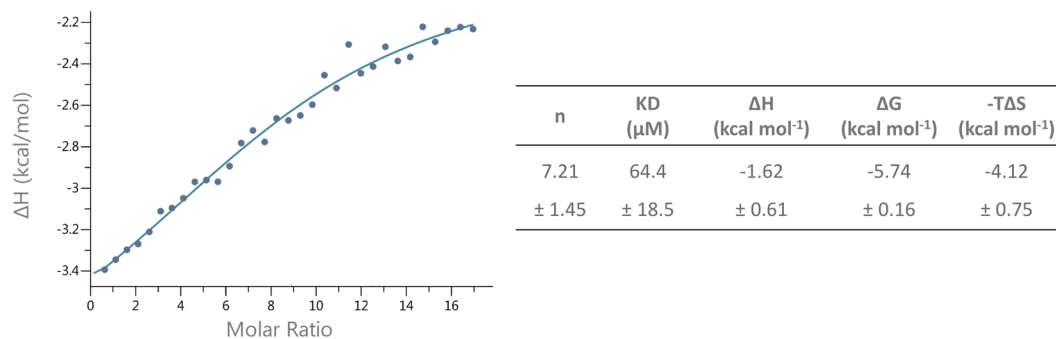
AGP (Fig. 1b, see also Table 2). KM4, which was raised against an AGP from *Avena sativa* which does not contain uronic acids<sup>26</sup>, showed no binding to the native, strongly charged *Z. marina* AGPs. After mild acid hydrolysis and especially after reduction of uronic acids, binding exceeded that of the original antigen (Fig. 1d). JIM8 and JIM13 are antibodies directed against AG glycan on AGPs but the precise epitope is not defined. Whereas JIM13 reacted with *Z. marina* AGPs, especially after mild acid hydrolysis, there was no interaction with JIM8 (Fig. 1e).

Combining the structural data of the polysaccharide moiety and antibody epitope binding studies, a predicted *Z. marina* AGP glycan structure is proposed (Fig. 2). This model incorporates both known structural features characteristic of angiosperm AGPs and features unique to *Z. marina* such as a very high degree of branching, consistent with the high amounts of 1,3,6-linked Galp residues and only traces of 1,6-linked Galp- containing side chains, and high amounts of terminal oligosaccharides rich in uronic acids, especially 1,4-linked GlcAp residues and terminal 4-OMe GlcAp. For comparison with *Z. marina* AGP an *A. sativa* AGP structure was added (see Supplementary Fig. S2).

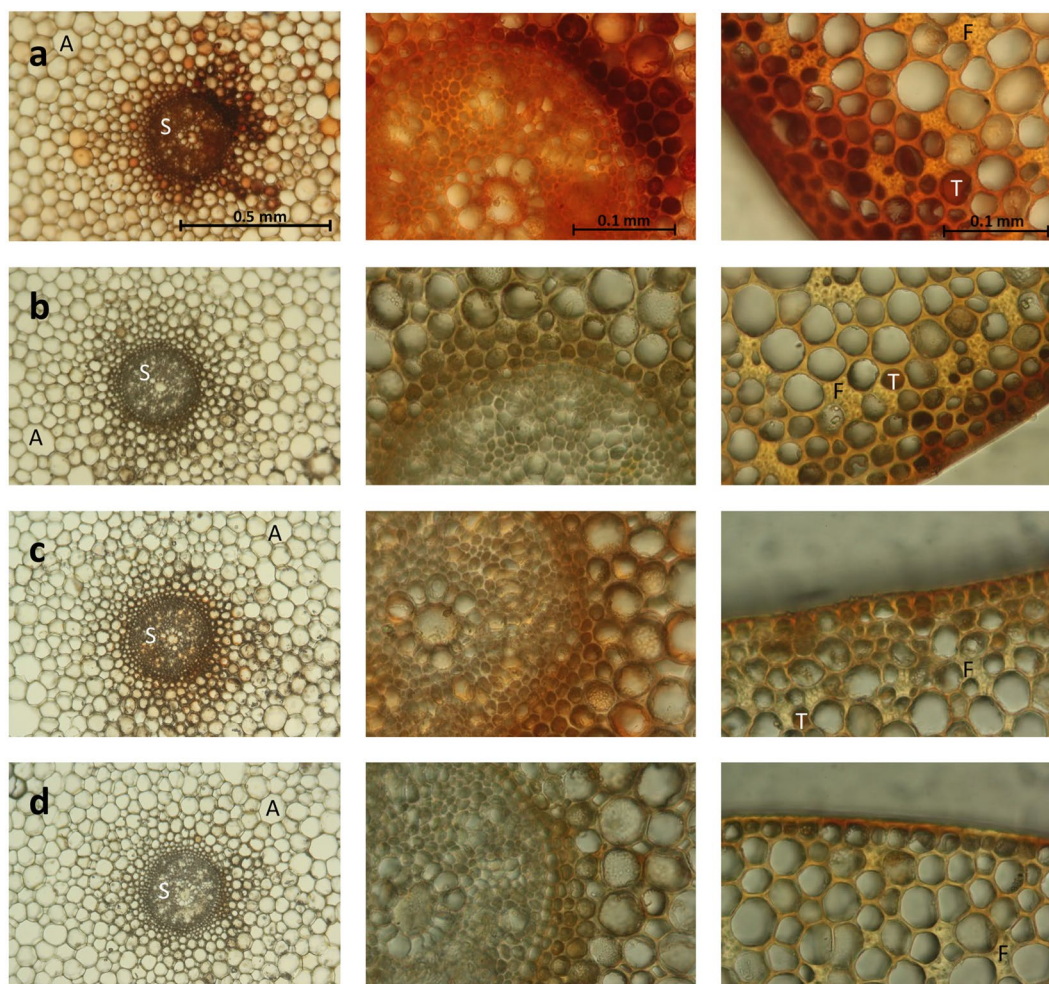
**Calcium-binding experiments by ITC.** ITC showed the calcium binding capacity of *Z. marina* AGP and gave information about its thermodynamic binding parameters (Fig. 3). One important value for binding studies is the dissociation constant ( $K_D$ ), which is in the micromolar range. The reaction is characterized as exergonic (negative  $\Delta G$ ), exothermic (negative  $\Delta H$ ) and mainly entropy-driven (high  $-T\Delta S$ ), consistent with the second law of the thermodynamics (Fig. 3). As a control experiment, AGP from fruits of *A. sativa*, which is free of uronic acids<sup>26</sup> (for structural proposal see Supplementary Fig. S2), was used in the same test system and showed no binding to  $Ca^{2+}$  in the ITC thermogram.

**Detection of AGPs and calcium-binding in tissues by light microscopy.** Cross sections of fresh rhizomes (Fig. 4a–d) showed the general anatomy evident for land plant rhizomes: epidermis, cortex and a central cylinder (stele). Furthermore, anatomical features described for other seagrass rhizomes<sup>27,28</sup> were detected. In the outer part of the parenchymatous cortex, tannin-containing cells as well as large fiber bundles are noticeable. Other typical elements are several air lacunae present in the cortex. Two opposite cortical vascular bundles, which have also been detected in rhizomes of *Zostera caulescens*<sup>28</sup> are also present (not shown). The red-coloured  $\beta$ GlcY was used to stain AGPs (Fig. 4a). AGPs were mainly detected in the stele and the surrounding inner layers of the cortex, the two cortical vascular bundles (not shown) as well as in the outer layers of the cortex. Staining with Alizarin S (Fig. 4c), which is an anthraquinone dye used to detect calcium deposits, revealed co-localization of calcium and AGPs.





**Figure 3.** Isothermal titration calorimetric detection of calcium-binding to *Zostera* AGP. Thermodynamic values of calcium-binding assay using ITC. Fitting curve using the one-site binding model. The values in the table were averaged from three measurements.



**Figure 4.** Co-localisation of calcium binding and occurrence of AGPs in *Zostera marina* rhizome detected by microscopy. (a) Red staining of AGPs by  $\beta$ GlcY in rhizome cross sections. (b) Negative control without  $\beta$ GlcY staining. (c) Rhizome cross section incubated in calcium chloride solution and stained with Alizarin S. (d) Negative control incubated in water and stained with Alizarin S. A = air lacunae, S = stele, F = vascular fibre bundles and T = tannin cells (according to<sup>27,28</sup>).

**Identification of AGP protein sequences in *Z. marina*.** The *Z. marina* genome contains 38,875 coding sequences (CDSs)<sup>1</sup>. Hydroxyproline-Rich Glycoproteins (HRGPs) and in particular, classical AGPs, were identified using the established motif and amino acid bias 'MAAB' pipeline<sup>29,30</sup>. Chimeric AGPs were identified by first screening the proteome for sequences with predicted AG regions, followed by filtering for sequences

containing at least one detectable protein domain (Pfam), and a signal peptide (SignalP 5.0). MAAB identified 13 sequences that are recognized as HRGPs (Fig. 5, Supplementary Fig. S3). The majority of these sequences are classical AGPs, 5 of which contain a GPI-anchor signal sequence (MAAB Class 1) and 4 non-GPI-AGPs (MAAB Class 4) (Fig. 5A). Additionally, there is one classical extensin (MAAB Class 2), two sequences with shared bias (MAAB Class 20 and 21), that contain both extensin and AGP glycomotifs, and a sequence from the <15% HRGPs (MAAB Class 24) (Fig. 5A). These proportions are similar to those found in the genomes of terrestrial monocots<sup>29,30</sup>.

In addition to the classical AGPs, there are a further 188 chimeric AGPs (33 protein kinase-like AGPs + 155 other chimeric AGPs) that contain one or more protein domains as identified in Pfam, a signal peptide and at least one AG-motif rich region (Fig. 5B). The most common chimeric AGPs observed are the protein kinase-like AGPs (PKGPs), copper oxidase plastocyanin/ENOD-like AGPs (PAGs), lipid-transfer xylogen-like AGPs (XYLPs) and fasciclin-like AGPs (FLAs) (Fig. 5B). The amount of glycosylation is predicted to vary between the highly glycosylated classical AGPs that contain AG glycomotifs throughout the protein backbone and the chimeric AGPs that are likely to be low to moderately glycosylated in different regions of the protein backbone (Fig. 5C). The chimeric AGPs also show high variation in domain organization, i.e. where the Pfam is located in relation to the AG region and how many of each, even within the same group (Supplementary Figs. S4 and S5).

The *Z. marina* AGP protein sequences are largely characteristic of known angiosperm AGP protein backbones suggesting the structural differences in glycan structures likely result from the glycan-active enzymes (glycosyltransferases (GTs) and methyltransferases) acting on the *Z. marina* AGPs. As a consequence the complement of GT and methyltransferase families implicated in AG glycan assembly were interrogated in the *Z. marina* genome.

**Glycan-active enzyme complement of *Z. marina*.** The glycan structures of AGPs are primarily assembled by GT enzymes from the GT14, GT31, GT61 and GT77 families and methyltransferases from the DUF579 family<sup>31–34</sup>. *Arabidopsis thaliana* sequences were used to search the *Z. marina* genome and identified 25 GT31s (galactosyltransferases (GalTs) and N-acetylgalactosaminyltransferases (GalNACTs), slightly higher than the number observed in other viridiplantae<sup>31</sup>. GT31 members are proposed to add the first Gal to Hyp as well as build the  $\beta$ -1,3 Gal backbone and  $\beta$ -1,6 Gal side-chains. When compared to the *A. thaliana* homologs (the only organism for which some GT31 functions have been assigned) there is a minor expansion in Hyp GalTs (Clade 7-V) which initiate glycosylation of the AGP protein backbone, and additional expansion of Clade 10-IV of unknown function (Supplementary Fig. S6). *Z. marina* has 14 GT14s, again, in line with the 10–18 found in other viridiplantae, with minor expansions in expansions in clade B5 (glucuronosyltransferases (GlcATs) and clade B2 of unknown function (Supplementary Fig. S7). The GT77 family also contains a number of well-characterised arabinosyltransferases (AraTs) and arabinofuranosyltransferases (ArafTs)<sup>35,36</sup>. Compared to *A. thaliana*, most clades contain similar numbers of homologs with clades D and E the most likely candidates for containing the relevant AGP-active 1,3-ArafTs, whereas the rhamnogalacturonan-II-active enzymes (clade B) are reduced (Supplementary Fig. S8).

Conversely, the GT61 and DUF579 families assessed had gene complements significantly differing from those found in other vascular plants. *Z. marina* has fewer GT61s of the A, D and F clades (Supplementary Fig. S8), thought to contain the relevant 1,5-AraTs, compared to the 3–10 usually found in vascular plants (and 20–50 in the Commelinidae land monocots)<sup>33</sup>. Similarly, there are fewer members of the DUF579 methyltransferase family (Fig. 6). Within this family, Clade III, has been shown to exhibit methylation of AGP glycans in *A. thaliana*<sup>34</sup>. *Z. marina* contains two sequences in this clade: KMZ57090.1 and KMZ73279.1 which we propose are the most likely candidates responsible for the extensive methylation of the AGP terminal GlcAp residues.

## Discussion

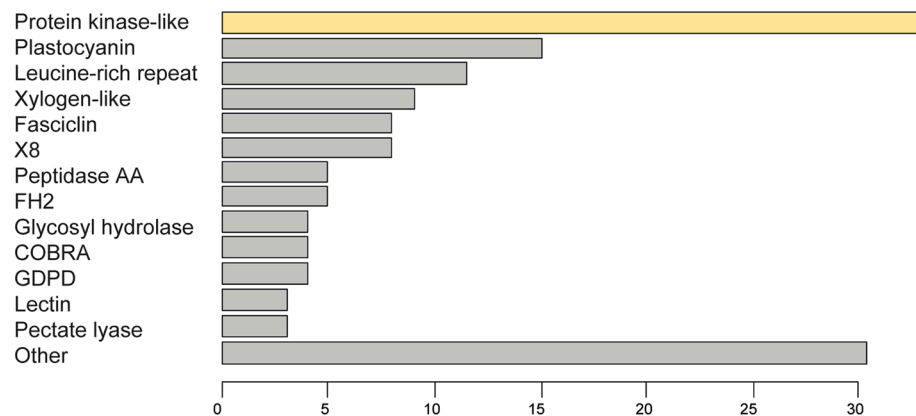
*Z. marina* is of major ecological importance<sup>37</sup> and the biomass is used as environmentally friendly and non-flammable insulating material, properties that are influenced by the composition of their cell walls. Despite this, knowledge of the composition of the cell walls within the polyphyletic group of seagrasses is largely restricted to the major polysaccharides that are unique to marine plant and algal species, and include an apiogalacturonan (“zosterin”)<sup>4,5,38</sup> and sulfated galactans<sup>3,39,40</sup>. Whether seagrass cell walls/cell surfaces also contain a similar complement of minor polysaccharides/glycoproteins, particularly those implicated as signalling molecules in terrestrial plants, such as the AGPs, was of particular interest from both an evolutionary and plant development perspective.

AGPs are key constituents of the extracellular matrix of flowering land plants with important functions in processes such as cell growth, cell proliferation, pattern formation, reproduction and plant-microbe interactions<sup>7,9–12,41</sup>. Interestingly, accumulation of AGPs has been observed in the media of salt-stressed suspension cultures of tobacco, tomato, *Acacia*, *A. thaliana*<sup>14,41,42</sup> and *Dactylis*<sup>43</sup>, supporting a role of AGPs in osmoregulation. In *Brassica*, a significant accumulation of AGPs in the xylem sap after 24 h of salt treatment was observed<sup>44</sup>. Studies on *Medicago sativa* leaves showed an increase of AGPs recognized by the JIM8 antibody in NaCl-treated plants<sup>45</sup>. When seagrasses evolved from monocotyledonous land plants and colonized the coastal marine ecosystems this was possibly the most severe change of habitat in evolution of flowering plants, mainly due to salt stress. Salinity impairs plant growth and development *via* both a water stress due to osmotic pressure and thus the reduction in water availability and also *via* a direct cytotoxicity due to excessive uptake of ions, mainly sodium (Na<sup>+</sup>) and chloride (Cl<sup>-</sup>) and a critical loss of potassium (K<sup>+</sup>). The proposed role of AGPs in salt stress made them an ideal candidate to investigate how changes in wall components could play a role during salt adaption to highly saline conditions.

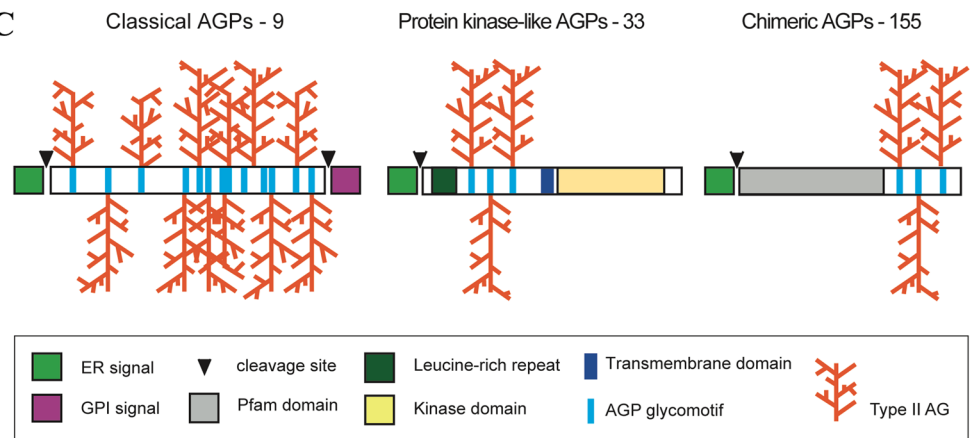
This study represents the first isolation and characterisation of AGPs in seagrasses. Interestingly, the amount of water-soluble AGPs in roots and rhizomes of *Z. marina* was around 20 fold higher than in leaves. One reason

ID		ext_sp	ext_tyr	prp	agp	past_percent	pvk_percent	psky_percent	p_percent	maab_class	maab_class_#
KMZ70582.1	ZOSMA_199G00190	4	0	0	32	56	34	42	24	AGP	4
KMZ66553.1	ZOSMA_294G00050	0	0	0	19	64	23	39	17	AGP	1
KMZ65764.1	ZOSMA_30G00450	0	0	0	18	58	32	35	21	AGP	1
KMZ63928.1	ZOSMA_38G00360	1	0	0	28	67	30	33	20	AGP	1
KMZ62892.1	ZOSMA_43G00740	4	3	0	51	49	37	42	29	AGP	4
KMZ59493.1	ZOSMA_68G00800	0	0	0	15	63	21	42	16	AGP	1
KMZ58528.1	ZOSMA_76G00830, partial	3	0	1	48	52	32	45	23	AGP	4
KMZ57387.1	ZOSMA_86G00220	5	0	0	51	61	40	42	27	AGP	4
KMZ56883.1	ZOSMA_8G00260	0	1	0	12	47	22	32	11	AGP	1
KMZ72364.1	ZOSMA_166G00640	47	50	0	16	49	66	72	35	CL-Extensin	2
KMZ74158.1	ZOSMA_133G00070	2	8	1	7	37	46	45	21	Extensin	20
KMZ66204.1	ZOSMA_2G02140	14	0	0	2	50	47	52	37	High-SPn	21
KMZ58455.1	ZOSMA_76G00110	0	1	0	16	47	19	31	11	<15% HRGP	24

### B Chimeric-AGPs

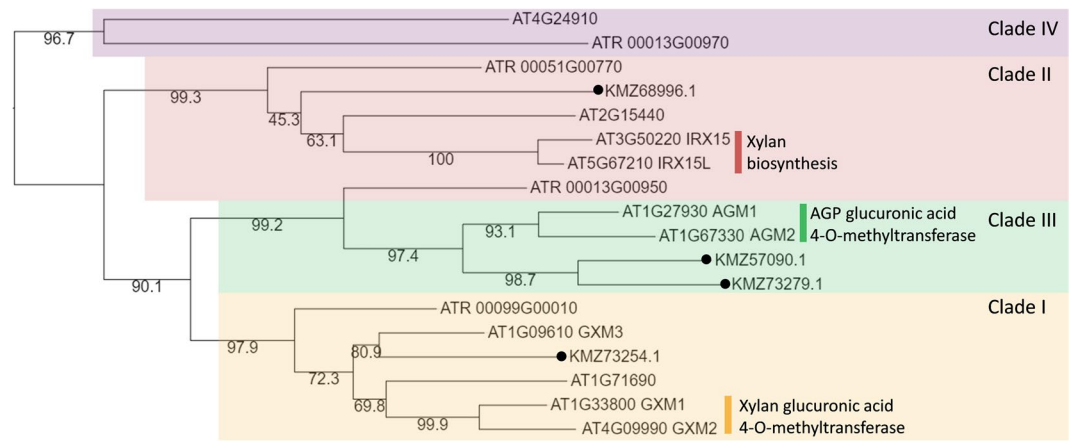


### C



**Figure 5.** HRGP and AGP classes identified in the *Zostera marina* proteome. (A) The majority of HRGPs identified by the MAAB pipeline are classical AGPs of Class 1: GPI-AGPs and Class 4: non-GPI-AGPs. The number of glycomotifs characteristic of extensins (ext\_sp and ext\_tyr), PRPs and AGPs in a given sequence is shown. The percentage amino acid bias characteristic of AGPs (PAST), PRPs (PVYK) and extensins (PSKY) is shaded in light green (0–25%), medium green (25–50%) and dark green (50–75%). (B) Chimeric AGPs, identified by first screening the genome for sequences that include a region with at least 3 AG glycomotifs and a Pfam domain (ragp package, GitHub). The major chimeric AGP classes with 3 or more representatives are shown with similar Pfam families grouped by clan. A full list of Pfam families and most common domain architectures are shown in Figs. S4 and S5. (C) Schematic of predicted AGP structures of classical AGPs and chimeric AGPs that include protein-kinase like AGPs and other chimeric AGPs in *Zostera marina*.





**Figure 6.** Phylogeny of DUF579 family members from *Zostera marina* (KM accession numbers, highlighted with black circles), and *Arabidopsis thaliana* (AT). Clades labelled as in<sup>34</sup>. Genes with known function indicated. For equivalent phylogenies of GT families, see Supplementary Figs. S6–S8.

for the higher content of AGPs in rhizomes and roots might be due to the fact that leaves persist for only one year whereas rhizomes and roots are perennial. If one function of AGPs is protection against salt, the longer lasting organs may have a higher AGP requirement. In the marine angiosperm *Ruppia* sulfated galactans possibly involved in salt protection are also mainly present in the rhizomes<sup>3</sup>.

The isolated AGP-fraction from *Z. marina* fulfills the general criteria which are typical for many AGPs like a large AG moiety that consists mainly of 3-, and 3,6-linked  $\beta$ -D-Galp residues, substituted primarily with terminal  $\alpha$ -L-Araf residues, a small protein part mostly rich in Hyp and the ability to be precipitated with Yariv phenylglycosides, for example, the  $\beta$ GlcY reagent. Molecular weight, protein content and amount of Hyp in *Zostera* are in a range reported for land plant AGPs. Similarly, the complement of genes in the draft *Z. marina* genome expected to encode proteins that would be precipitated by  $\beta$ GlcY is broadly in line with other vascular plants. Nine classical AGPs were identified, broadly similar to those observed in the commelinid monocots<sup>30</sup>. The chimeric AGPs (that contain one or more Pfam domains) are considerably more numerous than the classical AGPs and include the most common chimeric classes, FLAs, PAGs and XYPs and surprisingly, a large number of PK-like AGPs. Interestingly, PK-like AGPs were identified in the marine brown alga *Ectocarpus* and proposed to have roles as sensor molecules<sup>46</sup>.

Two main structural features were identified in *Z. marina* AGPs that are distinct from angiosperm land plant AGPs. Firstly, high branching combined with a high Araf content was observed, especially in AGPs isolated from the rhizome. For *A. thaliana*, it has been shown that Araf biosynthesis is important for salt stress tolerance. Mutations in a gene which is required for UDP-Ara biosynthesis led to reduced root elongation under high salinity and could be healed by application of either exogenous Ara or gum arabic, a commercially available gum from *Acacia senegal* rich in AGPs<sup>47</sup>. A survey investigating different resurrection plants from South Africa revealed that Ara-rich polymers plasticize resurrection plant cell walls in case of desiccation and facilitate rehydration<sup>48</sup>.

An even more striking feature of *Z. marina* AGPs are high amounts of GlcAps, especially 1,4-linked GlcAp and terminal 4-OMe GlcAp residues. The presence of significant levels of the 4-O-Me-GlcAp is not a common feature of land plant AGPs, but has been detected rarely, e.g. in AGPs from *Raphanus*<sup>49,50</sup>. The significant differences in content and composition of AGPs in *Z. marina* suggest specific functional adaptations to the marine environment. Given the unusual AG glycan structures, the genome was searched for key enzyme families known to be involved in AG synthesis. The genome contained a relatively standard set of sequences from the GT31 family (3- and 6- GalTs), GT77 (AraTs) and GT14 family (GlcATs), with members present from all the major clades characterised in each family. Conversely the GT61 (AraTs) and DUF579 (methyltransferases) families have markedly divergent distributions of members. This corresponds to the lack of 1,5-linked Araf residues (fewer GT61) and extensive methylation (DUF579) of GlcAp residues in *Z. marina*. The enzymes responsible for methylation of GlcAp in AGPs have recently been identified in *A. thaliana*<sup>34,51</sup>. Two sequences were identified as the most likely candidates responsible for the extensive methylation of the AGP terminal GlcAp residues in *Z. marina*. Since no gene family expansion was observed in DUF579 family Clade III, the markedly increased methylation observed in *Z. marina* AGPs is likely due to their unique upregulation (transcriptional or post-translational) in response to the exposure to a highly saline environment. In future, experimental verification of either altered transcript abundance or activity will further our understanding of this unique structural modification.

Uronic-acid rich residues are likely to lead to a polyanionic surface of *Z. marina* AGP which is likely of great importance with regard to adaption to the marine environment<sup>52</sup>. Important is the ability of uronic acids to bind  $\text{Ca}^{2+}$ , and it has been shown, that uronic acid containing AGPs bind  $\text{Ca}^{2+}$  more strongly than does pectin<sup>52</sup>. The authors identified stoichiometric  $\text{Ca}^{2+}$  binding by paired GlcA carboxyls present in AGPs and this model is consistent with our findings for *Z. marina* AGPs. It has been proposed that subtle structural variations of the  $\text{Ca}^{2+}$ -binding subunit of AGPs helps to discriminate effectively against binding of  $\text{Na}^{+}$ <sup>53,54</sup>. One possibility for this fine-tuning of ion-binding affinity might be O-methylation of different monosaccharides; e.g. the occurrence

of 3-Ome Rhap in moss AGPs<sup>55–57</sup> and especially 4-Ome GlcAp in *Z. marina* AGP. This confirms the results of Lamport and Várnai<sup>53</sup> for binding of different AGPs to Ca<sup>2+</sup> in a micromolar range. It is worth noting that a different methodology was used by these authors. We used the Chelex resin for Ca<sup>2+</sup> removal, which has high affinity to bivalent cations. Our control experiment with an uronic acid-free AGP from the monocotyledonous land plant *Avena sativa* (no binding to Ca<sup>2+</sup>) proved that uronic acid-rich terminal groups of *Z. marina* AGPs are essential for the Ca<sup>2+</sup>-binding properties. This finding is also supported by light microscopy of cross sections of the rhizomes, which revealed co-localisation of AGPs and Ca<sup>2+</sup>.

Ca<sup>2+</sup> ions play a crucial role in both the regulation of transport and exclusion of Na<sup>+</sup> and other mineral ions at the plasma membrane of plant cells, and for some time it was known that Ca<sup>2+</sup> protects an extremely salt-sensitive species (*Phaseolus vulgaris*) against damage caused by NaCl present in the medium<sup>58</sup>. Investigations on cotton supported the finding that Ca<sup>2+</sup> protects membranes from harmful effects of Na<sup>+</sup> and maintained membrane integrity<sup>59</sup>. As AGPs can be linked to the plasma membrane by a GPI-anchor, binding of Ca<sup>2+</sup> to AGPs might build a barrier protecting the plasma membrane against high Na<sup>+</sup>-concentrations in the marine environment.

Beside the possible barrier function of uronic-acid rich, Ca<sup>2+</sup>-binding AGPs close to the plasma membrane, another important mode of action of AGPs might be signalling. Signalling caused by AGPs might occur via either Ca<sup>2+</sup>-ions<sup>54</sup>, via oligosaccharides enzymatically cleaved from the AG glycan moieties of AGPs or by association of GPI-anchored AGPs with classical transmembrane proteins. It has been proposed, that Ca<sup>2+</sup>-release from AGPs is an auxin-dependent process via regulation of the plasma membrane (PM) H<sup>+</sup>-ATPase, which generates a lower periplasmic pH leading to dissociation of AGP-Ca<sup>2+</sup> carboxylates<sup>52,54</sup>. In *Torenia*, an AGP called AMOR has been identified which is responsible for cell-cell-communication in the pollen-pistil interaction. Interestingly, the terminal disaccharide 4-Ome GlcAp- $\beta$ -(1 → 6)-Galp is necessary and sufficient for activity<sup>60</sup>. The exact signalling pathways either by Ca<sup>2+</sup> or by oligosaccharides have yet to be elucidated for *Z. marina*.

## Conclusions

To date, the presence of AGPs has not been established in any seagrass species. We were able to isolate AGPs from *Z. marina* and structurally characterize these glycoproteins by analytical methods as well as by ELISA with different anti-AGP antibodies. The glycan structures exhibit unique features, including a high degree of branching and a high content of terminating 4-Ome GlcAp residues not known from land plants. These differences likely result from the glycan-active enzymes (GTs and methyltransferases). Bioinformatic investigation identified protein backbones belonging to 9 classical AGPs and a large number of chimeric AGPs. This group showed unusual dominance of PK-like AGPs, which are proposed to work as sensor molecules. For a deeper insight into structure-function correlation, the Ca<sup>2+</sup>-binding capacity of *Z. marina* AGPs was investigated by microscopy and ITC. In contrast to AGP from the land plant *A. sativa*, it showed binding in the range of the important Ca<sup>2+</sup> binding partner calmodulin. These findings shed light on the adaption process from land to the marine habitat.

## Methods

**Materials.** Fresh green, epiphyte-free samples of *Zostera marina* L. were collected after a stormy night at the beach near Olympic Centre Schilksee, Kiel (54°25'39.0"N 10°10'17.5"E) in March 2018. They were cleaned from coarse pollution, rinsed in tap water and freeze-dried. Half of the material was separated into the different plant organs (roots, rhizome and leaves) before freeze-drying.

**Isolation of high molecular weight fraction (HMF).** The dried plant material was milled in a MF 10 basic grinder (IKA-Werke GmbH & Co.KG, Staufen, Germany) with a sieve size of 1.0 mm. Ground material was then extracted with tenfold (w v<sup>-1</sup>) of 70% acetone (v v<sup>-1</sup>) for 22 h at 4 °C under constant movement on a shaker (Edmund Bühler GmbH, Bodelshausen, Germany). Acetone was removed by vacuum filtration. This procedure was repeated four times to ensure a successful removal of tannins and colored impurities. Afterwards an aqueous extraction (1:10 (w v<sup>-1</sup>)) of the air-dried plant material was carried out following the procedure of<sup>10</sup>. The resulting precipitate, the high molecular weight fraction (HMF), was separated by centrifugation at 19,000 g, 4 °C for 30 mins and freeze-dried. Samples of whole plant, rhizome, root and leaves were subjected to this same extraction procedure.

**Isolation of AGPs.** For purification of AGPs,  $\beta$ -glucosyl-Yariv reagent ( $\beta$ GlcY) was used and isolation was done according to<sup>61</sup>.

**Analysis of neutral monosaccharides.** Determination of the monosaccharide composition was carried out according to<sup>62</sup>. Either HMF (2–5 mg) or AGP (1–3 mg) fractions were mixed with 0.5 mg of the internal standard *myo*-inositol, then hydrolysed, reduced and acetylated. Identification and quantification of monosaccharides was performed with gas chromatography (GC) with parallel flame ionization detection (FID) and mass spectrometry (MS) (GC + FID: Agilent 7890B, Agilent Technologies, USA; MS: Agilent 5977B MSD, Agilent Technologies, USA; column: Optima-225, 25 m, 50  $\mu$ m, 0.25  $\mu$ m; helium flow rate: 1 ml min<sup>-1</sup>; temperature 230 °C; split ratio 30:1). A standard mixture was analysed, peak identification and quantification of known peaks was done through relative retention times and response factors, unknown peaks were identified by MS.

**Acid and alkaline partial hydrolyses of AGPs.** As described in<sup>63</sup> the samples were acid hydrolyzed in 12.5 mM aqueous oxalic acid (Ox) and heated under pressure for 5 h at 100 °C (Wheaton V-Vial; Bioblock Scientific, Thermolyne Corp., USA). Alkaline hydrolysis (AH) was performed as described in<sup>64</sup>.

**Reduction and labeling of uronic acids with NaBD<sub>4</sub>.** AGP samples (40 mg) were solubilized in 40 ml of deionized water and treated with N-cyclohexyl-N'-(2-morpholinoethyl) carbodiimidemethyl-*p*-toluenesulfonate and solutions of NaBD<sub>4</sub><sup>65</sup>. After this treatment the sample was purified through dialysis (MWCO 6–8 kDa) for 3 days at 4 °C against deionized water and freeze-dried.

**Linkage-type analysis.** Freeze-dried AGP samples (2–8 mg) of was dissolved in dimethylsulphoxide (DMSO) and methylated according to<sup>66</sup>. The partially methylated alditol acetates were extracted with dichloromethane and separated via GC (instrument: Agilent 7890B, Agilent Technologies, USA; column: Optima-1701–0.25 μm, Machery & Nagel, Düren, Germany; helium flow rate: 1 ml min<sup>-1</sup>; temperature: initial 170 °C for 2 min, with rate 1 °C min<sup>-1</sup> to 210 °C, then with rate 30 °C min<sup>-1</sup> until 250 °C is reached and following hold time of 10 min) and analyzed by MS and FID as above.

**Determination of Hydroxyproline (Hyp) content.** Hyp content was determined colorimetrically according to<sup>67</sup>.

**Elemental Analysis.** Quantitative determination of nitrogen was performed with HEKAtech CHNS Analyzer (HEKAtech GmbH, Wegberg, Germany) through burning in an excess of oxygen. The resulting combustion products were further analyzed. To calibrate the system aminobenzenesulfonamide was used.

**Size-exclusion chromatography (SEC).** SEC analysis was performed as described in<sup>68</sup>. For calculation of hydrodynamic volumes, different pullulans with the masses 85.3 × 10<sup>4</sup> Da, 38.0 × 10<sup>4</sup> Da, 18.6 × 10<sup>4</sup> Da, 2.37 × 10<sup>4</sup> Da, 1.22 × 10<sup>4</sup> Da (all from Shodex, Japan), 4.8 × 10<sup>4</sup> Da and 10.0 × 10<sup>4</sup> Da (both Honeywell Fluka, USA) were used. They were prepared and injected in the same way as the sample. A standard calibration curve was done and hydrodynamic volume of the samples were determined through elution time. For absolute molecular weight the peaks of MALS signal were integrated and results were fitted with a second order exponential function.

**Enzyme-linked immunosorbent assay (ELISA).** Binding of the different monoclonal AGP-antibodies JIM8, JIM13, LM2, LM6, KM1 and KM4 was investigated in an indirect ELISA. First the 96-well plates (Nunc, Nalge Nunc International, Roskilde, Denmark) were coated with different concentrations of either native AGP or partial hydrolysed AGP (0, 2.5, 5, 10, 25, 50 μg ml<sup>-1</sup> in double-distilled water, 100 μl per well) at 37.5 °C with open cover for 3 days. Afterwards the plates were washed three times with phosphate buffered saline (PBS) buffer (pH 7.4 with 0.05% Tween 20) and blocked with 0.1% (wv<sup>-1</sup>) BSA in PBS (pH 7.4, 200 μl per well, 1 h at 37.5 °C). The blocked plates were washed again with the same washing buffer as described above. Next 100 μl of the different antibodies (KM1 and KM4 diluted 1:10 (v v<sup>-1</sup>), JIM8, JIM13, LM2 and LM6 diluted 1:20 (v v<sup>-1</sup>)) was pipetted in each well and the plates were incubated for 1 h at 37.5 °C followed by a three-times washing step. The secondary antibody (either Anti-Rat-IgG or Anti-Mouse-IgG conjugated with alkaline phosphatase, Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany) was added in a dilution of 1:500 (v v<sup>-1</sup>) in PBS. After incubating and washing, color was developed by addition of 0.1 mg ml<sup>-1</sup> *p*-nitro-phenylphosphate in 0.2 M TRIS buffer (100 μl well<sup>-1</sup>). Absorption was measured at 405 nm in an ELISA reader (Tecan Spectra Thermo). Samples were tested in duplicate.

**Light microscopy.** To visualize AGPs in the rhizome, cross sections derived from handmicrotome (Euromex Microscopen BV, Arnhem, Netherlands) slicing of fresh rhizomes of *Z. marina* were stored in 80% (v v<sup>-1</sup>) ethanol. Afterwards slices were incubated in βGlcY solution (1 mg ml<sup>-1</sup> in 0.15 M NaCl) for 24 h and washed afterwards three times with 0.15 M sodium chloride solution (negative control without βGlcY).

For detection of calcium-binding regions in the rhizome, sections were either incubated with 25 mM calcium chloride solution for 15 min or with distilled water as a negative control. Afterwards both samples were treated with 2% m v<sup>-1</sup> Alizarin S reagent solution (pH 4) for 5 min. After washing three times with 80% (v v<sup>-1</sup>) ethanol, the sections were investigated by light microscopy (Carl Zeiss AG, Oberkochen, Germany, pictures acquired by Canon EOS 1000D, Canon AG, Tokyo, Japan).

**Calcium binding experiments by ITC.** Binding of Ca<sup>2+</sup> to *Z. marina* AGP was measured with isothermal titration calorimetry (ITC) on the instrument MicroCal PEAQ-ITC (Malvern Instruments Ltd., Malvern, UK). As a negative control Avena sativa AGP, free of uronic acids<sup>26</sup> was used. 280 μl of AGP samples in a concentration of 20 μM were deposited in the sample cell and titrated with 40 μl of calcium chlorid solution in a concentration of 2 mM. The injection pipett was stirred with 750 rpm, one single injection of 0.2 μl was made at the beginning and 32 injections of 1 μl followed. Each with a spacing of 150 s. The results were fitted in the MicroCal PEAQ-ITC analysis software (Malvern Instruments Ltd., Malvern, UK) and the K<sub>D</sub> value, as well as other thermodynamic values were calculated. To remove calcium contaminants in the AGPs before measurement, the samples were treated with Chelex sodium salt (Sigma-Aldrich Chemie GmbH) with a batch method and freeze-dried afterwards. Titrant and titrated solution were soluted in MES buffer solution (20 mM, pH 6) to avoid calcium interaction of conventional buffers and simulate physiological conditions in the plant cell. All experiments were performed in triplicate.

**Bioinformatics.** Sequences were gathered from the annotated ORFs of the *Z. marina* genome<sup>1</sup> and ZM genbank identifiers are used throughout. Sequence gathering of classical AGPs was performed using the MAAB pipeline<sup>29</sup> as implemented in the [R] language (ragp package, GitHub). Chimeric AGPs were identified by first screening the genome for sequences with predicted AG regions (ragp package, GitHub) with an AG region defined by ≥3 matches to the motif “[STA]P” within a 10 residue window. This was followed by filtering for those containing at least one detectable protein domain (Pfam), and a signal peptide (SignalP 5.0). PFAM domains

were annotated both by family and by clan (equivalent to superfamily). Domain structure was visualised using custom [R] scripts. Unaligned sequences for classical and chimeric AGPs are included as Supplementary Data files S1 and S2.

Glycosyltransferase (GT) enzymes from the GT14, GT31, GT61 and GT 77 families and methyltransferase enzymes from the DUF579 family were identified by searching the ORFs of the *Z. marina* genome by BLAST. For each enzyme family, all *A. thaliana* sequences of the respective family were used as queries and duplicate returned sequences removed. For each enzyme family, sequences were aligned using Clustal Omega, and prepared for phylogenetic analysis by trimming to the GT domain and applying the TrimAl-gappyout algorithm (which removes columns predicted to be phylogenetically uninformative<sup>69</sup>). The best fit substitution model was determined to be using Model J and a 1000 bootstrap maximum likelihood phylogeny were calculated with RAxML<sup>70</sup>. Aligned sequence sets for all enzyme families discussed are included as Supplementary Data S3.

Received: 6 February 2020; Accepted: 15 April 2020;

Published online: 19 May 2020

## References

- Olsen, J. L. *et al.* The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea. *Nature* **530**, 331–335 (2016).
- Lee, H. *et al.* The Genome of a Southern Hemisphere Seagrass Species (*Zostera muelleri*). *Plant Physiol.* **172**, 272–283 (2016).
- Aquino, R. S., Landeira-Fernandez, A. M., Valente, A. P., Andrade, L. R. & Mourão, P. A. S. Occurrence of sulfated galactans in marine angiosperms. *Evolutionary implications. Glycobiology* **15**, 11–20 (2005).
- Gloaguen, V. *et al.* Structural characterization and cytotoxic properties of an apiose-rich pectic polysaccharide obtained from the cell wall of the marine phanerogam *Zostera marina*. *J. Nat. Prod.* **73**, 1087–1092 (2010).
- Ly, Y. *et al.* Extraction, Isolation, Structural Characterization and Anti-Tumor Properties of an Apigalacturonan-Rich Polysaccharide from the Sea Grass *Zostera caespitosa* Miki. *Mar. Drugs* **13**, 3710–3731 (2015).
- Johnson, K. L., Jones, B. J., Schultz, C. J. & Bacic, A. Non-enzymic cell wall (glyco) proteins in The Plant Cell Wall (ed. Rose, J.) 111–154 (Blackwell Publishing, 2003).
- Ma, Y., Zeng, W., Bacic, A. & Johnson, K. AGPs through time and space in Annual Plant Reviews online (ed. Roberts, J.) 1–38 (John Wiley & Sons Ltd., 2018).
- Classen, B., Baumann, A. & Utermohlen, J. Arabinogalactan-proteins in spore-producing land plants. *Carbohydr. Polym.* **210**, 215–224 (2019).
- Seifert, G. J. & Roberts, K. The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* **58**, 137–161 (2007).
- Duchow, S., Dahlke, R. I., Geske, T., Blaschek, W. & Classen, B. Arabinogalactan-proteins stimulate somatic embryogenesis and plant propagation of *Pelargonium sidoides*. *Carbohydr. Polym.* **152**, 149–155 (2016).
- Goellner, E. M., Gramann, J. C. & Classen, B. Antibodies against Yariv's reagent for immunolocalization of arabinogalactan-proteins in aerial parts of *Echinacea purpurea*. *Planta Med.* **79**, 175–180 (2013).
- Ellis, M., Egelund, J., Schultz, C. J. & Bacic, A. Arabinogalactan-proteins. Key regulators at the cell surface? *Plant Physiol.* **153**, 403–419 (2010).
- Nguema-Ona, E., Vicré-Gibouin, M., Cannesan, M.-A. & Driouich, A. Arabinogalactan proteins in root-microbe interactions. *Trends Plant Sci.* **18**, 440–449 (2013).
- Lampart, D. T. A., Kieliszewski, M. J. & Showalter, A. M. Salt stress upregulates periplasmic arabinogalactan proteins. Using salt stress to analyse AGP function. *New Phytol.* **169**, 479–492 (2006).
- Kjeldahl, J. Neue Methode zur Bestimmung des Stickstoffs in organischen Körpern. *Z. Anal. Chem.* **22**, 366–382 (1883).
- Strasser, R. Plant protein glycosylation. *Glycobiology* **26**, 926–939 (2016).
- Singh, R. S., Kaur, N., Rana, V. & Kennedy, J. F. Pullulan. A novel molecule for biomedical applications. *Carbohydr. Polym.* **171**, 102–121 (2017).
- Yousefi A. R. & Razavi, S. M. A. Dilute Solution Properties of Emerging Hydrocolloids in *Emerging Natural Hydrocolloids. Rheology and Functions* (ed. Razavi, S. M. A.). 53–73 (John Wiley & Sons Ltd., 2019).
- Paulsen, B. S., Craik, D. J., Dunstan, D. E., Stone, B. A. & Bacic, A. The Yariv reagent: Behaviour in different solvents and interaction with a gum arabic arabinogalactanprotein. *Carbohydr. Polym.* **106**, 460–468 (2014).
- Yates, E. A. *et al.* Characterization of carbohydrate structural features recognized by anti-arabinogalactan-protein monoclonal antibodies. *Glycobiology* **6**, 131–139 (1996).
- Pennell, R. I. *et al.* Developmental Regulation of a Plasma Membrane Arabinogalactan Protein Epitope in Oilseed Rape Flowers. *Plant Cell* **3**, 1317–1326 (1991).
- Ruprecht, C. *et al.* A Synthetic Glycan Microarray Enables Epitope Mapping of Plant Cell Wall Glycan-Directed Antibodies. *Plant Physiol.* **175**, 1094–1104 (2017).
- Willats, W. G., Marcus, S. E. & Knox, J. P. Generation of monoclonal antibody specific to (1–5)-alpha-L-arabinan. *Carbohydr. Res.* **308**, 149–152 (1998).
- Classen, B., Csávás, M., Borbás, A., Dingermann, T. & Zündorf, I. Monoclonal antibodies against an arabinogalactan-protein from pressed juice of *Echinacea purpurea*. *Planta Med.* **70**, 861–865 (2004).
- Goellner, E. M. Untersuchungen zu Arabinogalactan-Proteinen aus *Triticum aestivum* L., *Secale cereale* L. und *Avena sativa* L.: PhD thesis. (CAU - Kiel University, Kiel, 2009)
- Goellner, E. M., Ichinose, H., Kaneko, S., Blaschek, W. & Classen, B. An arabinogalactan-protein from whole grain of *Avena sativa* L. belongs to the wattle-blossom type of arabinogalactan-proteins. *J. Cereal Sci.* **53**, 244–249 (2011).
- Cambridge, M. L., Fraser, M. W., Holmer, M., Kuo, J. & KENDRICK, G. A. Hydrogen sulfide intrusion in seagrasses from Shark Bay, Western Australia. *Mar. Freshwater Res.* **63**, 1027–1038 (2012).
- Kuo, J. & den Hartog, C. Seagrass morphology, anatomy, and ultrastructure in Seagrasses: biology, ecology and conservation (eds. Larkum, A. W., Orth, R., Duarte, C.M.) 51–87 (Springer Netherlands, Dordrecht, 2006).
- Johnson, K. L. *et al.* Pipeline to Identify Hydroxyproline-Rich Glycoproteins. *Plant Physiol.* **174**, 886–903 (2017).
- Johnson, K. L. *et al.* Insights into the Evolution of Hydroxyproline-Rich Glycoproteins from 1000 Plant Transcriptomes. *Plant Physiol.* **174**, 904–921 (2017).
- Egelund, J., Ellis, M., Doblin, M., Qu, Y. & Bacic, A. Genes and Enzymes of the GT31 Family. Towards Unravelling The Function(s) of the Plant Glycosyltransferase Family Members in Annual Plant Reviews: Plant Polysaccharides, Biosynthesis and Bioengineering, Volume 41 (ed. Peter Ulaskov), 213–234 (John Wiley & Sons, 2018).
- Ye, C.-Y., Li, T., Tuskan, G. A., Tschaplinski, T. J. & Yang, X. Comparative analysis of GT14/GT14-like gene family in *Arabidopsis*, *Oryza*, *Populus*, *Sorghum* and *Vitis*. *Plant Sci.* **181**, 688–695 (2011).
- Cenci, A., Chantret, N. & Rouard, M. Glycosyltransferase Family 61 in Liliopsida (Monocot). The Story of a Gene Family Expansion. *Front. Plant Sci.* **9**, 1843, <https://doi.org/10.3389/fpls.2018.01843> (2018).



34. Temple, H. *et al.* Two members of the DUF579 family are responsible for arabinogalactan methylation in *Arabidopsis*. *Plant Direct* **3**, e00117, <https://doi.org/10.1002/pld3.117> (2019).
35. Showalter, A. M. & Basu, D. Extensin and Arabinogalactan-Protein Biosynthesis. Glycosyltransferases, Research Challenges, and Biosensors. *Front. Plant Sci.* **7**, 814, <https://doi.org/10.3389/fpls.2016.00814> (2016).
36. Petersen, B. L., Faber, K. & Ulvskov, P. Glycosyltransferases of the GT77 family in Annual Plant Reviews: Plant Polysaccharides, Biosynthesis and Bioengineering, Volume 41 (ed. Peter Ulvskov) 305–320 (John Wiley & Sons, 2018).
37. Hemminga, M. A. & Duarte, C. M. Seagrasses in the human environment in Seagrass Ecology (eds. Hemminga, M. A. & Duarte, C. M.), 248–291 (Cambridge University Press, 2000).
38. Ovodov, Y. S., Mikheyskaya, L. V., Ovodova, R. G. & Krasnikova, I. N. The pectic substances of Zosteraceae Part V. Smith degradation of zosterine. *Carbohydr. Res.* **18**, 319–322 (1971).
39. Silva, J. M. C. *et al.* Biological activities of the sulfated polysaccharide from the vascular plant *Halodule wrightii*. *Rev. Bras. Farmacogn.* **22**, 94–101 (2012).
40. Kolsi, R. B. A. *et al.* Structural characterization and functional properties of antihypertensive *Cymodocea nodosa* sulfated polysaccharide. *Carbohydr. Polym.* **151**, 511–522 (2016).
41. Nothnagel, E. A. Proteoglycans and related components in plant cells. *Int. Rev. Cytol.* **174**, 195–291 (1997).
42. Olmos, E. L. G., De Garma, J., Gomez-Jimenez, M. C. & Fernandez-Garcia, N. Arabinogalactan Proteins Are Involved in Salt-Adaptation and Vesicle Trafficking in Tobacco by-2 Cell Cultures. *Front. Plant Sci.* **8**, 1092, <https://doi.org/10.3389/fpls.2017.01092> (2017).
43. Zagorchev, L., Petrova, S. L. & Odjakova, M. Arabinogalactan proteins in salt-adapted suspension cultures of *Dactylis glomerata*. *Gen. Appl. Plant Physiol.* **34**, 159–168 (2009).
44. Fernandez-Garcia, N. *et al.* Changes to the proteome and targeted metabolites of xylem sap in *Brassica oleracea* in response to salt stress. *Plant Cell Environ.* **34**, 821–836 (2011).
45. Boughanmi, N. *et al.* NaCl effect on the distribution of wall ingrowth polymers and arabinogalactan proteins in type A transfer cells of *Medicago sativa* Gabès leaves. *Protoplasma* **242**, 69–80 (2010).
46. Hervé, C. *et al.* 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**, 1428–1441 (2016).
47. Zhao, C. *et al.* Arabinose biosynthesis is critical for salt stress tolerance in *Arabidopsis*. *New Phytol.* **224**, 274–290 (2019).
48. Moore, J. P. *et al.* Arabinose-rich polymers as an evolutionary strategy to plasticize resurrection plant cell walls against desiccation. *Planta* **237**, 739–754 (2013).
49. Inaba, M. *et al.* L-Fucose-containing arabinogalactan-protein in radish leaves. *Carbohydr. Res.* **415**, 1–11 (2015).
50. Shimoda, R. *et al.* Enzymatic fragmentation of carbohydrate moieties of radish arabinogalactan-protein and elucidation of the structures. *Biosci. Biotech. Biochem.* **78**, 818–831 (2014).
51. Aalbers, F., Turkenburg, J. P., Davies, G. J. & Dijkhuizen, L. & van Lammerts Bueren, A. Structural and Functional Characterization of a Novel Family GH115 4-O-Methyl- $\alpha$ -Glucuronidase with Specificity for Decorated Arabinogalactans. *J. Mol. Biol.* **427**, 3935–3946 (2015).
52. Lampport, D. T. A., Tan, L., Held, M. & Kieliszewski, M. J. Phyllotaxis Turns Over a New Leaf - A New Hypothesis. *Int. J. Mol.* **21** (2020).
53. Lampport, D. T. A. & Várnai, P. Periplasmic arabinogalactan glycoproteins act as a calcium capacitor that regulates plant growth and development. *New Phytol.* **197**, 58–64 (2013).
54. Lampport, D. T. A., Várnai, P. & Seal, C. E. Back to the future with the AGP-Ca<sup>2+</sup> flux capacitor. *Ann. Bot.* **114**, 1069–1085 (2014).
55. Popper, Z. A. & Fry, S. C. Primary cell wall composition of pteridophytes and spermatophytes. *New Phytol.* **164**, 165–174 (2004).
56. Fu, H., Yadav, M. P. & Nothnagel, E. A. Physcomitrella patens arabinogalactan proteins contain abundant terminal 3-O-methyl-L-rhamnosyl residues not found in angiosperms. *Planta* **226**, 1511–1524 (2007).
57. Bartels, D. *et al.* Evolution of plant cell wall. Arabinogalactan-proteins from three moss genera show structural differences compared to seed plants. *Carbohydr. Polym.* **163**, 227–235 (2017).
58. Lahaye, P. A. & Epstein, E. Salt toleration by plants. *Enhancement with calcium*. *Science* **166**, 395–396 (1969).
59. Cramer, G. R., Läubli, A. & Polito, V. S. Displacement of ca by na from the plasmalemma of root cells. A primary response to salt stress? *Plant Physiol.* **79**, 207–211 (1985).
60. Mizukami, A. G. *et al.* The AMOR Arabinogalactan Sugar Chain Induces Pollen-Tube Competency to Respond to Ovular Guidance. *Curr. Biol.* **26**, 1091–1097, <https://doi.org/10.1016/j.cub.2016.02.040> (2016).
61. Classen, B., Mau, S.-L. & Bacic, A. The arabinogalactan-proteins from pressed juice of *Echinacea purpurea* belong to the hybrid class of hydroxyproline-rich glycoproteins. *Planta Med.* **71**, 59–66 (2005).
62. Blakeney, A. B., Harris, P. J., Henry, R. J. & Stone, B. A. A simple and rapid preparation of alditol acetates for monosaccharide analysis. *Carbohydr. Res.* **113**, 291–299 (1983).
63. Clarke, A. E., Gleeson, P. A., Jermyn, M. A. & Knox, R. B. Characterization and Localization of  $\beta$ -Lectins in Lower and Higher Plants. *Aust J. Plant Physiol.* **5**, 707 (1978).
64. Volk, R.-B., Blaschek, W. & Classen, B. Characterization of an arabinogalactan protein from the pressed juice of *Echinacea purpurea*. Investigations into the type of linkage between the protein and polysaccharide moieties. *J. Nat. Med.* **61**, 397–401 (2007).
65. Taylor, R. L. & Conrad, H. E. Stoichiometric depolymerization of polyuronides and glycosaminoglycuronans to monosaccharides following reduction of their carbodiimide-activated carboxyl group. *Biochemistry* **11**, 1383–1388 (1972).
66. Harris, P. J., Henry, R. J., Blakeney, A. B. & Stone, B. A. An improved procedure for the methylation analysis of oligosaccharides and polysaccharides. *Carbohydr. Res.* **127**, 59–73 (1984).
67. Stegemann, H. & Stalder, K. Determination of Hydroxyproline. *Clin. Chim. Acta* **18**, 267–273 (1967).
68. Happ, K. & Classen, B. 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, <https://doi.org/10.3390/plants8110460> (2019).
69. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl. A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973, <https://doi.org/10.1093/bioinformatics/btp348> (2009).
70. Stamatakis, A. RAXML version 8. A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313, <https://doi.org/10.1093/bioinformatics/btu033> (2014).

## Acknowledgements

The authors thank Prof. P. Knox (University of Leeds, UK) for providing antibodies LM2, LM6, JIM 8 and JIM13.

## Author contributions

L.P. and B.C. planned and designed the research. L.P. collected the plant material and performed the experiments. T.S. and K.J. performed the bioinformatic analyses. L.P., B.C., K.J. and A.B. analysed the data and wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-65135-5>.

**Correspondence** and requests for materials should be addressed to B.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020