Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

PORTLAND PRESS

Research Article

An *Arabidopsis thaliana* arabinogalactan-protein (AGP31) and several cationic AGP fragments catalyse the boron bridging of rhamnogalacturonan-II

Dayan Sanhueza[1,*], Rifat Ara Begum[1,†], Cécile Albenne[2,‡], Elisabeth Jamet[2] and Stephen C. Fry[1]

[1]The Edinburgh Cell Wall Group, Institute of Molecular Plant Sciences, The University of Edinburgh, Daniel Rutherford Building, The King's Buildings, Max Born Crescent, Edinburgh EH9 3BF, U.K.; [2]Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, CNRS, UPS, Toulouse INP, Auzeville-Tolosane, France

Correspondence: Stephen C. Fry (s.fry@ed.ac.uk)

OPEN ACCESS

Rhamnogalacturonan-II (RG-II) is a complex pectic domain in plant primary cell walls. *In vivo*, most RG-II domains are covalently dimerised via borate diester bridges, essential for correct cell-wall assembly, but the dimerisation of pure RG-II monomers by boric acid *in vitro* is extremely slow. Cationic 'chaperones' can promote dimerisation, probably by overcoming the mutual repulsion between neighbouring anionic RG-II molecules. Highly effective artificial chaperones include Pb^{2+} and polyhistidine, but the proposed natural chaperones remained elusive. We have now tested cationic peptide fragments of several *Arabidopsis thaliana* arabinogalactan-proteins (AGPs) as candidates. Fragments of AGP17, 18, 19 and 31 were effective, typically at ~25 μg/ml (9–19 μM), promoting the boron bridging of 16–20 μM monomeric RG-II at pH 4.8 *in vitro*. Native AGP31 glycoprotein was also effective, and hexahistidine was moderately so. All chaperones tested interacted *reversibly* with RG-II and were not consumed during the reaction; thus they acted catalytically, and may constitute the first reported boron-acting enzyme activity, an RG-II borate diesterase. Many of the peptide chaperones became less effective catalysts at higher concentration, which we interpret as due to the formation of RG-II–peptide complexes with a net positive charge, as mutually repulsive as negatively charged pure RG-II molecules. The four unique AGPs studied here may serve an enzymic role in the living plant cell, acting on RG-II within Golgi cisternae and/or in the apoplast after secretion. In this way, RG-II and specific AGPs may contribute to cell-wall assembly and hence plant cell expansion and development.

*Present address: Centro de Biotecnología Vegetal, Facultad de Ciencias de la Vida, Universidad Andrés Bello, Santiago 8370186, Chile
†Present address: Department of Biochemistry and Molecular Biology, University of Dhaka, Curzon Hall, Dhaka – 1000, Bangladesh
‡Present address: Laboratoire de Microbiologie et Génétique Moléculaires (LMGM), Centre de Biologie Intégrative (CBI), Université de Toulouse, CNRS, UPS, Toulouse, France

Received: 24 June 2022
Revised: 2 September 2022
Accepted: 2 September 2022

Accepted Manuscript online: 5 September 2022
Version of Record published: 23 September 2022

Introduction

Rhamnogalacturonan-II (RG-II) is a functionally important pectic domain of the plant primary cell wall. It is relatively small (~5 kDa in its monomeric form), but has an exceedingly complex and well-conserved structure. The anionic backbone of RG-II comprises seven to ten (1 → 4)-linked α-D-galacturonic acid (GalA) residues and is covalently contiguous with the backbones of homogalacturonan and RG-I. RG-II's backbone carries six side-chains (A–F), four of which are themselves anionic [1–3]. The whole RG-II molecule is thus highly negatively charged at physiological pH in the apoplast (often pH < 5 [4]) or Golgi cisternae (often pH ≃ 6 [5]). Intriguingly, two RG-II domains can be covalently cross-linked to each other via a borate diester bond between the apiose residue of their 'A' side-chains [6], leading to a molecular mass of ~10 kDa (Figure 1). The resulting dimerisation of RG-II can be demonstrated by column chromatography [7,8] or by polyacrylamide gel electrophoresis (PAGE) [9]. It is its ability to undergo boron bridging that supports the claim that RG-II is

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

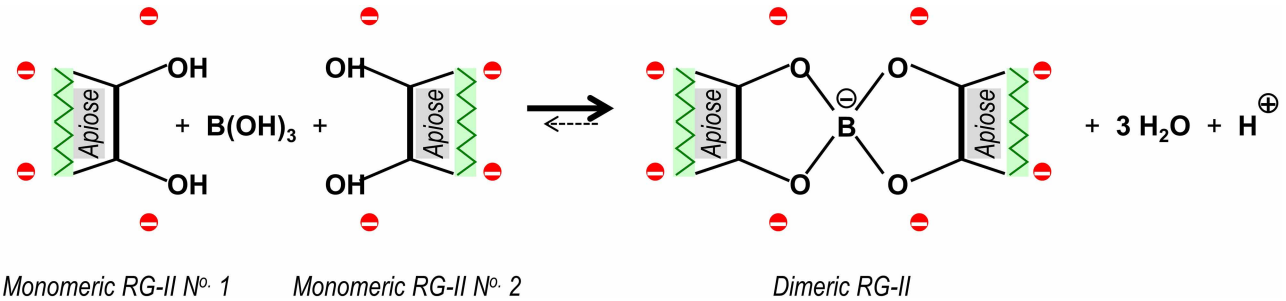

Figure 1. The essential reaction involved in boron bridging of RG-II.

Part of the neutral apiose residue of side-chain A (in each of two monomeric RG-II molecules) is shown. The charges of its four near-neighbouring anionic sugar residues are rendered as red circles (two α-GalA residues in the RG-II backbone plus an α-GalA and a β-GalA residue attached to the rhamnose adjacent to the apiose). Monomeric RG-II has a total of ∼11 additional anionic sugar residues, giving it a net charge of ∼−14 at the pH (4.8) used in our experiments. It is evident that the close approach of two RG-II monomers for boron bridging requires the overcoming of considerable electrostatic repulsion, and dimerisation introduces an additional negative charge on the previously neutral boron atom.

biologically important: failure to make these bridges, either because of boron deficiency or because of a mutation compromising the structure of side-chain 'A', results in defective cell growth and changes in cell-wall porosity [10–15].

Since RG-II is highly negatively charged, it may be difficult for pairs of RG-II molecules to approach closely enough to achieve boron bridging. Indeed, when an aqueous solution of RG-II is simply mixed with boric acid *in vitro*, almost no dimerisation occurs [16,17]). Boron bridging can, however, be very strongly enhanced in the presence of boric acid by the addition of appropriate cationic 'chaperones'. It is assumed that these cancel the negative charges of RG-II and thus permit two RG-II molecules to come close enough for boron bridging. Highly effective chaperones (in reaction mixtures containing 500 μM monomeric RG-II and 1.2 mM boric acid, buffered at pH 3.3) are the metal cations Pb^{2+}, Sr^{2+} and Ba^{2+} at 0.5 mM, whereas Ca^{2+}, Cd^{2+}, Cu^{2+}, Mg^{2+}, Ni^{2+} and Zn^{2+} are almost ineffective at this concentration) [16]. However, since these are not essential elements for plants, they cannot be the natural chaperones responsible for RG-II dimerisation *in vivo*. Ca^{2+} also works as a chaperone, but only at much higher (probably non-physiological) concentrations [16].

In attempts to discover the natural chaperone(s) responsible for dimerising RG-II *in vivo*, various other cations were tested [17]. A cationic extensin (glycoprotein rich in hydroxyproline tetraarabinoside) was effective, but the most successful chaperone found was polyhistidine (degree of polymerisation ∼100). Polyhistidine does not occur naturally; however, some arabinogalactan-proteins (AGPs) are rich in histidine and/or other cationic amino acid residues and may serve as natural chaperones, enabling the covalent reaction shown in Figure 1.

Like RG-II, AGPs have crucial biological functions in plants. For example, they appear to promote cell proliferation [18,19], cell expansion [20], xylem differentiation [21–23], somatic embryogenesis [24], pollen-tube growth [25], apoptosis [26], and cell–cell recognition [27,28]. Arabidopsis *agp19* mutants, for instance, have smaller, rounder, paler-green leaves than the wild-type, decreased stem growth, and fewer siliquae and seeds; and there were abnormalities in cell expansion, cell division and cell shape [29]. These abnormal phenotypes, especially in development, are compatible with the suggestion that some of the biological roles of AGPs and RG-II might share a common basis — for example, that certain AGPs might facilitate RG-II dimerisation.

The arabidopsis genome encodes 85 putative AGPs, some of which are highly cationic and all of which are processed (e.g. glycosylated) in the Golgi system [30]. Significantly, this is the same compartment where RG-II biosynthesis occurs [31,32] and the predominant site of boron bridging *in vivo* [33]. Furthermore, AGPs often acquire a glycosylphosphatidylinositol (GPI)-anchor, attaching them to lipid rafts of the plasma membrane [30,34,35] — another proposed site of RG-II dimerisation [36]. Thus, AGPs are at least transiently present in the right places to serve as chaperones responsible for RG-II cross-linking.

Therefore, we have now tested four cationic (His- or Lys-rich) peptide fragments of arabidopsis AGPs (AGP17, AGP18, AGP19 and AGP31; Table 1, Supplementary Figure S1) for 'polyhistidine-like' chaperone activity capable of dimerising RG-II *in vitro*. Other than these specific peptide fragments, AGP17, 18 and 19

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

Table 1 Cationic peptides and proteins exhibiting chaperone activity, catalysing RG-II dimerisation

Name	Sequence [1]	DP	Molecular mass (Da)	Predicted pI [2]	Net charge at pH 4.8 [3]	Optimal concentration (μg/ml) [4]	Completion of RG-II dimerisation [5]
Histidine	H	1	155	7.24	+0.9	50–1000‡	±
His6	HHHHHH	6	841	7.34	+6	400–1000‡	+
Poly-histidine	HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHH	* 106	* 14,555	7.77	+103	25	++++
AGP17p	KHKKKTKKHK	10	1291	8.18	+9	25–400	++
AGP18p	KHKKTTKKSKKH	12	1479	8.21	+8	25–400	++
AGP19p1	KHKRKHKHKRHHH	13	1794	8.84	+9	50	+++
AGP19p2	ASASKHKRKHKHKRHHHASAS	21	2427	8.84	+9	50	+++
AGP31p	YHHGHHHPHPPHHHHPHPHPHPHP	24	2935	7.96	+14	25–600	++
AGP31	EEVNHKTQTPSLAPAPAPYHHGHHHPHPPHHH HPHPHPHPHPPAKSPVKPPVKAPVSPPAKPPV KPPVYPPTKAPVKPPTKPPVKPPVSPPAKPPV KPPVYPPTKAPVKPPTKPPVKPPVYPPTKAPV KPPTKPPVKPPVYPPTKAPVKPPTKPPVKPPV SPPAKPPVKPPVYPPTKAPVKPPVSPPTKPPV TPPVYPPKFNRSLVAVRGTVYCKSCKYAAFNT LLGAKPIEGATVKLVCKSKKNITAETTTDKNG YFLLLAPKTVTNFGFRGCRVYLVKSKDYKCSK VSKLFGGDVGAELKPEKKLGKSTVVVNKLVYG LFNVGPFAFNPSCPK	335	† 35,914	9.14	+59	25	+++

*Mean of a polydisperse preparation;
†Calculated after removal of N-terminal signal sequence but without proline hydroxylation and glycosylation. The native glycoprotein was tested for chaperone activity;
‡No supra-optimal concentration was found;
[1] Amino acid residues with cationic and anionic side-chains are highlighted in shades of blue/cyan and red/magenta respectively;
[2] Isoelectric point predicted by IPC 2.0 ('peptide' if DP < 60, 'protein' if DP > 60): www.ipc2-isoelectric-point.org [47];
[3] Net charge at pH 5.5 predicted by IPC 2.0 ('peptide' if DP < 60, 'protein' if DP > 60): www.ipc2-isoelectric-point.org [47];
[4] Where a single value is given, this was a well-marked optimum and higher concentrations were less effective;
[5] After 16–24 h at optimal concentration.

are not cationic; however, much of the length of AGP31 is very rich in Lys, with many of the lysines occurring regularly every 4 or 8 residues (Table 1). There is evidence from *in-vitro* interactions on nitrocellulose membranes that AGP31 can bind to pectic components [37]. Proteins with His- and/or Lys-rich domains are not unique to arabidopsis, but also occur in most, probably all land plants. For example, proteins homologous to AGP31 can be found in flowering plants like *Gossypium hirsutum* (GhAGP31 [38]) and *Daucus carota* (DcAGP1 [39]). His/Lys domains can also be found in proteins of plants of more ancient lineages like the Equisetales (onekp|CAPN_scaffold_2004677_Equisetum_diffusum), Lycopodiales (onekp|PQTO_scaffold_2013256_Lycopodium_deuterodensum.2) and Psilotales (onekp|QVMR_scaffold_2003208_Psilotum_nudum). Interestingly, the presence of boron-bridged RG-II was described in all these plant orders [40].

The mRNA expression of *AGP17*, *AGP18* and *AGP19* is tissue-specific, suggesting roles in growth and development: for example *AGP17* is maximally expressed in seedlings, rosette-leaves, stems and flowers; *AGP18* in roots, rosettes, stems and flowers; and *AGP19* in roots, stems and flowers [29]. AGP31 is an abundant glycoprotein in cell walls of etiolated hypocotyls [41].

In this work, we tested the chaperone activity of the His- and/or Lys-rich peptides of the four mentioned AGPs as well as the native glycoprotein AGP31 and some related molecules (Table 1), and report that they do indeed promote the boron bridging of RG-II in the presence of boric acid *in vitro*. We discuss the possibility that these molecules could actually be the first enzymes involving boron compounds to be described.

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

Results

Histidine, hexahistidine and polyhistidine as potential chaperones in RG-II dimerisation

It was known that polyhistidine (DP ~ 10^2) at 100–700 μg/ml favoured the dimerisation of monomeric RG-II in the presence of boric acid [17]. In the present work, the dose–response relationship was investigated in greater detail (Figures 2a, 3a). In the absence of polyhistidine, boric acid caused little dimerisation. We consistently found that the optimal polyhistidine concentration was ~25–50 μg/ml (1.4–2.8 μM), which brought about the essentially complete dimerisation of a 7–14-fold molar excess of RG-II within 16 h of incubation. Higher concentrations (>200 μg/ml) gave much weaker dimerisation. Polyhistidine was effective only when the concentration of boric acid exceeded that of RG-II (~20 μM; Supplementary Figure S4).

Polyhistidine has a predicted pI of ~7.8 (Table 1), so at the pH of the gel buffer (8.8) it had a slight net negative charge (entirely due to the C-terminal −COOH group). This, together with its relatively large size (14.6 kDa), caused polyhistidine to migrate only a short distance into the gel. It stained weakly with Ag^+ (Figure 2a).

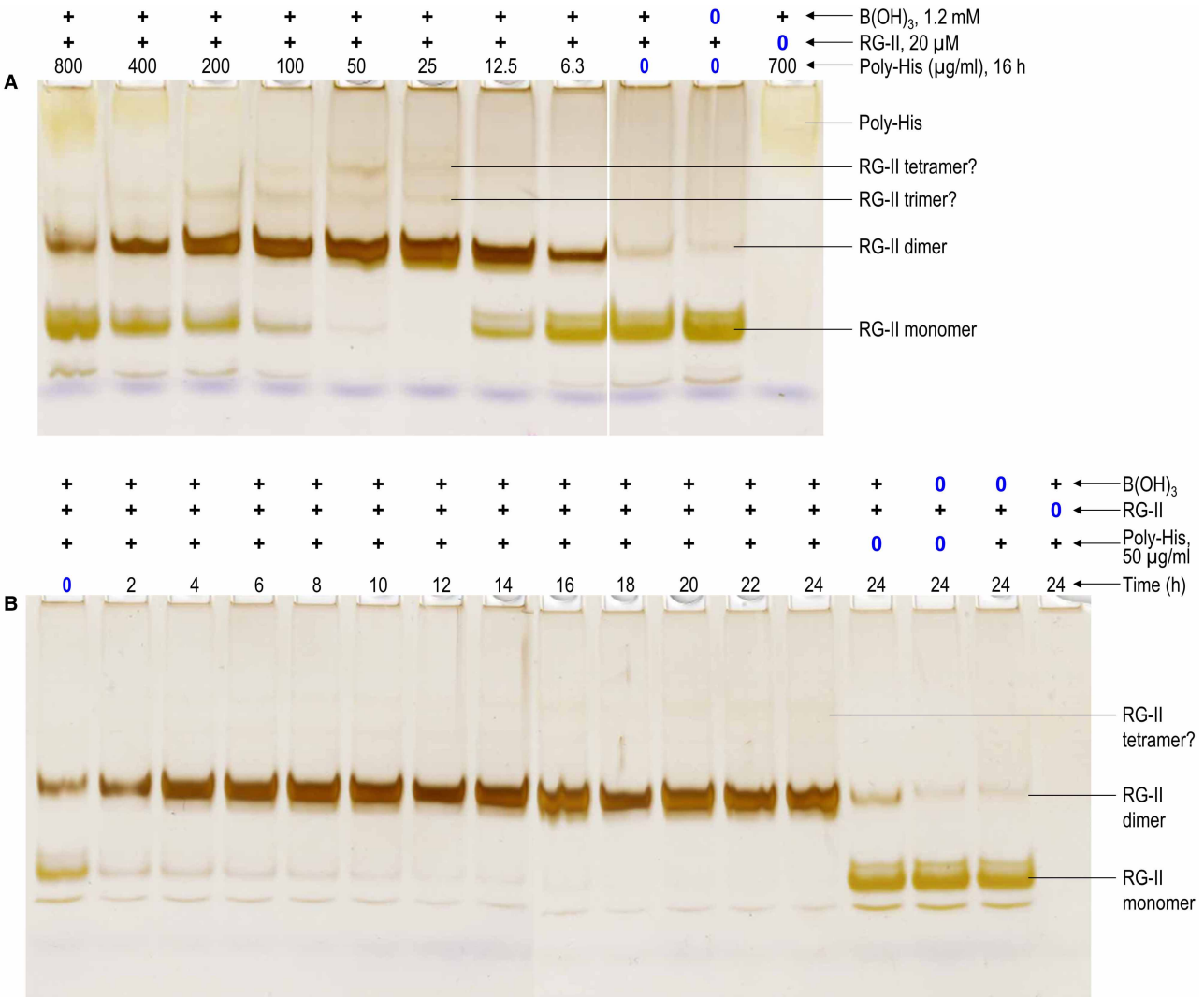

Figure 2. Polyhistidine as a cationic chaperone, catalysing the boron bridging of RG-II.

(**a**) The reaction mixture contained 100 μg/ml RG-II monomer (≈20 μM), 1.2 mM boric acid, 0–800 μg/ml (≈0–44 μM) polyhistidine.Cl^-, and 50 mM acetate (Na^+) buffer pH 4.8. Controls lacked boric acid or RG-II, as indicated above the gels. After 16 h, 0.8 μg of the RG-II was analysed by PAGE followed by silver staining. (**b**) As (**a**) but constant 50 μg/ml polyhistidine and varying incubation time (0–24 h). The bands are quantified in Figure 3.

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

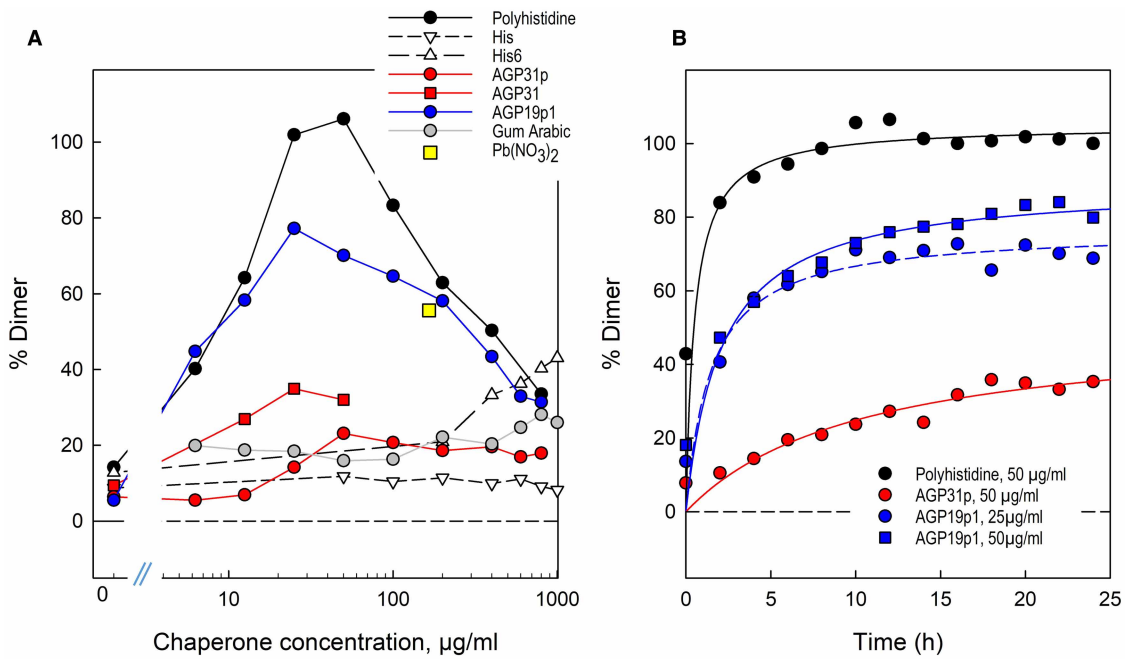

Figure 3. Quantification of chaperone-dependent RG-II dimerisation.

Silver-stained gels (shown in Figures 2, 4–7 and Supplementary Figure S6) were scanned and the luminosity of the monomer and dimer bands was estimated in Photoshop. The proportion of RG-II dimerised was thereby estimated (see Supplementary Figures S2, S3). In (**a**), the incubation time was 16 h [the single 0.5 mM lead nitrate sample is plotted at 166 µg/ml on the *x*-axis]. In (**b**), the curves are rectangular hyperbolae, fitted to the datapoints by SigmaPlot.

Interestingly, moderate concentrations of polyhistidine (25–200 µg/ml) provoked the formation of faint, slower-moving RG-II bands that might be interpreted as trimers and tetramers of RG-II (Figure 2a; also detectable in Supplementary Figure S5). These bands are unlikely to be RG-II–polyhistidine complexes because they became less abundant at the highest polyhistidine concentrations. We tried to elute the putative trimers and tetramers but they appear to change into dimers during the elution process. Nevertheless, it is difficult to envisage them being other than RG-II-derived, given the ingredients used in our reaction mixtures.

Time-course studies revealed that the dimerisation of RG-II caused by 50 µg/ml polyhistidine was initially very rapid, almost reaching completion in the first 2–4 h; thereafter, dimerisation was considerably slower, but did go to completion within ∼16 h (Figures 2b, 3b).

Although polyhistidine was an excellent chaperone able to boron-bridge RG-II, monomeric histidine (0.16 kDa; pI = 7.2) had only very slight chaperone activity compared with the histidine-free control. Any effect of monomeric histidine was approximately equal at all concentrations tested (50–1000 µg/ml; Figures 3a, 4a).

His_6, which is intermediate between histidine and polyhistidine, showed some chaperone activity, moderately promoting RG-II dimerisation, especially at concentrations of 400–1000 µg/ml; no supra-optimal concentration was found (Figures 3a, 4b). In this experiment, as in some others, a small amount of dimerisation occurred in the absence of chaperones, and even in the absence of deliberately added boron; however, this slight background activity does not invalidate the above main conclusions. Unlike histidine, His_6 was stainable on polyacrylamide gels, migrating between RG-II monomer and dimer though not modifying their mobilities. Thus the RG-II–His_6 complex (as with all RG-II–chaperone complexes tested in this work) was very short-lived compared with the electrophoresis run-time. The observed electrophoretic mobility towards the anode agrees with the prediction that His_6 (pI ≃ 7.3) has a net negative charge during PAGE (gel buffer pH 8.8).

AGP31 and its histidine-rich peptide as chaperones promoting RG-II

Although polyhistidine is an excellent chaperone, causing the boron bridging of RG-II, it is not a naturally occurring protein. However, plants do possess certain highly cationic (glyco)proteins, especially AGPs, that

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

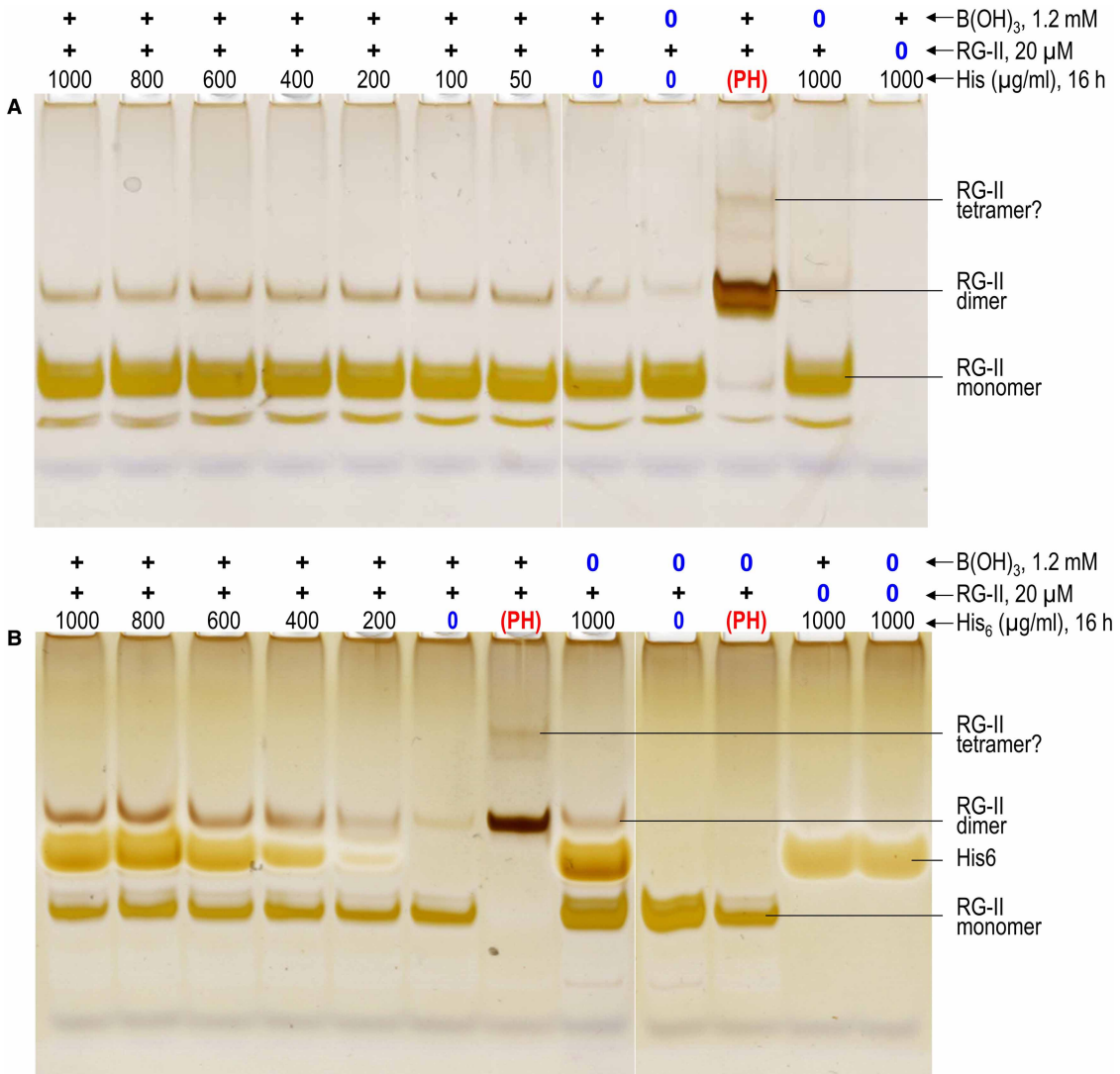

Figure 4. Concentration dependence of histidine and hexahistidine as potential chaperones.

The reaction mixtures contained 100 µg/ml RG-II monomer (≈20 µM), 1.2 mM boric acid, 0–1000 µg/ml histidine (0–6450 µM) (**a**) or $His_6.Cl^-$ (His_6; 0–966 µM) (**b**), and 50 mM acetate (Na^+) buffer pH 4.8. Controls lacked boric acid or RG-II, as indicated above the gels. Other controls contained 50 µg/ml polyhistidine.Cl^- (PH) instead of His or His_6. After 16 h at 20°C, 0.8 µg of the RG-II was analysed by PAGE followed by silver staining. The bands are quantified in Figure 3.

have sequences very rich in histidine. For example, arabidopsis AGP31 includes the sequence YHHGHHHPH PPHHHHPHPHPHPHP near its *N*-terminus (Supplementary Figure S1; Table 1; referred to here as AGP31p), which we synthesised and tested for chaperone activity. It was moderately effective at promoting RG-II dimerisation within 16 h (Figures 3b, 5b). The strongest effect was observed at ∼50 µg/ml; higher concentrations up to 800 µg/ml had very similar effects (Figures 3a, 5a). RG-II dimerisation by AGP31p was gradual over a 24 h period but never approached completion; over the first ∼12 h, dimerisation was approximately linear but reached only ∼30% completion within 12 h (Figures 3b, 5b). In a similar experiment, 50 µg/ml polyhistidine was more effective and faster acting (almost 100% dimerisation within 4 h; Figures 2b, 3a) than AGP31p.

Native AGP31 glycoprotein (∼50–90 kDa [41]) also exhibited chaperone activity, with an optimal concentration of ∼25 µg/ml (roughly 0.2–0.5 µM; Figures 6, 3a). The molar concentration is difficult to state with precision because of glyco-heterogeneity. However, the native glycoprotein is clearly effective at dimerising RG-II at a much lower molar concentration than the peptide, AGP31p (∼17 µM; Figure 5).

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

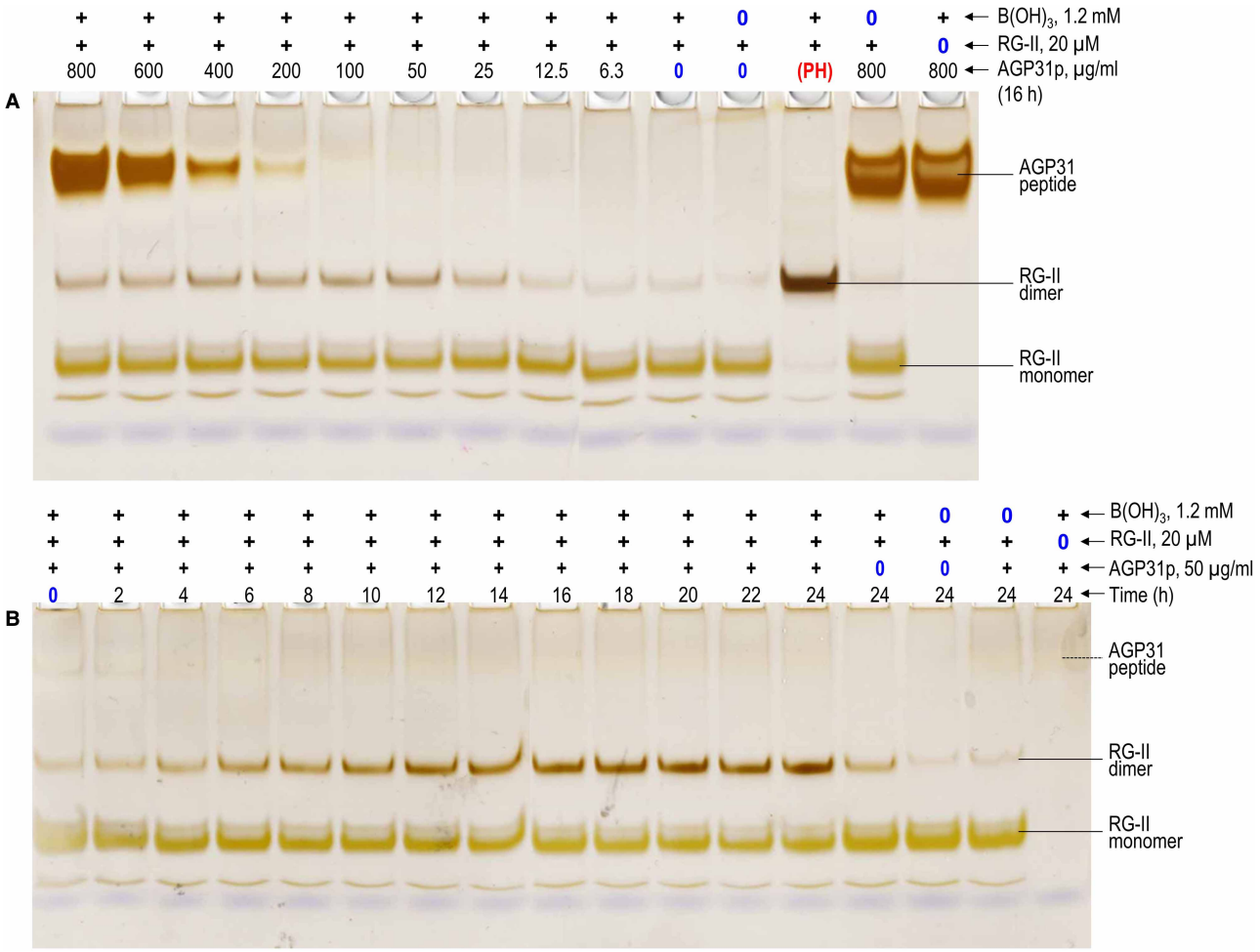

Figure 5. AGP31p as a cationic chaperone, catalysing the boron bridging of RG-II.

(**a**) The reaction mixtures routinely contained 100 µg/ml RG-II monomer (~20 µM), 1.2 mM boric acid, 0–800 µg/ml (0–273 µM) AGP31p, and 50 mM acetate (Na^+) buffer pH 4.8. Controls lacked boric acid or RG-II, as indicated above the gels. The polyhistidine.Cl^- (PH) for the sample in lane 12 was at 50 µg/ml. After 16 h, 0.8 µg of the RG-II was analysed by PAGE followed by silver staining. (**b**) As (**a**) but constant 50 µg/ml AGP31p and varying incubation time (0–24 h). The bands are quantified in Figure 3.

Cationic peptide fragments of AGP19 are also powerful RG-II chaperones

Another cationic AGP of arabidopsis is AGP19, whose sequence (Supplementary Figure S1) contrasts strongly with that of AGP31. An extremely cationic tridecapeptide fragment of AGP19 (AGP19p1; KHKRKHKHKRHHH; Table 1) was found to be an excellent chaperone, the optimal concentration for RG-II dimerisation being 25–50 µg/ml, leading to ~75% dimerisation (14–28 µM; Figures 3a, 7a). Higher concentrations were less effective, as noted for polyhistidine. As revealed by the dimer:monomer ratios visible in Figure 7a and Figure 3a, AGP19p1 was somewhat more effective than 0.5 mM $Pb(NO_3)_2$, which has been widely used as a [non-biological] model chaperone [42]. AGP19p1 was not detected by silver staining of gels because with a pI of 8.8 it would not be expected to migrate into the gel.

Although an effective chaperone at low concentrations, AGP19p1 did not take the dimerisation of RG-II to completion (Figures 3, 7), unlike polyhistidine. The dimerisation was rapid in first 2 h; thereafter, the process continued slowly (most easily seen by the gradual decrease in RG-II monomer) but never reached completion (Figures 3b, 7b,c).

The above peptide, AGP19p1, differs from native AGP19 in, among other ways, lacking the neutral, glycosylated hydrophilic residues which normally flank the highly cationic KHKRKHKHKRHHH sequence. We

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

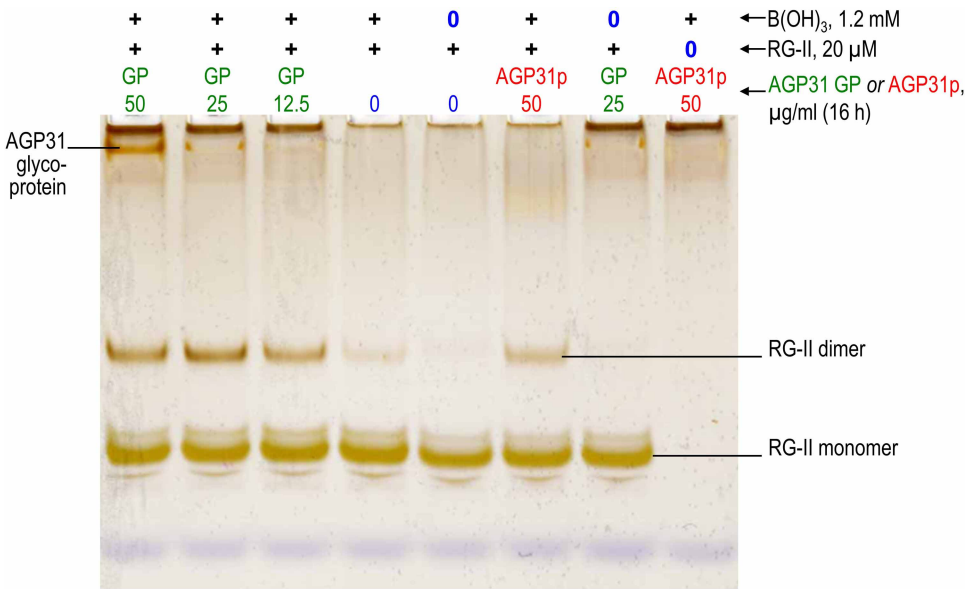

Figure 6. Native AGP31 glycoprotein compared with its peptide fragment AGP31p, as cationic chaperones, catalysing the boron bridging of RG-II.

Reaction mixtures contained 100 μg/ml RG-II monomer (≈20 μM), 1.2 mM boric acid, 0–50 μg/ml (i.e. up to roughly 0.7 μM) AGP31 glycoprotein (GP) or 50 μg/ml (17 μM) AGP31p, and 50 mM acetate (Na^+) buffer pH 4.8. Controls lacked boric acid or RG-II. After 16 h at 20°C, 0.6 μg of the RG-II was analysed by PAGE followed by silver staining. The loading in the left-hand lane contained 0.3 μg AGP31 glycoprotein. The bands are quantified in Figure 3.

therefore produced another AGP19 peptide (ASASTKHKRKHKHKRHHHASAS; AGP19p2; Table 1), whose two neutral ASAS flanks were designed to mimic the neighbouring hydrophilic hydroxyproline-containing sequences present in the full sequence of AGP19 (Supplementary Figure S1). The modified peptide, AGP19p2, also functioned as a chaperone, bringing about RG-II dimerisation, again with a distinct optimum concentration (~50 μg/ml; Figure 8c,e), but with little if any improvement over the 'naked' cationic sequence, AGP19p1. A direct comparison of AGP19p1 versus AGP19p2 is shown in Supplementary Figure S5.

Cationic peptides of AGP17, AGP18 and are also RG-II chaperones

Arabidopsis AGP17 and AGP18 (related to AGP19) also include highly cationic peptide stretches (Table 1). We found that 50 μg/ml AGP17p and AGP18p mimic AGP19p2 as chaperones, dimerising RG-II (Figure 8); the results are compared with those with Pb^{2+} and polyhistidine in Supplementary Figure S5.

AGP17p and AGP18p were effective over a broad range of concentrations, 12–400 μg/ml, but were less effective at the lowest and highest concentrations tested (6 and 800 μg/ml; Figure 8a,b,e). However, neither of them was as effective as AGP19p2 — most easily demonstrated by its ability (at 50 μg/ml) to effect the almost complete disappearance of RG-II monomer (Figure 8c,e). No concentration of AGP17p and AGP18p took RG-II dimerisation to completion within 16 h.

The controls (Supplementary Figure S5) show that the monomeric RG-II preparation lacked contaminating dimer. Polyhistidine was again able to dimerise RG-II completely and Pb^{2+} promoted it very well albeit not to completion. AGP peptides in the absence of boric acid did not form dimers.

Gum arabic AGP and its core protein have little chaperone activity

So far we have focused on AGPs with cationic stretches and observed good chaperone activity. We next tested a non-cationic AGP to discover whether the chaperone activity was due to the intrinsic nature of AGP or to the net positive charge. A much-studied and commercially available AGP is gum arabic. This is a glycoprotein composed of ~2% protein and ~98% carbohydrate. The protein backbone has roughly equal proportions of acidic and basic amino acid residues, and the carbohydrate chains contain ~20% glucuronic acid plus

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

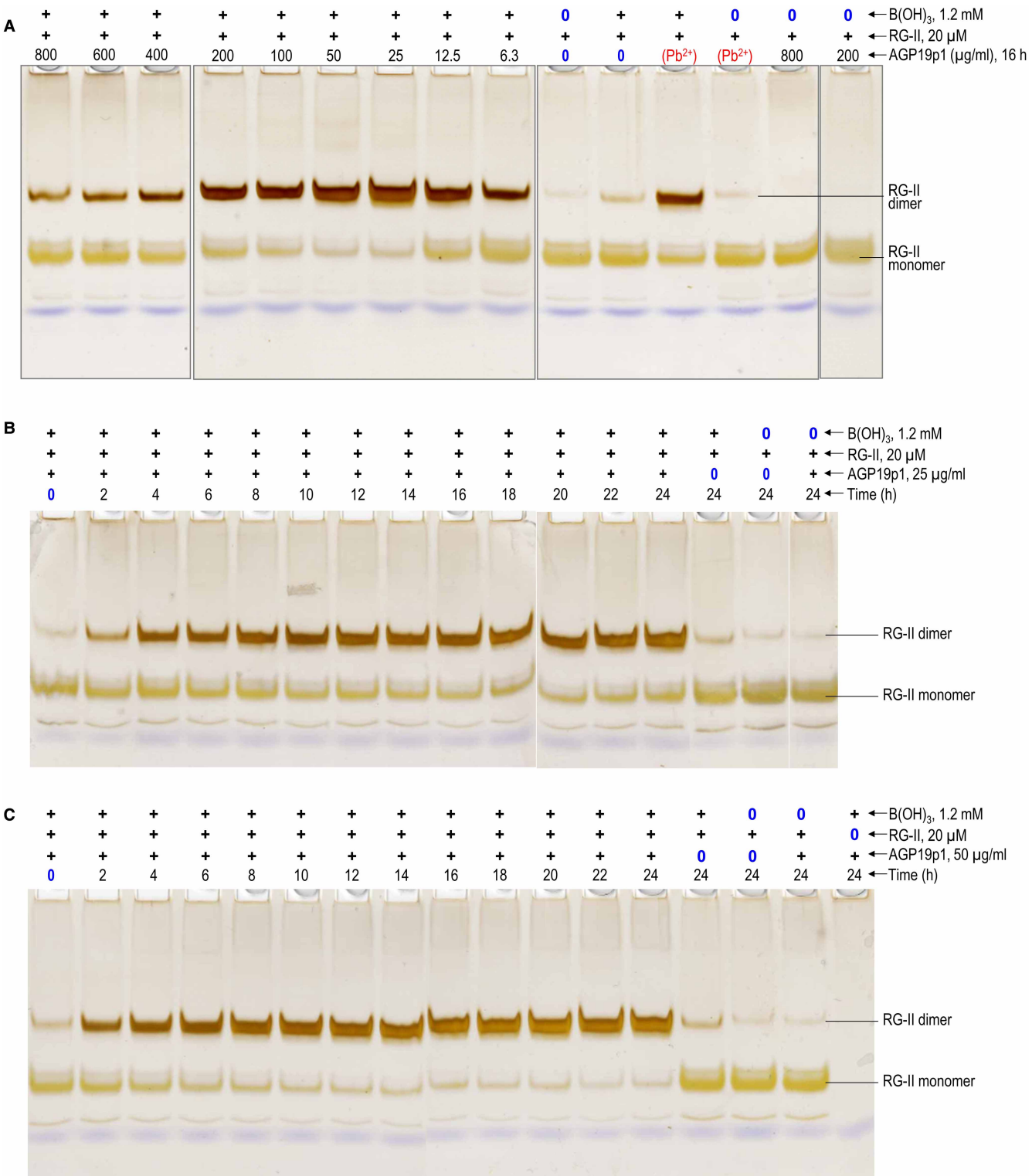

Figure 7. AGP19p1 action as an RG-II chaperone, catalysing the boron bridging of RG-II.

Reaction mixtures contained 100 μg/ml RG-II monomer (≈20 μM), 1.2 mM boric acid, 0–800 μg/ml (0–446 μM) AGP19p1 (tridecapeptide), and 50 mM acetate (Na^+) buffer, pH 4.8 at 20°C. (**a**) Various AGP19p1 concentrations [0–800 μg/ml (0–446 μM)]; always 16 h incubation. (**b**) 50 μg/ml (28 μM) AGP19p1, 0–24 h incubation; (**c**) 25 μg/ml (14 μM) AGP19p1, 0–24 h incubation. Controls lacked boric acid, RG-II or AGP19p1, or contained 0.5 mM Pb^{2+} as an alternative chaperone. After incubation, 0.8 μg of the RG-II was analysed by PAGE followed by silver staining. The bands are quantified in Figure 3.

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

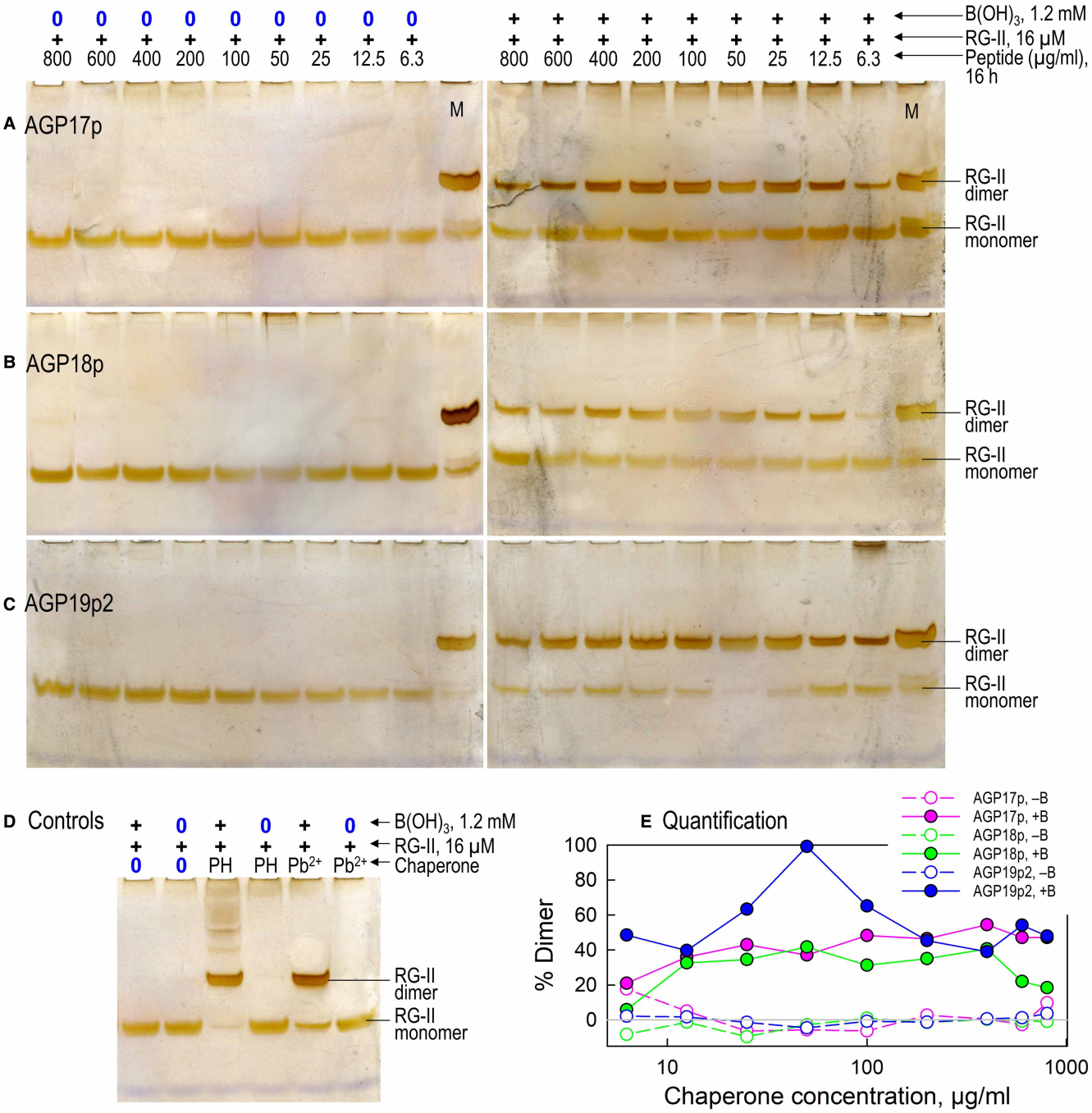

Figure 8. AGP17p, AGP18p and AGP19p2 as chaperones, catalysing the boron bridging of RG-II.

Reaction mixtures contained 80 μg/ml RG-II monomer (≈16 μM), 0 or 1.2 mM boric acid, 6.3–800 μg/ml* of an AGP peptide, and 50 mM acetate (Na^+) buffer, pH 4.8. Left-hand gels: controls lacking boric acid. Peptides were: a, AGP17p; b, AGP18p; c, AGP19p2. Panel (d) shows incubations conducted with 0.5 mM Pb^{2+}, 50 μg/ml polyhistidine (PH) or no chaperone instead of the AGP peptides. After 16 h at 20°C, 0.8 μg of the RG-II was analysed by PAGE followed by silver staining. Marker lanes (M) contained 1 μg monomeric RG-II + 1 μg dimeric RG-II. Panel (e) shows a quantification of the proportion of RG-II dimerised in the presence and absence of boric acid (filled and empty symbols respectively), as estimated from scans of the stained gels shown in (**a**)–(**c**). *The highest (800 μg/ml) peptide concentrations represent ~620 μM AGP17p, 540 μM AGP18p, and 320 μM AGP19p2.

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

4-*O*-methylglucuronic acid residues [43]; thus the overall glycoprotein is highly acidic. The native gum arabic glycoprotein had slight chaperone activity (Supplementary Figure S6a; Figure 3a) but fragments of the protein core, released by acid hydrolysis of the sugar residues, had almost no chaperone activity (Supplementary Figure S6b). The '1000*' loadings (right-hand lanes), contained gum arabic (or its protein core) plus polyhistidine, and show that these components did not produce bands that would interfere in the detection of RG-II monomer and dimer. However, it was not possible to quantify the dimeric bands in Supplementary Figure S6b owing to the presence of neighbouring gum arabic protein fragments.

Discussion

Size matters: large histidine peptides outperform small ones as chaperones

RG-II undergoes very little dimerisation when mixed with boric acid *in vitro* in the absence of chaperones, but several substances tested here were able to catalyse boron bridging. We confirmed the ability of polyhistidine and Pb^{2+} to act as cationic chaperones, assumed to allow polyanionic RG-II molecules to overcome their mutual electrostatic repulsion and approach closely enough for bridging via a borate diester bond.

The polyhistidine used (mean DP 106) was a highly effective chaperone, catalysing the complete dimerisation of RG-II (Figures 2a, 3a). At the other extreme, monomeric histidine did not function as a chaperone (Figure 4a). To explain this, consider that each RG-II molecule has a net charge of roughly −14 at pH 4.8, as used in our reaction mixtures. Presumably the RG-II–histidine ionic bonds were too short-lived to simultaneously cancel enough of the negative charges on the RG-II monomer (20 μM) to enable dimerisation, even in the presence of up to a 320-fold excess (6400 μM) of histidine. Indeed, RG-II–histidine ionic bonds may be compared with the RG-II–Na^+ ionic bonds which would have been present in all reaction mixtures owing to the 1250-fold molar excess (25 mM) of Na^+ from the acetate buffer. Thus small ions such as histidine and Na^+ do not work. Between polyhistidine and histidine, His_6 was a moderately effective chaperone at His_6:RG-II molar ratios between 20 : 1 and 50 : 1 (the highest ratio tested; Figures 3a, 4b).

Towards the discovery of natural chaperones

Although polyhistidine is an excellent artificial chaperone, this polypeptide does not occur naturally. As a step towards the discovery of natural chaperones, we therefore tested several His- and/or Lys-rich peptide sequences known from arabidopsis AGPs which are expected to be present in the apoplast and Golgi system — the feasible sites of RG-II dimerisation *in vivo*. We have shown that several such cationic peptides indeed possess chaperone activity, catalysing RG-II dimerisation: AGP17p, AGP18p, AGP19p1, AGP19p2 and AGP31p, whose sequence, size, isoelectric point and observed chaperone activity are listed in Table 1. Gum arabic (an AGP lacking a cationic stretch) and its peptide core exhibited little or no RG-II chaperone activity (Supplementary Figure S6). We conclude that the peptides of AGP17, 18, 19 and 31 owe their chaperone activity to their cationic nature rather than solely to any intrinsic activities as AGPs — for example an ability to bind 'Yariv antigens' (certain tri-β-glycosylated phenylazo compounds [45]).

AGP31p, a fragment of the glycoprotein AGP31 (which is abundant in rapidly growing etiolated hypocotyl cell walls), acted slowly as a chaperone (Figure 3b). It did not dimerise RG-II to completion and was thus not as effective as polyhistidine. However, AGP31 itself (the native glycoprotein) also acted as a chaperone, catalysing RG-II dimerisation to somewhat nearer completion than did the peptide at same (w/v) concentrations (12.5–50 μg/ml; Figuress 3a, 5a, 6). The optimal concentration of the native glycoprotein was 0.2–0.5 μM (25 μg/ml; Figure 3a), much lower than the tested molar concentrations of the peptides. Thus, the naturally occurring AGP31 glycoprotein is relatively effective as an RG-II chaperone even at very low molar concentrations (∼0.2–0.5 μM) which might be biologically relevant.

Native AGP31 is highly cationic (Table 1; Supplementary Figure S1). The molecular mass of the glycoprotein is higher than the quoted value of 35.9 kDa (Table 1) because many of the proline residues are hydroxylated and *O*-glycosylated [41], and at least one Asn residue is *N*-glycosylated [44]. Indeed, AGP31 exists *in vivo* as a broad range of sizes (∼50–90 kDa) [41]. Full-length AGP31 possesses numerous basic residues (50 Lys, 4 Arg, 15 His), whereas the acidic residues are 3 Asp, 5 Glu and at most 2 GlcA [41], supporting its description as a polycation. AGP31, synthesised in the endomembrane system, is ultimately targeted to the apoplast. Thus it exists in the right places to make a significant contribution to the boron bridging of RG-II *in vivo*. It is particularly relevant that AGP31 can physically interact *in vitro* with the pectic domains RG-I and homogalacturonan [37], though this has not yet been directly demonstrated for RG-II.

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

Other arabidopsis AGPs possessing cationic stretches are AGP17, AGP18 and AGP19. We found that fragments of these (AGP17p, AGP18p, AGP19p1 and AGP19p2) were effective chaperones catalysing RG-II dimerisation, typically at concentrations of 25–400 μg/ml (Figs. 7–8; Table 1).

The gene encoding the backbone of AGP19 is expressed (at the level of mRNA accumulation) in arabidopsis cell-suspension cultures and in all organs examined *in planta*, the order of mRNA abundance being stems > flowers > roots > leaves [29]. Its promoter drove GUS expression particularly in the vasculature of leaves, roots, stems, flowers, styles and siliques. An *agp19* null mutant had small, round, flat rosette-leaves, delayed growth, and short final lengths of hypocotyls and stems. On the abaxial leaf surface, the mutant had epidermal cells on average 0.73× the area of those of the wild-type; conversely, the mutant had 1.25× more epidermal cells per mm^2 of leaf [29], indicating that the mutation predominantly interferes in cell expansion. Likewise in the palisade and spongy mesophyll layers of leaves, the mutant had smaller cells (as seen in paradermal sections) and bigger air spaces. The expression profiles of AGP mRNAs and the phenotypes of *agp* mutants suggest key roles in cell expansion [20,29], which are compatible with AGPs' proposed roles in cell-wall biochemistry, such as catalysing the boron bridging of RG-II.

AGP17, AGP18 and AGP19 are each predicted to have a GPI anchor that will attach the glycoprotein to the outer face of the plasma membrane. This positioning is relevant because most RG-II dimerisation seems to occur either (a) very soon after secretion [9,36], thus probably in the vicinity of the plasma membrane, or (b) within the Golgi endo-membrane system [33], where AGPs are glycosylated. Therefore, at the likely sites of RG-II dimerisation, cationic AGP domains would be available to participate as chaperones catalysing boron bridging.

AGPs 17, 18 and 19 each possess one highly cationic domain: AGP17p, AGP18p and AGP19p respectively (Supplementary Figure S1, Table 1). Excluding these remarkable cationic domains, the glycoproteins differ from each other ionically — the basic:acidic amino acid ratios (after removal of the signal sequence) being 5 : 13, 4 : 10 and 1 : 0 in AGP17, AGP18 and AGP19 respectively. This difference may enable the three glycoproteins to serve different biologically relevant roles that involve ionic interactions, e.g. with RG-II or other pectic domains. The glycosylation of AGP19 does not appear to have been studied in detail, but extracted AGP19 has a wide range of molecular mass, ∼75–200 kDa (Figure 7 of Yang et al. [46]), indicating that it bears a very high proportion (∼80–90%) of carbohydrate. AGP17 is a glycoprotein of ∼80–150 kDa, ∼86% of the mass being carbohydrate (mainly Ara and Gal residues, with smaller proportions of Rha and GlcA [47]). The GlcA residues will decrease the pI of the whole glycoprotein and may thus also influence various biological functions *in vivo*. In the light of the observations on native AGP31 versus its peptide, AGP31p, it is possible that native AGP19 would be a better chaperone than the peptides (AGP19p1 and AGP19p2) tested so far. The same is also likely to apply to native AGP17 and 18.

The AGP and AGP peptides did not cause complete dimerisation of RG-II in our experiments (although AGP19p1 and AGP19p2 came close). Nevertheless, it is possible that they would be more effective under *in-vivo* conditions. For example, Ca^{2+} in the Golgi system or cell walls might serve as a cofactor, potentially synergising with AGPs. Also, we arbitrarily chose pH 4.8 for dimerisation experiments, whereas the natural pH at the specific site of boron bridge formation might differ from this. In addition, we used purified RG-II as substrate, whereas RG-II does not normally occur in free form *in vivo*; on the contrary, it is glycosidically linked to homogalacturonan and/or RG-I, which might help to manoeuvre the RG-II domains into an orientation favouring boron bridging. Finally, the natural boron 'donor substrate' may not be free boric acid, as used here, but rather an organic complex such as one with a glycosylinositol phosphorylceramide [36].

Chaperones interact *reversibly* with RG-II to facilitate boron bridging

The chaperones tested here did not cause any apparent loss of RG-II during PAGE. In contrast, polylysine, which was tested previously [17], did cause RG-II to 'disappear' during PAGE by forming a long-lived RG-II–polylysine complex that migrated towards the cathode (out of the gel). Furthermore, after incubation of RG-II with any of the chaperones tested here, both the dimeric RG-II and the remaining monomer had the same electrophoretic mobility as in the absence of chaperones, indicating that the RG-II–peptide ionic complex (proposed to be the species participating in boron bridge formation) was ephemeral, and dissociated during electrophoresis. Any irreversible RG-II–chaperone complex would be expected to differ from free RG-II in mobility during PAGE since all the peptide chaperones tested are at least 25% the molecular mass of an RG-II monomer. Thus all RG-II–chaperone interactions studied here are readily reversible; the complexes dissociated

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

during PAGE and would presumably also do so *in vivo*. This implies that the chaperones are not consumed when promoting RG-II dimerisation: they acted catalytically.

AGPs and their peptides as boron-acting enzymes

The above considerations support the notion that AGP31 and the cationic peptides tested here can serve as boron-acting enzymes. Crucially, AGP31 is a naturally occurring protein that we have now shown is capable of catalysing a chemical reaction (Figure 1) involving a change in covalent structure between substrates and products. The AGP and peptides were not used up during the reaction, as can be seen in those cases where the chaperone is detectable on silver-stained gels (AGP31, AGP31p, polyhistidine and His_6). The chaperones evidently bind the RG-II substrate reversibly and are released intact after catalysing the reaction. As expected for an enzyme, the larger chaperones (AGP31 and polyhistidine) can catalyse the boron-bridging reaction of an excess of RG-II: for example, 0.2–0.7 μM AGP31 functioned on 20 μM RG-II. In the case of the model chaperone, 1.4 μM polyhistidine was able to catalyse the complete dimerisation of a 14-fold molar excess of RG-II — behaviour typical of an enzyme. Furthermore, AGP19p2 (a fragment which may be a good model of an enzymic AGP) can also catalyse the almost complete dimerisation of RG-II (Figure 8e).

The reaction under investigation (Figure 1) occurs to a slight extent in the absence of a catalyst (e.g. Figures 3, 4); however, this property applies more extremely to many other well-accepted enzymic reactions such as those catalysed by carbonic anhydrase and superoxide dismutase. Moreover, the catalytic effect of AGP31 etc. on RG-II plus boric acid can be mimicked by simple inorganic agents such as $Pb(NO_3)_2$; but again the same can be said of the activities of many other enzymes e.g. carboxylesterases (mimicked by NaOH) and catalase (mimicked by MnO_2).

The reaction (Figure 1) catalysed by AGP31 and the peptides investigated here is of the type acid + alcohol → ester + H_2O. In the case of boric acid + RG-II, the reaction is essentially irreversible in the direction written here, unlike the corresponding reaction with most carboxylesterases, which is essentially irreversible in the opposite direction (ester + H_2O → acid + alcohol). However, enzymes do not dictate the direction of a reaction, so we suggest that this consideration should not affect the nomenclature of the enzyme activity proposed here: it is an esterase, more specifically an RG-II borate diesterase. It is of interest that the Enzyme Commission does not currently list any enzymes acting on boron compounds (https://iubmb.qmul.ac.uk/enzyme/EC3/).

Many peptide chaperones lose effectiveness at high concentration

Many of the chaperones tested, e.g. polyhistidine, exhibited an optimum concentration favouring boron bridging; higher and lower concentrations were less effective. At one extreme, boric acid plus RG-II in the complete absence of chaperones led to negligible dimerisation (e.g. Figure 2a), presumably because the negatively charged RG-II molecules (net charge approximately −14 at pH 4.8) mutually repelled. Conversely, in the presence of a high concentration of polyhistidine, two factors might impede dimerisation: (i) each RG-II molecule would become ionically bonded to an excess of the cationic peptide, forming a complex with a net positive charge, and such complexes would also mutually repel electrostatically, inhibiting dimerisation; (ii) only rarely would two RG-II molecules become simultaneously attached to the same polyhistidine molecule. However, at intermediate polyhistidine concentrations, there was an optimum where the complexes exhibited little electrostatic repulsion and two RG-II molecules would frequently become bonded side-by-side to the same polyhistidine molecule, enabling optimal boron bridging (Figure 2a).

His_6 has a net charge of ∼+5 at pH 4.8 so it is expected to bond ionically to RG-II under the incubation conditions used. We suggest that the RG-II–His_6 ionic linkage is much shorter-lived than the RG-II–polyhistidine linkage, and therefore at any moment the great majority of the RG-II molecules were not attached to enough His_6 molecules to acquire a net positive charge whose cationic repulsion would inhibit dimerisation. This would explain why even a very high His_6 concentration did not inhibit RG-II dimerisation.

Most of the AGP-based cationic peptides tested (AGP17p, AGP18p, AGP19p1, AGP19p2, and possibly AGP31p and the native AGP31 glycoprotein) also exhibited a similar phenomenon to that reported for polyhistidine, each exhibiting optimal concentrations (though sometimes a fairly broad range) under the experimental conditions used (Figs. 5–8; collated in Table 1 and Figure 3). This is likewise interpreted as evidence for an optimum, where the RG-II–chaperone complex has little or no net charge (and thus minimal electrostatic repulsion), and two RG-II molecules can simultaneously bind to the same glycoprotein or peptide molecule, as suggested above for polyhistidine.

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) 479 1967–1984
https://doi.org/10.1042/BCJ20220340

Why does chaperone-mediated dimerisation of RG-II decelerate during incubation *in vitro*?

The rate of RG-II dimerisation at optimal chaperone concentrations was often initially rapid, but later slowed down — in some cases eventually reaching completion, and in other cases not (Figure 3b). For example, the optimal polyhistidine concentration catalysed very rapid boron bridging during the first ∼2 h; thereafter the process slowed down but nevertheless reached completion. In other cases the deceleration was more extreme, and dimerisation never reached completion. For example, the boron bridging of RG-II by 50 μg/mg AGP19p1 was gradual: by 24 h, the system was approaching ∼80% complete dimerisation. At 25 μg/ml AGP19p1, ∼70% dimerisation had been achieved by 24 h and no further dimerisation appeared still to be occurring at that time.

It is unclear why the dimerisation sometimes stops before completion. We suspected that the reaction might be reaching an equilibrium, after which little or no further net dimerisation was occurring. However, when starting with pure dimer, we did not detect any appreciable monomerisation (even in the presence of chaperones), arguing against an equilibrium reaction. Furthermore, the observation that RG-II dimerisation *can* proceed to completion with certain chaperones argues against the hypothesis that an equilibrium is reached at which dimerisation equals re-monomerisation. And we assume that the peptides are stable during the (up to) 24 h incubations. One possibility is that dimeric RG-II has a stronger affinity for some chaperones than does the monomer so that the remaining monomer cannot be adequately chaperoned for boron bridging to continue.

Conclusions

Most of the pectic RG-II domains in living plants are covalently dimerised via borate diester bridges, essential for cell-wall assembly and correct porosity, but the formation of RG-II dimers from pure RG-II plus boric acid *in vitro* is extremely slow. Cationic chaperones greatly expedite boron bridging *in vitro*, probably by overcoming the mutual repulsion between neighbouring pairs of negatively charged RG-II molecules and/or by seating two RG-II molecules adjacent on the same chaperone molecule. Naturally relevant chaperones may include His-rich extensins [17], and highly effective artificial models include Pb^{2+} and polyhistidine [9,17,42].

We have now continued the search for naturally relevant chaperones, focusing on AGPs that are cationic overall and/or possess small, highly cationic domains. Cationic fragments of four such arabidopsis AGPs (AGP17, 18, 19 and 31) were effective, catalysing the boron bridging of RG-II *in vitro*. At least one of them (AGP31), was also effective as the native glycoprotein. We showed that all chaperones tested interact reversibly with RG-II, and so can be described as functioning catalytically — i.e. enzymically. To the best of our knowledge, such a role would be the first example of an enzyme activity (RG-II borate diesterase) acting on boron compounds.

The four unique AGPs studied here may serve a chaperoning role in the living plant cell, catalysing RG-II dimerisation within the Golgi cisternae and/or in the apoplast adjacent to the plasma membrane. In this way, RG-II and specific AGPs may be involved in guiding cell-wall assembly and hence cell expansion and development. Indeed, RG-II is present in the cell walls of all land-plants [40], and AGP-related sequences possessing cationic peptide domains also occur over a wide taxonomic range.

Materials and methods

Peptides and proteins

The peptides and artificial proteins used in this study (Table 1) were from Sigma–Aldrich (Poole, Dorset, U.K.) (polyhistidine; mean degree of polymerisation, DP, 106) and Neo Scientific (Cambridge, MA, U.S.A.) (hexahistidine; His_6), both as chloride salts, and Thermo Fisher Scientific (Inchinnan, Renfrew, U.K.) (AGP17p, AGP18p, AGP19p1 and AGP19p2). AGP19p2 was designed to add flanking hydrophilic (Ala-Ser-Ala-Ser) regions at both ends of the highly cationic tridecapeptide stretch, mimicking the biological glycoprotein sequence, where the positively-charged tridecapeptide is flanked by hydrophilic, glycosylated -Ala-Hyp- repeats. Gum arabic was from Sigma–Aldrich (G9752), as was L-histidine (H-8125). For preparation of de-glycosylated gum arabic, 500 mg gum arabic was subjected to acid hydrolysis in 2 M trifluoroacetic acid (TFA) at 100°C for 1 h, then dialysed in 12 kDa-cutoff dialysis tubing against several changes of water until all TFA had been removed, and finally dried; the yield was 14 mg. The native AGP31 glycoprotein was isolated from etiolated hypocotyls as described by Hijazi et al. [41]. The AGP31p cationic peptide was synthesised by Millegen (Labège, France).

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

RG-II isolation

Cultured *Rosa* cells ['Paul's Scarlet' rose, a complex hybrid; grown in boron-free medium [9] were rinsed in water, then alcohol-insoluble residue (AIR) was prepared by stirring in 75% ethanol at 20°C for 4–6 h twice. Pectin in the AIR was de-esterified with 1 M Na_2CO_3 at 4°C for 16 h, then the AIR suspension was slightly acidified with acetic acid, rinsed with water until neutral and freeze-dried. Endopolygalacturonase (10 U/ml; Megazyme, http://www.megazyme.com/) was added (~50 μl/mg AIR) and incubated at 20°C for 16 h in 50 mM acetate (Na^+, pH 4.8) [9]. The solubilised monomeric RG-II was purified by gel-permeation chromatography on Bio-Gel P-30 (Bio-Rad, Irvine, California, U.S.A.) in pyridine/acetic acid/water (1 : 1 : 98, containing 0.5% chlorobutanol), then dried *in vacuo*.

Gel electrophoresis

PAGE was performed according to Chormova et al. [9]. To prepare one 26.4% polyacrylamide gel (83 × 73 × 0.75 mm), we mixed 834 μl water, 834 μl 1.5 M Tris (Cl^-, pH 8.8), 3.33 ml 40% (w/v) acrylamide/bisacrylamide (29 : 1), 3.9 μl TEMED and 46.7 μl of freshly prepared 10% ammonium persulfate. The mixture was quickly poured, and a 10-tooth comb was inserted. The electrode buffer contained 50 mM Tris (free base) and 38 mM glycine, pH 9.2 (unadjusted). Samples (8–10 μl) were mixed with 2 μl loading buffer [0.63 M Tris–HCl containing 0.25% (w/v) bromophenol blue and 50% (v/v) glycerol, pH 8.8]. Electrophoresis was conducted at 200 V for 75 min. The gel was then fixed in ethanol/acetic acid/water (4 : 1 : 5) for 30 min, washed with water for 1 min three times, then treated successively with 400 μM sodium thiosulfate for exactly 1 min, water (3 × 1 min), freshly prepared 6 mM silver nitrate in 10 μM formaldehyde for 20 min, water (2 × 20 s) and 0.28 M Na_2CO_3 containing 8 μM sodium thiosulfate and 64 mM formaldehyde for 2–10 min. Colour development was stopped 30 s before the desired intensity had been reached by addition of 0.33 M Tris base in 2% v/v acetic acid for 10 min, and the gel was then immediately scanned. Colour development continued for ~30–60 s in the stopping solution, the background turning dark yellow.

Where possible, we semi-quantified the silver-stained bands so that the % dimerisation could be reported. The method is described in detail in the legend to Supplementary Figure S2. In brief, the stained gel was scanned and the image converted to greyscale in Adobe Photoshop 5.5, then a standard rectangle was placed on each RG-II band, within which the mean luminosity was read. To correct for the background, we also measured the mean luminosity in a neighbouring blank area of the gel. Each RG-II band's mean luminosity was subtracted from that of the nearby blank zone, giving 'Δ luminosity' value, which was proportional to the mass of RG-II in the band (e.g. Supplementary Figure S3). A given mass of dimer gives stronger silver-staining than the same mass of monomer on the same gel, so separate standard curves were prepared for monomer and dimer. A 'Δ luminosity' value of 10 corresponded to 0.156 ± 0.014 μg of RG-II dimer or 0.314 ± 0.042 μg of monomer (mean ± SE of 14 gels). For each gel, a factor was obtained for converting Δ luminosity to mass for monomer bands by reference to markers and/or any reaction mixtures in which essentially all the RG-II was still monomeric; and for dimers it was obtained by reference to markers and/or any reaction mixture in which almost all the RG-II had been dimerised. When no such dimer reference band was available on a particular gel, a comparably stained gel was used. It should be noted that the background staining of these gels is slightly uneven, so some faint bands ended up with an apparent negative mass, resulting in the '% dimerisation' being scored as slightly above 100% or below 0%, [e.g. in Figure 3, Figure 8e and Supplementary Figure S5b].

In-vitro dimerisation of RG-II

In studies of dimerisation *in vitro*, the RG-II monomer (16 or 20 μM) was incubated at 20°C for 16 h in an excess (1.2 mM) of H_3BO_3, with or without 0.5 mM $PbNO_3$, and with or without various peptides (Table 1) or native AGP31 glycoprotein (concentrations indicated on Figs.). All reaction mixtures were buffered at pH 4.8 with 50 mM acetate (Na^+); the optimal pH for dimerisation in the presence of peptides is 2.5–4.8 (Begum and Fry, in preparation). Samples were analysed by PAGE without further preparation; neither the presence of chaperones nor the buffer affected electrophoresis mobility.

Prediction of peptide and protein isoelectric points

The pI values of amino acid sequences were estimated (Supplementary Fig. S1, Table 1) by online services. Unless otherwise stated, we used IPC 2.0 (www.ipc2-isoelectric-point.org), which makes use of a combination of deep learning and support vector regression models [47]. Other websites ['Compute pI/Mw tool'

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

(https://web.expasy.org/compute_pi/), 'Bachem Peptide Calculator' (https://www.bachem.com/knowledge-center/peptide-calculator/), Prot pi (with 'ProMoST' data source for pK_a values; https://www.protpi.ch/Calculator/ProteinTool), and Prot pi (with 'native' data source for pK_a values)] gave rather different predictions (Supplementary Table S1).

Data Availability

All data relevant to this manuscript are provided within the paper as tables, figures and supplementary file. We have not included any structural/crystallographic data (for either macromolecular structures or small molecules), protein and nucleic acid sequence data other than what are already published on universally accessible websites, functional genomics and molecular interactions/proteomics/metabolomics data, computational models, or genetics data (genetic polymorphisms; genotype data).

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

S.C.F. thanks the BBSRC (UK; grant reference BB/H000690/1); D.S. thanks the Comisión Nacional de Investigación Científica y Tecnológica (Conicyt, Becas Chile Postdoctorado 2014, Chile) for financial support. R.A.B. thanks the Commonwealth Scholarship Commission for the award of a PhD Scholarship (reference BDCS-2016-64). E.J. and C.A. are thankful to CNRS and Paul Sabatier-Toulouse 3 University for supporting their work.

Open Access

Open access for this article was enabled by the participation of University of Edinburgh in an all-inclusive *Read & Publish* agreement with Portland Press and the Biochemical Society under a transformative agreement with JISC.

CRediT Author Contribution

Stephen C. Fry: Conceptualization, Resources, Supervision, Writing — original draft, Project administration. **Dayan Sanhueza:** Funding acquisition, Investigation, Methodology, Writing — review and editing. **Rifat Ara Begum:** Investigation, Methodology, Writing — review and editing. **Cécile Albenne:** Resources, Methodology, Writing — review and editing. **Elisabeth Jamet:** Resources, Supervision, Funding acquisition, Methodology, Writing — review and editing.

Abbreviations

AGP, arabinogalactan-protein; AGP17p etc., peptides of AGPs; AIR, alcohol-insoluble residue; DP, degree of polymerisation; GPI, glycosylphosphatidylinositol; PAGE, polyacrylamide gel electrophoresis; RG-II, rhamnogalacturonan-II.

References

1 Pabst, M., Fischl, M.R., Brecker, L., Morelle, W., Fauland, A., Köfeler, H. et al. (2013) Rhamnogalacturonan II structure shows variation in the side chains monosaccharide composition and methylation status within and across different plant species. *Plant J.* **76**, 61–72 https://doi.org/10.1111/tpj.12271

2 Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.N., Baslé, A., Gray, J. et al. (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* **544**, 65–70 https://doi.org/10.1038/nature21725

3 O'Neill, M.A., Black, I., Urbanowicz, B., Bharadwaj, V., Crowley, M., Koj, S. et al. (2020) Locating methyl-etherified and methyl-esterified uronic acids in the plant cell wall pectic polysaccharide rhamnogalacturonan II. *SLAS Technol.* **25**, 329–344 https://doi.org/10.1177/2472630320923321

4 Yu, Q., Tang, C. and Kuo, J. (2000) A critical review on methods to measure apoplastic pH in plants. *Plant and Soil* **219**, 29–40 https://doi.org/10.1023/A:1004724610550

5 Martinière, A., Bassil, E., Jublanc, E., Alcon, C., Reguera, M., Sentenac, H. et al. (2013) *In vivo* intracellular pH measurements in tobacco and arabidopsis reveal an unexpected pH gradient in the endomembrane system. *Plant Cell* **25**, 4028–4043 https://doi.org/10.1105/tpc.113.116897

6 O'Neill, M.A., Ishii, T., Albersheim, P. and Darvill, A.G. (2004) Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annu. Rev. Plant Biol.* **55**, 109–139 https://doi.org/10.1146/annurev.arplant.55.031903.141750

7 Matoh, T., Ishigaki, K., Ohno, K. and Azuma, J. (1993) Isolation and characterization of a boron-polysaccharide complex from radish roots. *Plant Cell Physiol.* **34**, 639–642 https://doi.org/10.1093/oxfordjournals.pcp.a078465

8 Matoh, T., Kawaguchi, S. and Kobayashi, M. (1996) Ubiquity of a borate-rhamnogalacturonan II complex in the cell walls of higher plants. *Plant Cell Physiol.* **37**, 636–640 https://doi.org/10.1093/oxfordjournals.pcp.a028992

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

9 Chormova, D., Messenger, D. and Fry, S.C. (2014) Boron bridging of rhamnogalacturonan-II, monitored by gel electrophoresis, occurs during polysaccharide synthesis and secretion but not post-secretion. *Plant J.* **77**, 534–546 https://doi.org/10.1111/tpj.12403
10 Liu, G., Dong, X., Liu, ., Wu, L., Peng, S. and Jiang, C. (2014) Boron deficiency is correlated with changes in cell wall structure that lead to growth defects in the leaves of navel orange plants. *Sci. Hortic.* **175**, 54–62 https://doi.org/10.1016/j.scienta.2014.06.036
11 O'Neill, M., Eberhard, S., Albersheim, P. and Darvill, A.G. (2001) Requirement of borate cross-linking of cell wall rhamnogalacturonan ii for arabidopsis growth. *Science* **294**, 846–849 https://doi.org/10.1126/science.1062319
12 Panter, P.E., Kent, O., Dale, M., Smith, S.J., Skipsey, M., Thorlby, G. et al. (2019) MUR1-mediated cell-wall fucosylation is required for freezing tolerance in *Arabidopsis thaliana*. *New Phytol.* **224**, 1518–1531 https://doi.org/10.1111/nph.16209
13 Sechet, J., Htwe, S., Urbanowicz, B., Agyeman, A., Feng, W., Ishikawa, T. et al. (2018) Suppression of Arabidopsis *GGLT1* affects growth by reducing the l-galactose content and borate cross-linking of rhamnogalacturonan-II. *Plant J.* **96**, 1036–1050 https://doi.org/10.1111/tpj.14088
14 Noguchi, K., Ishii, T., Matsunaga, T., Kakegawa, K., Hayashi, H. and Fijiwara, T. (2003) Biochemical properties of the cell wall in the *Arabidopsis* mutant *bor1-1* in relation to boron nutrition. *J. Plant. Nutr. Soil Sci.* **166**, 175–178 https://doi.org/10.1002/jpln.200390025
15 Iwai, H., Masaoka, N., Ishii, T. and Satoh, S. (2002) A pectin glucuronosyltransferase gene is essential for intercellular attachment in the plant meristem. *Proc. Natl Acad. Sci. U.S.A.* **99**, 16319–16324 https://doi.org/10.1073/pnas.252530499
16 O'Neill, M.A., Warrenfeltz, D., Kates, K., Pellerin, P., Doco, T., Darvill, A.G. et al. (1996) Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently cross-linked by a borate ester: *in vitro* conditions for the formation and hydrolysis of the dimer. *J. Biol. Chem.* **271**, 22923–22930 https://doi.org/10.1074/jbc.271.37.22923
17 Chormova, D. and Fry, S.C. (2016) Boron bridging of rhamnogalacturonan-II is promoted *in vitro* by cationic chaperones, including polyhistidine and wall glycoproteins. *New Phytol.* **209**, 241–251 https://doi.org/10.1111/nph.13596
18 Serpe, M.D. and Nothnagel, E.A. (1994) Effects of Yariv phenylglycosides on Rosa cell suspensions: evidence for the involvement of arabinogalactan-proteins in cell proliferation. *Planta* **193**, 542–550 https://doi.org/10.1007/BF02411560
19 Langan, K.J. and Nothnagel, E.A. (1997) Cell surface arabinogalactan-proteins and their relation to cell proliferation and viability. *Protoplasma* **196**, 87–98 https://doi.org/10.1007/BF01281062
20 Willats, W.G.T. and Knox, J.P. (1996) A role for arabinogalactan-proteins in plant cell expansion: evidence from studies on the interaction of (-glucosyl Yariv reagent with seedlings of *Arabidopsis thaliana*. *Plant J.* **9**, 919–925 https://doi.org/10.1046/j.1365-313X.1996.9060919.x
21 Gao, M. and Showalter, A.M. (2000) Immunolocalization of LeAGP-1, a modular arabinogalactan-protein, reveals its developmentally regulated expression in tomato. *Planta* **210**, 865–874 https://doi.org/10.1007/s004250050691
22 Motose, H., Sugiyama, M. and Fukuda, H. (2001) An arabinogalactan protein(s) is a key component of a fraction that mediates local intercellular communication involved in tracheary element differentiation of *Zinnia* mesophyll cells. *Plant Cell Physiol.* **42**, 129–137 https://doi.org/10.1093/pcp/pce014
23 Zhang, Y., Brown, G., Whetten, R., Loopstra, C.A., Neale, D., Kieliszewski, M.J. et al. (2003) An arabinogalactan protein associated with secondary cell wall formation in differentiating xylem of loblolly pine. *Plant Mol. Biol.* **52**, 91–102 https://doi.org/10.1023/A:1023978210001
24 Van Hengel, A.J., Van Kammen, A. and De Vries, S. (2002) A relationship between seed development, arabinogalactan-proteins (AGPs) and the AGP mediated promotion of somatic embryogenesis. *Physiol. Plant.* **114**, 637–644 https://doi.org/10.1034/j.1399-3054.2002.1140418.x
25 Wu, H., Wong, E., Ogdahl, J. and Cheung, A.Y. (2000) A pollen tube growth-promoting arabinogalactan protein from *Nicotiana alata* is similar to the tobacco TTS protein. *Plant J.* **22**, 165–176 https://doi.org/10.1046/j.1365-313x.2000.00731.x
26 Gao, M. and Showalter, A.M. (1999) Yariv reagent treatment induces programmed cell death in Arabidopsis cell cultures and implicates arabinogalactan protein involvement. *Plant J.* **19**, 321–331 https://doi.org/10.1046/j.1365-313X.1999.00544.x
27 Nothnagel, E.A. (1997) Proteoglycans and related components in plant cells. *Int. Rev. Cytol.* **174**, 195–291 https://doi.org/10.1016/S0074-7696(08)62118-X
28 Showalter, A.M. (2001) Arabinogalactan-proteins: structure, expression and function. *Cell. Mol. Life Sci.* **58**, 1399–1417 https://doi.org/10.1007/PL00000784
29 Yang, J., Sardar, H.S., McGovern, K.R., Zhang, Y. and Showalter, A.M. (2007) A lysine-rich arabinogalactan protein in Arabidopsis is essential for plant growth and development, including cell division and expansion. *Plant J.* **49**, 629–640 https://doi.org/10.1111/j.1365-313X.2006.02985.x
30 Showalter, A.M., Keppler, B., Lichtenberg, J., Gu, D. and Welch, L.R. (2010) A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins. *Plant Physiol.* **153**, 485–513 https://doi.org/10.1104/pp.110.156554
31 Egelund, J., Larsen Petersen, B., Saddik Motawia, M., Damager, I., Faik, A., Olsen, C.E. et al. (2006) *Arabidopsis thaliana* RGXT1 and RGXT2 encode Golgi-localized (1,3)-(-d-xylosyltransferases involved in the synthesis of pectic rhamnogalacturonan-II. *Plant Cell* **18**, 2593–2607 https://doi.org/10.1105/tpc.105.036566
32 Mohnen, D. (2008) Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.* **11**, 266–277 https://doi.org/10.1016/j.pbi.2008.03.006
33 Begum, R.A. and Fry, S.C. (2022) Boron bridging of rhamnogalacturonan-II in *Rosa* and arabidopsis cell cultures occurs mainly in the endo-membrane system and continues at a reduced rate after secretion. *Annals of Botany,* mcac119 https://doi.org/10.1093/aob/mcac119
34 Ellis, M., Egelund, J., Schultz, C.J. and Bacic, A. (2010) Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol.* **153**, 403–419 https://doi.org/10.1104/pp.110.156000
35 Zhou, K. (2019) Glycosylphosphatidylinositol-anchored proteins in *Arabidopsis* and one of their common roles in signaling transduction. *Front. Plant Sci.* **10**, 1022 https://doi.org/10.3389/fpls.2019.01022
36 Voxeur, A. and Fry, S.C. (2014) Glycosylinositol phosphorylceramides from *Rosa* cell cultures are boron-bridged in the plasma membrane and form complexes with rhamnogalacturonan II. *Plant J.* **79**, 139–149 https://doi.org/10.1111/tpj.12547
37 Hijazi, M., Roujol, D., Huan, N.K., Castillo, L.D.C., Saland, E., Jamet, E. et al. (2014) Arabinogalactan protein 31 (AGP31), a putative network-forming protein in *Arabidopsis thaliana* cell walls? *Ann. Bot.* **114**, 1087–1097 https://doi.org/10.1093/aob/mcu038
38 Gong, S.Y., Huang, G.Q., Sun, X., Li, P., Zhao, L.L., Zhang, D.J. et al. (2012) GhAGP31, a cotton non-classical arabinogalactan protein, is involved in response to cold stress during early seedling development. *Plant Biol. (Stuttg)* **14**, 447–457 https://doi.org/10.1111/j.1438-8677.2011.00518.x
39 Baldwin, T., Domingo, C., Schindler, T., Seetharaman, G., Stacey, N. and Roberts, K. (2001) DcAGP1, a secreted arabinogalactan protein, is related to a family of basic proline-rich proteins. *Plant Mol. Biol.* **45**, 421–435 https://doi.org/10.1023/A:1010637426934

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

Biochemical Journal (2022) **479** 1967–1984
https://doi.org/10.1042/BCJ20220340

40 Matsunaga, T., Ishii, T., Matsumoto, S., Higuchi, M., Darvill, A., Albersheim, P. et al. (2004) Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in pteridophytes, lycophytes, and bryophytes: implications for the evolution of vascular plants. *Plant Physiol.* **134**, 339–351 https://doi.org/10.1104/pp.103.030072

41 Hijazi, M., Durand, J., Pichereaux, C., Pont, F., Jamet, E. and Albenne, C. (2012) Characterization of the arabinogalactan protein 31 (AGP31) of *Arabidopsis thaliana*. *J. Biol. Chem.* **287**, 9623–9632 https://doi.org/10.1074/jbc.M111.247874

42 Ishii, T., Matsunaga, T., Pellerin, P., O'Neill, M.A., Darvill, A. and Albersheim, P. (1999) The plant cell wall polysaccharide rhamnogalacturonan II self-assembles into a covalently cross-linked dimer. *J. Biol. Chem.* **274**, 13098–13104 https://doi.org/10.1074/jbc.274.19.13098

43 Akiyama, Y., Eda, S. and Katō, K. (1984) Gum arabic is a kind of arabinogalactan–protein. *Agric. Biol. Chem.* **48**, 235–237 https://doi.org/10.1080/00021369.1984.10866126

44 Zeng, W., Ford, K., Bacic, A. and Haezlewood, J. (2018) *N*-Linked glycan micro-heterogeneity in glycoproteins of Arabidopsis. *Mol. Cell. Proteomics* **17**, 413–421 https://doi.org/10.1074/mcp.RA117.000165

45 Yariv, J., Rapport, M.M. and Graf, L. (1962) The interaction of glycosides and saccharides with antibody to the corresponding phenylazo glycosides. *Biochem. J.* **85**, 383–388 https://doi.org/10.1042/bj0850383

46 Yang, J., Zhang, Y., Liang, Y. and Showalter, A.M. (2011) Expression analyses of AtAGP17 and AtAGP19, two lysine-rich arabinogalactan proteins, in Arabidopsis. *Plant Biol.* **13**, 431–438 https://doi.org/10.1111/j.1438-8677.2010.00407.x

47 Sun, W., Xu, J., Yang, J., Kieliszewski, M.J. and Showalter, A.M. (2005) The lysine-rich arabinogalactan-protein subfamily in Arabidopsis: gene expression, glycoprotein purification and biochemical characterization. *Plant Cell Physiol.* **46**, 975–984 https://doi.org/10.1093/pcp/pci106

48 Kozlowski, L.P. (2021) IPC 2.0: prediction of isoelectric point and pK_a dissociation constants. *Nucleic Acids Res.* **49**, W285–W292 https://doi.org/10.1093/nar/gkab295

© 2022 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and distributed under the Creative Commons Attribution License 4.0 (CC BY).

An *Arabidopsis thaliana* arabinogalactan-protein (AGP31) and several cationic AGP fragments catalyse the boron bridging of rhamnogalacturonan-II

Dayan Sanhueza, Rifat Ara Begum, Cécile Albenne, Elisabeth Jamet, and Stephen C. Fry

Supplementary file

Contents

Figure S1. Sequences of selected arabidopsis arabinogalactan-proteins

Figure S2. Method for approximately quantifying silver-stained bands in Photoshop

Figure S3. Quantification of silver-stained RG-II bands after polyacrylamide gel electrophoresis

Figure S4. Effect of boron concentration on the polyhistidine-enhanced dimerisation of RG-II

Figure S5. Comparison of various cationic chaperones for catalysis of RG-II dimerisation

Figure S6. Deglycosylated and intact gum arabic as potential RG-II chaperones

Table S1. Variation in isoelectric point prediction between different online services

AGP17 pI 5.7, 13.55 kDa

MTRNILLTVTLICIVFITVGG

QSPATAPIHSPSTSPHKPKPTSPAISPAAPTPPESTEAPAKTPVEAPVEAPPSPTPASTPQISPPAPSP
EADTPSAPEIAPSADVPAPALTKHKKKTKKHKTAPAPGPASELLSPPAPPGEAPGPGPSDAFSPAADD
QS—GPI

GAQAISVVIQMVGAAAIAWSLLVLAF

AGP18 pI 7.6, 15.69 kDa

MDRNFLLTVTLICIVVAGVGG

QSPISSPTKSPTTPSAPTTSPTKSPAVTSPTTAPAKTPTASASSPVESPKSPAPVSESSPPPTPVPES
SPPVPAPMVSSPVSSPPVPAPVADSPPAPVAAPVADVPAPAPSKHKKTTKKSKKHQAAPAPAPELLGP
PAPPTESPGPNSDAFSPGPSADDQSG—GPI

AASTRVLRNVAVGAVATAWAVLVMAF

AGP19 pI 9.2, 16.58 kDa

MESNSIIWSLLLASALISSFSVNA

QGPAASPVTSTTTAPPPTTAAPPTTAAPPPTTTTPPVSAAQPPASPVTPPPAVTPTSPPAPKVAPVIS
PATPPPQPPQSPPASAPTVSPPPVSPPPAPTSPPPTPASPPPAPASPPPAPASPPPAPVSPPPVQAPS
PISLPPAPAPAPTKHKRKHKHKRHHHAPAPAPIPPS—GPI

PPSPPVLTDPQDTAPAPSPNTNGGNALNQLKGRAVMWLNTGLVILFLLAMTA

AGP31 pI 9.1, 35.91 kDa

MGFIGKSVLVSLVALWCFTSSVFT

EEVNHKTQTPSLAPAPAPYHHGHHHPHPPHHHHPHPHPHPHPPAKSPVKPPVKAPVSPPAKPPVKPPV
YPPTKAPVKPPTKPPVKPPVSPPAKPPVKPPVYPPTKAPVKPPTKPPVKPPVYPPTKAPVKPPTKPPV
KPPVYPPTKAPVKPPTKPPVKPPVSPPAKPPVKPPVYPPTKAPVKPPVSPPTKPPVTPPVYPPKFNRS
LVAVRGTVYCKSCKYAAFNTLLGAKPIEGATVKLVCKSKKNITAETTTDKNGYFLLLAPKTVTNFGFR
GCRVYLVKSKDYKCSKVSKLFGGDVGAELKPEKKLGKSTVVVNKLVYGLFNVGPFAFNPSCPK

Figure S1. Sequences of selected arabidopsis arabinogalactan-proteins.
Mature proteins after cleavage of proposed signal peptide and insertion of proposed GPI (glycosyl-phosphatidylinositol) anchor. Isoelectric point predictions (by IPC 2.0) and molecular mass are after removal of the N-terminal signal peptide, and without glycosylation. **Underlined, bold** stretches are the peptides tested in the present work. *Italic* stretches are proposed to be removed in the mature proteins. N = putative Asn *N*-glycosylation site. Colour code:

P	single proline	S,T	serine & threonine	H	histidine
PP	double proline	D	aspartic acid	K	lysine
PPP	triple proline	E	glutamic acid	R	arginine

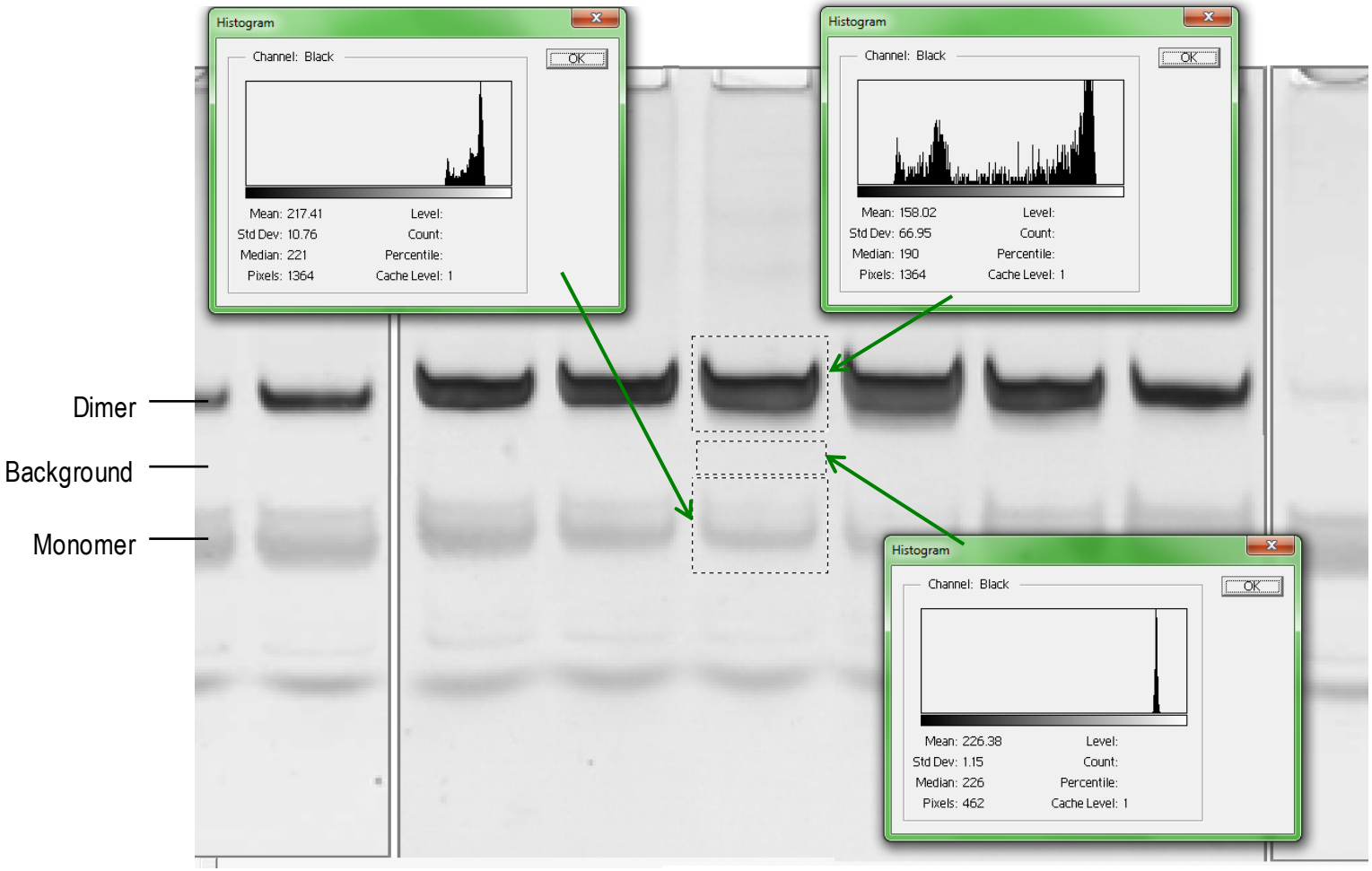

Figure S2. Method for approximately quantifying silver-stained bands in Photoshop.

As an example, this diagram shows part of Fig. 7a, indicating how the bands in the 50 µg/ml AGP19p1 lane were semi-quantified. In Adobe Photoshop 5.5, the scan was converted to greyscale, then a standard rectangle was placed on each RG-II band, and the histogram facility was applied, giving a mean luminosity (which is expressed by the software per unit area; in this case, 217.41 and 158.02 for the monomer and dimer, respectively). The rectangle always had the same area for every RG-II band on a given gel, and was always large enough to accommodate the whole band. Each reading thus gave the mean luminosity of the RG-II band plus an area of background gel. To correct for the background, we also placed a rectangle on the mid-point between each pair of RG-II bands and recorded this mean luminosity reading (226.38 in this case). In some cases, the background rectangle was smaller than the rectangle used for the RG-II bands, as illustrated here, so that no RG-II tail was included. The mean luminosity is not affected by this difference in area. The mean luminosity value returned for pure white and black zones of an image are 255.0 and 0.0 respectively. Each RG-II band's mean luminosity was subtracted from its nearest background, and the result (Δ luminosity) was used as a measure of the mass of RG-II in the band. For each gel, a separate conversion factor was determined. As shown in Fig. S3, mass was approximately proportional to the Δ luminosity value. A given mass of dimer gives a much higher value than the same mass of monomer on the same gel, so separate conversion factors were used for monomers and dimers on each gel.

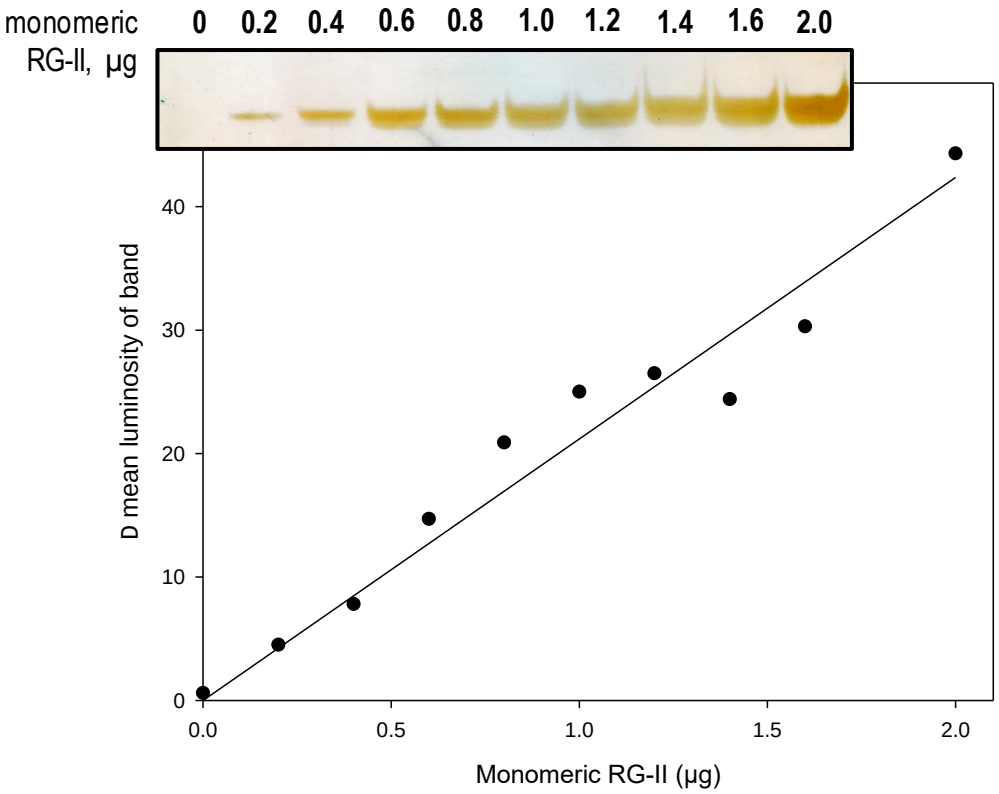

Figure S3. Quantification of silver-stained RG-II bands after polyacrylamide gel electrophoresis.

Various loadings of monomeric rhamnogalacturonan-II, purified from arabidopsis cell-suspension cultures [9], were analysed by PAGE and silver-stained. In Photoshop, the image was converted to greyscale and the mean luminosity within a standard rectangle was measured and corrected for the nearest-neighbour background area of the gel (see Fig. S2).

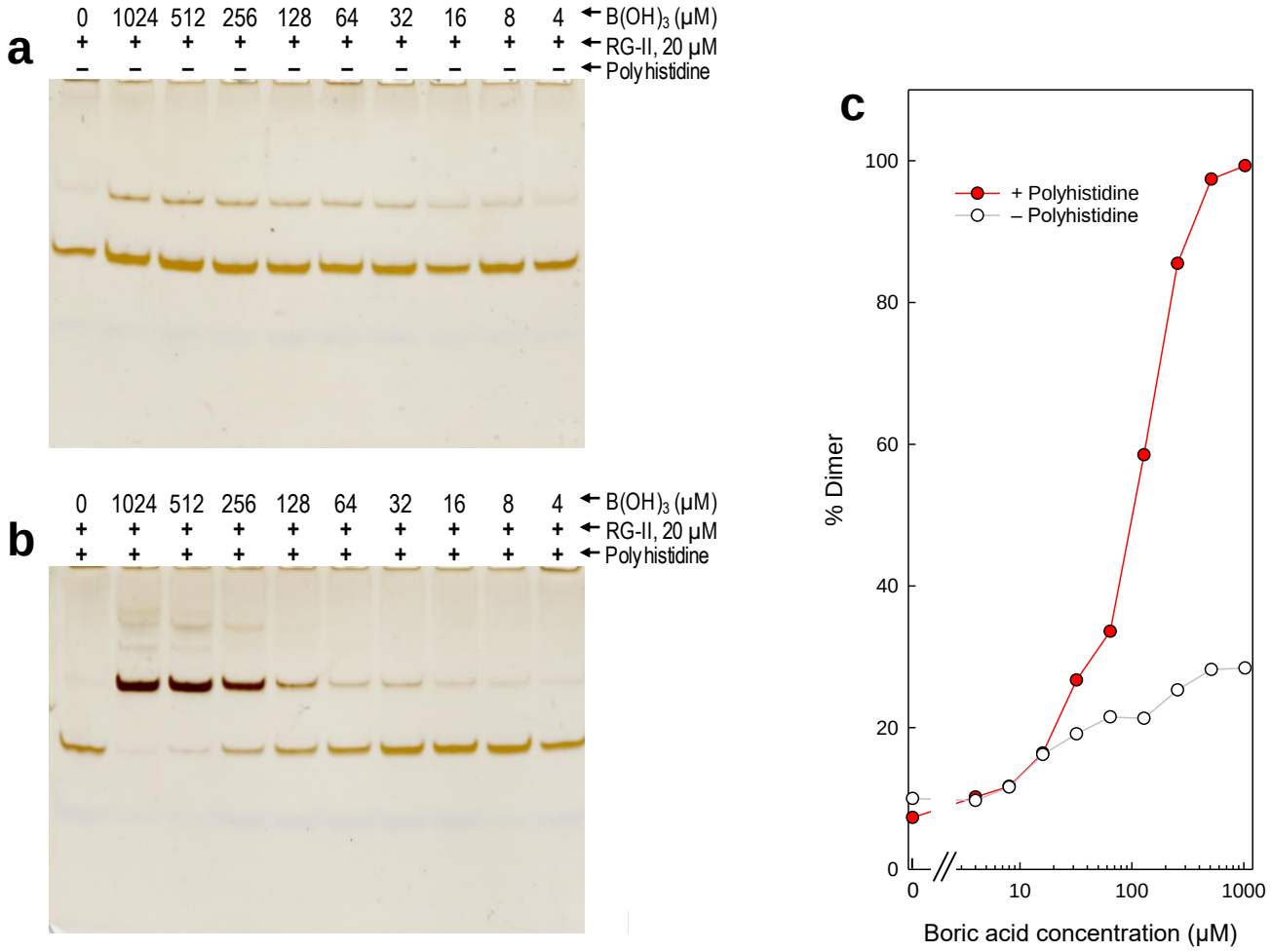

Figure S4. Effect of boron concentration on the polyhistidine-enhanced dimerisation of RG-II.

Reaction mixtures contained 50 µg/ml monomeric RG-II (20 µM) with 0–1024 µM boric acid and 50 mM acetate (Na^+, pH 4.8), without (a) or with (b) polyhistidine.Cl^- (50 µg/ml; 2.8 µM). After 4 h at 20°C, 0.8 µg of the RG-II was analysed by PAGE followed by silver staining. (c) Quantification of the extent of dimerisation.

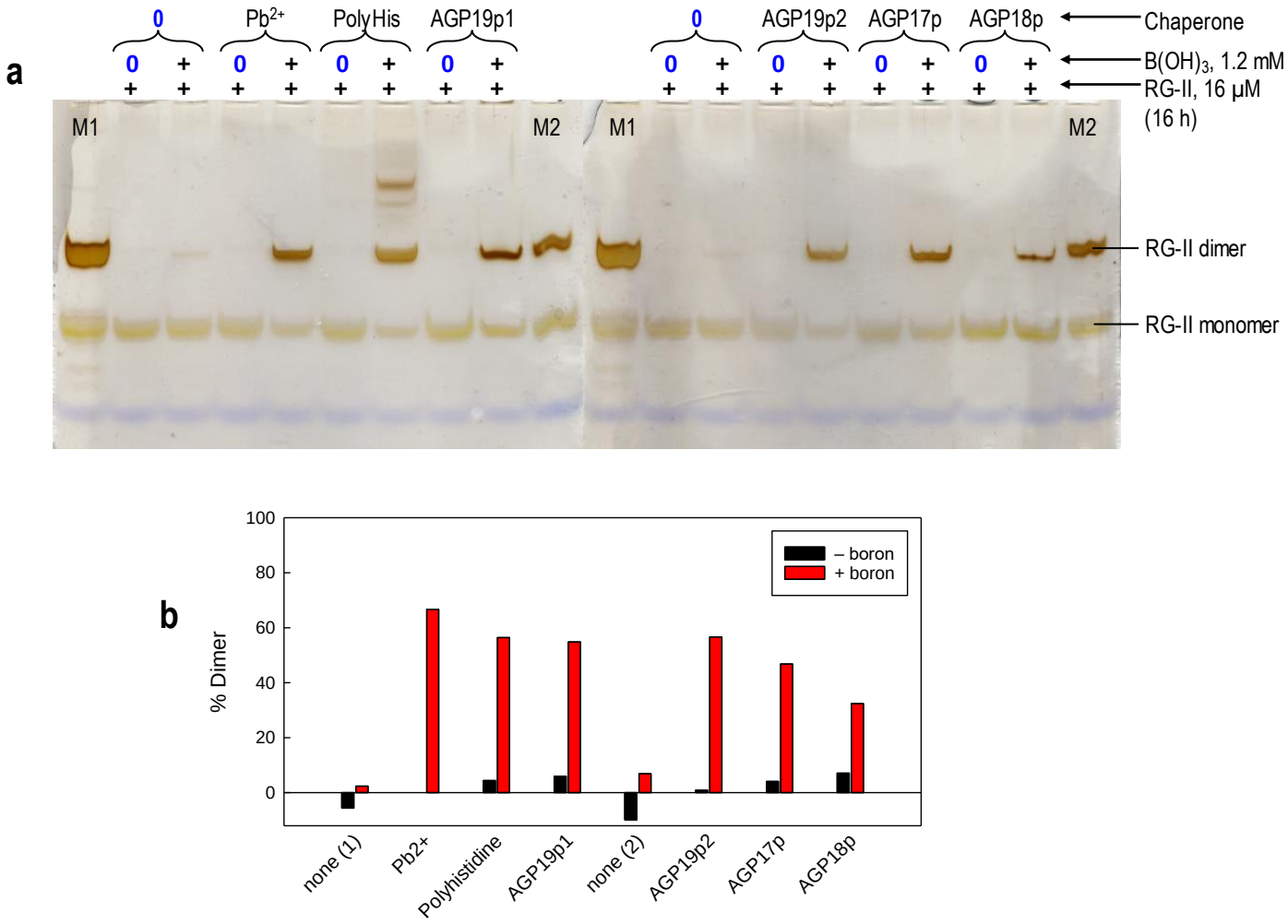

Figure S5. Comparison of various cationic chaperones for catalysis of RG-II dimerisation.

Reaction mixtures contained 80 µg/ml monomeric RG-II with 0 or 1.2 mM boric acid and 50 mM acetate (Na^+, pH 4.8) plus a chaperone: 0.5 mM Pb^{2+} or 50 µg/ml* polyhistidine, AGP17p, AGP18p, AGP19p1 or AGP19p2. Controls lacked any chaperone. After 16 h at 20 °C, 0.8 µg of the RG-II was analysed by PAGE followed by silver staining. (a) Stained gel, (b) quantification of the % dimerisation. Markers were: M1, 1.2 µg monomeric RG-II + 1.2 µg dimeric RG-II; M2, 0.8 µg monomeric RG-II + 0.8 µg dimeric RG-II.

*The 50 µg/ml peptide solutions represent 2.7 µM poly-His, 39 µM AGP17p, 34 µM AGP18p, 28 µM AGP19p1 and 20 µM AGP19p2.

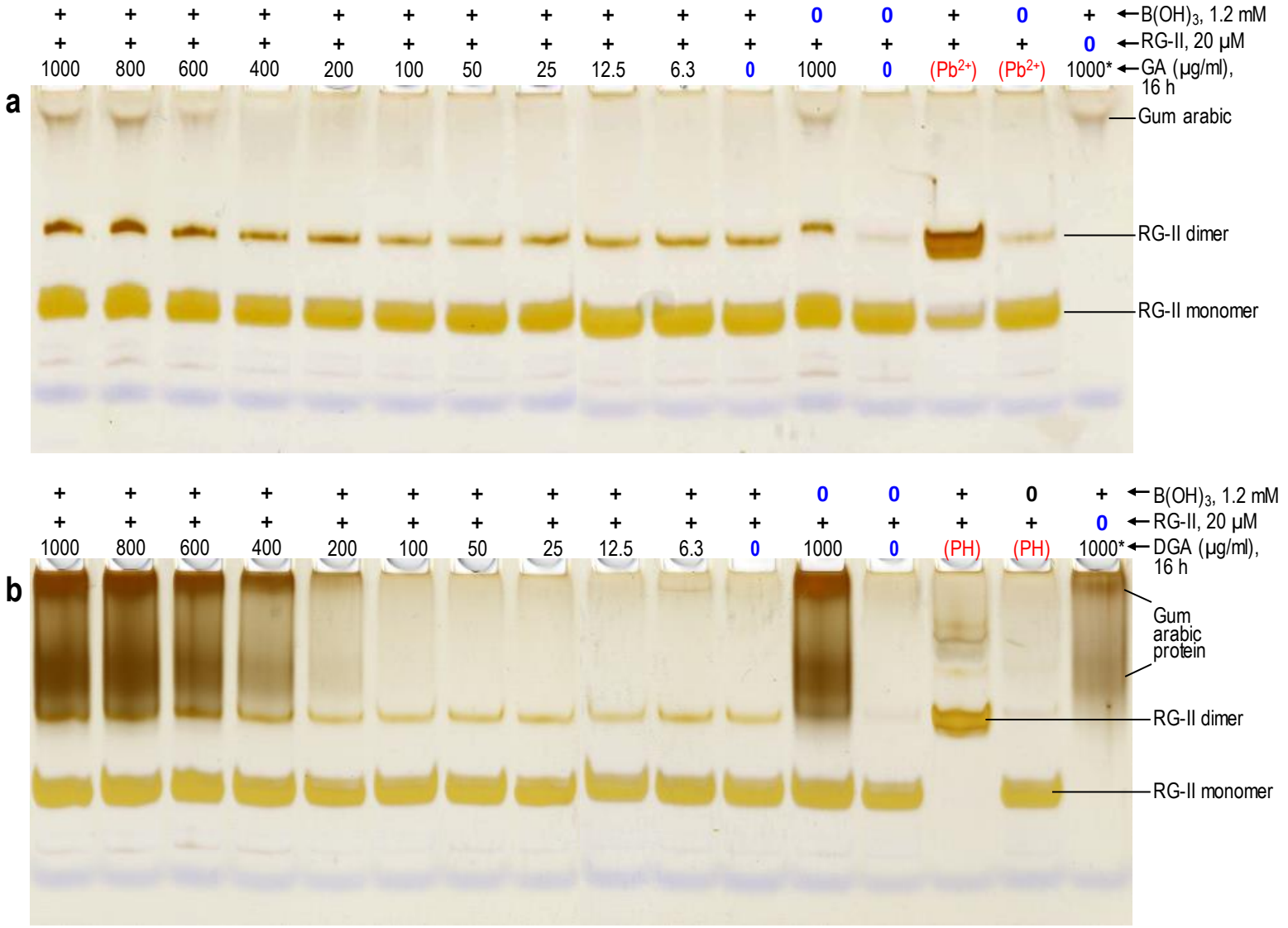

Figure S6. Deglycosylated and intact gum arabic as potential RG-II chaperones.

Reaction mixtures contained 100 µg/ml RG-II monomer (≈ 20 µM), 1.2 mM boric acid, 0–1000 µg/ml of **(a)** gum arabic (GA) or **(b)** the protein core of deglycosylated gum arabic (GAPr), and 50 mM acetate (Na^+) buffer, pH 4.8. Controls lacked boric acid or RG-II, or contained 0.5 mM Pb^{2+} or 50 µg/ml polyhistidine (PH) as alternative chaperones. After 16 h at 20 °C, 0.8 µg of the RG-II was analysed by PAGE followed by silver staining. The right-hand lanes (marked 1000*) contained 1000 µg/ml GA or GAPr plus 50 µg/ml polyhistidine. The bands in (a) are quantified in Fig. 3a.

Table S1. Variation in isoelectric point prediction between different online services*

Name	Sequence[1]	pI (Expasy)[2]	pI (Bachem)[3]	pI (ProtPi) (ProMoST)[4]	pI (ProtPi) (Native)[5]	pI (IPC 2.0)[6]
Histidine	H	6.74	7.90	6.59	6.59	7.24
His_6	HHHHHH	7.21	8.19	7.26	7.48	7.34
Poly-histidine	HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HH	8.10	8.94	8.16	8.64	7.77
AGP17p	KHKKKTKKHK	10.78	11.52	10.71	11.20	8.18
AGP18p	KHKKTTKKSKKH	10.78	11.52	10.61	11.23	8.21
AGP19p1	KHKRKHKHKRHHH	12.04	12.56	12.51	12.53	8.84
AGP19p2	ASASKHKRKHKHKRHHHASAS	12.04	12.56	12.51	12.54	8.84
AGP31p	YHHGHHHPHPPHHHHPHPHPHPHP	7.44	8.32	7.32	7.75	7.96
AGP31	EEVNHKTQTPSLAPAPAPYHHGHHHPHPPHHHHPHPHPHPHPPAKSPVKPPV KAPVSPPAKPPVKPPVYPPTKAPVKPPTKPPVKPPVSPPAKPPVKPPVYPPT KAPVKPPTKPPVKPPVYPPTKAPVKPPTKPPVKPPVYPPTKAPVKPPTKPPV KPPVSPPAKPPVKPPVYPPTKAPVKPPVSPPTKPPVTPPVYPPKFNRSLVAV RGTVYCKSCKYAAFNTLLGAKPIEGATVKLVCKSKKNITAETTTDKNGYFLL LAPKTVTNFGFRGCRVYLVKSKDYKCSKVSKLFGGDVGAELKPEKKLGKSTV VVNKLVYGLFNVGPFAFNPSCPK	10.19	10.66	10.00	10.64	9.14

*Calculated after *in-silico* removal of the putative *N*-terminal signal sequence but without proline hydroxylation and glycosylation.

[1] Amino acid residues with cationic and anionic side-chains are highlighted in shades of blue and red respectively.

[2] Isoelectric point estimated by 'Compute pI/Mw tool' (https://web.expasy.org/compute_pi/).

[3] Isoelectric point estimated by 'Bachem Peptide Calculator' (https://www.bachem.com/knowledge-center/peptide-calculator/).

[4] Isoelectric point estimated by Prot pi (with 'ProMoST' data source for pK_a values) (https://www.bachem.com/knowledge-center/peptide-calculator/).

[5] Isoelectric point estimated by Prot pi (with 'native' data source for pK_a values) (https://www.bachem.com/knowledge-center/peptide-calculator/).

[6] Isoelectric point estimated by IPC 2.0 (peptide if DP<60, protein if DP>60) (www.ipc2-isoelectric-point.org).