

## Synthesis and Solution-Phase Conformation of the RG-I Fragment of the Plant Polysaccharide Pectin Reveals a Modification-Modulated Assembly Mechanism

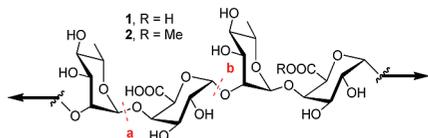
Eoin M. Scanlan,<sup>†</sup> Mukram M. Mackeen,<sup>†</sup> Mark R. Wormald,<sup>‡</sup> and Benjamin G. Davis<sup>\*†</sup>

Chemistry Research Laboratory, Department of Chemistry, University of Oxford, Mansfield Road, Oxford OX1 3TA, U.K., and Glycobiology Institute, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, U.K.

Received October 26, 2009; E-mail: Ben.Davis@chem.ox.ac.uk

Pectins are complex polysaccharides in which four primary domains have been described: homogalacturonan (HGA), rhamnogalacturonan I (RG-I), rhamnogalacturonan II (RG-II), and xylogalacturonan (XGA).<sup>1,2</sup> Modifications of the carbohydrate backbones of RG-I and HGA allow the plant cell to react to changing conditions such as aging and host–pathogen interactions.<sup>3</sup> Two important modifications involve methyl esterification of the galacturonic acid residues and the formation of calcium bridges between pectin chains.<sup>2</sup> The HGA region of pectin is highly methyl-esterified, especially as it emerges from production in the Golgi apparatus. These regions are then de-esterified in the plant cell wall by the action of plant pectinmethylsterases (PMEs). Although heavily studied for the HGA region of pectin,<sup>4</sup> methyl esterification of galacturonic acid (GalA) has also been reported in RG-I.<sup>5</sup> These modifications alter the 3D structure of the pectic matrix and may play a fundamental role in cell growth and defense mechanisms.<sup>3</sup>

Here we report the synthesis and detailed conformational study of a synthetically pure fragment of RG-I (**1**) and its methyl ester RG-I-OMe (**2**) and the effect of esterification upon calcium binding. Experimental 3D structures determined by NMR spectroscopy reveal that a single methyl ester critically and dramatically modulates calcium binding and the subsequent biophysical properties of this key plant polysaccharide.



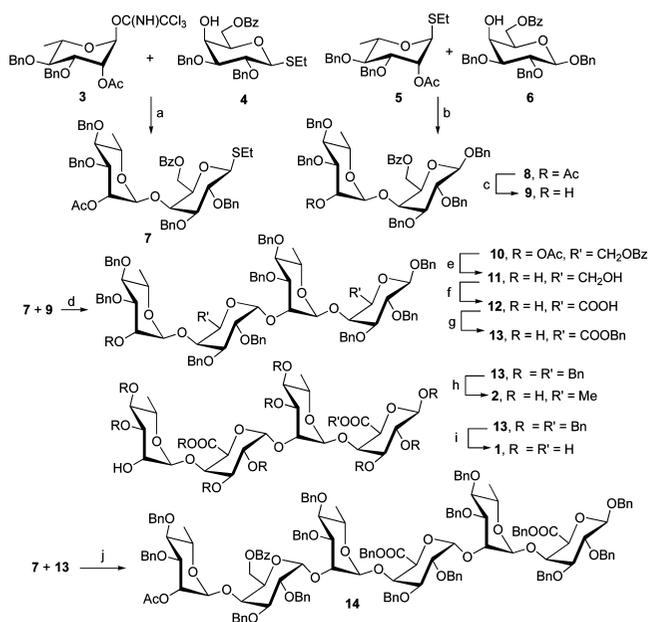
**Figure 1.** RG-I tetrasaccharide unit **1** and corresponding methyl ester **2**.

Analysis of the RG-I structure (Figure 1) initially suggested stereoselective oligomer assembly<sup>6–10</sup> through a late-stage coupling of an L-rhamnosyl (L-Rha)-terminated glycosyl donor fragment to a D-GalA-terminated acceptor fragment (assembly **a** in Figure 1); this logic was suggested by a typically strong preference in L-rhamnosylation for the required  $\alpha$ -L-Rha linkage in RG-I, which contrasts with the typically poor selectivity in the formation of the *cis* linkage<sup>11</sup> of  $\alpha$ -D-GalA. However, despite the synthesis of a variety of appropriate building blocks [see the Supporting Information (SI)], their use in convergent disaccharide-to-disaccharide (2 + 2) coupling surprisingly failed. Consequently, counter to traditional glycosylation logic,<sup>11</sup> we assembled RG-I via late-stage  $\alpha$ -*cis*-galactosylation (assembly **b** in Figure 1) using a latent/quasiorthogonal donor–acceptor assembly strategy.<sup>12–15</sup>

Monosaccharide building blocks **3–6** were readily prepared from the parent sugars L-rhamnose and D-galactose (see the SI). Coupling of the rhamnosyl trichloroacetimidate donor **3** and the galactosylthioglycoside acceptor **4** (itself a latent donor) furnished the  $\alpha$ -linked

disaccharide **7**, which was employed directly as a disaccharide donor unit for assembly of tetrasaccharide **10** and hexasaccharide **14** (Scheme 1). Disaccharide acceptor **9** was prepared by coupling of rhamnosyl thioglycoside **5** and galactosyl acceptor **6** followed by selective deprotection of the C-2' acetyl of **8**. The key (2 + 2) coupling between **7** and **9** proceeded smoothly, giving tetrasaccharide **10** in 83% yield. Despite the relative *cis*-1,2 configuration, **10** was formed as the desired  $\alpha$ -anomer with apparently complete stereoselectivity. Primary (OH-6, OH-6'') and secondary (OH-2'') hydroxyl groups were revealed in one pot upon treatment with freshly prepared sodium methoxide. Selective oxidation of OH-6 and OH-6'' to carboxylate using sequential TEMPO/NaClO<sub>2</sub> and NaOCl successfully converted Gal residues to GalA. Finally, following benzyl esterification to aid purification, careful control of the deprotection conditions allowed selective access to both **1** and **2** from **13**. Thus, treatment under catalytic hydrogenation conditions furnished methyl ester **2** as the major product when MeOH was used as the solvent and the free acid **1** in THF/H<sub>2</sub>O solvent. The potential for iterative extension of the pectin chain was demonstrated through (4 + 2) glycosylation to give **14**.

### Scheme 1<sup>a</sup>



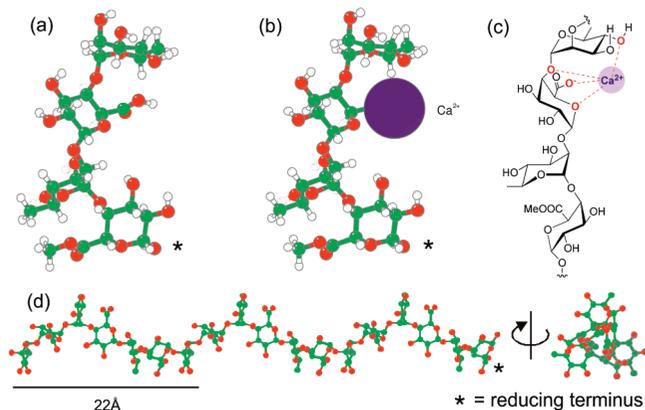
<sup>a</sup> Reagents and conditions: (a) TMSOTf, DCM, 0 °C, 65%; (b) NIS, TMSOTf, DCM, 0 °C, 75%; (c) NaOMe, MeOH/THF, rt, 84%; (d) NIS, TMSOTf, DCM, 0 °C, 83%; (e) NaOMe, MeOH/THF, rt, 78%; (f) TEMPO, NaClO<sub>2</sub>, NaOCl (two stages), 0 °C to rt; (g) Cs<sub>2</sub>CO<sub>3</sub>, BnBr, DMF, rt, 72% over two steps; (h) Pd/C, H<sub>2</sub>, MeOH, rt, 44%; (i) Pd/C, H<sub>2</sub>, THF/H<sub>2</sub>O, rt, 62%; (j) NIS, TMS-OTf, DCM, 0 °C, 68%.

The conformations, carboxylate pK<sub>a</sub> values, and calcium binding properties of these synthetic samples of **1** and **2** were investigated by NMR

<sup>†</sup> Department of Chemistry.  
<sup>‡</sup> Department of Biochemistry.

spectroscopy. The  $pK_a$  values of the single acid residue in **2** and the two carboxylic acid residues in **1** were 3.6, indicating no interactions between the two titratable groups in **1**. The value of 3.6 is similar to those previously reported for isolated pectin and polygalacturonic acid.<sup>16</sup>

Conformations of **1** and **2** in both water and 800 mM  $\text{Ca}^{2+}$ (aq) were determined using cross-linkage rotating-frame nuclear Overhauser effects (ROEs) and molecular modeling. These revealed very similar arrangements for **1** and **2** (Figure 2) in which the two carboxy groups are placed on opposite sides, consistent with the  $pK_a$  results.



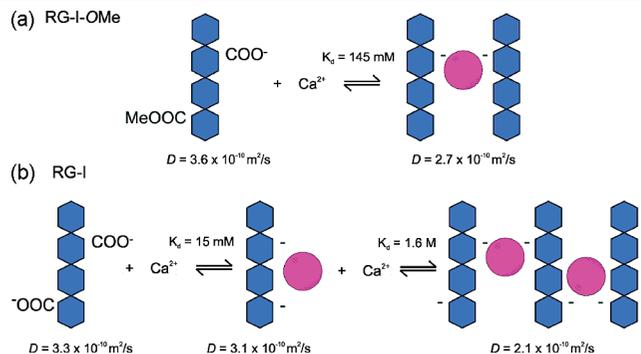
**Figure 2.** Conformation of **2** in the absence (a) and presence (b) of  $\text{Ca}^{2+}$ . (c)  $\text{Ca}^{2+}$  binding site in **2**. (d) Model of the helical structure of RG-I pectin.

A pioneering molecular modeling study carried out on the RG-I disaccharide linkages of Rha $\alpha$ 1–4Gal $\alpha$  and Gal $\alpha$ 1–2Rha suggested three potential conformers for each linkage.<sup>17</sup> Our conformational experimental results are consistent with only one of the three conformers for each linkage. Interestingly, this is the high-energy conformer [ $(\phi, \psi) = (-40^\circ, -10^\circ)$ ] in the relaxed potential maps for the Gal $\alpha$ 1–2Rha linkage and not one of the predicted low-energy conformers. The tetrasaccharide model was then used to build an 18-mer model that showed threefold symmetry ( $3_1$ ) and a pitch of 22 Å containing two repeating units per helical turn (Figure 2d). This right-handed glycan helix therefore provides an effective and repeating lateral display of putative liganding carboxylates.

Calcium binding to **2** showed a single binding curve with a  $K_D$  value of 145 mM at pH 5.6. The presence of 800 mM calcium had no effect on the observed pattern of ROEs. The pattern of chemical shift changes resulting from calcium addition was consistent with a single binding site on **2** (Figure 2b). Calcium binding was also associated with a decrease in the diffusion coefficient,  $D$  (Figure 3), as measured by DOSY spectroscopy, consistent with dimer formation to produce a neutral complex.

Calcium binding to **1** is a more complex multievent process. There is an initial relatively<sup>18–20</sup> high affinity binding with a  $K_D$  of 15  $\text{mM}^{21–23}$  that is not accompanied by a decrease in diffusion coefficient; this is followed by further, more complex changes in chemical shift that can be approximated by a 1.6 M  $K_D$  binding curve and are associated with a significant change in diffusion. For the higher-affinity binding, the 10-fold increase in binding relative to **2** is consistent only with interaction of the bound calcium with both carboxylate groups of **1**, which required a conformational change to bring these two groups to the same side of the molecule, resulting in a neutral monomer complex (“monocomplex”). The lower affinity binding is then associated with aggregation. The unsaturated binding curve resembles that for the previously reported multichain association of polygalacturonic acid in the presence of  $\text{Ca}^{2+}$ .<sup>24</sup> This Ca-binding model for **1** and **2** is summarized in Figure 3.

Rees<sup>25</sup> and colleagues have proposed the “egg-box model” for the occurrence of dimers and multichain association during  $\text{Ca}^{2+}$  binding



**Figure 3.** Working model for switching by methyl (de)esterification **1**  $\rightleftharpoons$  **2**/(a)  $\rightleftharpoons$  (b) in heteropolymeric sugar unit self-assembly. Two very different  $\text{Ca}^{2+}$ -binding modes of (b) RG-I **1** and (a) RG-I-OMe **2** are interconverted.

based on homopolymeric structures (e.g., homogalacturonic acid) and a requirement of  $\sim 14$  residues.<sup>26</sup> Our results show that through heteropolymeric units it is possible to form discrete oligomers for a much shorter pectin chain based on the RG-I-OMe tetrasaccharide. Moreover, the nonesterified RG-I tetrasaccharide displays multimetric interactions between chains initiated by monocomplexation with  $\text{Ca}^{2+}$ . A switch between these two dramatically different states is modulated simply by methyl (de)esterification, a known<sup>2,5</sup> dynamic postbiosynthetic regulatory modification in plants.

In summary, syntheses of the RG-I tetrasaccharide and its methyl ester RG-I-OMe have revealed the first experimental conformational structures. These show that a single methyl ester generates significant differences in biophysical behavior, without altering the conformation, through the modulation of self-assembly. The combined effect of Ca(II) and methylation revealed here suggests a concerted molecular basis for two major dynamic modifications (Ca binding and pectin methyl esterification)<sup>2</sup> in plants.

**Supporting Information Available:** Experimental details, extended synthetic discussion, characterization data, and spectra. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- Ridley, B. L.; O'Neill, M. A.; Mohnen, D. *Phytochemistry* **2001**, *57*, 929.
- Caffall, K. H.; Mohnen, D. *Carbohydr. Res.* **2009**, *344*, 1879.
- Protsenko, M. A.; Buza, N. L.; Krinitzyna, A. A.; Bulantseva, E. A.; Korableva, N. P. *Biochemistry (Moscow)* **2008**, *73*, 1053.
- Goldberg, R.; Prat, R.; Morvan, C. *Carbohydr. Polym.* **1994**, *23*, 203.
- Rihouey, C.; Morvan, C.; Borissova, I.; Jauneau, A.; Demarty, M.; Jarvis, M. *Carbohydr. Polym.* **1995**, *28*, 159.
- In regard to synthetic studies on pectin regions (RG-I/II, HGA), see the SI for an extended discussion and refs 7–10 for leading examples of protected RG-I oligosaccharide precursors/analogues.
- Rich, J. R.; McGavin, R. S.; Gardner, R.; Reimer, K. B. *Tetrahedron: Asymmetry* **1999**, *10*, 17.
- Mariyama, M.; Takeda, T.; Shimizu, N.; Hada, N.; Yamada, H. *Carbohydr. Res.* **2000**, *325*, 83.
- Reiffarth, D.; Reimer, K. B. *Carbohydr. Res.* **2008**, *343*, 179.
- Nemati, N.; Karapetyan, G.; Nolting, B.; Endress, H.-U.; Vogel, C. *Carbohydr. Res.* **2008**, *343*, 1730.
- Demchenko, A. V. *Synlett* **2003**, 1225.
- Boons, G. J. *Tetrahedron* **1996**, *52*, 1095.
- Boons, G. J.; Grice, P.; Leslie, R.; Ley, S. V. *Tetrahedron Lett.* **1993**, *34*, 8523.
- Boons, G. J.; Isles, S. *Tetrahedron Lett.* **1994**, *35*, 3593.
- Cao, S. D.; Gan, Z. H.; Roy, R. *Carbohydr. Res.* **1999**, *318*, 75.
- Ralet, M. C.; Dronnet, V.; Buchholt, H. C.; Thibault, J. F. *Carbohydr. Res.* **2001**, *336*, 117.
- Cros, S.; Garnier, C.; Axelos, M. A.; Imbert, A.; Perez, S. *Biopolymers* **1996**, *39*, 339.
- Typical sugar–Ca  $K_D$  values are  $>20 \text{ mM}$  (see refs 19 and 20).
- Gould, R. O.; Rankin, A. F. *J. Chem. Soc. D* **1970**, 489.
- Angyal, S. J. *Adv. Carbohydr. Chem. Biochem.* **1989**, *47*, 1.
- Intracellular  $[\text{Ca}^{2+}]$  levels in plants reach up to  $\sim 100 \text{ mM}$  (see refs 22 and 23).
- Tuteja, N.; Sopory, S. K. *Plant Signal. Behav.* **2008**, *3*, 525.
- White, P. J.; Broadley, M. R. *Ann. Bot.* **2003**, *92*, 487.
- Rinaudo, M.; Ravanat, G.; Vincendon, M. *Makromol. Chem.* **1980**, *181*, 1059.
- Grant, G. T.; Morris, E. R.; Rees, D. A. *FEBS Lett.* **1973**, *32*, 195.
- Powell, D. A.; Morris, E. R.; Gidley, M. J.; Rees, D. A. *J. Mol. Biol.* **1982**, *155*, 517.

JA9090963