ELSEVIER

Carbohydrate Research 306 (1998) 243–255

CARBOHYDRATE RESEARCH

Morphology of galactomannans: an X-ray structure analysis of guaran [1]

Rengaswami Chandrasekaran [a,*], Akella Radha [a], Kenji Okuyama [b]

[a] *Whistler Center for Carbohydrate Research, Smith Hall, Purdue University, West Lafayette, IN 47907-1160, USA*

[b] *Faculty of Technology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184, Japan*

Received 13 May 1997; accepted 26 July 1997

Abstract

The three-dimensional structure of a galactomannan has been determined by careful analysis of the X-ray data obtained from polycrystalline and well oriented fibers of guaran. The polymer forms a flat 2-fold helix of pitch 10.38 Å. The galactosyl side-chain hydrogen bonds to the mannan backbone intramolecularly and provides structural stability. Four helices are packed in an orthorhombic unit cell of dimensions $a = 9.26$, $b = 30.84$ and c (fiber axis) = 10.38 Å according to space group symmetry $I222$. Each chemical repeat consisting of 1.0 mannosyl and 0.6 galactosyl units is associated with two structured water molecules which are responsible for the formation and stability of sheets of guaran helices in the crystalline lattice. The crystallographic R-value is 0.185 for the best model. The guaran molecular structure and packing arrangement presented here are readily adoptable to the entire galactomannan family.

Keywords: Guaran; Galactomannan; X-ray diffraction; Three-dimensional structure

1. Introduction

The worldwide industrial importance of the family of galactomannans from plant seeds hardly needs to be emphasized. The polysaccharide has a mannan backbone which is randomly substituted with (1 → 6)-linked α-D-galactosyl units. Depending on its source and the galactose/mannose (G/M) ratio, there is a gradation in the rheological properties of galactomannans. For example, as the G/M ratio increases from 0.3 in locust bean gum (also known as carob gum), to 0.45 in tara gum to 0.6 in guaran, all three water-soluble, the solubility and viscosity increase. These polysaccharides are extensively used in the food industries for their thickening effects in soups and sauces. The paper and textile industries are also beneficiaries of their rheological properties. The ability of guaran to sustain high viscosity at elevated temperatures is greatly exploited in oil field recovery. Thus, the variability in galactose composition and distribution along the mannan main chain is the origin for observed rheological differences of the galactomannans [1].

* Corresponding author. Tel.: +1-765-494-4923; fax: +1-765-494-7953; e-mail: chandrar@foodsci.purdue.edu.

[1] Dedicated to Professor Roy Whistler, the pioneer in structural studies of guaran.

0008-6215/98/$19.00

In order to understand the molecular basis of the functional properties of these interesting polymers, it is helpful to know their molecular structures as well as intermolecular interactions at atomic resolution. Toward this goal, some preliminary investigations [2–4] have reported that despite differences in their species, the X-ray diffraction patterns from oriented and/or polycrystalline fibers of up to eight galactomannans, having G/M ratios in the range of 0.33 to 0.93, indicate 2-fold helix symmetry and pitch of about 10.3 Å, similar to that of mannan [5]. Their unit cell dimensions are remarkably similar and the volume of the orthorhombic cell in every case is large enough to accommodate four helices, related by space group symmetry. Unfortunately, precise structural details have not been reported for any of these galactomannans.

We have now used good quality diffraction patterns obtained from polycrystalline and well-oriented fibers of guaran and successfully determined both the molecular structure and packing arrangement in the unit cell. Ordered water molecules are found to promote lateral organization of guaran helices. The repeating unit of this helix corresponds to

$$\begin{array}{c}\rightarrow 4)\text{-}\beta\text{-D-Man-}(1\rightarrow \\ \qquad\quad 6 \\ \qquad\quad \uparrow \\ \qquad\quad 1 \\ (\alpha\text{-D-Gal})_x\end{array}$$

where the subscript x, the G/M ratio, is 0.6 for guaran; hence, occupancy of 1.0 and 0.6 were used for Man and Gal, respectively, in structure factor calculations. For stereochemical purposes, the description of the galactomannan in this presentation is based on a tandem 6-linked Gal substitution on every residue of its mannan backbone. Hence, the structural results are applicable to the entire family of galactomannans.

2. Experimental data

Material.—A sample of commercially sold guar gum, under the trade name 'Jaguar A-40-F' (Stein-Hall, New York), was dissolved in distilled water (3 mg/mL) at room temperature. The cold water insoluble pellet was removed after centrifugation [6]. Following phenol treatment to remove traces of protein impurity, the solution was passed through a 5 μm pore filter and then dialyzed against distilled water. Finally, the sample was lyophilized.

Fiber preparation.—Droplets of the polysaccharide solution (1 mg/mL) were placed in the gap ($\sim$ 2 mm) between beaded glass rods in a fiber puller and allowed to dry slowly under controlled relative humidity (r.h.), usually in the range of 70% to 95%. Upon reaching a semi-solid state, the fiber was gradually stretched to a length of 4 mm or more. The density of a fiber, when required, was measured by the flotation method using a mixture of bromobenzene and chlorobenzene.

Diffraction pattern and unit cell contents.—Pinhole collimated and Ni-filtered CuKα radiation (λ = 1.5418 Å) from a Picker microfocus X-ray generator operated at 40 kV and 6 mA was used to irradiate a guaran fiber and record its diffraction pattern on a flat-film camera. During X-ray exposure, the fiber chamber was continuously flushed with a stream of helium gas that had earlier bubbled through a chosen saturated salt solution so as to retain the guaran sample at a desired r.h. and also to minimize undesirable background air scattering. The fiber was dusted with calcite powder (characteristic *d*-spacing 3.035 Å) for internal calibration, i.e., accurate determination of fiber-to-film distance ($\sim$ 3.9 cm).

A typical diffraction pattern shown in Fig. 1 confirms that the specimen has very good polycrystallinity and orientation. There are a total of 20

Fig. 1. X-Ray diffraction pattern from a polycrystalline and well oriented fiber of guaran (81% r.h.) which was tipped approximately 9° towards the incident beam in order to register the meridional reflection on the second layer line.

non-meridional and 2 meridional reflections up to 2.5 Å resolution. All these spots are indexable with an orthorhombic unit cell of dimensions $a = 9.26$, $b = 30.84$ and c (fiber axis) $= 10.38$ Å. The estimated standard deviations are 0.13, 0.22, and 0.05 Å, respectively. The two meridional reflections on the 2nd and 4th layer lines support the notion that the polysaccharide conforms to 2-fold helix symmetry and a pitch of 10.38 Å. The observed and calculated interplanar spacings are listed in Table 1. The systematic absences for $h + k + l =$ odd reflections are consistent with space group $I222$. According to the measured fiber density (1.44 g/mL), the unit cell has room only for four one-turn guaran helices I to IV, as well as four water molecules associated with each helix, which are related by 222 symmetry. Fig. 2 shows these locations and symmetry operations for this space group, No. 23 in ref. [7].

X-ray intensity data.—The diffraction pattern was scanned on an Optronics rotating drum microdensitometer and digitized at 100 μm intervals. A Lexidata graphics station connected to a VAX system was used to display the pattern. Steps relating to background estimation and removal, integrated intensity and Lp factor for each observed Bragg reflection were performed according to published procedures [8,9] and thereby the observed structure amplitudes were computed. The lowest measured intensity was assigned as the threshold value for each of the 15 faint reflections that are too weak to be seen in the diffraction pattern.

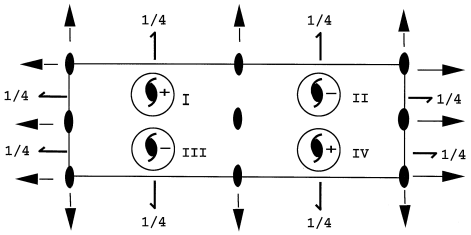

Fig. 2. Symmetry elements for space group $I222$ marked for one unit cell. The a and b axes are down and across the page, respectively. Roman letters refer to the four guaran molecules whose helix axes coincide with crystallographic 2_1 symmetry along c; + and − define the up and down polarities.

Table 1
Intensity, Bragg angle, Miller indices, and measured (d_o) and calculated (d_c) interplanar spacings of the observed reflections in the diffraction pattern of guaran

Reflection	Intensity	2θ (deg)	d_o Å	h	k	l	d_c Å
1	S	5.8	15.36	0	2	0	15.40
2	S	11.4	7.74	0	4	0	7.70
3	M	13.1	6.74	1	3	0	6.89
4	M	17.1	5.19	0	6	0	5.13
5	VSD	19.6	4.52	2	0	0	4.63
				2	2	0	4.45
6	S	23.1	3.85	0	8	0	3.85
7	W	9.3	9.49	0	1	1	9.85
8	M	14.2	6.23	1	2	1	6.32
9	W	17.4	5.09	1	4	1	5.15
10	M	21.7	4.10	1	6	1	4.12
11	M	32.6	2.75	3	4	1	2.71
12	W	33.3	2.69	2	9	1	2.67
13	VS	16.9	5.24	0	0	2	5.19
14	S	20.0	4.43	0	4	2	4.31
15	W	21.6	4.11	1	3	2	4.15
16	M	24.3	3.66	1	5	2	3.65
17	M	26.8	3.32	0	1	3	3.44
18	M	28.1	3.17	1	2	3	3.18
19	W	34.5	2.60	2	3	3	2.68
20	W	35.6	2.52	1	8	3	2.48
21	M	34.9	2.57	0	0	4	2.60
22	M	36.5	2.46	0	4	4	2.46

VS, S, M, and W denote very strong, strong, medium and weak, respectively. D stands for diffuse.

3. Structure analysis

Molecular modelling and refinement.—The linked-atom least-squares program [10] was used to construct a guaran 2-fold helix of pitch 10.38 Å whose backbone resembles that of mannan or cellulose [11]. Pyranose rings in the standard $^4C_1$ conformation [12] were assigned to Man and Gal residues. The six major conformation angles varied were (ϕ_1, ψ_1, χ_1) in the main chain and (ϕ_2, ψ_2, χ_2) in the side chain. Note that the galactose residue orientation relative to mannose is controlled by χ_1, ϕ_2 and ψ_2. Two-dimensional energy maps for the Man–Gal disaccharide in the (ϕ_2, ψ_2) plane for the three staggered domains of χ_1 were helpful to assign their probable starting values in the side chain. The bond angles τ_1 and τ_2 at the glycosidic bridge oxygen atoms O-4 and O-6, for the $1 \rightarrow 4$ and $1 \rightarrow 6$ linkages, respectively, were fixed at 116.5°. Orientation (μ), and three positioning parameters ($u = 0.25$, $v = 0.25$, w), expressed in fractions of a, b and c, respectively, define the location of the first helix, μ and w being the only two variable packing parameters. Once their values were known, the equivalent positions for the space group symmetry were used to generate all four helices in the unit cell, and the packing arrangement was scrutinized. In the final

round of refinement, both sugar rings in the repeating unit were flexed by varying their endocyclic torsion and bond angles subject to ring closure constraints; in addition, τ_1 and τ_2 were allowed to vary.

The function minimized by the LALS program is

$$\Omega = \Sigma w_m(\Delta F_m)^2 + \Sigma e_i({}_0\theta_i - \theta_i)^2 + \Sigma k_j({}_o d_j - d_j)^2 + \Sigma \lambda_h G_h$$

The first term minimizes the sum of squares of differences between observed (F_o) and calculated (F_c) X-ray structure amplitudes. An unobserved reflection was included in the refinement only when its calculated structure amplitude was greater than that observed. The second term helps to retain a conformation or bond angle θ_i near its standard or expected value ${}_0\theta_i$. The third term optimizes hydrogen bonds and reduces steric compression within and between molecules. The last term imposes constraints (G_h) for helix connectivity and sugar ring closure; it vanishes when they are fully satisfied. The parameters w_m, e_i and k_j are the individual weights associated with observations in the first three categories in that order, and λ_h is the Lagrange multiplier for the hth constraint. In the initial stages of refinement of the hydrated polymer structure, water smeared scattering factors [13] were used to calculate the structure factors.

Location of water molecules.—Three-dimensional difference Fourier maps, in sections at intervals of $c/32$, were computed with $2F_o - F_c$ as coefficients in which F_c and its phase were derived from the current model corresponding to normal scattering factors [14]. Peaks of positive electron density surrounding the guaran helices were tested for water molecules, each satisfying at least two links either among themselves or with a polymer chain. Inclusion of a water molecule depended mostly on the relative drop in the crystallographic R-value $(= \Sigma||F_o| - |F_c||/\Sigma|F_o|)$.

4. Results

Molecular structure and packing arrangement.—The starting 2-fold helical model was a generic galactomannan chain, characterized by χ_1 in the gauche-minus domain, that was constructed for visualizing a xanthan: galactomannan complex [15]. Keeping this helix as a rigid body and atom O-1 of mannose to define the x-axis of the helix, the best values for μ and w were selected by examining the variation of contacts between the chains in the unit cell in the entire 2-dimensional plane at sufficiently close intervals (10° for μ and 0.05 for w). The position for minimum steric compression (120°, 0.35) was equally

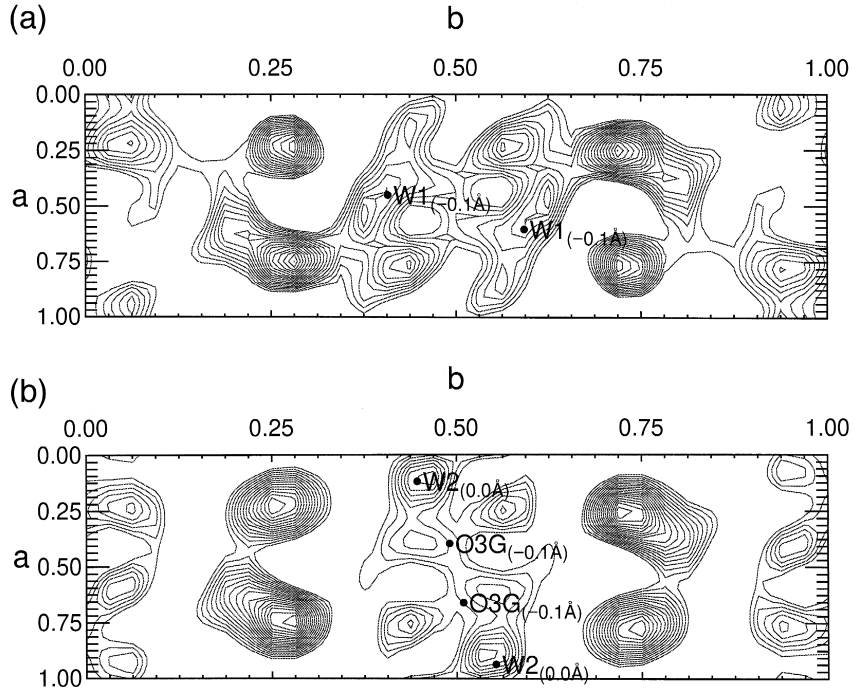

Fig. 3. Two sections of three-dimensional $(2F_o - F_c)$ map at levels (a) $29c/32$ and (b) $31c/32$ showing $W1$ and $W2$, respectively. The actual z-displacement of each labeled atom from its section is given in parentheses.

satisfactory in terms of the X-ray fit ($R = 0.28$). Refinement of this crystal structure (namely, varying the 6 molecular and 2 packing parameters) along with the X-ray scale factor produced a converged model free from any unacceptable contacts, displaying several intra and inter-chain hydrogen bonds and with a significantly lower *R*-value (0.24). An interesting observation was that each helix was interacting with the remaining three helices in the unit cell.

Aborted models.—Since there is no steric restriction for generating a galactomannan helix that exploits the gauche plus and trans domains for χ_1, we considered both of them as alternatives and performed the preliminary packing analysis as before. The contact search alone was adequate to declare either model as far inferior to that just described (χ_1 = gauche minus case). The same conclusion was independently reached on the basis of $R = 0.34$ and 0.36, respectively, for the losers. Consequently, we made no attempts to resurrect these two possibilities

Table 2
Cartesian and cylindrical polar coordinates of a repeating unit of guaran

Group	Atom	*x* (Å)	*y* (Å)	*z* (Å)	*r* (Å)	ϕ (°)
Mannose	C-1	−0.0633	−0.5720	1.2747	0.5755	−96.31
	C-2	0.5249	−1.5223	2.3093	1.6103	−70.98
	C-3	0.0483	−1.0336	3.6685	1.0347	−87.32
	C-4	0.3186	0.4175	4.0436	0.5252	52.65
	C-5	−0.0620	1.3387	2.8894	1.3401	92.65
	C-6	0.5309	2.7249	3.0283	2.7761	78.98
	O-1	0.4457	−0.7852	0.0000	0.9029	−60.42
	O-2	1.9430	−1.5269	2.1912	2.4712	−38.16
	O-3	0.6548	−1.8160	4.6989	1.9304	−70.17
	O-4	−0.4457	0.7852	5.1900	0.9029	119.58
	O-5	0.4902	0.6553	1.7535	0.8184	53.20
	O-6	−0.4711	3.7056	3.2937	3.7354	97.25
	H-1	−1.1614	−0.5095	1.2892	1.2682	−156.31
	H-2	0.2001	−2.5502	2.0898	2.5580	−85.51
	H-3	−1.0425	−1.1561	3.7405	1.5567	−132.04
	H-4	1.3900	0.5307	4.2653	1.4879	20.90
	H-5	−1.1565	1.4396	2.8461	1.8466	128.78
	H-61	1.0670	2.9887	2.1048	3.1735	70.35
	H-62	1.2712	2.7273	3.8418	3.0090	65.01
Galactose	C-1	−0.6518	4.7977	2.4547	4.8418	97.74
	C-2	0.3178	5.8936	2.8771	5.9022	86.91
	C-3	0.2319	6.8893	1.7304	6.8932	88.07
	C-4	0.6132	6.3598	0.3542	6.3893	84.49
	C-5	−0.4044	5.2759	0.0148	5.2914	94.38
	C-6	−0.0744	4.5465	−1.2702	4.5471	90.94
	O-1	−0.4711	3.7056	3.2937	3.7354	97.25
	O-2	−0.0803	6.4122	4.1410	6.4127	90.72
	O-3	1.0911	8.0000	1.9947	8.0741	82.23
	O-4	1.9331	5.8206	0.3772	6.1332	71.63
	O-5	−0.3353	4.3714	1.1280	4.3842	94.39
	O-6	1.0747	3.7127	−1.1262	3.8651	73.86
	H-1	−1.6969	5.1358	2.5124	5.4089	108.28
	H-2	1.3255	5.4677	2.9918	5.6261	76.37
	H-3	−0.7988	7.2640	1.6464	7.3078	96.28
	H-4	0.5740	7.1797	−0.3780	7.2026	85.43
	H-5	−1.3995	5.7302	−0.1010	5.8986	103.72
	H-61	−0.9357	3.9346	−1.5763	4.0443	103.38
	H-62	0.1279	5.2786	−2.0660	5.2801	88.61
Water	*W*1	1.6008	4.8692	5.6192	5.1256	71.80
	*W*2	−1.4681	6.0742	6.4290	6.2491	103.59

The coordinates of the first repeat of helices I to IV in the unit cell are: (I) $0.25a + x$, $0.25b + y$, $0.35c + z$; (II) $0.25a + x$, $0.75b - y$, $0.65c - z$; (III) $0.75a - x$, $0.25b + y$, $0.65c - z$; (IV) $0.75a - x$, $0.75b - y$, $0.35c + z$.
Additional repeats can be generated by applying helix symmetry to each of them.

and they were comfortably rejected with greater than 99.5% confidence [16].

Development of the crystal structure.—The first difference Fourier map for the winning (χ_1 = gauche minus) model was very informative. Examination of the sections at various levels of z confirmed that the orientations of the galactosyl residue and its terminal hydroxyl group were correct. As expected, because of partial occupancy, the side chain atoms uniformly appeared at a lower electron density than the main chain atoms. The vicinity of atoms O-2 and O-3 of Gal also contained two positive peaks (Fig. 3) that could be readily identified with the two water molecules, $W1$ and $W2$, that were expected from fiber density. These two guest molecules in the asymmetric unit follow the same 2_1 helix symmetry as the polymer chain because of space group equivalence. Inclusion of these one after the other in structure factor calculation with normal atomic scattering factors produced a substantial drop in R, to 0.25 and 0.21, respectively, from 0.30 for the polymer chains alone. As the unit cell contents had all been accounted for, the second difference map for the amended crystal structure exposed no new peaks to be concerned with. A final round of crystal structure refinement with flexible sugar rings, and normal scattering factors, helped to improve the X-ray fit substantially. At the end of quick convergence, neither the molecular and packing parameters, nor the water positions shifted significantly from the values prior to refinement. This best model, with $R = 0.185$, is sound on steric grounds. This X-ray fit accounts for the 20

Table 3
Observed and calculated structure amplitudes for guaran for (a) even and (b) odd layer lines

a

h	k	$l = 0$	2	4
0	0	**M** *[1278]*	**M** *[115]*	**M** *[38]*
0	2	246 *388*	[60] *[36]*	[91] *[52]*
1	1	[71] *[57]*		
0	4	249 *282*	291 *222*	260 *178*
1	3	119 *141*	132 *49*	[262] [193]
0 1	6 5	411 *372*	251 *245*	
2 2	0 2	803 *845*	[140] *[120]*	
1 2 0	7 4 8	373 *336*		
2	6	(185) *(190)*		

b

h	k	$l = 1$	3
0	1	48 *63*	157 *109*
0	3	[51] *[13]*	[78] *[67]*
1	0	[64] *[45]*	
1	2	255 *193*	306 *188*
0 1	5 4	90 *96*	[148] *[65]*
2 1 0 2	1 6 7 3	258 *274*	219 *199*
2 1 0	5 8 9	(193) *(257)*	363 *313*
2	7	(145) *(171)*	[305] *[121]*
3 3	0 2	(184) *(215)*	
1 3	10 4	237 *191*	
0 2	11 9	221 *174*	
3	6	[220] *[67]*	

In each box, the observed amplitude is given in the first line and the calculated amplitude in the next line in italics. The curved and square brackets respectively refer to the below threshold refections included in and rejected from the least-squares refinement. M refers to meridional reflection. The calculated structure amplitudes include a temperature factor with $B = 4.0$ Å^2.

Table 4
Major conformation and bridge bond angles (and e.s.d.) in degrees in guaran and mannan helices

Parameter	Guaran	Mannan I	Mannan II	Remarks
ϕ_1(O-5M–C-1M–O-4′M–C-4′M)	−100(3)	−90	−88	$\beta(1\rightarrow4)$
ψ_1(C-1M–O-4′M–C-4′M–C-5′M)	−149(3)	−149	−153	$\beta(1\rightarrow4)$
χ_1(C-4M–C-5M–C-6M–O-6M)	−109(4)	180	−23	hydroxymethyl
τ_1(C-1M–O-4′M–C-4′M)	118(2)	117	117	bridge
ϕ_2(O-5G–C-1G–O-6M–C-6M)	32(3)			$\alpha(1\rightarrow6)$
ψ_2(C-1G–O-6M–C-6M–C-5M)	−123(3)			$\alpha(1\rightarrow6)$
χ_2(C-4G–C-5G–C-6G–O-6G)	69(4)			hydroxymethyl
τ_2(C-1G–O-6M–C-6M)	117(2)			bridge

The letters M and G, and symbol ′ refer to mannose, galactose, and the reducing end, respectively.

observed and 4 out of 15 unobserved reflections. The atomic coordinates of the final model (after applying $\mu = 119.6°$ for helix I) are given in Table 2. The observed and calculated structure amplitudes are listed in Table 3.

Morphological features.—The main chain of the galactomannan helix has an unmistakable fully extended mannan-like conformation. This can be readily verified from the major conformation angles of guaran listed in Table 4 along with those of mannan I and mannan II [11] for comparison. Using helix I as reference, Table 5 contains the details of a set of 17 attractive interactions per repeating unit that stabilize the guaran crystal structure. They are sequentially numbered in italic type in the second column for easy identification. As shown in Fig. 4, the familiar O-3H · · · O-5 hydrogen bond (*1*) connects adjacent mannose residues to provide structural stability. By adopting a *tg* orientation, i.e., trans to O-5 and gauche to C-4 ($\chi_1 = -109°$), atom O-6 of the backbone is 3.7 Å away from the helix axis. This enables the galactosyl unit to stick out and swing toward the reducing end such that its atom O-3, at 8.1 Å, describes the periphery of the helix. The two conformation angles ϕ_2 and ψ_2 (Table 4) are crucial for such a disposition that aligns the galactose ring roughly coplanar with the mannose ring to which it is attached. Atom O-3 of the adjoining mannose residue at the non-reducing end is just at hydrogen bonding distance (*2*) from atom O-5 of galactose; this constellation is strengthened by yet another hydrogen bond O-6G · · · O-3M (*3*), because of the *gg* orientation adopted by atom O-6G. As a bonus, this orientation is particularly beneficial for an intraresidue

Table 5
Attractive interactions within and among guaran helices and those involving water molecules

Type	Interaction	Atom X	Atom Y	X · · · Y (Å)	Precursor P	P–X · · · Y (°)
Intrachain	*1*	O-3M	O-5M	2.77	C-3M	100
	2	O-3M	O-5G	3.04	C-3M	143
	3	O-6G	O-3M	2.64	C-6G	85
	4	O-6G	O-4G	2.72	C-6G	81
Interchain	*5*	O-2G	O-2G(II)	2.60	C-2G	112
	6	O-2G	O-3G(II)	2.65	C-2G	137
	7	O-3G	O-3G(IV)	2.51	C-3G	114
	8	O-4G	O-3G(IV)	2.78	C-4G	109
	9	O-4G	O-4G(III)	2.95	C-4G	103
	10	O-6G	O-4G(III)	2.98	C-6G	89
	11	O-6G	O-6G(III)	2.48	C-6G	144
Water bridges	*12*	O-2G	*W*1	2.72	C-2G	96
	13	O-3G(II)	*W*1	2.69	C-3G(II)	154
	14	O-4G(III)	*W*1	2.71	C-4G(III)	121
	15	O-2G	*W*2	2.70	C-2G	148
	16	O-3G(II)	*W*2	2.89	C-3G	80
	17	C-1G(III)	*W*2	2.96	O-6M	132

Unless otherwise labeled in parentheses, all atoms belong to helix I.

O-6G · · · O-4G hydrogen bond (*4*). Thus, the creation of three-center hydrogen bonds at O-3 of mannose, two of them with neighboring galactose towards the non-reducing end, is responsible for local stability. Propagation of these interactions in a 2-fold helical fashion generates a nearly planar galactomannan chain whose wings (galactose residues) span almost 16 Å from left to right (Fig. 4).

Guaran–guaran interactions.—The crystalline lattice is an array of alternately up and down-pointing polymer chains located at intervals of $a/2$ and $b/2$. Of the four in the unit cell, I and IV are up, while II and III are down. With the intermolecular separation, i.e., distance between the helix axes, of I from II, or III from IV, being $b/2$ (15.42 Å), there are considerable hydrogen bonding interactions among their galactosyl side chains. This results in the formation of a hydrogen bonded sheet in the *bc*-plane. The two sheets in the unit cell, passing through $a/4$ and $3a/4$, are related by 222 symmetry. As most of the polysaccharide atoms enrich these sheets (Fig. 5), the 200 reflection is very strong. Atoms O-2 and O-3 give rise to a pair of bifurcated hydrogen bonds *5* and *6*. Similarly, since I and IV are only 16.1 Å apart along the diagonal of the *ab*-plane, atoms O-3 and O-4 generate two hydrogen bonds *7* and *8*. Finally, I and III are separated by $a/2$ (4.63 Å) whose atoms O-4 and O-6 add three more hydrogen bonds *9*, *10*, and *11*. In other words, every hydroxyl group in the galactose residue is a donor for an interchain hydrogen bond, at least once.

Water bridges.—The two water molecules per chemical repeat, $W1$ and $W2$, are anchored to atom O-2G by hydrogen bonds *12* and *15*, respectively.

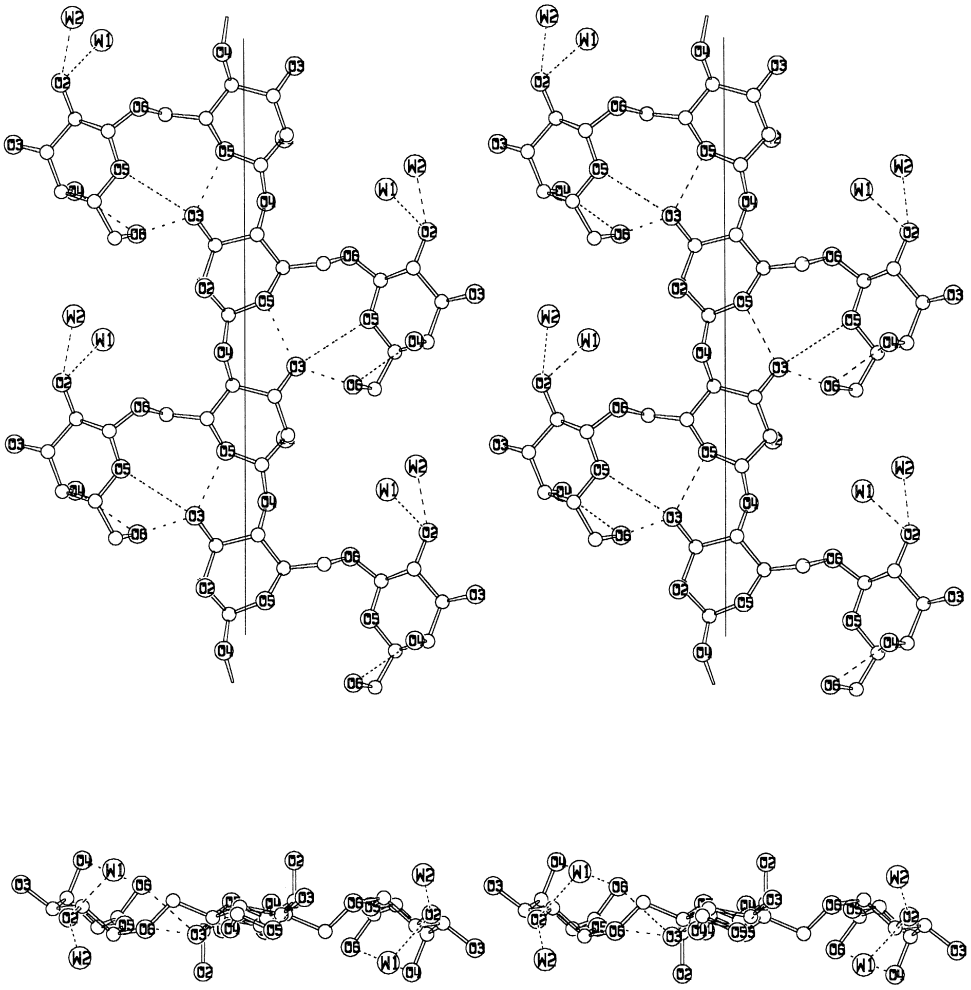

Fig. 4. Two mutually perpendicular stereo views of the guaran helix. Two turns are shown at the top. The lamellar helix of width 16 Å is stabilized by a series hydrogen bonds (dashed lines) involving both main and side chain atoms. The two water molecules/repeat are closer to galactose than mannose residue. The vertical line, helix axis, measures $2c$. An axial projection of one turn at the bottom reiterates the flatness of the galactomannan chain; its thickness is only 4 Å.

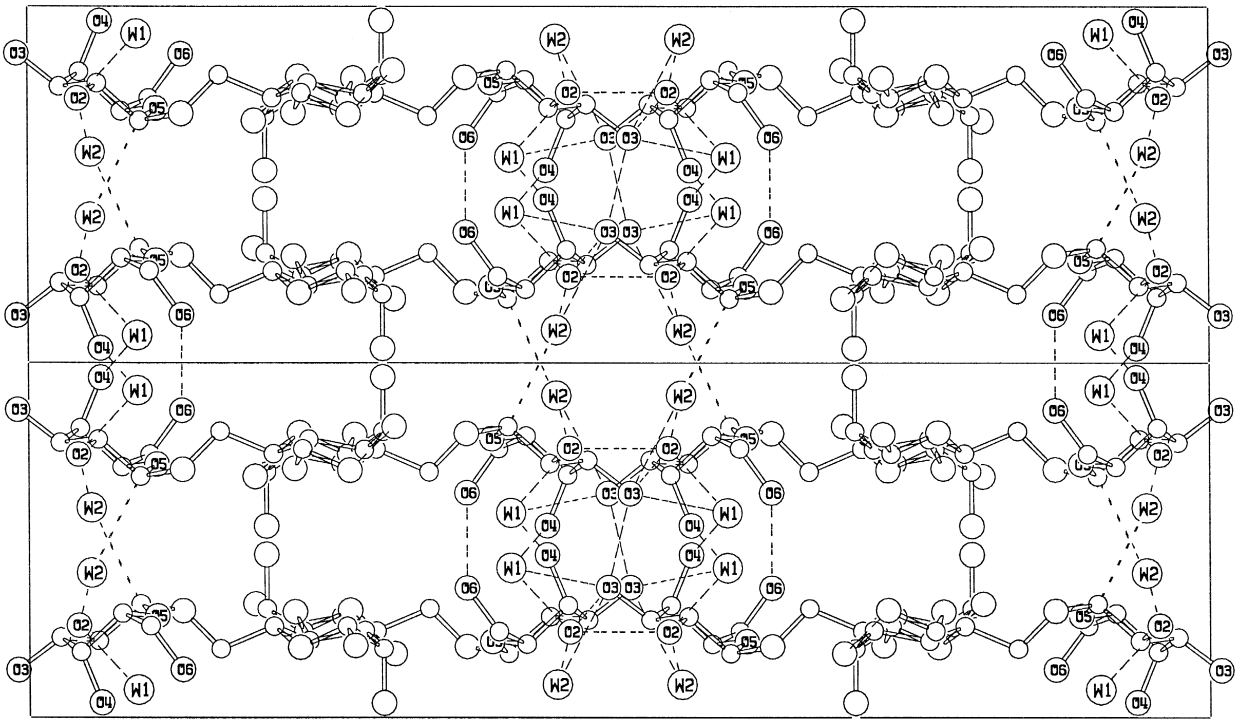

Fig. 5. Packing arrangement, as viewed down the *c*-axis, of the galactomannan helices in two unit cells (2 *a* is down and *b* across the page). Both direct and water mediated hydrogen bonds (dashed lines) involving galactose hydroxyl groups (labeled) are responsible for the association of the sheets of helices in pairs within and between unit cells.

They themselves are not connected owing to a long, 3.3 Å, separation. However, *W*1 is also able to link with atoms O-3G of II and O-4G of III through *13* and *14* and thus is able to stay in position. Likewise, *W*2 further links with not only O-3G of II (*16*), but also forms a hydrogen bond with C-1G of III (*17*).

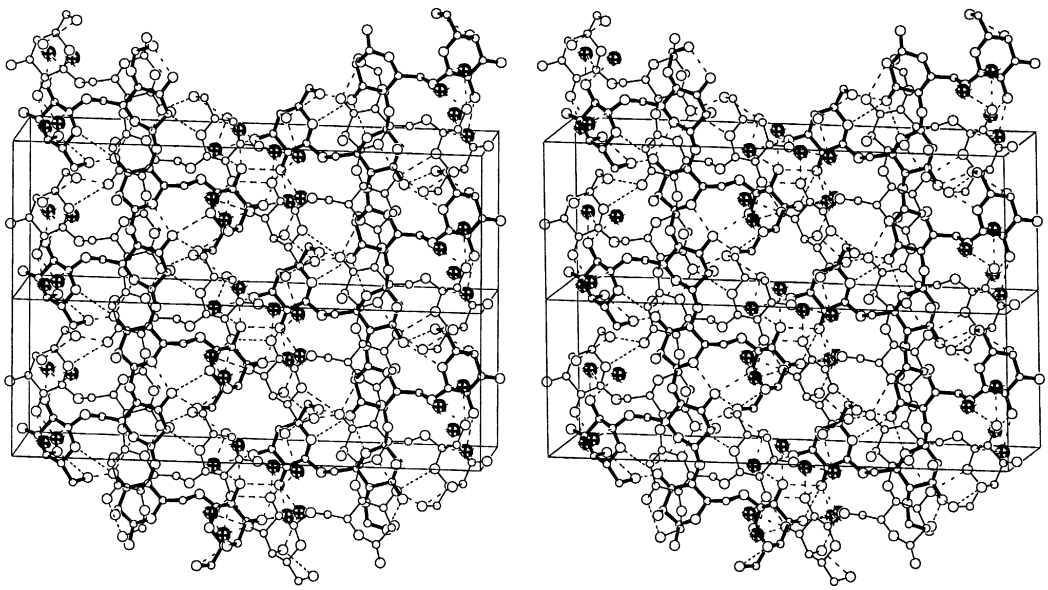

Fig. 6. A stereo view, roughly normal to the *c*-axis, of the dense packing and intermolecular interactions of helices in two unit cells. The sheet behind (thick bond) is $a/2$ from the sheet in front (thin bond). Hydrogen bonds (dashed lines) between helices and those involving water molecules (crossed circles) are extensive. While there is considerable stacking between the front and back mannan chains, the galactose residues, surrounded by water molecules, form a column in the middle and display little overlap.

Overall, as shown in Fig. 5, $W1$ and $W2$ are crucial for connecting adjacent sheets of galactomannan chains within, as well as between, unit cells essentially through the galactosyl units in the side chains and reinforcing the three-dimensional structure.

Thus, the lateral organization of galactomannan helices in the solid state is dominated by an extensive network of hydrogen bonds, as shown in Fig. 6. If the hydroxyl groups of peripheral galactose residues dominate in these interactions, so do the two water molecules and their symmetry mates. These observations suggest that depletion of galactose residues in their vicinity would allow the water molecules to wander away, leading to disruption of crystallinity.

5. Discussion

The space group riddle.—It is well over a decade since galactomannans of varying G/M ratios were all shown to pack in isomorphous unit cells [2] similar to that presented here. Four years went by before the *d*-spacings of nearly 20 reflections in the diffraction pattern of (highly substituted) fenugreek gum were first reported [3]. Relying on $h00$, h odd, and $0k0$, k odd reflections as absences, and on the assumption that the galactomannans were unable to preserve the formal 2-fold helix symmetry, the space group was suggested as $P2_12_12$ and not $P2_12_12_1$. Just two years ago, similar *d*-spacings for another series of galactomannans appeared that echoed the same appraisal [4]. In both cases, the authors based their judgment on a single datum of what they believed to be a faint meridional reflection on the third layer line that corresponded to an interplanar spacing of 3.46 Å. To date, no structural details have come from these two groups or others.

On the other hand, our guaran diffraction pattern (Fig. 1) has revealed more information. Common to several patterns are the elongated spot shapes away from the center. Unlike the sharp and small spots obtained for the calcium salt of welan that represents large crystallites in the fiber [17], the guaran pattern has elliptical spots of increasing dimensions with decreasing *d*-spacings which are diagnostic of smaller crystallites and larger disorientation in the specimen. It is, hence, not surprising to see arcs instead of spots on third and fourth layer lines. In particular, on the third layer line, the first non-meridional reflection (17 in Table 1) is a case in point; it is so wide as to trespass across the meridian into the adjoining quadrant. This persists independent of fiber tilt used to record an alleged meridional reflection on a chosen layer line. Such observations are not unusual to fibrous polymers. The diffraction pattern of mannan II (2-fold helix of pitch 10.35 Å) has a sharp meridional reflection on $l = 2$, but the 013 reflection (*d*-spacing 3.4 Å) appears as a long horizontal arc centered on the meridian [5]. Its unit cell is roughly half that of guaran along b and its crystal structure is compatible with space group $I222$ [18]. Gellan forms a 3-fold helix and produces sharp meridional reflections on $l = 3$, 6 and 9; yet, faint intensity is seen on $l = 5$, 8 and 10; the space group is $P3_1$ [19]. Poly d(IC) · poly d(IC) and related D-DNA structures are right-handed eight-fold helices of pitch 24.2 Å and they crystallize in space group $P4_1$ [20]. Apart from the expected meridional reflection on $l = 8$, the 7th layer line displays a strong non-meridional 017 reflection (*d*-spacing = 3.4 Å) that extends prominently to the meridian; this was earlier erroneously interpreted in terms of a seven-fold left-handed helix [21]. A-DNA is an 11-fold helix of pitch 28.2 Å; the real meridional reflection on $l = 11$ is much weaker than the 109 reflection (*d*-spacing = 3.1 Å) that crosses the meridian; the space group is $C2$ [22]. The list goes on and on, but these examples attest to some implicit structural features, such as stacking effects between planar bases or sugar rings at the molecular level, or else, disorientation and related properties at the macroscopic level. Both are prominently expressed in the diffraction patterns and hence caution is essential in interpreting artifacts.

Our careful examination of the entire set of measured *d*-spacings and their Miller indices for guaran (Table 1) has revealed that the assignment of 003 as an observed reflection for galactomannans as previously reported [3,4] is a crystal clear case of mistaken identity. Instead, the full data on systematic absences and observed reflections are unequivocally consistent with a body centered unit cell that we have used to establish the first galactomannan structure that retains perfect 2-fold screw symmetry. Needless to say, the space group is no longer a puzzle.

Relationship to mannan II.—The influence of side chains on the stability of branched polysaccharide helices is usually a topic of much enigma. Galactomannans [2–4] and xanthan [23–25] are prime examples of polysaccharides having a long and frustrating record of indecisive structural investigations. Our present results on guaran structure have demonstrated that monosaccharide side chains, namely, galactose residues, are systematically hydrogen bonded to the mannan main chain (Fig. 4) so as to create a flat, only 4 Å thick, galactomannan molecule that is al-

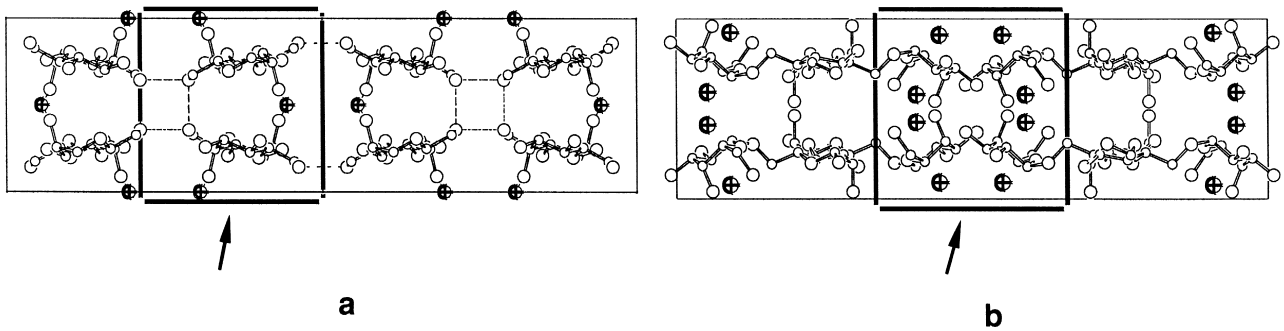

Fig. 7. Comparison of the packing arrangement in (a) two unit cells of mannan II [18] with that in (b) one unit cell of galactomannan. See text for explanation.

most twice as wide as an isolated mannan helix itself. As mentioned before, the side chains of adjacent helices further interact favorably (Fig. 6) to form a sheet parallel to the *bc*-plane. It is surprising that atoms in the backbone are not in hydrogen bonding positions with neighboring helices.

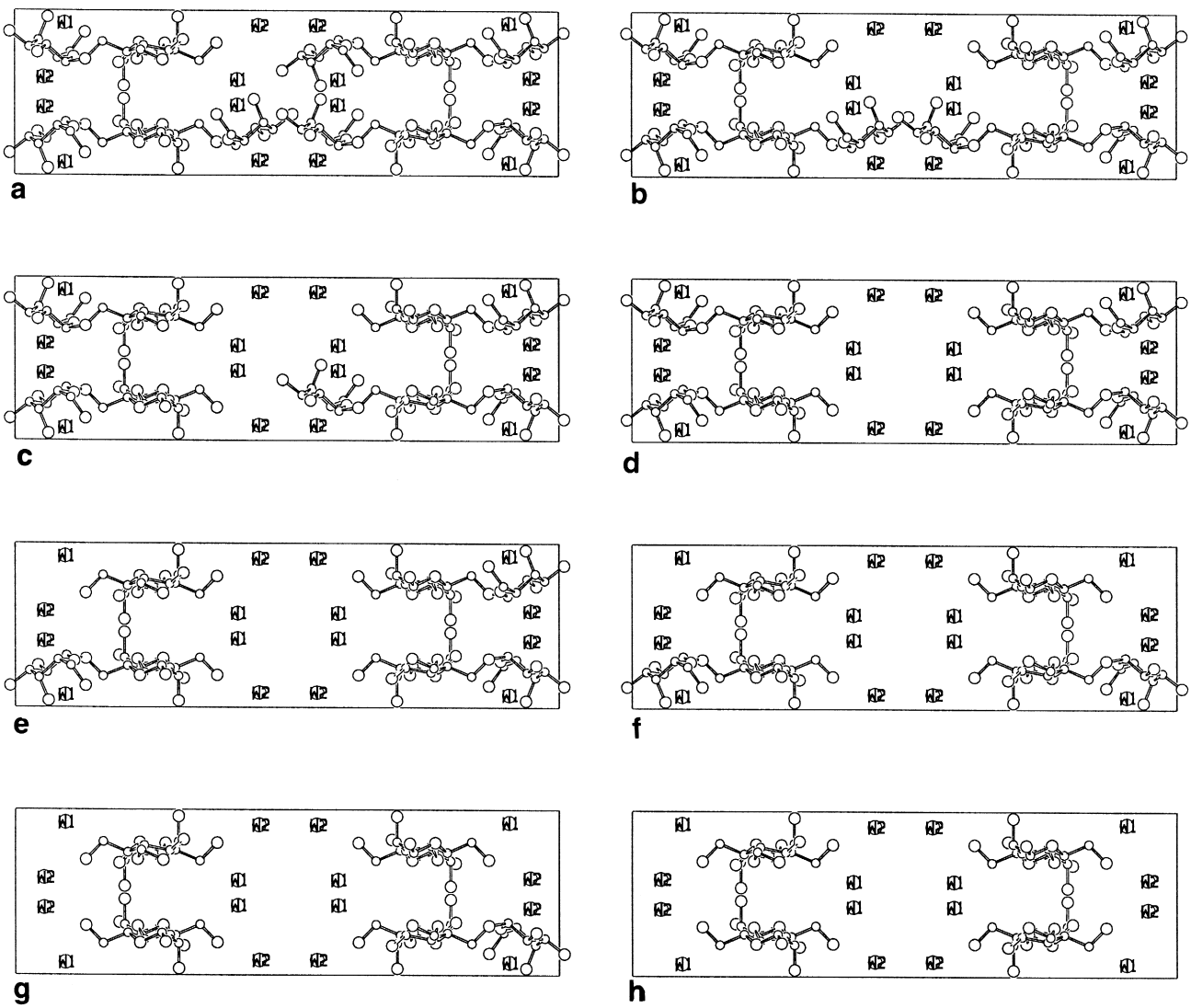

Fig. 8. Panels (a) to (h) correspond to sequential removal of one to eight side chains from the galactomannan unit cell. These static pictures help to visualize the formation of holes at sites where galactose residues are absent when the galactomannan chains associate.

In the mannan II structure, mannose residues give rise to two interhelical hydrogen bonds: O-6H · · · O-6 and water mediated O-2H · · · W · · · O-2 [18]. The former is feasible because of the low $\chi_1 = -23°$ orientation for atom O-6. This is ruled out in the galactomannan structure as O-6 is substituted; as a result, its C-6–O-6 bond is further rotated by $-86°$. The water bridge in mannan II connects helices $a/2$ apart, but this cannot be copied in the case of galactomannan because the distance between the relevant O-2 atoms, 6.5 Å, is long.

Beyond space group identity, there are considerable similarities, and differences, in the unit cell contents of galactomannan compared to that of mannan II. Some of them can be seen from the unit cell projections along the *c*-axis (Fig. 7). Using the boxed regions at the arrows as reference sites, the two unit cells of mannan II (Fig. 7a) are nearly superposable on one unit cell of guaran (Fig. 7b). There is also good match on either side of the box where mannan chains are present. Within the box, however, the mannan chains in Fig. 7a have the right width to cover the galactose rings from adjacent galactomannan chains and the associated water molecules in Fig. 7b which are juxtaposed. This is fully understandable in terms of their respective widths. The implication of this observation is that the column of galactose residues and water molecules running parallel to the *c*-axis in the galactomannan unit cell (Fig. 6) can be substituted by a pair of antiparallel mannan chains that would almost mimic the mannan II structure.

Hydration of galactomannan helices.—The X-ray diffraction pattern from guaran has been instrumental in visualizing the three-dimensional structure of a prototype galactomannan. When translated, the G/M ratio of 0.6 for guaran represents occupancy factors 1.0 and 0.6 for mannose and galactose, respectively, in the chemical repeat of the X-ray structure. In a real situation, however, the distribution is not regular; at any given site, a galactosyl unit is either present or absent, but nothing in between. In the prototype structure, there are a total of eight side chains present in the unit cell. A pictorial representation of the unit cell contents in the *c*-axis projection for different G/M ratios will be informative. Assuming that crystallinity is preserved, the eight panels in Fig. 8 depict sequentially the effect of removing one, two, up to all eight galactose residues in the unit cell. For example, when the first galactose is removed (Fig. 8a), a gap in the middle is immediately noticeable; its size increases as additional galactose residues are pulled out and when four are taken out, the hole at the center (Fig. 8d) contains only the structured water molecules. As mentioned earlier, it is wide enough to accommodate a pair of mannan chains. Removal of the remaining galactose residues creates holes on the left and right side of the polymer chains. But at every stage, as the water molecules that were once bound to the galactose residues become unstable, the galactomannan chains will start losing their original crystalline arrangement. Such a disruption is the beginning for surrounding water in the environment to gush into the voids. This is the cause for observed hydration. Once initiated, melting of the crystalline region will enhance solubility and hence viscosity.

By the time all the galactose residues are missing (Fig. 8h), only four mannan chains are left, apart from the water molecules. The latter are likely to go astray. Thus, this panel as the final packing arrangement is unrealistic for mannan since gradual lateral contraction (up to 7.1 Å) along the *b*-axis would occur during the depletion of galactose residues.

This schematic representation helps to realize that fenugreek gum (G/M = 0.9), for example, might have far fewer holes between its chains, yet leaving room for water molecules to enter. This will explain its solubility in water. A galactomannan with a very low G/M ratio, such as carob gum, would be intermediate between guaran and mannan in solubility [6].

Cross-linking of galactomannan chains and their derivatives in the presence of borates and transition metal ions modifies rheological properties. While guaran itself gives rise to a highly viscous aqueous solution, cross-linking leads to gelation [1]. Such derivatives are widely used in hydraulic fracturing for oil and gas recovery, carpet dyeing in the textile industry, and sizing in paper industry, to name a few. Structural and modelling studies are under way to unravel the molecular origin of these fascinating physical properties of galactomannans in general, and guaran in particular.

Acknowledgements

This research was supported in part by the Industrial Consortium of the Whistler Center for Carbohydrate Research, and Nutrasweet Kelco, a Unit of Monsanto.

References

[1] H. Maier, M. Anderson, C. Karl, K. Magnuson, and R.L. Whistler, in R.L. Whistler and J.N. BeMiller

(Eds.), *Industrial Gums. Polysaccharides and Their Derivatives*, Academic Press, New York, 1993, pp. 181–226.
[2] Y.Y. Chien and W.T. Winter, *Macromolecules*, 18 (1985) 1357–1359.
[3] B.K. Song, W.T. Winter, and F.R. Taravel, *Macromolecules*, 22 (1989) 2641–2644.
[4] V.J. Kapoor, H. Chanzy, and F.R. Taravel, *Carbohydr. Polymers*, 27 (1995) 229–233.
[5] E. Frei and R.D. Preston, *Proc. R. Soc. London, Ser. B*, 169 (1968) 127–145.
[6] P.A. Hui and H. Neukom, *TAPPI*, 47 (1964) 39–42.
[7] *International Tables for X-ray Crystallography*, Vol. I, Kynoch Press, England, 1969, pp. 109.
[8] R.P. Millane and S. Arnott, *J. Appl. Crystallogr*., 18 (1985) 419–423.
[9] R.P. Millane and S. Arnott, *J. Macromol. Sci. Phys.*, 24 (1985) 193–227.
[10] P.J.C. Smith and S. Arnott, *Acta Crystallogr., Sect. A*, 34 (1978) 3–11.
[11] R. Chandrasekaran, *Adv. Carbohydr. Chem. Biochem*., 52 (1997), 311-439.
[12] S. Arnott and W.E. Scott, *J. Chem. Soc., Perkin Trans. 2*, (1972) 324–335.
[13] R. Chandrasekaran and A. Radha, *J. Biomol. Struct. Dynamics*, 10 (1992) 153–168.
[14] *International Tables for X-ray Crystallography*, Vol. IV, Kynoch Press, England, 1974, pp. 99–101.
[15] R. Chandrasekaran and A. Radha, *Carbohydr. Polymers*, 32 (1997) 201-208.
[16] W.C. Hamilton, *Acta Crystallogr.*, 18 (1965) 502–510.
[17] R. Chandrasekaran, A. Radha, and E.J. Lee, *Carbohydr. Res.*, 252 (1994) 183–207.
[18] R.P. Millane and T.L. Hendrixson, *Carbohydr. Polym.*, 25 (1994) 245–251.
[19] R. Chandrasekaran, L.C. Puigjaner, L.C. Joyce, and S. Arnott, *Carbohydr. Res.*, 181 (1988) 23–40.
[20] S. Arnott, R. Chandrasekaran, L.C. Puigjaner, J.K. Walker, I.H. Hall, and D.L. Birdsall, *Nucl. Acids Res.*, 11 (1983) 1457–1474.
[21] H.R. Drew and R.E. Dickerson, *EMBO J.*, 1 (1982) 663–667.
[22] R. Chandrasekaran, M. Wang, R.-G. He, L.C. Puigjaner, M.A. Byler, R.P. Millane, and S. Arnott, *J. Biomol. Struct. Dynamics*, 6 (1989) 1189–1202.
[23] G. Holzwarth and E.B. Prestridge, *Science*, 197 (1977) 757–759.
[24] R. Moorhouse, M.D. Walkinshaw, and S. Arnott, *ACS Symp. Ser.*, 45 (1977) 90–102.
[25] R.P. Millane, T.V. Narasaiah, and S. Arnott, in V. Crescenzi, I.C.M. Dea, S. Paoletti, S.S. Stivala, and I.W. Sutherland (Eds.), *Biochem. Biotechnol. Ind. Polysaccharides*, Gordon and Breach, New York, 1989, pp. 469–478.