ELSEVIER

Carbohydrate Research 311 (1998) 203–217

CARBOHYDRATE RESEARCH

The preferred conformations of the four oligomeric fragments of Rhamnogalacturonan II

Karim Mazeau, Serge Pérez *

Centre de Recherches sur les Macromolécules Végétales, CNRS, BP 53, 38041 Grenoble Cedex 9, France[1]

Received 25 February 1998; accepted 15 July 1998

Abstract

Rhamogalacturonan II (RG-II) is a structurally complex pectic mega-oligosaccharide that is released enzymatically from the primary cell wall of higher plants. RG-II contains 28 monosaccharide units (MW≅6KDa) which belong to 12 different families of glycosyl residues, including very unusual ones such as Kdo, Dha, aceric acid, and apiose. Eighteen different disaccharide segments can be identified, and so far the primary structure has not yet been determined. These monomeric units are arranged into four structurally well-defined oligosaccharide side chains, linked to a pectic backbone made up of 1,4-linked α-D-galactosyluronic acid residues. The specific attachment sites of these four side-chains on the pectic backbone remains to be elucidated. The present work presents a three-dimensional database of all the monosaccharide and disaccharide components of RG-II. The conformational behavior of D-Api*f* and L-AceA*f* monosaccharide has been assessed through computations performed with the molecular mechanics program MM3 using the flexible residue approach. For each furanosyl residue, energies of various envelope and twist conformers were systematically calculated as a function of the puckering parameters Q and ϕ. Energy minima are observed in both the Northern and Southern zones of the conformational wheel of each monosaccharide. As for the constituting segments, the conformational behaviour of 18 different disaccharides was evaluated using the flexible residue procedure of the MM3 molecular mechanics procedure. For each disaccharide, the adiabatic energy surface, along with the locations of the local energy minima and drawings of the conformations of each local minimum located in the energy maps have been established. The geometries of the minima and the potential energy surfaces of the different fragments were included in the database of the POLYS, a program for building oligo and polysaccharides. All these results were used for the generation, prior to a complete optimization, of the complete structure of each fragment of RG-II. It is shown that both A and B fragments are very flexible about the two sidechain glycosidic linkages which are closest to the backbone. The remaining part of the sidechain is rigid for the heavily branched A fragment, it is flexible for the more linear B fragment. The lowest energy conformer of each fragment results in good exposure of the hydroxyl groups of the apiosyl residues. Some possible implications of these features in boron complexation are presented.

Keywords: Rhamnogalacturonan II; Kdo; Dha; Aceric acid; Apiose Conformation; MM3; Furanosyl; Puckering parameters

* Corresponding author. Tel.: 33 476 03 76 30; Fax: 33 476 03 76 29; e-mail: perez@cermav.grenet.fr

[1] Associated with University Joseph Fourier.

0008-6215/98/$—see front matter
PII S0008-6215(98)00190-6

1. Introduction

Rhamnogalacturonan-II (RG-II) is a pectic mega-oligosaccharide that is released from the primary cell walls of higher plants by treatment with endopolygalacturonase. RG-II has been isolated from the walls of suspension-cultured sycamore cells [1], Douglas fir [2], rice [3], onion [4], kiwi fruits [5,6], red wine [7,8] and from Pectinol AC [9,10]. Unlike most of the known complex carbohydrates, whose structure depends on many factors, RG-II has been shown to have virtually the same structure in every plant from which it has been isolated. It is most unusual to have such a well-defined and highly conserved carbohydrate structure in the plant kingdom.

RG-II contains L-Rha, L-Ara, L-Fuc, D-Gal, D-GalA and D-GlcA, and rare monosaccharides such as 2-*O*-methyl-L-fucose, 2-*O*-methyl-D–Xyl, apiose, aceric acid, Kdo and Dha. RG-II has a backbone composed of at least seven (1→4)-linked α-D-GalA residues [11] to which four structurally well-identified side chains are attached [12]. The carboxyl groups of some of the D-GalA residues are esterified with methyl groups. Eighteen different disaccharide segments can be identified in RG-II. While almost fully characterized, its complete structure remains to be fully established. In particular, the specific sites of attachment of the four side chains along the homogalacturonan backbone and the specific location of the methyl–ester groups are still unknown.

The biological function of RG-II has not been established yet. Recent reports have shown that it is the most likely candidate responsible for borate complexation [13] to the extent that most of the boron would be bound to RG-II in the cell wall [14]. RG-II exists as both a monomer and a dimer that is cross-linked by a borate di-ester linkage. This linkage is stable, the dimer accounts for 98% of the RG-II in some red wines. Since boron is an essential micronutrient for plant [15], this strongly suggests an obvious biological importance for RG–II. It has also been shown to enhance the expression of the macrophage Fc receptor [16].

Since RG-II may be involved in a series of important biological phenomena, an understanding of the molecular mechanism of recognition is essential. Unfortunately, such a large and complex molecule is difficult to analyze by experimental methods. Nevertheless, detailed structural information can be complemented by computational conformational analysis. The aim of this work is the characterization of the conformations of each fragment of RG-II since the primary structure of the fragments is still not established. This will provide a starting point towards the characterization of the conformational properties, dynamics and boron interaction of the whole structure.

The treatment of RG-II at that level of complexity can be undertaken using the following route: (a) identification of monosaccharide constituents; (b) identification of disaccharide segments of RG-II; (c) molecular modeling of each segment, and description of the potential energy surface; (d) collection of the 3-D structures of low energy conformers; (e) building realistic possible starting conformations from assembly of the low energy conformers; (f) energy minimization on the whole structure. The present work describes the results of the application of such a procedure to the four oligomeric fragments that constitute the side chains of RG-II.

2. Materials and methods

The schematic representation for the hypothetical structure of RG-II is displayed in Fig. 1. As recently reported [8] RG-II contains 12 different glycosyl residues: L-Rha*p*, L-Ara*f*, and L-Ara*p*, L-Fuc*p*, D-Gal*p*, D-GalA*p* and D-GlcA*p*, and rare monosaccharides such as 2-*O*-methyl-L-Fuc, 2-*O*-methyl-D-Xyl, 3-C-(hydroxymethyl)–D-*glycero*-tetrose or apiose (D-Api*f*), 3-C-carboxy-5-deoxy-L-Xyl or aceric acid (L-AceA*f*), 3-deoxy-D-*manno*-2-octulosonic acid (Kdo*p*) and 3-deoxy-D-lyxo-2-heptulosaric acid (D-Dha*p*). Eighteen different disaccharide segments can be identified in RG-II. For the sake of clarity, these disaccharides moieties have been labeled RG-II*i*, with $i = 1, 18$, as shown in Table 1. RG-II has four structurally well defined oligosaccharides side chains (namely A, B, C, and D) that are attached to a backbone composed of at least seven (1→4)-linked α-D-GalA residues.

Nomenclature.—The recommendation and symbols proposed by the Commission on Biochemical Nomenclature are used throughout this paper [17].

Building blocks.—RG-II contains 12 different monomeric units as shown in Fig. 2. Nine of them are pyranoid; these are D-Gal*p*, D-GalA*p*, D-GlcA*p*, L-Ara*p*, D-Xyl*p*, L-Fuc*p*, L-Rha*p*, D-Dha*p* and Kdo*p*. The starting geometry and conformation of each of these nine monosaccharides was

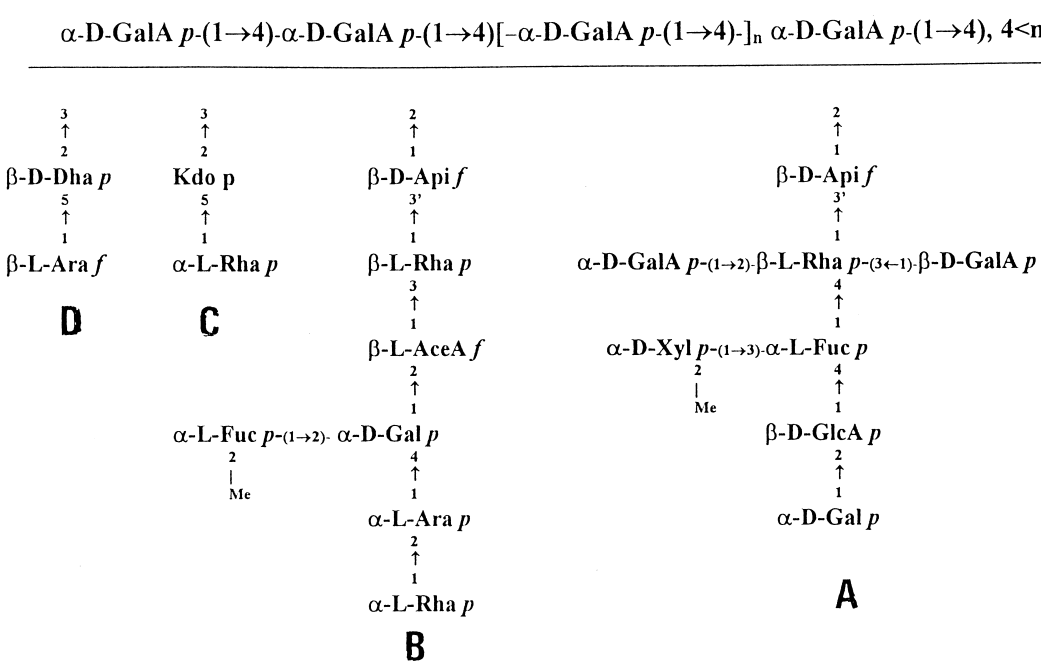


Fig. 1. Primary structure of RG-II (from ref. [8]).

Table 1
Nomenclature of all the disaccharides

β-D-Api*f*(1→2)-α-D-GalA*p*	RGII-01
β-L-Rha*p*(1→3)-β-D-Api*f*	RGII-02
β-D-GalA*p*(1→3)-β-L-Rha*p*	RGII-03
α-D-GalA*p*(1→2)-β-L-Rha*p*	RGII-04
α-L-Fuc*p*(1→4)-β-L-Rha*p*	RGII-05
α-D-Xyl2Me*p*(1→3)-α-L-Fuc*p*	RGII-06
β-D-GlcA*p*(1→4)-α-L-Fuc*p*	RGII-07
α-D-Gal*p*(1→2)-β-D-GlcA*p*	RGII-08
β-L-AceA*f*(1→3)-β-L-Rha*p*	RGII-09
α-D-Gal*p*(1→2)-β-L-AceA*f*	RGII-10
α-L-Fuc2Me*p*(1→2)-α-D-Gal*p*	RGII-11
α-L-Ara*p*(1→4)-α-D-Gal*p*	RGII-12
α-L-Rha*p*(1→2)-α-L-Ara*p*	RGII-13
β-L-Ara*f*(1→2)-α-L-Rha*p*	RGII-14
Kdo*p*(2→3)-α-D-GalA*p*	RGII-15
α-L-Rha*p*(1→5)-Kdo*p*	RGII-16
β-D-Dha*p*(2→3)-α-D-GalA*p*	RGII-17
β-L-Ara*f*(1→5)-β-D-Dha*p*	RGII-18

taken the MONOBANK data base [18] or constructed. Each residue has been submitted to a complete optimization through MM3(92) [19,20]. The first five monosaccharides display $^4C_1$ conformation, L-Fuc*p* and L-Rha*p* are in a $^1C_4$ conformation. The two others have their ring in the $^2C_5$ form as indicated by crystal structures [21,22] and homonuclear NMR coupling constants [23]. Nevertheless, eventually some conformational changes may be expected to occur and they can be fully accounted for.

Three among the twelve different monosaccharides, which are present in RG-II, are furanoids. These are L-Ara*f*, D-Api*f* and L-AceA*f*. Up to now, only the L-Ara*f* has been subjected to a thorough conformational investigation [24]. Flexible ring shapes can be conveniently defined by the so-called pseudorotation parameters. Their description involves amplitudes, which measure the deviation from planarity, and phase angles that describe the type of distortion. One amplitude parameter and one phase angle are needed to characterize five-membered ring puckering. The Cremer–Pople parameters [25,26] (Q, ϕ) are used throughout this study. The relationship between the Cremer–Pople parameters and the 20 different twist (T) and envelope (E) shapes that can be adopted by a five-membered ring are identified on the pseudorotation wheel.

As starting models for the molecular mechanics calculations, 121 five-membered rings were built following the procedure described for the conformational study of fructofuranose [27,28] and arabinofuranose [24]. For each of the 10 envelope shapes, six rings were built using six displacements (0.3–0.8 Å) of one atom above or under the ring plane. In the case of the twist shapes, two atoms were displaced (six displacement values from 0.1 to

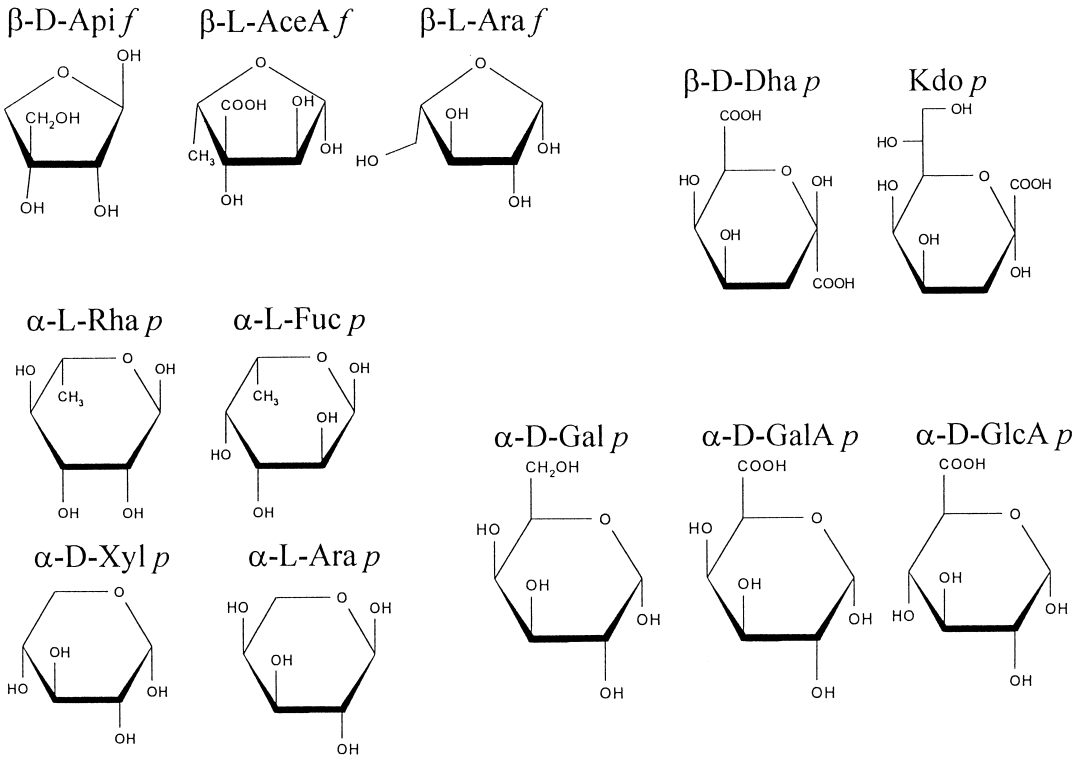


Fig. 2. Schematic representation of all the constitutive monomeric units occurring in RG–II.

0.5 Å) in opposite directions. The planar ring was also constructed. Therefore 121 starting models, on the pseudorotational wheel, were built for each relaxed map. The generation of the 121 starting models was repeated 54 and 81 times, to account for all the combinations arising from the different orientations of the lateral groups of L-AceA*f* and D-Api*f*, respectively. Different orientations of the pendant groups were considered as starting conformations for L-AceA*f* and D-Api*f*, respectively. The molecular mechanics program MM3(92) was used for the energy calculations. The three atoms that define the plane were made coincident with the *x*–*y* plane; optimizations were then performed by allowing the Cartesian coordinates of each atom to vary except for the *z* coordinate of the five ring atoms. For each so optimized conformation, the Q and ϕ parameters were recalculated. Adiabatic potential energy surfaces were plotted as a function of the puckering parameters by using only the lowest energy conformer at each Q and ϕ coordinate. The adiabatic maps were constructed from the optimization of 6534 and 9801 different conformers of L-AceA*f* and D-Api*f*, respectively.

Glycosidic linkages.—The global shape of a disaccharide is mainly governed by the rotations about the glycosidic linkage [29]. The relative orientations of saccharide units are expressed in terms of the glycosidic linkage torsion angles Φ and Ψ which have the definition Φ = O-5–C-1–O–1–C′-*x* and Ψ = C-1–O-1–C'_x–C'_{x-1} for a 1→*x* linkage. An additional rotatable bond (ω) occurs in particular in the 1→3′ glycosidic linkage. All three staggered positions of the torsion angle ω are considered.

The conformational space of each disaccharide segment of RG-II (18 disaccharides) was explored in a systematic fashion by stepping the glycosidic Φ, Ψ and ω torsion angles in 20° increments over the whole angular range. Adiabatic contour maps were drawn for each disaccharide by using different orientations of the pendent groups (and different conformations of the five-membered rings when necessary). At each conformational microstate a geometry optimization was performed by allowing coordinates of each atom to vary except those defining the Φ and Ψ torsion angles. Only the lowest energy conformer at each (Φ, Ψ) point is used for the Ramachandran-like contour plots. This procedure has the advantage of overcoming the well known multiple-minima problem of the potential energy hypersurface due to specific

orientations of the pendant groups and of fully describing the conformational flexibility around the glycosidic linkages.

The exact positions of the different minima were determined after additional minimizations by removing the constraints on the (Φ and Ψ) torsion angles.

Oligomeric fragments.—Information from the maps computed for the parent disaccharides was used to construct oligomeric fragments of RG-II in different conformations. All the combinations of the energy-minima of the disaccharides were used as starting points for unconstrained minimizations. Starting geometries of the oligomers were constructed by using the POLYS program [30].

Force field.—Geometry optimization was performed using the molecular mechanics program MM3 [19,20]. This force field contains a correction for the anomeric effect and has been shown to be especially adapted for the study of carbohydrates. The block-diagonal minimization method, with the default energy convergence criterion of 0.00008**n* kcal mol^{-1} per five iterations, *n* being the number of atoms, was used for grid point optimizations. To mimic a hydrated environment of the molecules, a dielectric constant, $\varepsilon = 78.5$, was used in all the calculations. The iso-energy contours were drawn by interpolation of 1 kcal/mol and the 10 kcal/mol contour was selected as the outer limit. Contoured maps were generated with a home made program. All calculations were carried out on a network of Silicon Graphics Indigo workstations.

3. Results and discussion

Conformational puckering of the five-membered rings.—The 3-C-hydroxymethyl-D-*glycero*-tetrose or β-D-Api*f*, having the trivial name apiose, belongs to the class of branched-chain monosaccharides. In the open chain form, the carbon C-3 atom is not asymmetric: two of the substituants on C-3 atom are hydroxymethyl groups. These two hydroxyls can, in principle, participate in the formation of acetals with the aldehyde group at C-1. As a result, the two diastereoisomeric forms of furanose are therefore possible, e.g. D-erythro- and L-threo- configurations (Fig. 3). However, since the L-threo-compound has never been observed in Nature, the present conformational analysis was confined to the D-erythro- form of apiose.

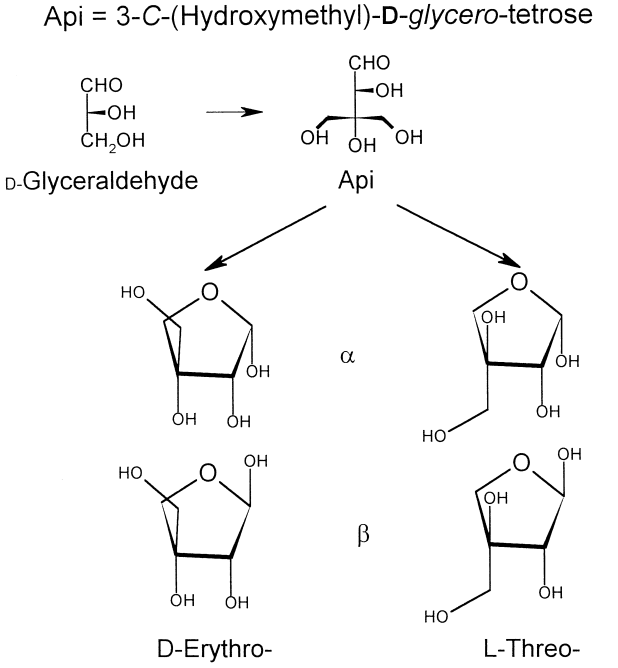


Fig. 3. Schematic representation of Apiose in the open chain form along with the two possible ring forms.

The (Q,ϕ) adiabatic maps of β-D-Api*f* and β-L-AceA*f* are displayed on Figs. 4 and 5, respectively. Two low energy regions, called North (N) and South (S) can be located on these maps. The accessible area with the lowest energy is located on the South domain of each map. Three pseudorotamers: $^4T_3$, E_3 and $^2T_3$ are enclosed within the 1 kcal/mol limit. After unconstrained optimization, all these structures converge to the E_3 shape; it is located at $Q = 0.36$ and $\phi = 107°$ coordinates for Api*f* and at $Q = 0.38$ and $\phi = 107°$ for AceA*f*. The North low-energy part of the maps is higher in energy; this narrow accessible area is delimited by the third contour for Api*f*. The only conformer is $^3E_2$ and after unconstrained optimization, this conformer converges to the same shape with $Q = 0.36$ and $\phi = 287°$. The North minimum is 2.6 kcal/mol higher in energy than the South one. Both minima of Api*f* are displayed on Fig. 4. The location of the North low-energy conformations of AceA*f* is quite different. This domain encloses the ^2E, $^2T_3$, E_3, $^4T_3$, E_4 and $^4T_0$ segment of the pseudorotational itinerary. After optimization, the minimum adopts the E_3 shape, it is located at $Q = 0.38$ and $\phi = 287.1°$ coordinates. The North minimum is, in relative energy, at 0.64 kcal/mol as compared to the South minimum. Both minima are displayed on Fig. 5.

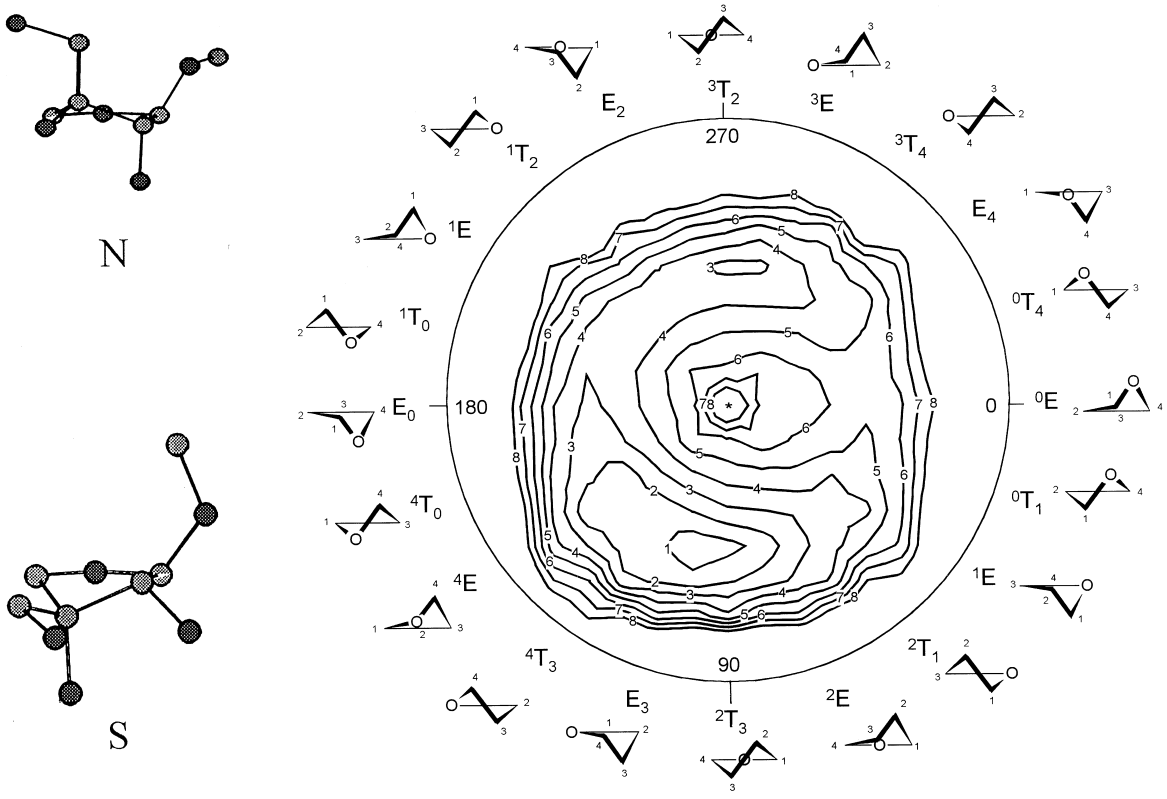


Fig. 4. Pseudorotational wheel of apiose ring encompasses the 20 twist and envelope shapes. Adiabatic pseudorotational energy surface of Apiose as a function of the puckering parameters Q and ϕ. Isoenergy contours are drawn by interpolation of 1 kcal/mol, with respect to the absolute minimum along with a representation of the low-energy conformers. The molecular drawings of the low energy conformers North, and South, are presented.

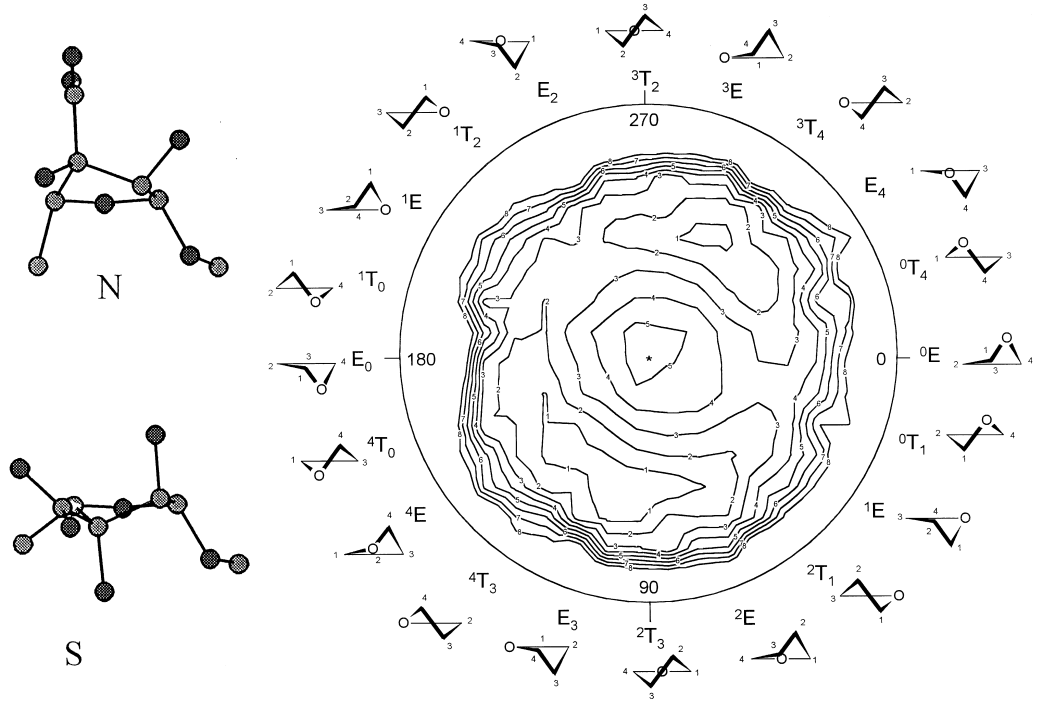


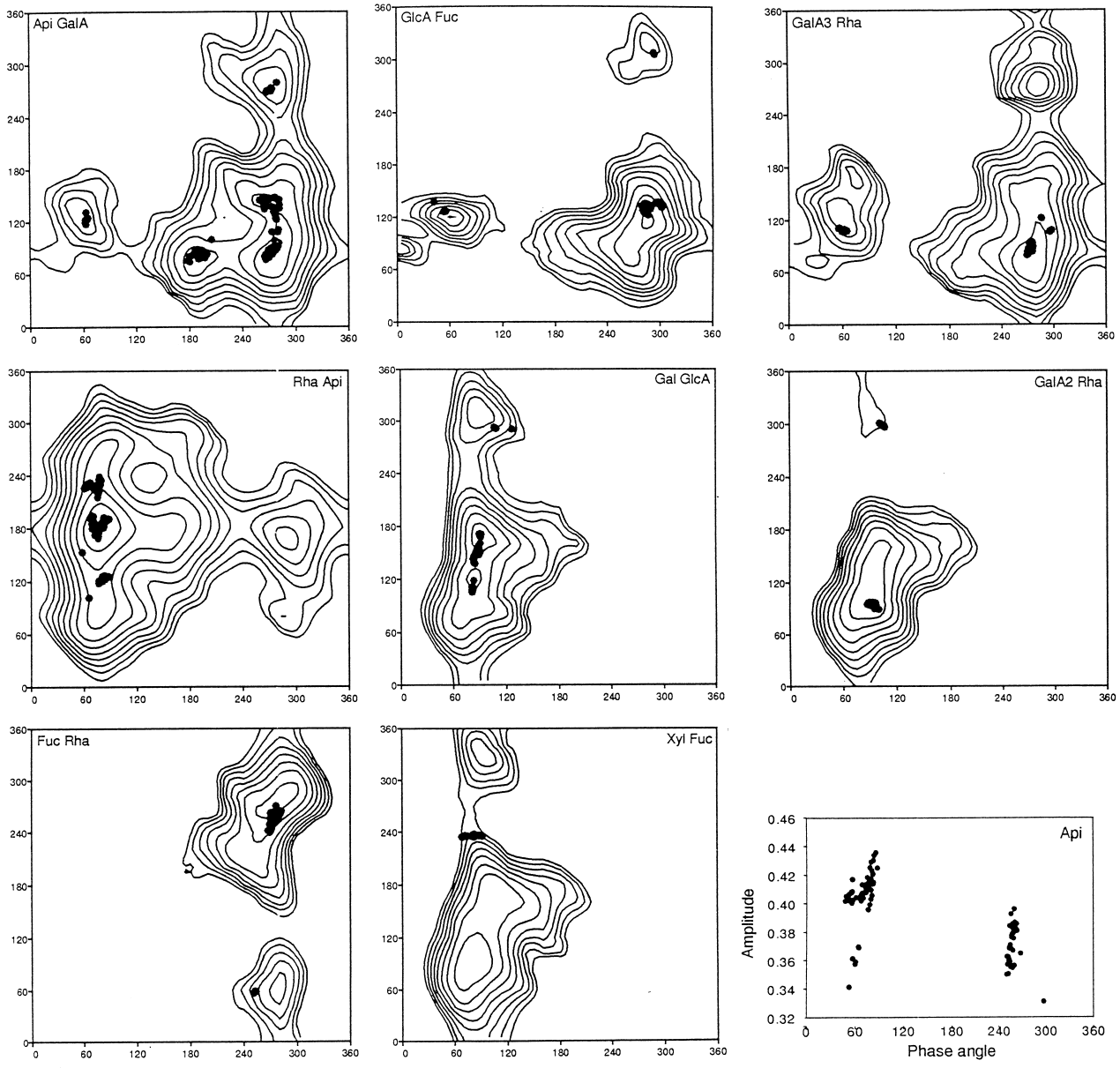
Fig. 5. Pseudorotational wheel of aceric acid ring encompasses the 20 twist and envelope shapes. Adiabatic pseudorotational energy surface of aceric acid as a function of the puckering parameters Q and ϕ. Isoenergy contours are drawn by interpolation of 1 kcal/mol. with respect to the absolute minimum along with a representation of the low–energy conformers. The molecular drawings of the low energy conformers North, and South, are presented.

As is generally observed, furan rings such as aceric acid and apiose adopt two main conformations located in both North and South part of the pseudorotational wheel [34]. A delicate balance between several effects determines the populations of these conformations. In the first, the glycosidic C-1–O bond would prefer a *quasi*-axial orientation as predicted by the anomeric effect. This criterion is satisfied by the E_2 and, to a lesser extent by the E_3 conformations of Api*f* (anomer β, configuration D-erythro type). The aceric acid unit has been investigated as the β anomer of the L configuration. The axial or *quasi*-axial orientation of the glycosidic oxygen is, in this case obtained by the ^3E or $^3T_4$ of the North part of the pseudorotational wheel and by the E_3 or $^4T_3$ of the South one. Contrary to the glycosidic oxygen, the other ring substituants would adopt a *quasi*-equatorial disposition. On the C-2, C-3 and C-4 carbon centers, steric interactions are dominant; the preferred puckerings are the ones that avoid the 1,3 syn-diaxial steric clashes between bulky side chains. The lowest energy conformer of Api*f* has both the hydroxymethyl group and the secondary hydroxyl O-2–H axial; the situation is reversed for the second conformation in which O-3–H is the only side chain in the axial orientation. As for AceA*f*, the lowest energy conformer has the carboxylic group and the methyl group in the equatorial orientation. Both secondary hydroxyl groups are axial while in the second minimum this situation is reversed: carboxylic and methyl groups in the axial orientation and both secondary hydroxyl groups are equatorial.

The puckering energy surface of β-L-Ara*f* has already been reported [24]. It has been shown that two minima do occur on this surface. The lowest energy conformer being the North one, it has a $^3T_2$ shape with puckering parameters of $Q=0.41$ and $\phi=271.1°$. The South minimum adopts a $^2T_3$ conformation with $Q=0.36$ and $\phi=60.5°$ and its energy is 1.3 kcal/mol higher than the North one.

Potential energy surfaces of the disaccharides.— The set of the 18 adiabatic maps that have been computed for all the disaccharide segments constitutes the database underlying all the structural features of RG-II. None of these adiabatic maps had been described previously. They provide a graphical description of potential energy changes as a function of the relative orientation of the contiguous units at the glycosidic linkage. They also indicate the shape and the position of the energy wells along with the location of the energy minimum. The routes for inter-conversion between conformers and the heights of the transition states are also indicated. For each disaccharide, the adiabatic energy surface, along with the locations of the local energy minima and drawings of the conformation of each local minimum located on the energy maps have been established. They are presented in a graphical format, in our website at http://www.cermav.cnrs.fr. A somewhat more limited presentation of the data is given in Fig. 6 for the potential energy surfaces and in Table 2 for the locations of the glycosidic linkages and the puckering of the furanoside constituents.

The MM3 force field is known to correctly reproduce the exo-anomeric effect occurring in the saccharidic structures discussed herein. This effect favors the *gauche* orientations of the aglycone of the glycosidic C-1–O-1 bond (torsion angle Φ). Molecular orbital calculations on 2-methoxyoxane [29], a model compound of carbohydrates, have shown that the energy profile of the rotation around the C-1–O-1 bond of the β anomer have three minima. The lowest energy conformer being at Φ value of 300° whereas minor conformations being at 60 and 180°. The energy profile of the α configuration of 2-methoxyoxane is quite different; it shows a minimum at Φ value of 60° and another minimum at 180°, higher in energy. The conformational potential of the Φ torsion angle is strongly influenced by stereoelectronic effects. On the contrary to this angle, the rotational profile of the torsion angle Ψ depends mainly on steric constraints. A straightforward consequence is the following: The conformational preference of all disaccharides whose non-reducing end is a β-D or α-L pyran (the glycosidic C-1–O-1 bond is then in the equatorial position) is *synclinal*. For the same reason the favoured orientation of the Φ angle of all disaccharides having the non-reducing pyran monomer in the α-D or β-L configuration (axial glycosidic bond) is $+$*synclinal*. Almost all structures obey this general rule. In two particular cases the corresponding potential energy surfaces exhibit unusual Φ values for the low energy conformers. The first concerns the disaccharides that contain furan rings, on the non-reducing end. The five-membered ring flexibility induces strong changes in the orientation of the C-1–O-1 bond. This bond does not have any exact axial or equatorial character. The second case concerns disaccharides having Kdo*p* and Dha*p* at the non-reducing end. In these monomers the presence of a carboxylic group on



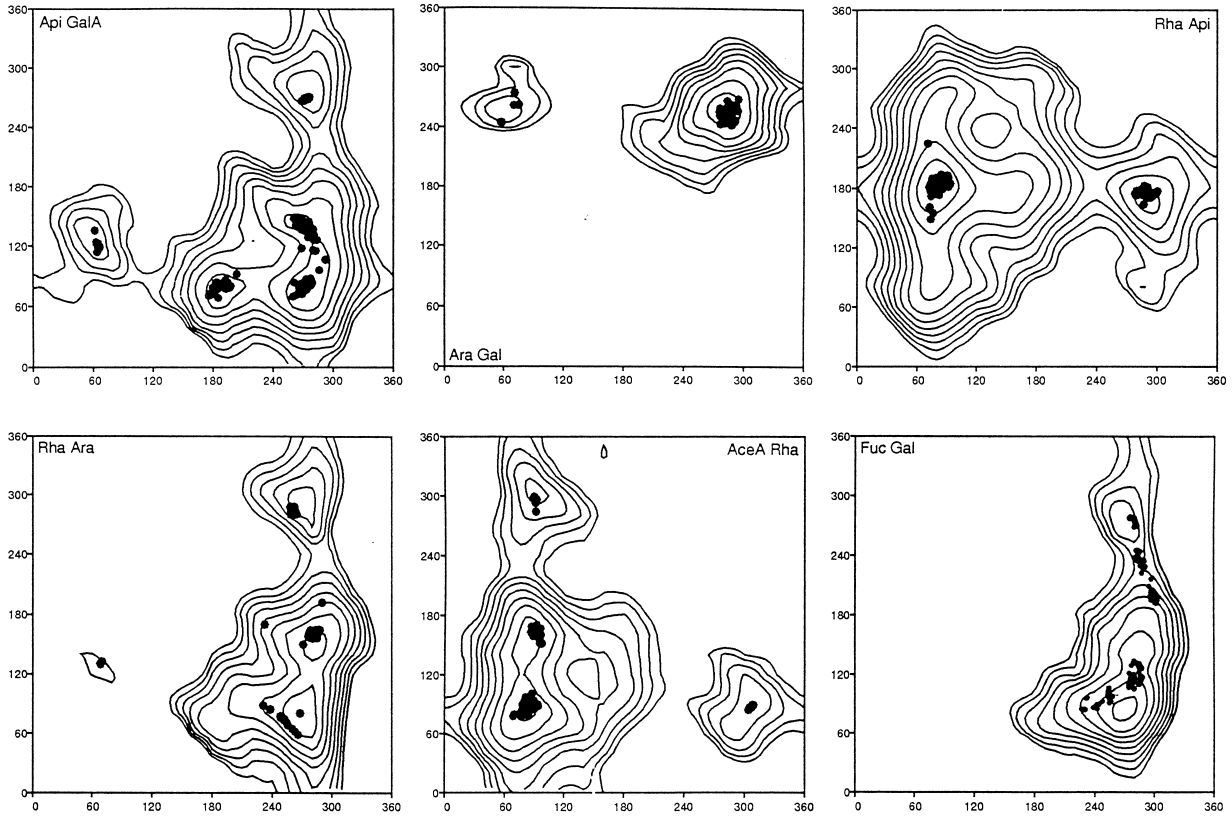
(a)

Fig. 6. Adiabatic potential energy surfaces of all the disaccharide segments found in the four branches of RG-II, as a function of the Φ and Ψ torsion angles. The isoenergy contours are drawn with interpolation of 1 kcal/mol above the minimum of each map. The dots correspond to conformations occurring at each linkage when the disaccharide moiety belongs to one of the four branches of RG-II (see text for explanation). (a) The A branch; (b) The B branch; (c) The C branch; (d) The D branch.

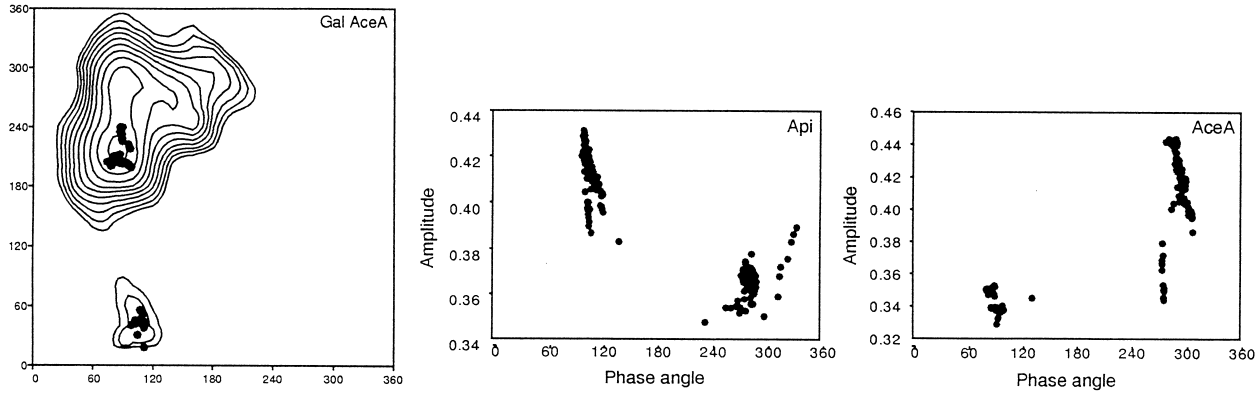
the anomeric center can perturb the 'normal' rotational profile of the Φ torsion angle.

In agreement with the exo-anomeric effect, the calculated Φ, Ψ potential energy surfaces of most of the disaccharides have the lowest-energy area at Φ values corresponding to the *synclinal* orientation (+60° or −60°). The disaccharides RGII-03, RGII-07, RGII-05, RGII-11, RGII-12, RGII-13 and RGII-16 whose non-reducing monomer is β-D-GalA*p*, β-D-GlcA*p*, α-L-Fuc*p*, α-L-Ara*p* and α-L-Rha*p*, respectively, favor the *−synclinal* orientation of the Φ angle (−60°); minor conformations display the *+synclinal* domain (+60°). As for the disaccharides RGII-04, RGII-06, RGII-08, RGII-10, and RGII-02 whose non-reducing monomer are either α-D-GalA*p*, α-D-Xyl*p*, α-D-Gal*p* or β-L-Rha*p*, they favor the *+synclinal* orientation of the Φ angle (+60°).

The Φ, Ψ map of RGII-15, having Kdo*p* as non reducing end shows a very large accessible surface



(b)

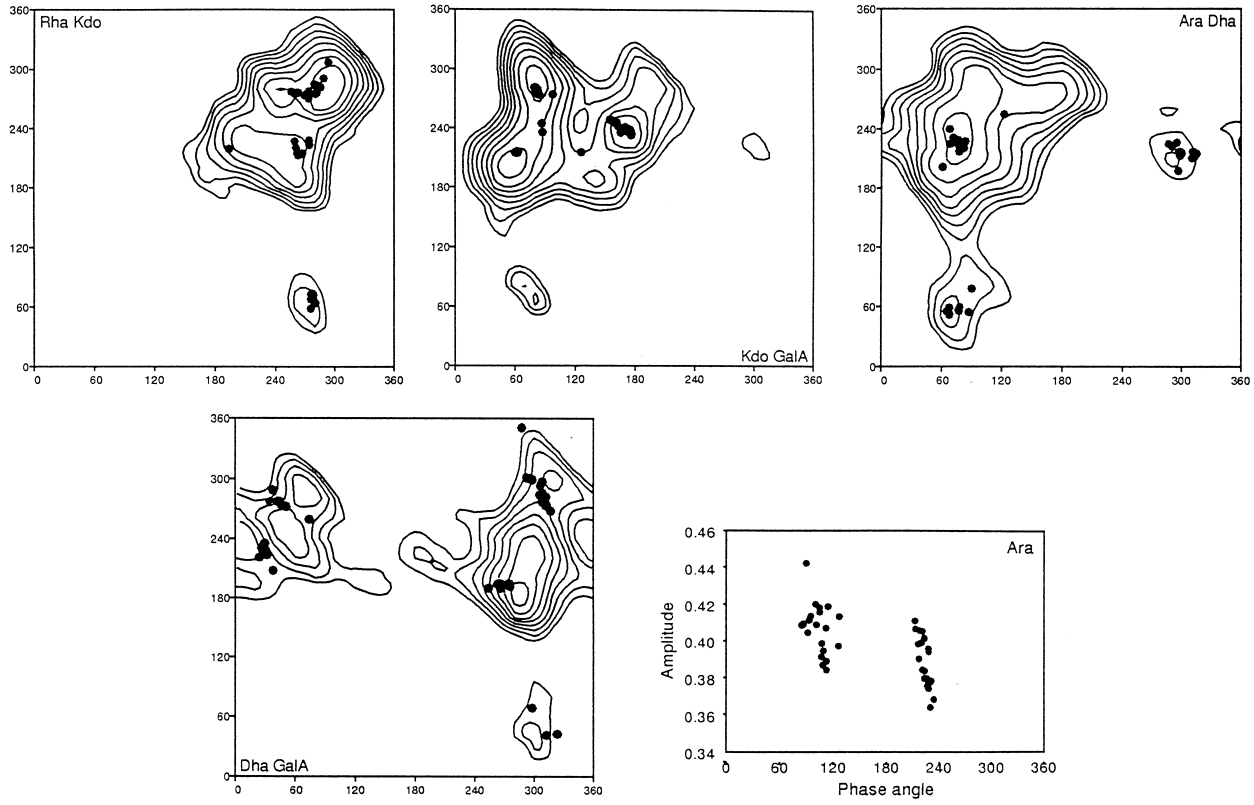


(c)

Fig. 6.—*continued*

extended in both Φ and Ψ directions. The two lowest energy conformers have the *+synclinal* orientation of the Φ torsion angle; this is an expected behavior for a α-D–type molecule. However, other stable conformers do occur with the *antiperiplanar* orientation of the Φ angle. The effect of the carboxylic group at the anomeric position on the conformational properties of the glycosidic bond is less pronounced for RGII-17 having Dha*p* at non-reducing end. The corresponding potential energy surface is very extended along the Φ direction, reflecting the influence of the carboxylic group. However, the *antiperiplanar* orientation of the Φ angle while being 10 kcal/mol lower in energy does not correspond to an energy minimum. Furthermore, low-energy conformers that span the *−synclinal* domain of the Φ torsion angle are dispersed, having values of 280, 300 and 320°.

The low-energy areas of the potential energy surfaces of disaccharide segments containing furan



(d)

Fig. 6.—*continued*

rings at the non-reducing end position, i.e., RGII-01, RGII-09, RGII-14 and RGII-18 show unexpected behaviour. The presence of the very flexible furan ring results in a global increase of the total accessible space. This indicates a high flexibility for these disaccharides. The map of RGII-01 shows that the flexibility along both Φ and Ψ torsion angles is not restricted. A low-energy region can be found at Φ values of 180°. The potential energy surfaces of RGII-09, RGII-14 and RGII-18 have the lowest energy region located within the *+synclinal* domain of the Φ angle; this region is extended towards the *antiperiplanar* domain of this angle.

The rotational freedom around the second glycosidic bond, represented by the torsion angle Ψ, is in general mostly influenced by steric interactions. In the disaccharides RGII-01, RGII-03, RGII-06, RGII-08, RGII-09, RGII-11, and RGII-13, the Ψ angle can span the whole angular range within an energy barrier of 10 kcal/mol suggesting a greater conformational freedom than that of the Φ torsion angle. In general, the two *gauche* or all three staggered (*gauche* and *trans*) orientations correspond to energy minima. The structures RGII-02, RGII-04, RGII-05, RGII-07, RGII-10, RGII-14, RGII-15, RGII-16, RGII-17, and RGII-18 have energy barriers between the different staggered positions of the Ψ torsion angle slightly higher than the upper iso-contour limit. Some rare cases such as RGII-12 in which Ψ adopts only one position, are also found.

A quantitative estimate of the flexibility of the different disaccharides is given by the calculated partition function (**q**) whose values are reported in Table 2. For RGII-12, **q** had a value of 13.5 deg^2; it is the less flexible disaccharide. Whereas for the most flexible one, **q** reaches a value of 35.7 deg^2 (RGII-01). As was previously suggested, all the disaccharides that contain a furanose ring on the non-reducing end position have large values of **q**, indicating the conformational flexibility of these segments.

The results presented here show that the disaccharides have different conformational flexibility depending on the monomeric units and on the linkage type. From the different low-energy area of each potential energy surface, it is possible to extract and characterize the conformational features of the

Table 2
Torsion angle values (°) across the glycosidic bonds, puckering parameters (Å, °), MM3 energies (kcal/mol) and calculated partition function (deg^2) of the different minima of all the disaccharides

	Φ	Ψ	ω	Q	ϕ	E	q
RGII01: β-D-Api*f*-(1→2)-α-D-GalA*p*							35.7
RGII01-1	280	80		0.40	116	33.3	
RGII01-2	260	140		0.41	101	33.6	
RGII01-3	180	80		0.41	107	34.6	
RGII01-4	280	280		0.41	111	36.1	
RGII01-5	60	120		0.37	109	37.3	
RGII02: β-L-Rha*p*-(1→3′)-β-D-Api*f*							25.9
RGII02-1	80	180	180	0.41	111.05	29.7	
RGII02-2	300	160	180	0.40	114.39	34.6	
RGII03: β-D-GalA*p*-(1→3)-β-L-Rha*p*							31.6
RGII03-1	280	60				24.2	
RGII03-2	280	280				25.1	
RGII03-3	60	120				27.2	
RGII03-4	80	160				27.8	
RGII04: α-D-GalA*p*(1→2)-β-L-Rha*p*							18.4
RGII04-1	80	100				30.1	
RGII04-2	100	300				37.4	
RGII05: α-L-Fuc*p*-(1→4)-β-L-Rha*p*							20.4
RGII05-1	280	280				25.6	
RGII05-2	280	60				28.5	
RGII06: 2-*O*-methyl-α-D-Xyl*p*-(1→3)- α-L-Fuc*p*							25.3
RGII06-1	80	100				27.3	
RGII06-2	100	320				30.5	
RGII06-3	280	140				35.9	
RGII07: β-D-GlcA*p*-(1→4)-α-L-Fuc*p*							20.3
RGII07-1	280	120				24.0	
RGII07-2	60	120				27.5	
RGII07-3	280	320				29.7	
RGII08: α-D-Gal*p*-(1→2)-β-D-GlcA*p*							25.5
RGII08-1	80	120				25.7	
RGII08-2	100	160				25.7	
RGII08-3	100	300				28.6	
RGII09: β-L-AceA*f*-(1→3)-β-L-Rha*p*							28.4
RGII09-1	80	80		0.91	182	28.8	
RGII09-2	80	160		0.87	183	29.5	
RGII09-3	100	300		0.84	184	32.8	
RGII09-4	300	100		0.90	182	33.4	
RGII10: α-D-Gal*p*-(1→2)-β-L-AceA*f*							18.0
RGII10-1	80	220		0.41	93	11.7	
RGII10-2	100	60		0.37	82	19.5	
RGII11: α-L-Fuc*p*2OMe-(1→2)-α-D-Gal*p*							18.0
RGII11-1	260	80				25.9	
RGII11-2	260	280				30.4	
RGII12: α-L-Ara*p*-(1→4)-α-D-Gal*p*							13.5
RGII12-1	280	260				21.3	
RGII12-2	60	260				26.3	
RGII13: α-L-Rha*p*-(1→2)-α-L-Ara*p*							28.2
RGII13-1	260	80				22.4	
RGII13-2	280	140				22.7	
RGII13-3	260	300				25.8	
RGII13-4	60	120				29.7	
RGII14: β-L-Ara*f*-(1→2)-α-L-Rha*p*							21.5
RGII14-1	80	80		0.41	219	29.2	
RGII14-2	80	320		0.40	224	35.2	
RGII14-3	320	100		0.40	228	36.0	

continued

Table 2—*contd*

	Φ	Ψ	ω	*Q*	ϕ	*E*	*q*
RGII15: Kdo*p*-(2→3)-α-D-Gal*p*							21.4
RGII15-1	60	200				32.4	
RGII15-2	80	280				32.5	
RGII15-3	160	240				34.3	
RGII15-4	180	220				34.9	
RGII15-5	140	180				36.0	
RGII15-6	80	60				38.0	
RGII15-7	300	220				39.2	
RGII16: α-L-Rha*p*-(1→5)-Kdo*p*							18.5
RGII16-1	300	280				31.2	
RGII16-2	260	280				32.6	
RGII16-3	260	220				33.2	
RGII16-4	280	60				37.5	
RGII17: β-D-Dha*p*-(2→3)-α-D-GalA*p*							28.4
RGII17-1	280	180				27.5	
RGII17-2	60	240				31.2	
RGII17-3	80	280				31.4	
RGII17-4	320	300				31.5	
RGII17-5	300	40				33.6	
RGII18: β-L-Ara*f*-(1→5)--D-Dha*p*							28.4
RGII18-1	80	240		0.42	214	31.6	
RGII18-2	60	60		0.40	227	36.9	
RGII18-3	300	220		0.41	219	38.4	

energy minima. Selected geometrical parameters along with the energies of the different energy minima of all disaccharides are presented in Table 2. Some minima are stabilized by inter-residue hydrogen bonds between the different hydroxyl groups. In these cases, intramolecular hydrogen bonds may be formed as a further stabilization of an already existing low energy conformation. This is not a general feature because we have used a large dielectric constant with the aim to simulate a hydration effect.

Conformational analysis of the A-D fragments of RG-II.—The different geometries of the minima were combined in order to generate models of the complete structure of each fragment. These fragments were then subjected to a complete energy minimization. One should note that, while poorly represented if not absent from the equilibrium mixture, all secondary minima of the disaccharides were considered, as hydrogen bonding or stacking interactions between distant residues can be created because of favorable folding (i.e., long range interactions). For the same reason the North conformer of Api*f* was considered. Fragments A and B are very large and a simple combinatorial analysis predicts that the number of conformers that have to be studied is 25,920 and 15,360 for the fragment A and B, respectively. All these conformers were studied.

The C and D fragments are composed of two residues each. They were studied as side chains of the main (1→4)-linked α-D-GalA*p* backbone. Two regular conformations corresponding to the 2_1 and 3_1 helicoidal conformations were considered for the backbone. Such a choice is consistent with X-ray diffraction data in the solid state and NMR studies in solution [32,33]. Based on this necessary simplification, 784 and 225 geometry combinations for fragments C and D, respectively, were investigated.

To present the results, those Φ, Ψ data, calculated for each fragment, with relative energy of less than 10 kcal/mol have been projected onto each map of the corresponding disaccharides. They are represented as dots on Fig. 6, along with the puckering parameters of the five membered rings. A molecular drawing of the lowest energy conformer of each fragment is given in Fig. 7.

The projection of the calculated data onto the maps of the corresponding parent disaccharides makes it clear that most of the minima, principal and/or secondary, are sampled by the glycosidic bonds of each fragment, suggesting a general conformational flexibility. However, the oligomer effect may cause constraints that induce slight deviations away from the exact Φ, Ψ position of the energy minima that were calculated for the disaccharides.

Within an energy window of 10 kcal/mol, 108, 240, 31 and eight different conformers have been

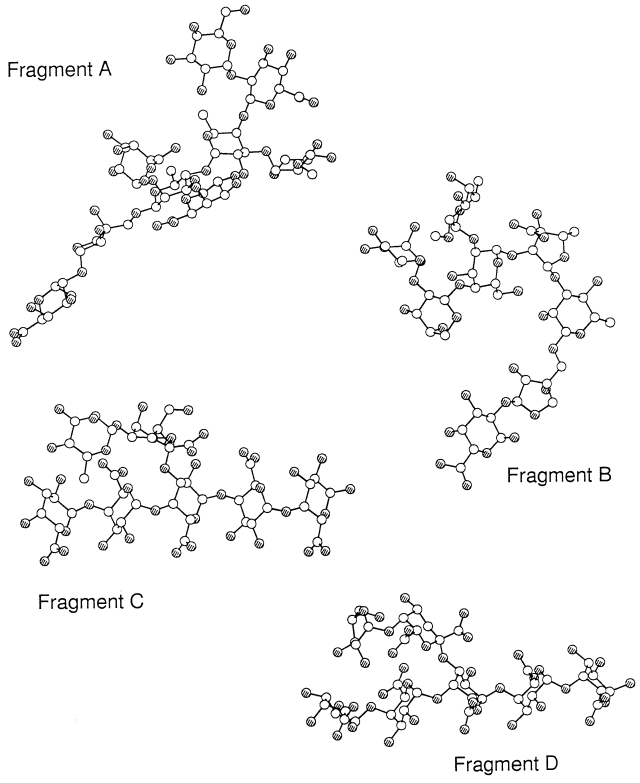


Fig. 7. Molecular drawings of the lowest energy conformer of the four different fragments.

calculated for fragment A, B, C and D, respectively. This substantial reduction in the number of conformations suggests first that all the possible combinations of the minima of the parent disaccharides are not compatible with the oligomerization (within a reasonable energy value). Secondly, this also suggests that all fragments are flexible and can adopt many different shapes. Not surprisingly, the B fragment is more flexible than the others are. A reasonable explanation is that the fragment B is almost linear (except one branch) and in addition the central furan L-AceA*f* residue enhances the conformational freedom of this chain. On the contrary, fragment A is heavily branched: the central L-Rha*p* residue is tetrasubstituted and the neighboring L-Fuc*p* sugar is trisubstituted. Steric hindrance of bulky monomers therefore prevents some combinations to be created within a reasonable internal energy window. Fragments C and D are composed of two side chain residues each; they cannot adopt a wide range of different shapes as compared to fragments A and B.

One has to note that the junction zone of both A and B fragments is constituted by the same trisaccharide: β-L-Rha*p*-(1→3′)-β-D-Api*f*-(1→2)-α-D-GalA*p*. From the disaccharide maps, it is evident that these two glycosidic bonds display a large conformational flexibility not only because the upper 10 kcal/mol limit covers a very large Φ, Ψ surface of each map, but also because the lowest energy contours are either numerous (five different regions in the RGII–01 map) or even very extended (the third contour covers 200 and 130° range on Φ and Ψ direction, respectively, in the RGII-02 map). The calculated **q** values are of 25.9 and 35.7 deg^2 for RGII-02 and RGII-01 disaccharides, respectively. In addition, the central D-Api*f* residue has two distinct minima and the three rotatable bonds between L-Rha*p* and D-Api*f* residues enhance the flexibility of the trimer. All the five minima that were calculated for the RGII-01 disaccharide are sampled by the oligomers A and B. Moreover, most of the calculated torsion angles of the A and B oligomers are located within the first energy iso-contour of the potential energy surface of the corresponding dimmer. This indicates that the conformational freedom of this glycosidic bond is not influenced by the structure as a whole. Such a behaviour might be due to its end-chain location. The calculated data reveal that the lowest energy conformations do occur within the two lowest energy areas. Departure from these areas is easy for the B fragment; it results in an energy loss of 4 kcal/mol and 1.5 kcal/mol for fragments A and B, respectively. This glycosidic bond is flexible about the Φ angle. Whereas only the lowest energy area is visited by the glycosidic bond between residues L-Rha*p* and D-Api*f* in fragment A (Fig. 6(a)) (which is mostly limited by the Φ torsion angle) large variations of the Ψ angle give rise to three different minima. In fragment B, both minima of the RGII-02 are sampled, suggesting here again a greater conformational flexibility of this fragment with respect to the A one. The torsion angle ω, linking the L-Rha*p* and the D-Api*f* residues also reflects the difference in flexibility of both fragments. In the fragment A, this angle adopts only one of the three staggered orientations. However, calculated data covers a broad range of values: from 48 to 138°. In the fragment B, all three possible positions are sampled by this torsion angle and the three lowest energy conformers of this fragment are differentiated by the values of this torsion angle. Within the 10 kcal/mol limit, the *+gauche* orientation is in the 49 to 68° range; the *−gauche* one between 291 and 324° and the trans orientation is between 177 and 197°. D-Api*f* residues of both fragments are capable of exploring

the two North and South conformations. In the fragment A, the puckering amplitude shows small variations ($\delta Q = 0.09$ at max) whereas large variations are observed for the phase angle ϕ. The phase angle of the South minimum of D-Api*f* varies within 45 and 42° for fragments A and B, respectively. The phase angle of the North minimum of D-Api*f* varies within 10 and 99° for fragments A and B, respectively.

The remaining glycosidic bonds of fragment A are very constrained. For example, the Φ, Ψ values of the glycosidic bond linking 2-O-Me-Xyl*p* to the L-Fuc*p* residues are all located on the conformational barrier between the two minima that were calculated for the model disaccharide (RGII-06). This part of the molecule has one well-defined low-energy shape and the folding is stabilized by numerous hydrogen bonds between distant residues. Departure from this folding implies breaking this hydrogen-bonding scheme, which would result in a loss of the stabilization.

Fragments B, C and D are extremely flexible over the whole structure as demonstrated by the large number of energetically equivalent conformers. In the case of fragment B, all glycosidic bonds but the one between L-Ara*f* and D-Gal*p* residues have at least two competitive conformations. This is supported by the flexibility of the L-AceA*f* five membered ring in which the phase angle displays a broad range of values around both N (33°) and S (51°) puckerings.

4. Conclusions

The conformational study of two rare furanosyl monosaccharides and of 18 disaccharide fragments, which constitute the RG-II mega-oligosaccharide, has been presented. Along with previous knowledge about the conformation of the other 10 monosaccharides, the present set of information constitutes the data bank required to characterize the structural features and the dynamic of RG-II. Precise information about the low energy domains, along with the description of the features displayed by all the local minima have been provided. A net consequence of these results is the fairly high number of stable conformers for each disaccharide. The construction of the four oligomeric fragments that constitute the side chains of RG-II has been completed. The many possible structures have been explored and submitted to an extensive energy minimization. Even though there exists a substantial reduction of the number of viable conformations for each fragment, their number amounts to 108, for A, 240 for B, 31 for C and 8 for D. Nevertheless, some generalized pictures have emerged from the present study:

(a) The L-Rhap–D-Api*f*–D-GalA*p* junction sequence of fragment A and B is extremely extended in its lowest energy conformation.
(b) Both hydroxyl groups of the central D-Api*f* residue protrude from the chain.
(c) In fragment A, the remaining structure is branched and this causes a sharpening of the energy surfaces of the glycosidic linkages and results in a conformationally well-defined and quite rigid structure.
(d) This particular folding of the fragment A is stabilized by numerous hydrogen bonds that are created between non-neighboring residues.
(e) Fragment B is very flexible, note that this fragment has four methyl groups (2-*O*-MeFuc*p*, L-AceA*f* and two L-Rha*p* residues). In the lowest energy conformer all these hydrophobic substituants are located at the surface of the molecule. This strongly hydrophobic surface may play a role in the boron complexation.
(f) The lowest energy minima of fragments C and D extend along the longitudinal axis of the (1→4)-linked α-D-GalpA*p* backbone, towards the non-reducing end.
(g) Both fragments are stabilized by hydrogen bonding and by hydrophobic contacts between the two residues of the chain and the backbone. As for the fragment C, the conformation of the glycosidic bond of the branching point (Kdo*p*–D-GalA*p*) is stabilized by hydrogen bonds between the carboxylic group of Kdo*p* and two contiguous D-GalA*p* residues. The L–Rha*p* residue is bridged with another D-GalA*p*. The hydrogen-bonding interactions that stabilize the lowest energy conformation of fragment D involve the terminal L-Ara*f* residue and galacturonic acid residues of the backbone.

In the absence of further information about the primary structure of RG-II it would be hopeless to envisage building up all the low energy conformers that can be displayed for either the monomer or the dimmer. Nevertheless, some results from the

present work indicate that the lowest energy conformers of fragment A and B all exhibit a good exposure of the hydroxyl groups of the D-Api residues. Therefore, despite the relatively buried location of these apiosyl residues at the vicinity of the pectic backbone, there exists an exposure of their two key hydroxyl groups which may be compatible with the stereochemical requirements for making a borate-diol ester formation.

5. Supplementary material

The set of the 18 adiabatic maps that have been computed for all the disaccharide segments found in RG-II is available at the following address: http://www.cermav.cnrs.fr/databank/RG2. A schematic representation of each disaccharide is provided, along with the geographical location of each local minimum in the potential energy surface and molecular drawings of the corresponding conformation.

Acknowledgements

The authors are very much indebted to Drs. Patrice Pellerin and Catherine Hervé du Penhoat for simulating discussions and careful reading of the manuscript.

References

[1] A.G. Darvill, M. McNeil, and P. Albersheim, *Plant Physiol*., 62 (1978) 418–422.
[2] J.R.Thomas, M. McNeil, A.G. Darvill, and P. Albersheim, *Plant physiol*., 83 (1987) 659–671.
[3] J.R. Thomas, A.G. Darvill, and P. Albersheim, *Carb. Res*., 185 (1989) 261–277.
[4] S. Ishii, *Phytochemistry*, 21 (1982) 778–780.
[5] R.J. Redgwell, L.D. Melton, and D.J. Brasch, *Carb. Res*., 182 (1988) 241–258.
[6] R.J. Redgwell, L.D. Melton, D.J. Brasch, and J.M. Coddington, *Carb. Res*., 226 (1992) 287–302.
[7] T. Doco, J.M. Brillouet *Carb. Res*., 243 (1993) 333–343.
[8] P. Pellerin, T. Doco, S. Vidal, P. Williams, J.M. Brillouet, and M.A. O'Neill, *Carb. Res*., 290 (1996) 183–197.
[9] M.W. Spellman, M. McNeil, A.G. Darvill, P. Albersheim, and K. Henrick, *Carb. Res*., 122 (1983) 115–129.
[10] W.S. York, A.G. Darvill, M. McNeil, and P. Albersheim, *Carb. Res*., 138, (1985), 109–126.
[11] T.T. Stevenson, A.G. Darvill, and P. Albersheim, *Carb. Res., 182 (1988) 207–226*.
[12] V. Puvanesarajah, A.G. Darvill, and P. Albersheim, *Carb. Res*., 218 (1991) 211–222.
[13] M. Kobayashi, T. Matoh, and J.I. Azuma, *Plant Physiol*. 110 (1996) 1017–1020.
[14] T. Matoh, S. Kawaguchi, and M. Kobayashi, *Plant Cell. Physiol*. 37 (1996) 636–640.
[15] B.J. Shelp, in U.C. Gupta (Ed.), *Boron and its role in crop production*, CRC press, Boca Raton, 1993, pp 53–85.
[16] K.S. Shin, H. Kiyohara, T. Matsumoto, and H. Yamada, *Carb. Res*., 300 (1997) 239–249.
[17] IUPAC-IUB Joint Commission on Biochemical Nomenclature, *Eur. J. Biochem*. 131 (1983) 5–7.
[18] S. Pérez and M.M. Delage, *Carbohydr. Res*., 212 (1991) 253–259.
[19] N.L. Allinger, Y.H. Yuh, and J.-H. Lii, *J. Am. Chem. Soc*., 111 (1989) 8551–8566.
[20] N.L. Allinger, M. Rhaman, and J.-H. Lii, *J. Am. Chem. Soc*., 112 (1990) 120–140.
[21] V. Mikol, P. Kosma, and H. Brade, *Carb. Res*., 263 (1994) 35–42.
[22] C. Kratky, D. Stix, and F.M. Unger, *Carb. Res*., 92 (1981) 299–304.
[23] K. Bock, J.U. Thomsen, P. Kosma, R. Christian, O. Holst, and H. Brade, *Carb. Res*., 229 (1992) 213–224.
[24] S. Cros, C. Hervé du Penhoat, S. Pérez, and A. Imberty, *Carb. Res*., 248 (1993) 81–93.
[25] D. Cremer and J.A. Popple, *J. Am. Chem. Soc*., 97 (1975) 1354–1358.
[26] D. Cremer and J.A. Popple, *J. Am. Chem. Soc*., 97 (1975) 1358–1367.
[27] A.D. French and V. Tran, *Biopolymers*, 29 (1990) 1599–1611.
[28] S. Pérez, C. Meyer, A. Imberty, and A.D. French, in M. Mathlouthi, J.A. Kanters and G.G. Birch (Eds.) Sweet Taste Chemoreception, Elsevier, London, 1993, pp 57–73.
[29] I. Tvaroska and T. Bleha, *Adv. in Carb. Chem. and Biochem*., 47 (1989) 45–123.
[30] S.B. Engelsen, S. Cros, W. Mackie, and S. Pérez, *Biopolymers*, 39 (1996) 417–433.
[31] S. Pérez, M. Kouwijzer, K. Mazeau, and S.B. Engelsen, *J. Mol. Graphics*, 14 (1996) 307–321.
[32] C. Gouvion, K. Mazeau, A. Heyraud, F.R. Taravel, and I. Tvaroska, *Carbohydr. Res*., 261 (1994) 187–202.
[33] S. Cros, C. Hervé du Penhoat, N. Bouchemal, H. Ohassan, A. Imberty, and S. Pérez, *Int. J. Biol. Macromol*., 14 (1992) 313–320.