Carbohydrate Research, 204 (1990) 27–36
Elsevier Science Publishers B.V., Amsterdam

RAMM — a new procedure for theoretical conformational analysis of carbohydrates*

Tibor Kožár, František Petrák, Zuzana Gálová,
Department of Biophysics, Institute of Experimental Physics, Slovak Academy of Sciences, 043 53 Košice (Czechoslovakia)

and Igor Tvaroška
Institute of Chemistry, Slovak Academy of Sciences, 842 38 Bratislava (Czechoslovakia)

(Received September 22nd, 1989; accepted for publication, January 29th, 1990)

ABSTRACT

The development is described of a new molecular-mechanics procedure which is designed for investigation of the molecular structure and for the prediction of minimum-energy conformations of flexible saccharide systems. In place of the optimization of energy with respect to the Cartesian coordinates of the atoms, used in conventional molecular mechanics, the minimization is carried out on internal geometrical parameters, namely, bond lengths, bond angles, and torsional angles. For this purpose, the non-derivative method of conjugate direction based on the Powell–Zangwill algorithm is used, resulting in a considerable computational advantage for large systems. Another principal feature of the procedure is the algorithm for a determination of the energetically favored orientations of pendant groups, based on a random-walk technique in a given conformation, independent of the initial starting geometry. The performance of the program RAMM using this procedure has been illustrated by comparing (φ, ψ) maps obtained with different approaches utilizing the MM2 set of potential functions for a model disaccharide. It is shown that the shape of the (φ, ψ) maps and the relative energies of conformers are significantly affected by orientation of pendant groups.

INTRODUCTION

Over the previous decade it has been recognized that the conformations (three-dimensional shapes) of carbohydrates partly determine the biological functions of these compounds. Therefore, insight into the conformational properties of carbohydrates is crucial to a complete understanding of their functions in living organisms. Among other structurally oriented methods, conformational-energy calculations have been extensively used to describe the shape of mono-, oligo-, and poly-saccharides[1]. An enormous amount of experimental evidence derived from crystal-structure elucidation of saccharides that has been accumulated over several decades supports the view that, in general, internal geometrical parameters (bond lengths, bond angles, and torsion angles) of saccharides may be divided into rigid monomeric residues and flexible glycosidic linkages. This has led to the major computational approach in use today, namely, calculations of two-dimensional [or three-dimensional, in the case of a (1→6) linkage]

* Presented at EUROCARB V, the Vth European Carbohydrate Symposium, Prague, August 21–25, 1989.

0008-6215/90/$ 03.50

conformational-energy maps with the rigid monosaccharide entities. However, surveys of known crystal structures[2-4] and quantum-chemical calculations[5-7] have revealed considerable alterations occurring in ring geometry with change of orientation about a glycosidic linkage. Therefore, as a more, appropriate approach, calculations have been suggested[8,9] that entail optimization of all internal coordinates, and recently such calculations have been carried out[10]. Although this approach has contributed significant insight into energetics, local stability, and flexibility of carbohydrates, it has encountered some difficulty when applied to the problem of predicting stable conformations. The complication lies in the existence of a very large number of local minima on the multidimensional potential-energy surface of a carbohydrate, especially those connected with pendant groups. This is documented in the case of maltose[8-13], where the orientations of the hydroxymethyl groups are very important parameters governing the shape of the two-dimensional, conformational-energy maps.

This problem has been referred to as the multiple-minima problem in proteins, and it is still one of the major obstacles in that field. Within the past two decades, several algorithms have been proposed to overcome the multiple-minima problem and to locate local minima (see for example ref. 14).

Herein is described a new algorithm which is intended to avoid the influence of the starting orientations of the pendant groups on the location of the minima, and its performance is illustrated on an example of a (1→2)-linked disaccharide.

RESULTS AND DISCUSSION

Description of the RAMM program. — A complete description of the new program package, RAMM (*RA*ndom *M*olecular *M*echanics), which includes molecular mechanics, solvent effects, and the calculation of charges, volumes, and surfaces, will be published elsewhere[15]. Here, we describe only the basic features which differ from those of a standard molecular-mechanics program.

The program uses the bond lengths, bond angles, and torsional angles as input parameters. With these values as the initial parameters, the minimum-energy conformation of a molecule is calculated by minimizing the total energy with respect to the bond lengths, bond angles, and torsion angles. A choice of the number and type of the optimized or fixed geometrical parameters is very simple; it is done by including or removing the parameter from a table. Thus, very different optimizations can be carried out with different constraints. For example, the optimization with two fixed torsion angles requires the fixing of only two variables, whereas the same optimization in Cartesian-coordinate space requires restriction of the movement of 8 atoms (24 Cartesian coordinates).

In the energy minimization, the non-derivative method of conjugate directions with the algorithm of Powell–Zangwill[16,17] is used. Although the minimization algorithm, which uses the first and second derivatives of the energy function with respect to each Cartesian coordinate, converges faster than non-derivative algorithms, the flexibility of the algorithm based on internal coordinates makes it more efficient, especially in

the case of large flexible oligosaccharide molecules, where restricted optimization is required in order to have the problem at manageable dimensions. The same algorithm we have used was very successful in conjunction with the quantum-chemical PCILO method[18-20].

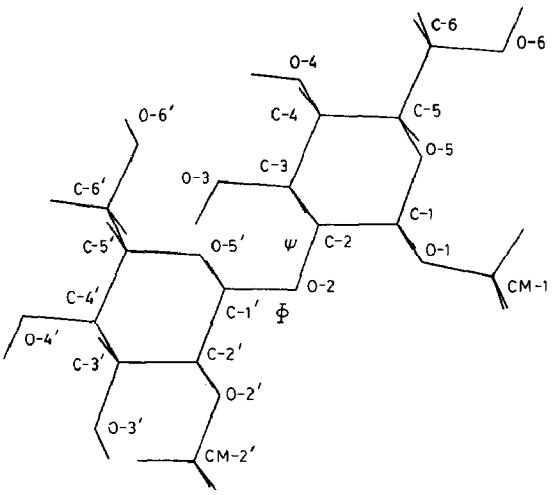
In the present stage, the program uses the potential functions and parameterization of the MM2 (see ref. 21) and MM2CARB (ref. 20) methods. However, other potential functions can be used with very simple modifications of the program.

Conformational analysis of oligosaccharides presents a combinatorial problem. The analysis involving 360-degree rotations in D degree increments about N different bonds, requires the examination of $(360/D)^N$ conformations. The number of conformations increases exponentially with the number of bonds being rotated. For example, if D is 20° and N is 10 (which corresponds to 10 pendant groups in an ordinary disaccharide) then 18^{10} conformations must be examined, but if D is increased to 60 degrees, the number of conformations is still very high (6^{10}) for each combination of φ and ψ. The sheer magnitude of such a problem precludes use of even the fastest molecular-mechanics method. Because of the multi-dimensional problem the optimization in Cartesian coordinates of atoms is insufficiently efficient to find the global minimum, and the optimized structure corresponds to the nearest local minimum of the initial structure. To remedy this shortcoming, the development was undertaken of a so-called random walk (RW) technique to scan the conformational space of pendant groups and to find the conformer having the lowest energy. The RW technique allows the decrease of the colossal number (18^{10}) of conformers to a reasonable quantity and permits the conformational analysis to be conducted on everyday computers (even on personal computers). The RAMM program uses a random walk on torsional angles, with several possible options. It is expected that the combination of RW with consequent minimization of energy on internal coordinates is the most powerful feature of the newly developed RAMM program.

Randomization had already been used for model analysis of polysaccharide structures in the PS79 method of Zugenmaier and Sarko[22]. The starting structures for geometry refinement in the foregoing method are generated in a random manner. The values of the optimized internal coordinates in PS79 are convergent to the preset standard values in the given limits. The PS79 procedure does not solve the multiminima problem of the rotatable, pendant groups, where the limit of the torsion angle χ(i) is from $-180°$ to $+180°$. To find the most advantageous orientation of the pendant groups by PS79, it is necessary to input their corresponding standard values while the values of torsion angles χ(i) and ω are obtained as direct results of the RAMM procedure.

The RAMM methodology has been built up especially for purposes of conformational analysis. Therefore it includes, in addition to the calculation of two-dimensional maps, the possibility of calculating selected multidimensional maps. This option is important for conformational analysis of complex molecules; for example, glycoconjugates. Here, the only restrictions come from the computer-time consumption.

Some features of the RAMM program are illustrated by the conformational



1

Fig. 1. Atom numbering and torsion angles of the molecule of methyl *O*-(2-*O*-methyl-β-D-glucopyranosyl)-(1→2)-α-D-glucopyranoside (**1**).

analysis of a model disaccharide. The results obtained also demonstrate the influence of the orientations of the pendant group on the conformational properties of the given molecule.

Conformational-energy maps of a (1→2)-linked disaccharide. — The labeling of the atoms is shown in Fig. 1, a schematic representation of the disaccharide molecule methyl *O*-(2-*O*-methyl-β-D-glucopyranosyl)-(1→2)-α-D-glucopyranoside (**1**). The orientation at the glycosidic linkage is described by a set of two torsion angles:

$\varphi = \tau$ [C-2′–C-1′–O-2–C-2]

$\psi = \tau$ [C-1′–O-2–C-2–C-1]

The orientation of the hydroxymethyl groups is described by a torsion angle of $\omega = \tau$ [C-4–C-5–C-6–O-6]. The orientations of the hydroxylic hydrogen atoms are characterized by a torsion angle of $\chi(\mathrm{i}) = \tau$ [C_{i-1}–C_i–O_i–H-O_i]. The torsion angles of the methyl groups are defined in the following way: χ(CM-1) = τ [C-2–C-1–O-1–CM-1] and χ(CM-2′) = τ [C-1′–C-2′–O-2′–CM-2′]. The values of torsion angles are referred to as either ±*sc* and *ap*, corresponding to values of ±(60 ±30)° and (180 ±30)° for staggered conformations, or ±*ac* and *sp* of ±(120 ±30)° and (0 ±30)° for eclipsed conformations.

The computational strategy employed was as follows. As the first step, the so-called rigid (φ, ψ) map was calculated by simple rotation around the C-1′–O-2 and O-2–C-2 bonds, the ring geometry being considered invariant. In the next step the best orientation of each of the pendant groups was looked for by applying the RW technique. From the large number of conformers generated for a given value of φ and ψ, that having the lowest energy was selected. Finally, the energies of conformers were minimized for every grid point of the (φ, ψ) map with respect to the torsion angles ω and χ(i).

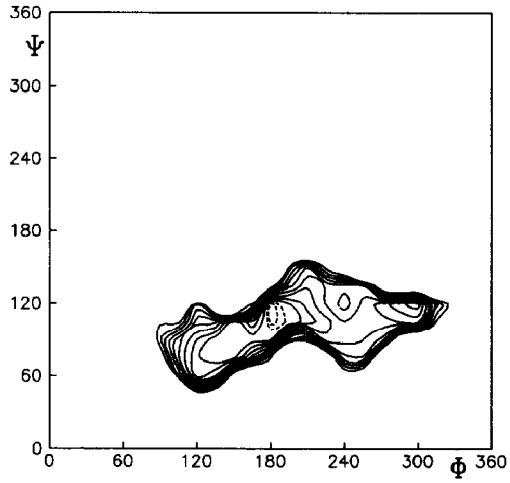


Fig. 2. Conformational-energy map of **1** in rigid geometry. Dashed lines represent the main minimum region (energy levels 15 and 25 kJ/mol). Given with solid lines are contours in 25 kJ/mol intervals from 50 to 250 kJ/mol.

This resulted in the so-called partially relaxed conformational-energy map. For all methods, the starting geometry of D-glucose units was taken from former MM2 optimization of di-O-methylated D-glucose, and the standard MM2 parametrization was used for all of the calculations.

The rigid conformational-energy map shown in Fig. 2 is very restricted. One of the reasons is steric interactions between two methyl groups that are close to the glycosidic linkage.

To compare the (1→2)-disaccharide structure modelled with the experimental data for carbohydrates possessing an equatorial (1→2)-glycosidic linkage and a pyranoid ring in the $^4C_1$ form, the geometries of all available relevant structures were retrieved from the Cambridge Structural Database. Our list did not include erroneous or disordered structures. The structures retrieved and their reference codes are as follows. BUFTIF: methyl 3,4-*O*-isopropylidene-2,6-di-*O*-(2,3,4,6-tetra-*O*-acetyl-*β*-D-galactopyranosyl)-*α*-D-galactopyranoside ethyl acetate solvate[23]; SOPROS: *O*-*β*-D-glucopyranosyl-(1→2)-*α*-D-glucose monohydrate[24] (sophorose). Their (φ, ψ) values are similar: = (−178°,122°); SOPROS = (162°,98°).

The calculated minimum-energy conformer is at $\varphi = 180°$ and $\psi = 100°$, and the corresponding experimental values for SOPROS are $\varphi = 162°$ and $\psi = 98°$. The other calculated minima are at $\varphi = 300°$, $\psi = 120°$, but with the energy 36 kJ/mol higher than the minimum, and at $\varphi = 240°$ and $\psi = 80°$ and relative energy of 67 kJ/mol.

The conformational map obtained using the RW technique is shown in Fig. 3. A net consequence of the RW method is that the reorientation of the pendant groups allows for more conformational states to be reached. In the case of **1** this flexibility generates significant enlargement of the energy domains in the φ, and especially along the ψ, direction. In the rigid approximation, the repulsive terms of van der Waals

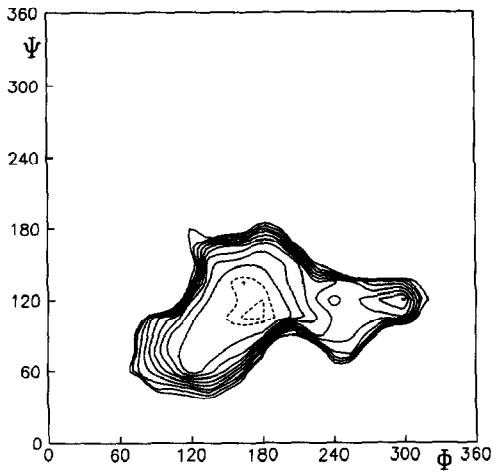


Fig. 3. Conformational-energy map after random walk. Energy contours are equivalent to Fig. 2.

interactions between pendant groups dominate while introducing the reorientations of their positions relieved most of the steric conflicts through variations of atomic position. The low-energy region is larger than it was for rigid geometry. Similar features were observed in the calculation of relaxed maps for oligosaccharides. In the present work, the other geometrical parameters were not relaxed. It may be expected that such relaxation will result in further enlargement of the low-energy minimum. The main energy minimum was moved during RW calculation to slightly different values of φ and ψ (160° and 140°).

The partially relaxed conformational map is shown in Fig. 4. The optimization of

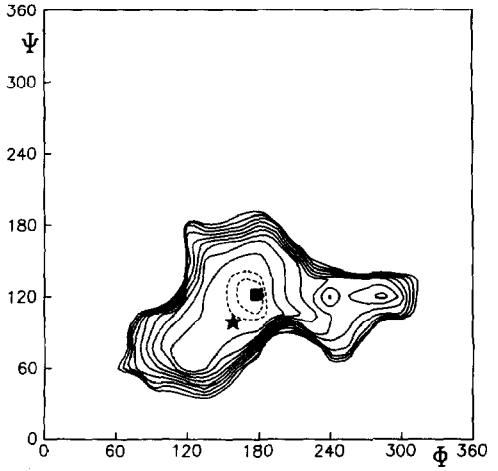


Fig. 4. Conformational-energy map in semirigid geometry. Energy contours are equivalent to Fig. 2. The included experimental data designated by ■ and ★ correspond to the structures retrieved from the CSD with reference codes BUFTIF and SOPROS.

torsion angles did not significantly influence the overall shape of the map obtained from RW, although the energies were lower. The (φ, ψ) values corresponding to the global energy minimum are 180° and 120°, which are in agreement with the experimental data of BUFTIF (−178° and 122°). Also, SOPROS falls in the low-energy region of the map. The order of the other minima differs from that from the rigid map; the minimum at φ = 280° and ψ = 120° is energetically less favorable (69 kJ/mol) than the minimum at φ = 120° and ψ = 80° (with an energy of 22 kJ/mol) above the main minimum. This minimum has not been registered on the map from rigid geometry.

Flexibility of the pendant groups. — Recent calculations of the relaxed potential energy surfaces for disaccharides[11–13] show that some fluctuations of only 15° occurred about ideal staggered orientations. A transition between different orientations has never been observed. The results of the present work show a striking flexibility of pendant groups. A distribution of torsion angles describing the positions of the H-(O-4′), O-6′, and CM-2′ atoms is given in Figs. 5–7. The orientations of pendant groups, as obtained from the random walk and consequent optimization, are grouped into three regions corresponding to staggered positions and three groups of eclipsed orientations.

The calculated distributions of torsion angle not only show dependence on (φ, ψ), but also differences between the groups. Properties of groups at the same position on both rings, except the methyl groups in **1**, exhibit similar features. The random walk magnitudes of torsion angles are spread throughout all of the conformational space. The distributions in Figs. 5–7 indicate substantial differences inherent in the position of the pendant group on the ring. There is great freedom of motion associated with hydroxyl groups, as 360° of rotation is possible for all hydrogen atoms according to the random walk [for an illustration, see the values of the H-(O-4′) torsion angle in Fig. 5]. The optimized orientations of pendant groups are grouped mainly into three regions

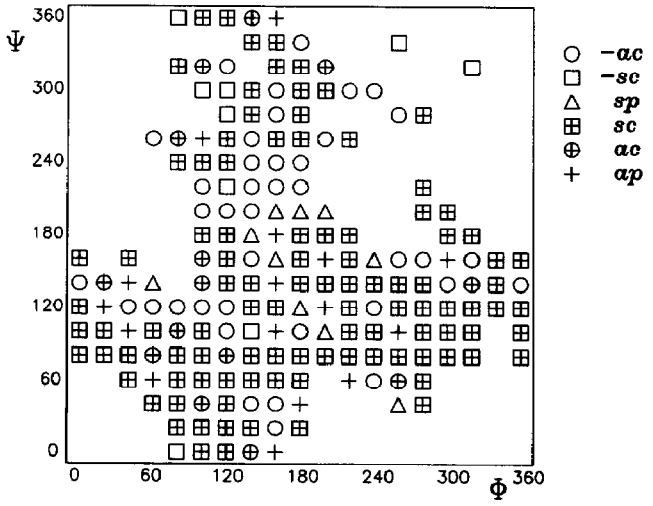


Fig. 5. Conformational dependence of the torsion angle of the hydroxylic H-(O-4′) hydrogen atom that resulted from the random walk.

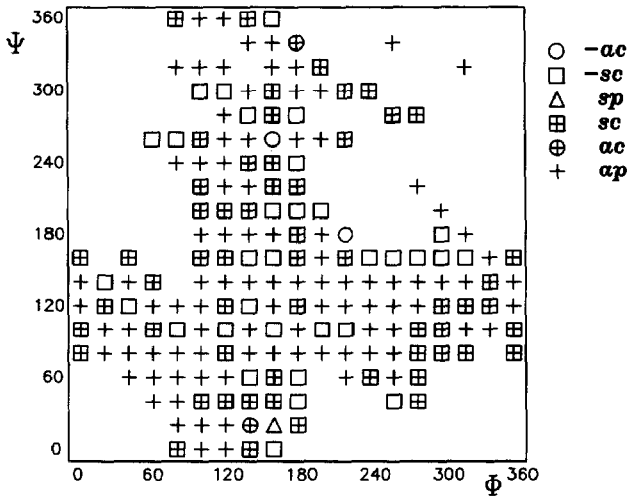


Fig. 6. Conformational dependence of the torsion angle ω of the hydroxymethyl group of the nonreducing ring that resulted from the random walk and consequent optimization.

corresponding to staggered positions. In the case of H-(O-4′), the *sc* orientation is the most frequent.

Conformational flexibility is also a characteristic feature in the case of the hydroxymethyl groups. For these groups (see Fig. 6), two positions are favored: *sc* (ω = 60°) and *ap* (ω = 180°). The $-sc$ orientations were found to be the most stable for only a few grid points of the conformational-energy map. It was found that the proportion of $-sc$ is larger for the hydroxymethyl group on the nonreducing ring.

In contrast, the rotation of the methyl groups is more restricted (see Fig. 7).

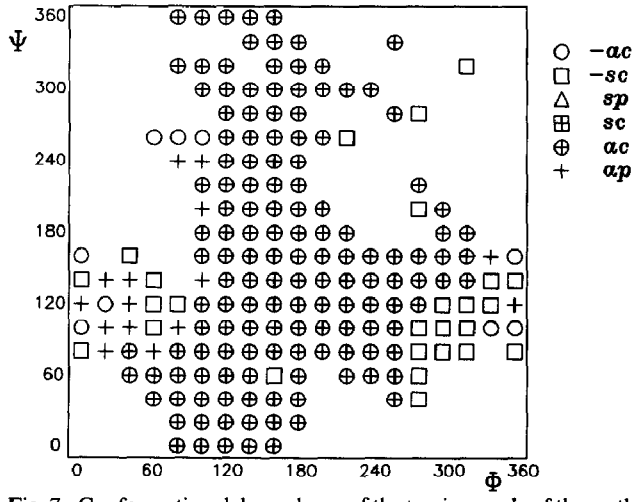


Fig. 7. Conformational dependence of the torsion angle of the methyl group, CM-2′, that resulted from the random walk and optimization.

Although, from the random-walk method, different orientations between 120 and 300° are available for the methyl group on the first ring, the optimized orientations are restricted to two regions, namely, around 120° (*ac*) and 300° (−*sc*). The *ac* conformations are favored. The flexibility of the second methyl group (not shown) is even more restricted than that of the first one, and the optimized orientations are almost exclusively at *ap*.

CONCLUSIONS

Despite the fact that all the calculations were performed with rigid-ring geometry, they document the importance of the pendant-group orientations in modelling studies. The work presented demonstrates that many transitions between different orientations occur during rotation about φ and ψ. With respect to rigid-modelling studies, the present results indicate that flexible pendant groups are a necessity for more realistic descriptions of the conformational behavior of oligosaccharides. Therefore, the relaxation of all internal coordinates through a complete energy-minimization that permits the pendant groups to adjust their orientation in response to φ, ψ constitutes an important step towards a realistic description of disaccharide conformation.

The methodologically new procedure implemented in the program RAMM is a very powerful and convenient tool for the conformational analysis of carbohydrates that allows taking into account the flexibility of pendant groups.

REFERENCES

1 I. Tvaroška, *Pure Appl. Chem.*, 61 (1989) 1201–1216.
2 G. A. Jeffrey and A. D. French, *Molecular Structure by Diffraction Methods*, Vol. 6, The Chemical Society, London, 1978.
3 A. D. French and V. C. Murphy, *Carbohydr. Res.*, 27 (1973) 391–406.
4 B. Fuchs, L. Schleifer, and E. Tartakovsky, *Nouv. J. Chim.*, 8 (1984) 275–278.
5 I. Tvaroška and T. Kožár, *Carbohydr. Res.*, 90 (1981) 173–185.
6 T. Kožár and I. Tvaroška, *Theor. Chim. Acta*, 53 (1979) 9–19.
7 I. Tvaroška, *Theoretical Chemistry of Biological Systems*, in G. Náray-Szabó (Ed.), Elsevier, Amsterdam, 1986, pp. 283–348.
8 I. Tvaroška and S. Pérez, *Carbohydr. Res.*, 149 (1986) 389–410.
9 K. Rasmussen, *Potential Energy Functions in Conformational Analysis, Lecture Notes in Chemistry*, Vol. 37, Springer-Verlag, Berlin, 1985.
10 M. Ragazzi and D. R. Ferro, *Carbohydr. Res.*, in press.
11 S. N. Ha, L. J. Madsen, and J. W. Brady, *Biopolymers*, 27 (1988) 1927–1952.
12 V. Tran, A. Buleon, A. Imberty, and S. Pérez, *Biopolymers*, 28 (1989) 679–690.
13 A. D. French, *Carbohydr. Res.*, 188 (1989) 206–211.
14 K. D. Gibbson and H. A. Scheraga, in R. H. Sarma and M. H. Sarma (Eds.), *Structure*, Vol. 1, *From Proteins to Ribosomes*, Academic Press, 1988, pp. 67–94.
15 RAMM, Technical Manual.
16 M. J. D. Powell, *Comput. J.*, 7 (1964) 155–162.
17 W. I. Zangwill, *Comput. J.*, 10 (1967) 293–296.
18 I. Tvaroška and T. Kožár, *J. Am. Chem. Soc.*, 102 (1980) 6929–6936.
19 I. Tvaroška and T. Kožár, *Theor. Chim. Acta*, 70 (1986) 99–114.
20 I. Tvaroška and L. Václavík, *Carbohydr. Res.*, 160 (1987) 137–149.
21 U. Burkert and N. L. Allinger, *Molecular Mechanics*, Am. Chem. Soc., Washington, D.C., 1982.

22 P. Zugenmaier and A. Sarko, *ACS Symp. Ser.*, 141 (1980) 225–237.
23 J. D. Hoogendorp, A. J. De Kok, and C. Romers, *Carbohydr. Res.*, 114 (1983) 169–180.
24 J. Ohanessian, F. Longchambon, and F. Arene, *Acta Crystallogr., Sect. B*, 34 (1978) 3666–3671.