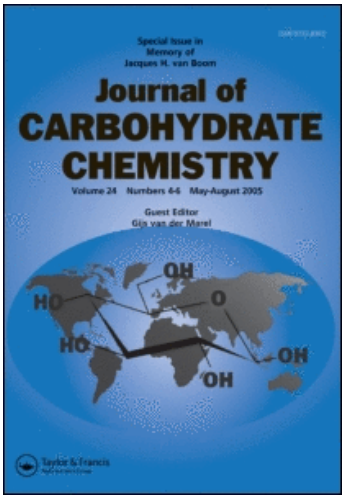



## Journal of Carbohydrate Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713617200

# Conformational Analysis of *C*-Disaccharides using Molecular Mechanics Calculations

Emmanuel Mikros; George Labrinidis; Serge Pérez

**To cite this Article** Mikros, Emmanuel , Labrinidis, George and Pérez, Serge(2000) 'Conformational Analysis of *C*-Disaccharides using Molecular Mechanics Calculations', Journal of Carbohydrate Chemistry, 19: 9, 1319 — 1349
**To link to this Article: DOI:** 10.1080/07328300008544154
**URL:** http://dx.doi.org/10.1080/07328300008544154

J. CARBOHYDRATE CHEMISTRY, 19(9), 1319–1349 (2000)

# CONFORMATIONAL ANALYSIS OF *C*-DISACCHARIDES USING MOLECULAR MECHANICS CALCULATIONS.

Emmanuel Mikros,[a] George Labrinidis[a] and Serge Pérez[b]

[a]Department of Pharmacy, University of Athens, Panespistimiopoli, Zografou, GR-15771, Athens, Greece.
[b]Centre de Recherches sur les Macromolécules Végétales,[1] CNRS, BP 53X, 38041 Grenoble cedex, France.



*Received October 15, 1999 - Final Form October 5, 2000*

## ABSTRACT

Relaxed-residue energy maps based on the MM3 force field were computed for the methyl glycosides of eight *C*-linked D-glucosyl disaccharides: the two-bond axial-equatorial linked disaccharides β–kojibioside [(1→2)α-], β–nigeroside [(1→3)α–] and β–maltose [(1→4)α–], the two-bond equatorial-equatorial linked disaccharides β–sophoroside [(1→2)β–], β–laminarabioside [(1→3)β–], β–cellobioside [(1→4)β–] and the three-bond-linked (1→6) disacharides *C*-isomaltoside and *C*-gentiobioside. Optimized structures were calculated on a 20° grid spacing of the torsional angles about the *C*-glycosidic bonds and the final isoenergy surfaces were based on 11664 conformations, for the two-bond-linked disaccharides and 69984 conformations for the three-bond-linked disaccharides. Boltzmann-weighted $^3J$ coupling constants were calculated and compared to the experimental values. They are satisfactory except for maltose where hydrogen bonds cause an over-estimation of the energy differences between the conformers. The energy maps are similar to maps of the corresponding *O*-disaccharides, but there are differences in the locations and the relative energies of the minima. The preferred conformations of the *C*-glycosidic bonds are as if they were conforming to the *exo*-anomeric effect but are closer to staggered conformations than shown by the MM3 results for the O-linkages.

## INTRODUCTION

Disaccharide analogues in which the glycosidic oxygens are replaced by a methylene group have been the subject of widespread interest both as carbohydrate mimics as well as potential enzyme inhibitors. The first synthesis of a *C*-disaccharide, namely, β(1→6)-*C*-disaccharide (β-D-Glc*p*-$CH_2$ (1→6) α-D-Glc*p*-OMe), was performed by Rouzaud and Sinaÿ.[2] Since then, several approaches to *C*-disaccharides and analogues have been reported.[3]

That work contributed to the renewal of the relationship of carbohydrate chemistry to glycobiology. Oligosaccharide mimics are potential inhibitors for the biosynthetic pathways to glycoproteins. They may find applications as antibacterial,[4a] antiviral,[4b] antitumoral[4c] or fertility control agents[4d] as well as inhibitors of α–amylases,[4e] sucrase and maltase.[4f] Because of such a broad range of potential applications, the preferred conformations of *C*-disaccharides have been the subject of several investigations. In a number of contributions,[5] Kishi and coworkers have postulated that the *C*-disaccharides mimic the conformational behavior of the corresponding *O*-glycosides. These conclusions are based on a semiquantitative analysis of NMR data, mainly coupling constants. However, conformational differences between *C*- and *O*-glycosides have been claimed to occur in the case of lactose when investigated via the combination of NMR NOE and J data) and molecular mechanics calculations.[6] This was further exemplified by the study of the bio-active conformation of *C*-lactoside bound to ricin-B[7a] and β-galactosidase.[7b]

In this study, molecular mechanics calculations have been used to study the conformational characteristics of *C*-disaccharides having: two-bond-α-linked glucose residues, kojibioside (α-D-Glc*p*-$CH_2$ (1→2) β–D-Glc*p*-OMe) (**1**), nigeroside (α-D-Glc*p*-$CH_2$ (1→3) β–D-Glc*p*-OMe) (**2**), maltoside (α-D-Glc*p*-$CH_2$ (1→4) β–D-Glc*p*-OMe) (**3**); the corresponding two-bond-β-linked sophoroside (β-D-Glc*p*-$CH_2$ (1→2) β–D-Glc*p*-OMe) (**4**), laminarabioside (β-D-Glc*p*-$CH_2$ (1→3) β–D-Glc*p*-OMe) (**5**), cellobioside (β-D-Glc*p*-$CH_2$ (1→4) β–D-Glc*p*-OMe) (**6**); and three-bond-linked glucose residues, isomaltoside (α-D-Glc*p*-$CH_2$ (1→6) β–D-Glc*p*-OMe) (**7**), in the α configuration, and gentiobioside (β-D-Glc*p*-$CH_2$ (1→6) β–D-Glc*p*-OMe) (**8**), in the β configuration (Fig 1). The investigation of the conformational features of *C*-disaccharides provides a way to

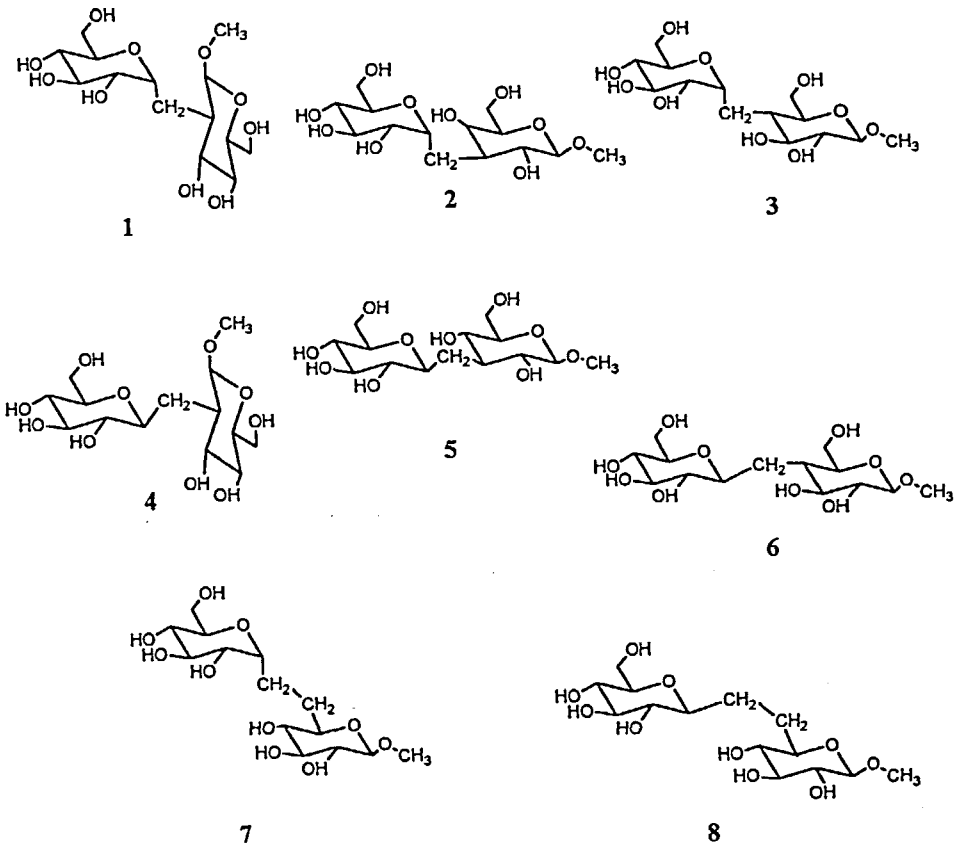



**Figure 1.** Schematic representation of the eight *C*-disaccharides, along with the torsional angles of interest.

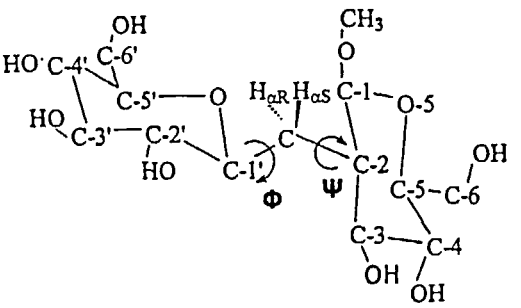


study the steric interactions around the glycosidic linkages of carbohydrates, in the absence of electronic stabilization arising from the *exo*-anomeric effect. Comparisons are made between the modeling results for *C*- and *O*-disaccharides.

## RESULTS AND DISCUSSION

### Nomenclature

Labeling of the atoms follows the IUPAC Nomenclature for Carbohydrates.[8] The carbon atom at the *C*-glycosidic bond is labeled Cα, whereas the two hydrogen atoms are labelled HαR and HαS, corresponding to the proR and proS assignments, respectively.



Scheme 1

The torsion angles around the glycosidic linkage (Fig. 1) for the (1→2), (1→3) and (1→4)-disacharides are defined as (Scheme 1):

$\Phi$ = O5'-C1'-C$\alpha$-C$_i$ ; $\Psi$ = C1'-C$\alpha$-C$_i$-C$_{i+1}$

where i = 2, 3, 4, depending on the disaccharide studied.[9]

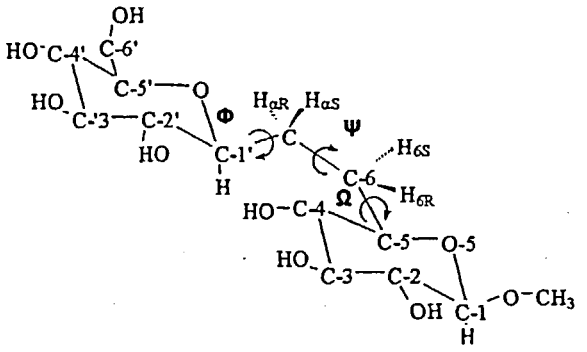
The orientations of the hydroxymethyl groups are referred to as either *gauche-gauche (gg)*, *gauche-trans* (*gt*) or *trans-gauche* (*tg*). In this terminology, the dihedral angle O5-C5-C6-O6 is considered first and the dihedral angle C4-C5-C6-O6 is considered second.

In the case of (1→6) *C*-disaccharides, the following definitions are used (Scheme 2): [10]

$\Phi$ = O5'-C1'-C$\alpha$-C6; $\Psi$ = C1'-C$\alpha$-C6-C5 and $\Omega$ = C$\alpha$-C6-C5-O5

**Computational Methods.**

The conformational behavior of the eight *C*-disaccharides was investigated using the MM3 force field.[11] This is a highly detailed force field, which in addition to the classical terms, includes anisotropy of hydrogens, corrections for stereoelectronic effects, cross-terms such as torsion-stretch, torsion-bend and bend-bend interactions, a Buckingham-type potential for nonbonded interactions and explicit terms for hydrogen bonding. This force field has been compared favorably to other molecular mechanics force fields[12] and has been successful in modeling many disaccharides. The single exception appears to be sucrose, where inadequate parameterization for overlapping




**Scheme 2**



anomeric effect has been claimed to cause error in the calculated energy. MM3, and most other force fields have not performed well with the hydroxymethyl group.[13]

Thirty-six starting structures were used at each Φ, Ψ point for each *C*-disaccharide having two-bond-linked glucose residues. Taking the *gg*, *gt* and *tg* conformers of both hydroxymethyl groups into account yielded 9 combinations. The starting positions of the secondary hydroxyl groups were set as *c* (clockwise) or *r* (anticlockwise), resulting in four combinations: *cc*, *cr*, *rc* and *rr*. The relaxed maps were computed using rigid rotation followed by harmonic constraint minimization in 20° increments for Φ and Ψ spanning the whole angular range. This approach resulted in the minimization of (18 x 18 x 9 x 4 = 11664) conformations for each disaccharide. For each of the three-bond linked disaccharides, 216 starting structures were calculated as the third torsional angle, Ω, was varied in 20° increments over its full angular range. Therefore, 69984 conformations were minimized for each disaccharide. The bulk dielectric constant (ε) was set to 4.0 as this intermediate value models the crystalline glucopyranosyl ring better than either ε = 1.5 (vacuum) or ε = 80 (water).[14] Convergence was accepted when the iterative decrease in energy ΔE was less than N x 0.00008 kcal/mol, where N was the number of atoms in the molecule. From these relaxed energy maps, adiabatic surfaces were built by selecting the lowest energy structure for a given Φ, Ψ point.

For each local low-energy region of a map, the conformation that contributed the lowest energy grid point was submitted to a further, non-restricted full matrix

minimization in order to find the local minimum. If the structure remained within the local well, the lowest energy structure was taken as the local minimum structure. Usually, the final locations of minima were found within the local regions, although they could have torsional angles several degrees removed from those of the grid-point structure.

Coupling constants, $^3J_{H,H}$ for vicinal hydrogen atoms of a H-C-C-H segment were calculated using a Karplus type equation[15] with Haasnoot-Altona parameterization.[16] It accounts for the J dependence on the dihedral angle of the H-C-C-H fragment, and on the electronegativities and orientations of the α and β substituents. The derived general form of the Altona equation used here was:

$$^3J_\theta = A\cos(2\theta) + B\cos(\theta) + C\sin(2\theta) + D$$

where θ corresponds to the different dihedrals : Φ(R), Φ(S) ... Ω(R), Ω(S)

The parameters A, B, C and D for the kojibiose, nijerose and maltose torsional angles are as following:

| | Φ(R) | Φ(S) | Ψ(R) | Ψ(S) |
|---|---|---|---|---|
| A | 4.781 | 4.781 | 5.356 | 5.356 |
| B | -0.990 | -0.990 | -0.990 | -0.990 |
| C | 1.157 | 0.950 | -0.104 | 0.104 |
| D | 5.912 | 5.912 | 6.232 | 6.232 |

The corresponding parameters for sophorose, laminarabiose and cellobiose are:

| | Φ(R) | Φ(S) | Ψ(R) | Ψ(S) |
|---|---|---|---|---|
| A | 4.781 | 4.781 | 5.356 | 5.356 |
| B | -0.990 | -0.990 | -0.990 | -0.990 |
| C | -0.950 | -1.157 | -0.104 | 0.104 |
| D | 5.912 | 5.912 | 6.232 | 6.232 |

Finally the parameters used for the (1→6) disaccharides are:

a) *C*-isomaltose

| | Φ(R) | Φ(S) | Ψ(RR) | Ψ(RS) | Ψ(SR) | Ψ(SS) | Ω(R) | Ω(S) |
|---|---|---|---|---|---|---|---|---|
| A | 4.781 | 4.781 | | 6.535 | | | 4.781 | 4.781 |
| B | -0.990 | -0.990 | | -0.730 | | | -0.990 | -0.990 |
| C | -0.950 | -1.157 | | -0.073 | | | -0.950 | -1.157 |
| D | 5.912 | 5.912 | | -6.675 | | | 5.912 | 5.912 |

b) *C*-gentiobiose

| | Φ(R) | Φ(S) | Ψ(RR) | Ψ(RS) | Ψ(SR) | Ψ(SS) | Ω(R) | Ω(S) |
|---|---|---|---|---|---|---|---|---|
| A | 4.781 | 4.781 | | 6.535 | | | 4.781 | 4.781 |
| B | -0.990 | -0.990 | | -0.730 | | | -0.990 | -0.990 |
| C | 0.950 | 1.157 | | -0.073 | | | -0.950 | -1.157 |
| D | 5.912 | 5.912 | | -6.675 | | | 5.912 | 5.912 |

Because macroscopic properties arise from the balanced contribution of each element of an ensemble, averaging of $^3J$ coupling constants was performed on the total

population of 11664 conformers in the (1→2)-, (1→3)- and (1→4)-disaccharides and the 69984 conformers in the (1→6) disaccharides. Assuming that the entropy differences among the different conformers are negligible, the probability $P_i$ of the *i*th conformer depends upon its relative energy $E_i$ : $P_i = \exp(-E_i/RT) / \Sigma (\exp(-E_i/RT))$. Ensemble averages were calculated from the distribution according to $\langle^3J\rangle = \Sigma P_i J_\theta$

For each disaccharide, the $\langle^3J\rangle$ coupling constants were calculated and compared to the experimental values when these were available from the literature.

## Conformational analysis

### TWO-BOND, α-LINKED, AXIAL-EQUATORIAL *C*-DISACCHARIDES.

***C*-kojibioside** - The MM3-generated adiabatic map of *C*-kojibiose (α-D-Glc*p*-$CH_2$ (1→2) β-D-Glc*p*-OMe) is shown in Fig. 2. A deep trough of low energy conformations runs in the Ψ-direction and is centered about the perfectly staggered conformation at Φ = 60°. The four minima, the three lowest of which are from this trough, are displayed in Fig. 2. Their conformational features, and steric energy values are listed in Table 1. The global minimum (A) is located at Φ, Ψ = 77 .3°, -73.8° ; it is stabilized by an inter-residue hydrogen bond between O-3 and O-5'. Conformer B occurs at Φ, Ψ = 58.9°, -161.0°. A small energy difference exists between these two conformers. Two additional minima with relatively higher energies are found. One of these (C) is a shallow depression with each of the glycosidic torsional angles close to perfectly staggered positions with Φ, Ψ = 175.5°, -58.3°. This conformation is less than 2.0 kcal/mol higher than the global minimum. It is stabilized an inter-residue hydrogen bond between O-2' and O-3. The other conformer (D) exists at perfectly staggered conformation in Φ, with Ψ close to 40°.

Almost all the initial orientations of the hydroxymethyl groups are represented on the final map (Table 2) but the *gtgt* orientation contributes more than 86% of the lowest energy structures. The *cc*, *rc* and *rr* orientations are favored to similar extents. The MM3-generated map of *O*-kojibiose was reported previously.[15] Similarities and differences between the *O*- and *C*- disaccharides can be found. The general map characteristics are similar. This is not unexpected since the aspects of the outer contour reflect the importance of the axial-equatorial glycosidic linkage in determining conformation.

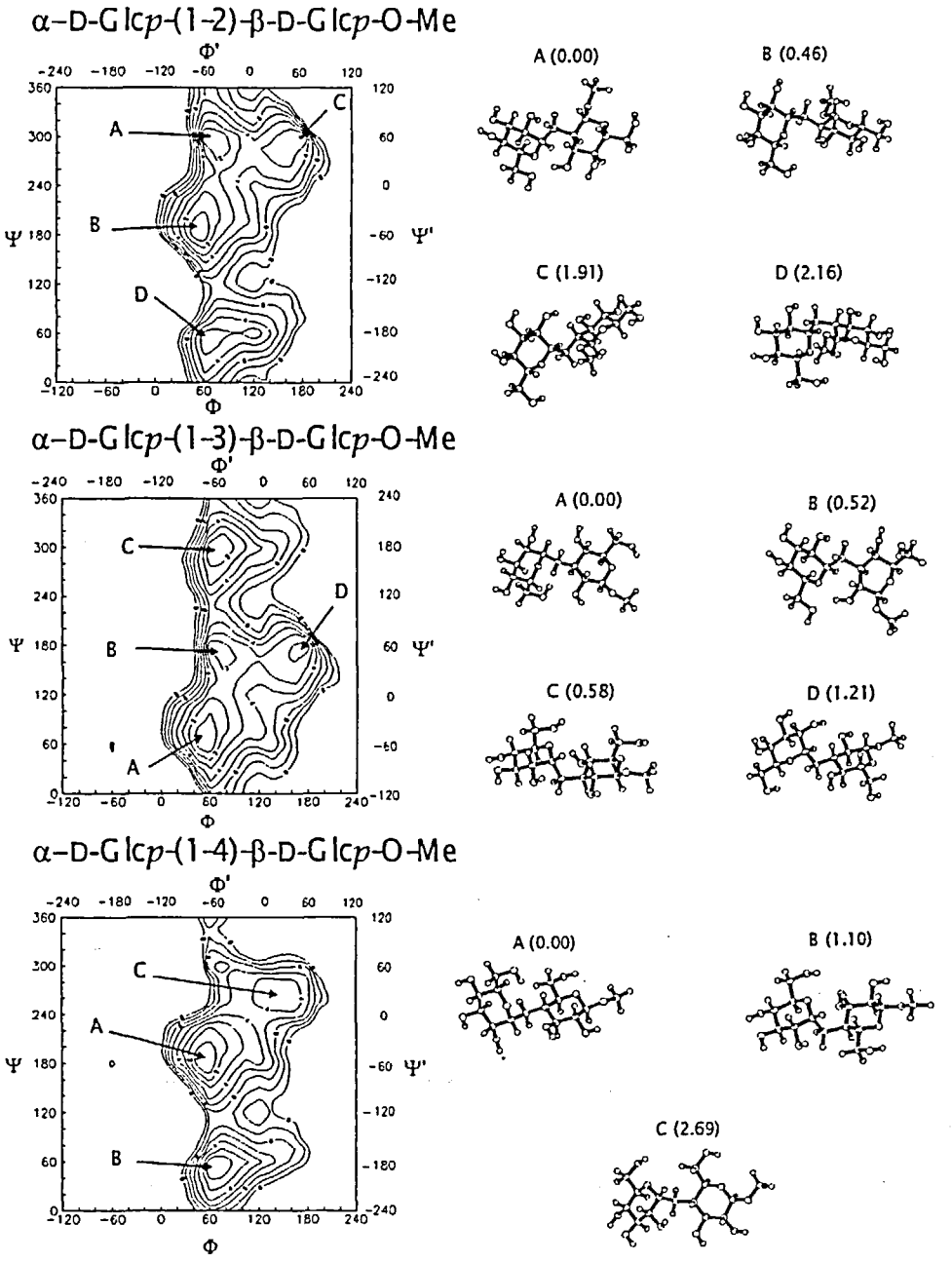



**Figure 2.** Adiabatic relaxed map of *C*-kojibiose, *C*-nigerose and *C*-maltose in the MM3(92) force field as a function of the Φ and Ψ glycosidic torsion angles. Iso energy contours are drawn in 1 kcal/mol intervals out to the 8 kcal/mol from the global minimum. The lowest energy conformation of each domain has been represented by a ball and stick model, and its location is indicated on the map. Φ' and Ψ' stands for the dihedral angles H1'-C1'-Cα-Ci (i=2, 3, 4) and Hi-Ci-Cα–Ci', respectively.

**Table 1.** Energy minima for MM3-generated relaxed-residue analysis of two-bond-linked disaccharides. Comparison with correspondent *O*-glycosides[13,14]

| Disaccharide | | *C*-Glycosides | | | | *O*-Glycosides | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Conformer | Location of Minimum | | Energy | Rel. Energy | Location of Minimum | | Energy | Rel. Energy |
| | | Φ | Ψ | | | Φ | Ψ | Kcal / mol | |
| Kojibiose | A | 77.3 | -73.8 | 32.43 | 0.00 | 86.2 | -78.0 | 20.78 | 0.00 |
| | B | 58.9 | -160.6 | 32.89 | 0.46 | 85.3 | -146.4 | 22.23 | 1.45 |
| | C | 175.5 | -58.3 | 34.34 | 1.91 | 177.5 | -56.9 | 24.64 | 6.67 |
| | D | 60.6 | 39.5 | 34.59 | 2.16 | 92.9 | 67.9 | 27.45 | 3.86 |
| Nigerose | A | 58.7 | 80.5 | 33.40 | 0.00 | 84.8 | 118.0 | 22.49 | 0.77 |
| | B | 77.1 | 163.7 | 33.92 | 0.52 | 84.0 | 166.3 | 21.42 | 0.00 |
| | C | 77.8 | -59.4 | 33.99 | 0.59 | 97.0 | -53.6 | 24.47 | 3.05 |
| | D | 175.2 | -179.4 | 34.61 | 1.21 | 177.0 | 183.3 | 26.22 | 4.80 |
| Maltose | A | 54.1 | -174.6 | 31.85 | 0.00 | 66.0 | -167.4 | 22.05 | 0.00 |
| | B | 79.3 | 62.7 | 32.95 | 1.10 | 92.0 | 69.5 | 24.74 | 2.69 |
| | C | 142.7 | -96.7 | 34.55 | 2.70 | 120.6 | -96.3 | 24.09 | 2.04 |
| | | | | | | 91.7 | 147.5 | 22.22 | 0.17 |
| Sophorose | A | -57.1 | -62.1 | 29.78 | 0.00 | -72.0 | -128.3 | 21.45 | 0.00 |
| | B | -78.1 | 57.0 | 29.97 | 0.19 | -84.8 | 49.2 | 22.73 | 1.28 |
| | C | 73.7 | -74.9 | 30.43 | 0.65 | | | | |
| | D | -62.8 | -138.1 | 30.78 | 1.00 | -86.0 | -163.4 | 21.48 | 0.03 |
| | E | 58.8 | -119.1 | 31.62 | 1.84 | 59.8 | -110.7 | 23.25 | 1.80 |
| | F | -167.0 | -175.0 | 31.70 | 1.92 | | | | |
| | G | 65.7 | 75.8 | 33.30 | 3.52 | | | | |
| Laminarabiose | A | -81.4 | 62.3 | 29.53 | 0.00 | -85.0 | 78.0 | 20.15 | 0.00 |
| | B | -56.4 | 177.9 | 29.82 | 0.29 | -72.0 | 111.0 | 20.18 | 0.03 |
| | C | -79.2 | -60.7 | 30.65 | 1.12 | -83.8 | -71.2 | 23.76 | 3.61 |
| | D | 73.9 | 164.3 | 31.53 | 2.00 | | | | |
| | E | 159.4 | 101.8 | 31.80 | 2.27 | | | | |
| | F | 60.4 | 119.9 | 32.07 | 2.54 | 57.3 | 132.4 | 24.34 | 4.19 |
| | G | 62.9 | -42.2 | 33.08 | 3.55 | | | | |
| Cellobiose | A | -63.5 | 63.9 | 29.60 | 0.00 | -85.6 | 52.3 | 22.34 | 2.08 |
| | B | -80.0 | -179.6 | 30.32 | 0.72 | -84.0 | -166.2 | 20.26 | 0.00 |
| | C | 59.8 | -119.7 | 30.54 | 0.94 | 59.4 | -119.2 | 23.85 | 3.59 |
| | D | -63.4 | -99.0 | 30.57 | 0.97 | -69.5 | -130.7 | 20.29 | 0.03 |
| | E | -159.0 | -158.0 | 31.86 | 2.26 | | | | |
| | F | 61.35 | 81.25 | 33.71 | 4.11 | | | | |

**Table 2.** Starting structures contributing to MM3-generated relaxed-residue Φ,Ψ disacharide maps.

| Disaccharide | Complete Map | | | | | Lowest 8 kcal/mol | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Hydroxy-Methyl | | Secondary Hydroxyls | | % Area of map | Hydroxy-methyl | | Secondary hydroxyls | |
| Sophorose | *gggt* | 1.85% | *cc* | 12.96% | 62.04% | *gggt* | 0.49% | cc | 8.45% |
| | *gtgg* | 0.31% | *cr* | 40.12% | | *gtgt* | 94.52% | cr | 36.81% |
| | *gtgt* | 91.36% | *rc* | 24.38% | | *tggt* | 4.97% | rc | 26.86% |
| | *gttg* | 1.23% | *rr* | 22.53% | | | | rr | 27.86% |
| | *tggt* | 5.25% | | | | | | | |
| Laminarabiose | *gggt* | 3.08% | *cc* | 13.58% | 59.87% | *gggt* | 1.03% | cc | 12.37% |
| | *gtgg* | 2.16% | *cr* | 37.65% | | *gtgg* | 1.03% | cr | 35.05% |
| | *gtgt* | 92.28% | *rc* | 13.89% | | *gtgt* | 96.91% | rc | 13.40% |
| | *gttg* | 1.85% | *rr* | 34.87% | | *gttg* | 1.03% | rr | 39.17% |
| | *tggt* | 0.62% | | | | | | | |
| Cellobiose | *gggg* | 1.54% | *cc* | 16.36% | 55.55% | *gggg* | 1.66% | cc | 11.67% |
| | *gggt* | 3.70% | *cr* | 43.21% | | *gggt* | 1.66% | cr | 37.22% |
| | *ggtg* | 0.62% | *rc* | 8.02% | | *ggtg* | - | rc | 10.55% |
| | *gtgg* | 12.03% | *rr* | 32.41% | | *gtgg* | 12.77% | rr | 40.55% |
| | *gtgt* | 25.92% | | | | *gtgt* | 34.44% | | |
| | *gttg* | 3.08% | | | | *gttg* | 2.77% | | |
| | *tggg* | 4.63% | | | | *tggg* | 3.33% | | |
| | *tggt* | 38.88% | | | | *tggt* | 36.67% | | |
| | *tgtg* | 9.57% | | | | *tgtg* | 6.67% | | |
| Kojibiose | *gggg* | 0.31% | *cc* | 6.48% | 37.65% | *gggt* | 0.82% | cc | 0.82% |
| | *gggt* | 2.77% | *cr* | 26.85% | | *gtgt* | 93.44% | cr | 26.23% |
| | *ggtg* | 0.31% | *rc* | 34.56% | | *tggt* | 5.73% | rc | 27.05% |
| | *gtgg* | 0.31% | *rr* | 32.09% | | | | rr | 45.90% |
| | *gtgt* | 86.72% | | | | | | | |
| | *gttg* | 2.16% | | | | | | | |
| | *tggt* | 7.09% | | | | | | | |
| | *tgtg* | 0.31% | | | | | | | |
| Nigerose | *gggg* | 0.31% | *cc* | 3.70% | 38.89% | *gggt* | 0.79% | cc | 3.96% |
| | *gggt* | 2.47% | *cr* | 5.55% | | *gtgg* | 22.22% | cr | 0.79% |
| | *ggtg* | 1.23% | *rc* | 19.14% | | *gtgt* | 75.39% | rc | 23.81% |
| | *gtgg* | 16.97% | *rr* | 71.61% | | *tggt* | 1.58% | rr | 71.43% |
| | *gtgt* | 70.68% | | | | | | | |
| | *gttg* | 3.70% | | | | | | | |
| | *tggt* | 4.63% | | | | | | | |
| Maltose | *gggt* | 1.85% | *cc* | 3.39% | 31.17% | *gtgg* | 2.97% | cc | 0.99% |
| | *ggtg* | 0.31% | *cr* | 33.02% | | *gtgt* | 81.18% | cr | 17.82% |
| | *gtgg* | 2.47% | *rc* | 12.34% | | *gttg* | 9.90% | rc | 8.91% |
| | *gtgt* | 73.15% | *rr* | 51.23% | | *tggt* | 5.94% | rr | 72.27% |
| | *gttg* | 9.87% | | | | | | | |
| | *tggg* | 0.31% | | | | | | | |
| | *tggt* | 11.115 | | | | | | | |
| | *tgtg* | 0.93% | | | | | | | |

The single trough of low energy conformers found in the Ψ- direction indicates that there is freedom of rotation about this equatorial bond. In both cases, the troughs have all their global minima at Φ values between 60° and 180°. In the case of *O*-kojibiose, one notices the importance of the *exo*-anomeric effect in favoring the occurrence of conformations for Φ values between 60 to 80° and in working against the staggered position for Φ values close to 180°.

In the case of *C*-kojibiose, the values about Φ angles are more centered around Φ = 60°. Moreover, the orientation of Φ = 180°, as in conformer D, is likely to be more easily populated than in the case of *O*-kojibiose. In the *O*-kojibiose map, 25% of the structures have relative energies lower than 8 kcal/mol instead of 38% in *C*-kojibiose (Table 2). The energy barrier between the A and D minima is ~5 kcal/mol in *C*-kojibiose and the corresponding barrier in *O*-kojibiose is ~7 kcal/mol. The relative magnitudes of these differences suggest that the *C* disaccharide has more conformational freedom.



**C-nigeroside** - The adiabatic map of *C*-nigerose (α-D-Glc*p*-$CH_2$-(1→3)-β-D-Glc*p*-OMe) generated by MM3 is shown in Fig. 2, and the optimized minima are listed in Table 1. All but one of the *C*-nigerose minima are contained in a single low energy region extending vertically across the map in the Ψ-direction. They occur within an energy window of 1.2 kcal/mol. The wells that contain the four low energy conformers have similar shapes and dimensions. The global minimum (A) is at Φ, Ψ = 58.7°, 80.5° and occurs without any inter-residue hydrogen bond. In contrast, each of the three other low energy conformers have an intramolecular hydrogen bond when a bulk dielectric constant ε = 4 is used.

The type B conformation (Φ, Ψ = 77.1°, 163.7°) contains an inter-ring hydrogen bond formed between O-2 and O-5'. The conformation of the second lowest energy minimum (conformer C) located at Φ, Ψ = 77.8°, 80.5° accommodates an O-6'...O-4 hydrogen bond. The fourth conformer, D, is located in the staggered 180° region with Φ= 175.2° and Ψ = -179.4°, and is stabilized by the occurrence of a hydrogen bond between O-2' and O-4. Several starting conformations contribute to the final map (Table 2). The dominant hydroxymethyl orientation is *gtgt* with *gtgg* representing a considerable proportion (16%). The *rr* orientation is the most important set of secondary hydroxyl group orientations.

The adiabatic map of *O*-nigerose[15] also exhibits a vertical extension across the map in the Ψ-direction. It contains three minima corresponding to the *C*-disaccharide low energy conformers A: Φ, Ψ = 85°, 118°, B: Φ, Ψ = 84°, 166°, C: Φ, Ψ = 97°, -54°.

These locations are within 20° of those calculated for the *C*-disaccharide. There is a difference in the relative energies in the first two minima. In *O*-nigerose, the global minimum corresponds to the B conformer which is ~0.8 kcal/mol more stable than the second conformer, the global minimum for *C*-nigerose (Table 1). A plateau region occurs at approximately 180°, 180°, on which a fourth energy minimum exist as a shallow depression with a high relative energy. This is significantly different than for *C*-analogues, where a well localized, low energy conformation occurs for this perfectly staggered orientation.

The energy barrier between B and C minima in the O-nigerose is lower than 6 kcal/mol while in *C*-nigerose the corresponding energy barrier is lower than 5 kcal/mol. The total area surrounded from the 8 kcal contour line is 30% in O-nigerose while it is the 38.9% (Table 2) in the *C*-nigerose map suggesting again that the *C*-disaccharide has greater conformational freedom.

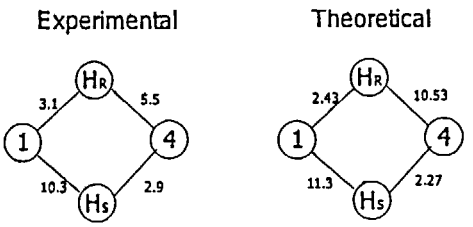


***C*-maltoside** - The MM3-generated adiabatic map of *C*-maltose (α-D-Glc*p*-$CH_2$ (1→4) β-D-Glc*p*) is shown in Fig. 2, along with the location and the representation of the low energy conformers. Geometric details for the optimized minima are listed in Table 1. Three distinct low energy minima can be found within the domain delineated by the 8 kcal/mol iso-energy contour. The lowest energy conformer occurs at Φ, Ψ = 54.1°, -174.6° and has no inter-residue hydrogen bonds. That conformation has both of its hydroxymethyl groups located on the same side of the molecule. A second low energy region exists, with a similar value of Φ but with Ψ near the staggered 60° potential well (conformer B at Φ, Ψ = 79.3°, 62.7°). The energy difference between these minima is about 1 kcal/mol. The third conformer, C, occurs at Φ, Ψ = 142.7°, -96.7°, with an energy 2.7 kcal/mol above that of conformer A.

The *gtgt* orientation of hydroxymethyl groups dominates the starting conformations that yield the lowest energies (Table 2) as it does for kojibiose and nigerose. Mixed orientations, such as *gttg* and *tggt*, are also important. Starting secondary hydroxyl orientations are *rr*, *cr*, and *rc* with *rr* most prelevant. The area enclosed by the 8 kcal/mol contour contains 31.2% of the whole map.

**Table 3.** Calculated vicinal coupling constants across the C1'-Cα-Ci bridge of the two-bond-linked disaccharides.

| Disaccharide | Coupling Constants | | | |
|---|---|---|---|---|
| | $J_{\Phi(R)}$ | $J_{\Phi(S)}$ | $J_{\Psi(R)}$ | $J_{\Psi(S)}$ |
| Kojibiose | 2.67 | 10.92 | 9.46 | 4.23 |
| Nijerose | 2.77 | 10.77 | 9.25 | 3.54 |
| Maltose | 2.43 | 11.30 | 10.53 | 2.27 |
| Sophorose | 10.37 | 2.42 | 3.32 | 8.26 |
| Laminarabiose | 10.74 | 2.07 | 5.86 | 7.10 |
| Cellobiose | 10.21 | 2.00 | 4.80 | 4.34 |




**Scheme 3**



The homonuclear coupling constants calculated are listed in Table 3. They can be compared (Scheme 3) with the coupling constants that have been measured across the axial *C*-glycosidic linkage ($J_{\Phi r} = 3.1$, $J_{\Phi s} = 10.3$ ; $J_{\Psi r} = 5.5$, $J_{\Psi s} = 2.9$).[4b]

A clear difference is observed between the calculated and the experimental value for the $J_{\Psi r}$ coupling constant. This suggests that the orientations of the *C*-aglyconic bond 60° and/or -60° are more populated in solution. That means that the energy differences between the A, B and C conformers are not well predicted and the areas around B and/or C are more populated. It is interesting to note that when the C3 hydroxyl group is protected (peracetyl -*C*-disaccharide) or completely removed (3-deoxy-*C*-maltoside) the vicinal coupling constants $J_{\Psi r}$ around the *C*-aglyconic bond are 7.1 and 11.4 Hz, respectively.[3b] This suggests that an interaction of the C3 hydroxyl group stabilizes either conformer B or C or both. Indeed, when the energy minimum structures were

optimized with a dielectric constant $\varepsilon = 1$, the calculated relative energies were: A, 0.1 kcal/mol ; B, 0 kcal/mol ; and C, 4.2 kcal/mol, because the B conformer was stabilized by an inter-residue hydrogen bond between O-6' and O-3. It is reasonable to suppose that this stabilization could occur under certain circumstances in solution resulting in a higher population for this conformer. The average coupling constants calculated based on the probability distribution of the three low energy conformers are: $J_{\Phi_T} = 1.6$ Hz, $J_{\Phi_S} = 11.0$ Hz, $J_{\Psi_T} = 6.7$ Hz, $J_{\Psi_S} = 3.1$ Hz, in quite good agreement with the experimental values. These results show the limitations of calculations performed in vacuum without taking into account explicit solvent molecules.

The adiabatic map of *O*-maltose[15] again has the same general shape but in the same area as the global minimum there is another low energy conformer that does not exist on the *C*-disaccharide map. The total area within the 8 kcal/mol contour is only 27.8% of the whole map. In contrast with all other disaccharides in this study, the energy barrier between A and B observed on the *O*-maltose map is lower (<5kcal/mol) than for the C-maltoside (<6 kcal/mol).

The general map characteristics for the three disaccharides studied are similar. Each map has a single trough of low energy conformers in the Ψ-direction which indicates the freedom of rotation about this equatorial bond. There is also a minimum with a high relative energy having $\Phi \sim 180°$, which is only a plateau for the *O*-disaccharides. The 8 kcal/mol contour line encompasses a higher percentage of Φ,Ψ space for the *C*-disaccharides and the energy barriers between minima are in general lower for the C-disaccharides.

### TWO-BOND, β-LINKED, DIEQUATORIAL *C*-DISACCHARIDES

***C*- Sophoroside** - The MM3 map of *C*-sophoroside (β-D-Glc*p*-$CH_2$ (1→2) β-D-Glc*p*-OMe) is shown in Fig. 3, along with four local minima. Table 1 lists the relative MM3 energies and details of the geometries of the optimized minima. Two main perpendicularly intersecting troughs appear on the map. The lowest energy minimum is located in a perfectly staggered conformation at Φ, Ψ = -57.1°, -62.1°. Conformer B occurs at Φ, Ψ = -78.1°, 57° and it is stabilized by an inter-residue hydrogen bond between O3 and O5'. Conformers C and E have relative energies lower than 2 kcal/mol and lie in a region that extends in the Ψ direction from approximately 0° to -60°.

Conformer C occurs at $\Phi, \Psi = 73.7°, -74.9°$ and is stabilized by a hydrogen bond between O3 and O5'. Conformer E is located at an eclipsed orientation in $\Psi = -119.1°$ with $\Phi = 58.8°$ and has a hydrogen bond between O2' and O1. In the same $\Psi$ direction, there is another conformer, G, with $\Phi, \Psi = 65.7°, 75.8°$. Its relative energy is 3.5 kcal/mol, despite stabilization by a hydrogen bond between O3 and O2'. Two additional minima are located on a plateau that extends in the $\Phi$ direction from approximately -180° to -60° with a relative energy less than 3 kcal/mol. Conformer D, with $\Phi, \Psi = -62.8°, -138.1°$, is stabilized by a hydrogen bond between O6' and O1. Conformer F has $\Phi, \Psi = -167°, -175°$. Table 2 shows the importance of individual starting conformations to the final points for the whole map as well as for the region enclosed by the 8 kcal/mol contour lines. Only the *gt* orientations of the hydroxymethyl groups contribute to the map points with relative energies <8 kcal/mol, although several combinations of initial orientations of the secondary hydroxyl groups were used.

The shapes of the *O*-sophorose map[16] and the *C*-sophorose map are similar, although differences are observed. The minima found on the *O*-disaccharide map correspond to *C*-disaccharide conformers A, D, B and E. Conformer D on the *O*-sophorose map lies in the same low energy region as minimum A, less than 1 kcal/mol higher in energy. For *C*-sophorose, D is a shallow depression with a high relative energy. Another difference between the two maps is that the staggered conformers C, F and G do not occur for *O*-sophorose. For *O*-sophorose, conformer F is on a slowly descending plateau while the other two conformers C and G do not exist at all within the 8 kcal/mol contour line. For *C*-sophorose, the A and B minima are shifted to $\Phi,\Psi$ values that are close to staggered conformations; for *O*-sophorose, $\Phi$ angles for A and B are -72° and -86° respectively, showing that the main perpendicularly intersecting trough is shifted from the staggered conformation.

Although the shapes of the surfaces are similar, the structures having relative energies lower than 8 kcal/mol cover 62% of the map for *C*-sophorose while they represent only 45% of the map for *O*-sophorose. The energy barrier in the $\Phi$ direction is less than 5 kcal/mol for the *C*-disaccharide while it is larger than 7 kcal/mol for *O*-sophorose. The above two results suggest that *C*-sophorose has greater flexibility.

**C-Laminarabioside** - The adiabatic map of *C*-laminarabioside (β-D-Glc*p*-$CH_2$ (1→3) β-D-Glc*p*-OMe) generated by MM3 is shown in Fig. 3, and the details describing

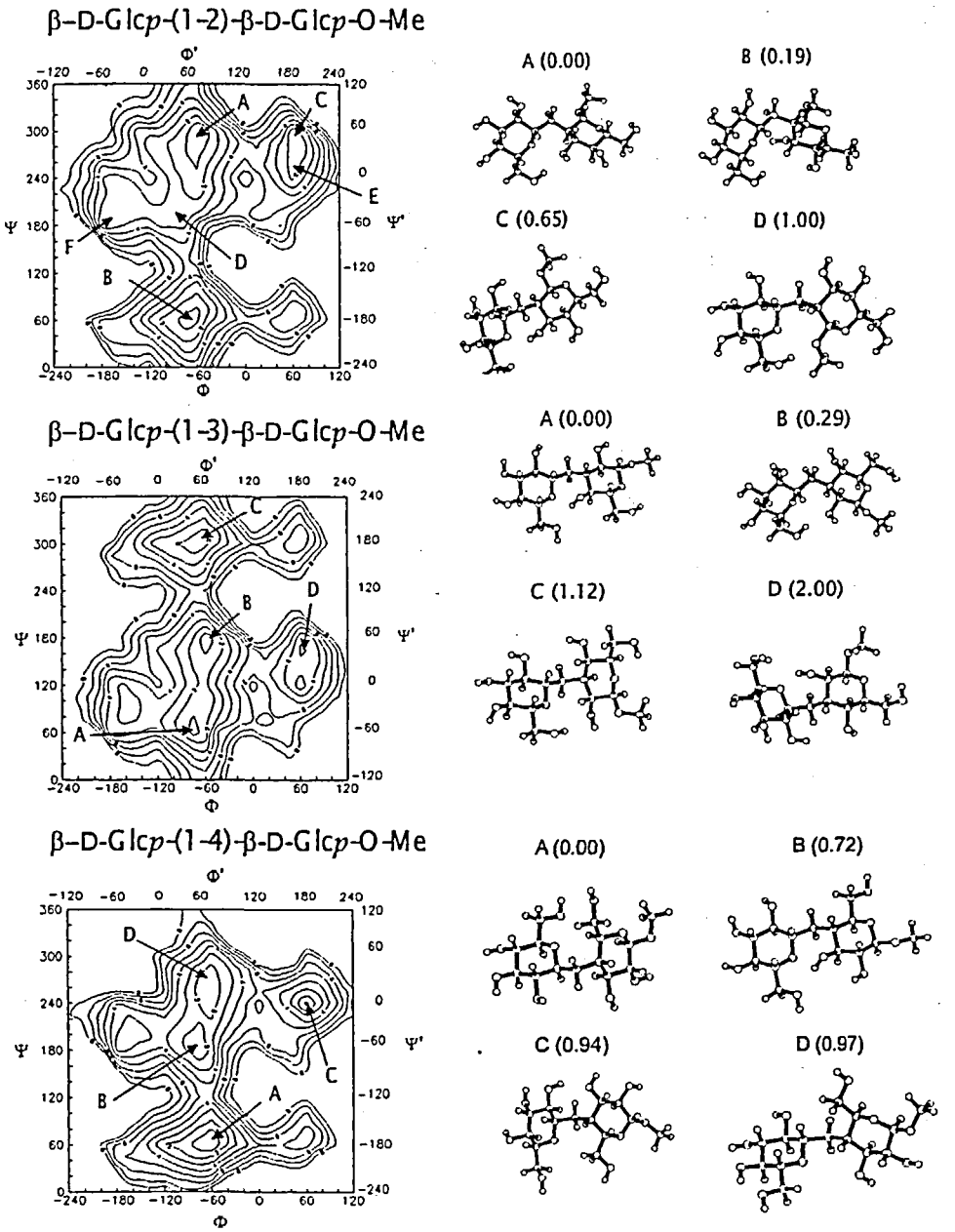



**Figure 3.** Adiabatic relaxed map of *C*-sophorose, *C*-laminarabiose and *C*-cellobiose in the MM3(92) force field as a function of the Φ and Ψ glycosidic torsion angles. For details see Fig 2.

the geometry of the optimized minima are listed in Table 1. The three lowest energy minima A, B and C are located in a region parallel to the Φ axis centered around Φ = -60°, as would be expected if the exo-anomeric effect were operative. Two minima are located in a region less than 2 kcal/mol above the global minimum parallel to the Ψ axis between 60° and 180°; the global minimum A at Φ = -81.4° and Ψ = 62.3° is stabilized by a hydrogen bond between O4 and O6'; conformer B occurs at Φ, Ψ = -56.4°, 177.9° with only slightly higher energy (0.29 kcal/mol). The third minimum, C, is located at Φ, Ψ = -79.2°, -60.7°. Conformer D has relatively higher energy (1.12 kcal/mol) with Φ, Ψ = 73.9°, 164°. Conformer E with Φ,Ψ = 159.4°, 101.8° is stabilized by a hydrogen bond between O2' and O2. The minima F and G found in the same trough as the D conformer, parallel also to the Ψ axis centered around Φ = 60° and have relatively high energies (> 2 kcal/mol); conformer F has an almost eclipsed conformation with Ψ = 119.9 (Φ= 60.4°), and is stabilized by a hydrogen bond between O4 and O2'. Conformer G has Φ, Ψ = 62.9°, -42.2°, and is stabilized by hydrogen bond between O5' and O4.

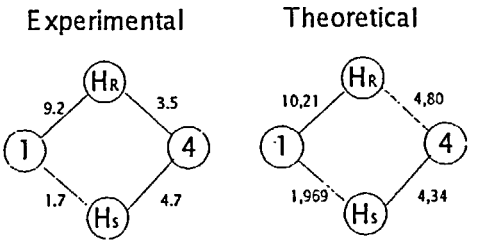


Table 2 shows the statistics concerning the individual starting conformations contributing to the final points for the whole map as well as for the region enclosed by the 8 kcal/mol contour lines. Almost all of the final map points, with relative energies < 8 kcal/mol adopt the *gt* orientation for both hydroxymethyl groups. Again all four combinations of secondary hydroxyl groups contribute to the final map.

In the map of *O*-laminarabiose,[16] there are also two perpendicularly intersecting low energy troughs and minima A, B, C and F are located in similar regions. In *C*-laminarabiose, the additional minima that occur are shifted to staggered conformations in regions where *O*-laminarabiose has a relatively high energy plateau (E) or does not have minima (D, G). For *C*-laminarabiose, the conformations with Erel < 8 kcal/mol represent 60% of the map, while in *O*-laminrobiose, they represent only 35% of the map.

Energy barriers between the lower energy intersecting trough (Φ = -60°) and the higher energy one (Φ = 60°) are lower for *C*-laminarabiose (<6 kcal/mol) than for *O*-laminarabiose (> 8kcal/mol). The above data suggest that there is greater flexibility in the Φ direction for *C*-laminarabiose than for the *O*-disaccharide.

**_C_-Cellobioside** - The adiabatic map of *C*-cellobioside (β-D-Glc*p*-$CH_2$ (1→4) β–D-Glc*p*-OMe) is shown in Fig. 3 along with four of the optimized minima. The location




**Scheme 4**

of the minima and the energy differences are also summarized in Table 1. As seen for both of the above equatorial - equatorial linked disacharides, the map has two perpendicular low energy troughs along the Ψ axis. The global minimum is at a perfect staggered conformation with Φ, Ψ = -63.5°, 63.9°. Two other low energy minima B and D are located in a low energy region between Ψ 180° and 300°. Conformer B is at Φ, Ψ = -80°, -179.6°, stabilized by a hydrogen bond between O3 and O6', while conformer D is at Φ,Ψ = -63.5°, -99.0°. Two other low energy conformers are found along the Ψ axis with Φ ≈60°. Conformer C, with low relative energy, is at Φ = 59.8° with Ψ eclipsed (-119.7°) and is stabilized by a hydrogen bond between O4 and O2'. The less stable conformer F, at Φ, Ψ = 61.4°, 81.3° is in the same Ψ direction as C and has a hydrogen bond between O5' and O3. Finally, conformer E is located at Φ,Ψ = -159°, -158° and is stabilized by a hydrogen bond between O6 and O2'.

Different starting conformations contribute to the final *C*-cellobioside map (Table 2). The *gtgt* and *tggt* starting orientations of the hydroxymethyl groups occur more frequently. The points with Erel < 8 kcal/mol cover 56% of the total map and the *gtgt* and *tggt* initial orientations contribute 70%. All four combinations of secondary hydroxyl groups contribute to the final map with the *rr* and *cr* orientations occurring most frequently. The four different homonuclear $^3J$ coupling constants between the Cα protons and H1' and H4 respectively were calculated as summarized in Table 3. They are in (Scheme 4) quite good agreement with the experimental values, [5b] indicating that the theoretical analysis successfully predicts the experimental data.

The *O*-cellobiose map[17] also exhibits the same general shape but the conformation around the *C*-aglyconic bond is rather different from that described for *O*-cellobiose. The

central lowest energy region for *O*-cellobiose contains conformers that correspond to B and D in *C*-cellobiose. However, their energies are clearly higher (>2 kcal/mol) than that for conformer A, the global minimum in *C*-cellobiose. Therefore, these calculations indicate that there are notable differences between the *C*- and *O*- 1→4 equatorial-equatorial linked disaccharides and these differences are similar to those previously reported for *C*- and *O*- lactose.[6] Additionally, the *O*-cellobiose map contains fewer energy minima; conformer E is on a descending plateau while conformer F does not exist within the 8 kcal/mol contour line. The points of the map with Erel < 8 kcal/mol represent only 30% of the total for *O*-cellobiose. The energy barriers along the Ψ axis, e.g., between A and B conformers, are less than 5 kcal/mol in *C*-cellobiose while in *O*-cellobiose the corresponding energy barrier is ≈ 6 kcal/mol. The same is true for conformational changes along the Φ axis where the energy barrier between conformer C and D is less than 5 kcal/mol in *C*-cellobiose while the corresponding energy barrier in *O*-cellobiose is about 1 kcal/mol higher. The above results indicate that conformational changes in *C*-disaccharides are easier and *C*-cellobiose is more flexible than *O*-cellobiose.



The general map characteristics are similar for all three equatorial - equatorial linked disaccharides. Two low energy troughs occur on each map and energy barriers between the troughs are similar and lower than 6 kcal/mol. The low energy trough is centered at Φ ≈ -60°, in agreement with the geometry predicted by the *exo*-anomeric effect. Common differences with the corresponding *O*-disaccharides for all three *C*-glucosides are the number of minima in the staggered conformations, the lower energy barriers between conformers, and the higher percentages of Φ,Ψ space within the 8 kcal/mol contours. Significant differences are observed in the case of the 1→4 linkage where the global minima are different for the *C*- and *O*- disaccharides.

### THREE BOND LINKED *C*-DISACCHARIDES

***C*-isomaltoside** - Three dimensional energy maps of *C*-isomaltoside (α-D-Glc*p*-$CH_2$ (1→6) β-D-Glc*p*-OMe) are shown in Fig. 4. All of the points on the maps having energies less than a) 1 kcal/mol, b) 3 kcal/mol, c) 5 kcal/mol and d) 8 kcal/mol above the global minimum are shown in the four 3D presentations. The three corresponding contour diagrams are shown in Fig. 5 for staggered conformations of Ω along with the corresponding minima. Eleven minima have energies within the 8 kcal/mol contour, and the exact location of each is summarized in Table 4.

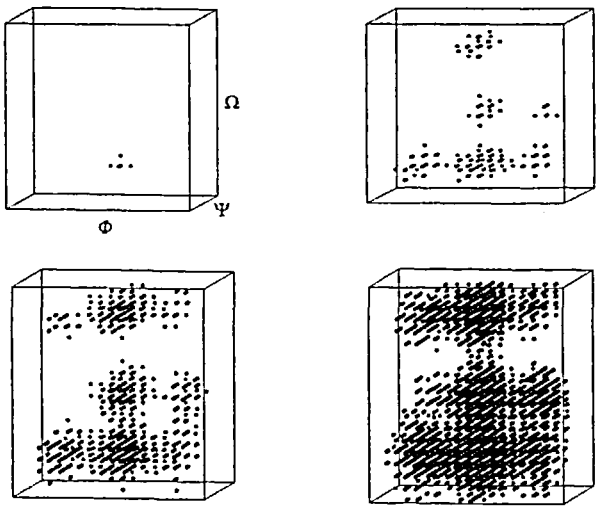



**Figure 4.** Relaxed-residue map of *C*-isomaltoside where the distributions of the points having energy a) lower than 1 kcal/mol b) lower than 3 kcal/mol c) lower than 5 kcal/mol d) lower than 8 kcal/mol are shown as a function of Φ, Ψ and Ω. The Ψ and Ω axis run from 0° to 360° and Φ runs from -120° to 240°.



All the minima have perfectly staggered conformations about the three bonds described by the Φ, Ψ and Ω dihedrals. The Φ angles for the three lowest energy conformers (A, B, I) are ≈ 60°, in agreement with what would be predicted by the *exo*-anomeric effect. Other minima have Φ angles at near 180° (C, D, F, G). The minima found have the angle Ω in staggered orientations of 60° (*gt*), -60 (*gg*) and 180° (*tg*) in order of increasing energy. The global minimum is significantly lower in energy than the second minimum (1 kcal/mol), indicating that the population of the region around this conformer is considerably higher than the others. The structures having less than 8 kcal/mol relative energy represent the 39.1% of the 5832 points that contribute to the construction of the final map. Table 6 gives statistics about the contributions of the initial conformers to the adiabatic energy surfaces. Seventy-eight % of the conformers have the hydroxymethyl group of the non reducing ring in the *gt* orientation. However, a considerable percentage of conformers (21 %) adopts the *tg* orientation. The preferred orientation of the secondary hydroxyl groups is *rr* (76%).

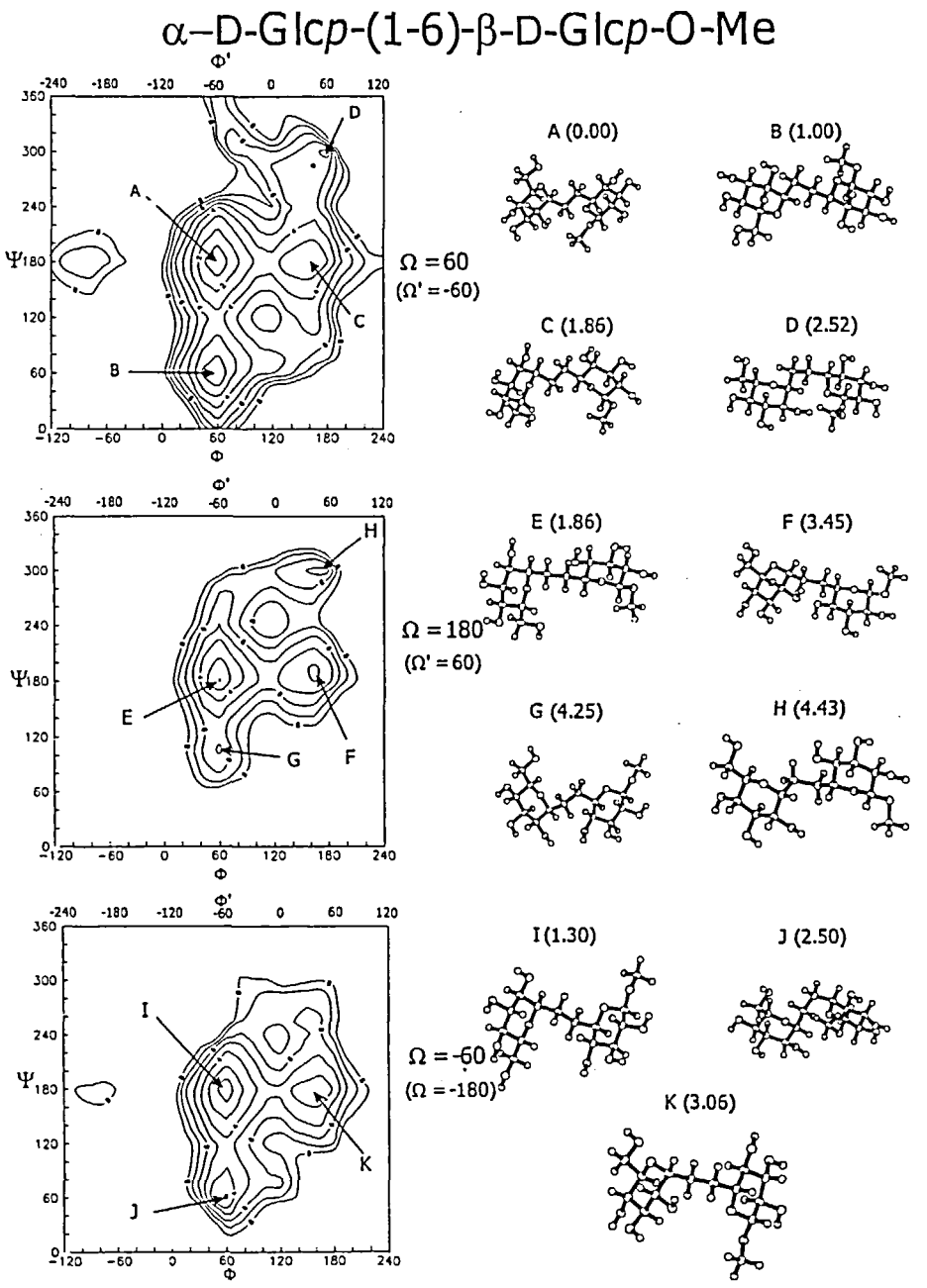


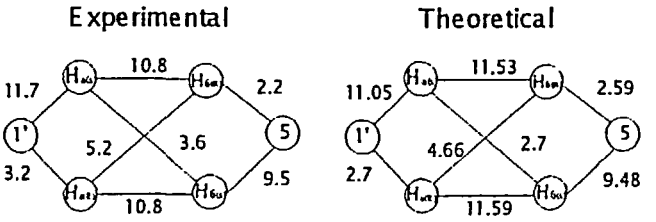

**Figure 5.** Contour maps for *C*-isomaltoside at constant Ω: (a) Ω = 60°, (b) Ω = 180°, (c) Ω = -60°. The Φ, Ψ planes are contoured at 1 kcal/mol increments to 8 kcal/mol above the global minimum. The low energy conformers A-I indicated on the maps are represented by ball and stick models. Φ' and Ω' stands for H1-C1'-Cα-C6 and H5-C5-C6-Cα, respectively.

**Table 4.** MM3-calculated minima for *C*-isomaltoside and *C*-gentiobioside. Comparison with the *O*-glycosides[17]

| Disaccharide | | *C*-Glycosides | | | | *O*-Glycosides | | | |
|---|---|---|---|---|---|---|---|---|---|
| | conformer | Location of Minimum | | | Rel. Energy | Location of Minimum | | | Rel. Energy |
| | | Φ | Ψ | Ω | | Φ | Ψ | Ω | |
| Gentiobiose | A | -62.6 | -179.7 | 62.2 | 0.00 | -77.5 | 178.4 | 68.9 | 0.00 |
| | B | -61.4 | -79.5 | 54.3 | 0.72 | -75.5 | -83.9 | 81.3 | 1.85 |
| | C | 58.7 | 179.8 | 62.0 | 1.00 | 64.3 | -168.8 | 69.8 | 3.51 |
| | J | -62.4 | 179.3 | -57.5 | 1.11 | -77.4 | 179.2 | -64.4 | 0.47 |
| | D | -61.4 | -79.5 | 60.9 | 1.20 | | | | |
| | G | -59.8 | -174.7 | 176.5 | 1.38 | -76.4 | 177.3 | 173.3 | 1.75 |
| | H | -61.7 | -57.4 | 177.7 | 1.68 | -84.8 | -82.3 | 179.0 | 2.09 |
| | E | -164.3 | 179.6 | 64.1 | 1.69 | -88.9 | 74.9 | 63.8 | 1.75 |
| | F | -177.4 | -57.4 | -60.7 | 1.73 | | | | |
| | K | 59.0 | 179.1 | -56.6 | 2.10 | 63.9 | -168.6 | -63.7 | 4.06 |
| | I | 63.4 | -162.4 | 176.5 | 2.53 | 60.1 | -178.1 | 168.4 | 5.34 |
| | L | -177.1 | 161.8 | -62.8 | 2.67 | | | | |
| | M | -63.6 | -97.3 | -60.2 | 3.25 | 68.6 | -71.5 | 178.5 | 7.42 |
| Isomaltose | A | 57.5 | -179.7 | 62.3 | 0.00 | 75.8 | -173.8 | 69.5 | 0.00 |
| | B | 58.4 | 61.8 | 58.4 | 1.00 | 78.5 | 83.1 | 66.0 | 1.71 |
| | I | 57.6 | 179.6 | -57.3 | 1.30 | 76.8 | -175.1 | -62.1 | 0.12 |
| | C | 157.9 | -179.7 | 62.1 | 1.86 | | | | |
| | E | 60.3 | -176.9 | 177.9 | 1.86 | 76.5 | 179.7 | 180.9 | 1.23 |
| | J | 55.5 | 64.5 | -65.5 | 2.50 | 75.4 | 100.4 | -64.6 | 2.79 |
| | D | 171.6 | -57.0 | 73.0 | 2.52 | | | | |
| | K | 158.4 | 178.5 | -58.4 | 3.06 | | | | |
| | F | 161.2 | -175.1 | 176.1 | 3.45 | | | | |
| | G | 159.1 | -57.9 | 178.5 | 4.25 | | | | |
| | H | 61.8 | 99.6 | 178.5 | 4.43 | | | | |
| | | | | | | 90.8 | -69.7 | 80.6 | 2.118 |

**Table 5.** Calculated glycosidic coupling constants for *C*-isomaltoside and *C*-gentiobioside.

| Disaccharide | Coupling Constants | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $J_{\Phi(R)}$ | $J_{\Phi(S)}$ | $J_{\Psi(RR)}$ | $J_{\Psi(RS)}$ | $J_{\Psi(SR)}$ | $J_{\Psi(SS)}$ | $J_{\Omega(R)}$ | $J_{\Omega(S)}$ |
| Isomaltose | 2.70 | 11.05 | 4.66 | 11.59 | 11.53 | 2.70 | 2.59 | 9.48 |
| Gentiobiose | 9.59 | 2.95 | 4.94 | 10.07 | 10.05 | 4.40 | 2.80 | 9.80 |




**Scheme 5**

The eight $^3J$ coupling constants characterizing the three bond angles were calculated. They are summarized in Table 5 and are shown in Scheme 5 in juxtaposition with experimental values given by Kishi *et al.*[4a] The theoretical values are in good agreement with the experimental data showing that the conformational analysis reflects the equilibrium existing between the different structures in solution.

Compared with the corresponding map of the *O*-disaccharide,[19] the general shape is similar. However, the global minimum on the *O*-isomaltose map is located at $\Omega$= -170°. The second minimum on the *O*-isomaltose map is only slightly higher in energy than the first minimum (0.2 kcal/mol). It is found at $\Omega$ = 57° and is quite similar to the global minimum of *C*-isomaltose. The conformers found in the *C*-glycoside map with $\Phi$ = 180° are not present on the *O*-disaccharide map. Instead, there are only descending plateaus in the same regions. For *O*-isomaltose, the structures found within the 8 kcal/mol contour represent only 18.8% of the total and the energy barriers in the $\Omega$ direction are lower (<4 kcal/mol) than for *C*-isomaltose (< 5kcal/mol). In the *O*-disaccharide, the $\Phi$ angle for all low energy conformers have values between 74° and 102°, altered from staggered

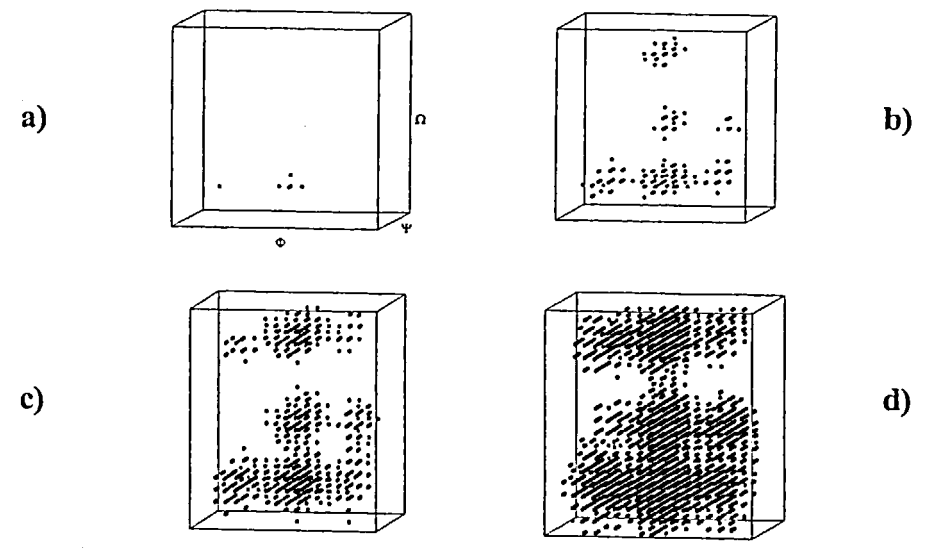



**Figure 6.** Relaxed-residue maps of *C*-gentiobioside where the distribution of the points having energy a) lower than 1 kcal/mol, b) lower than 3 kcal/mol, c) lower than 5 kcal/mol, d) lower than 8 kcal/mol are shown as a function of Φ, Ψ and Ω. The Ψ and Ω axis run from 0° to 360° and Φ runs from -120° to 240°.



geometry because of the *exo*-anomeric effect. In *C*-isomaltose, the Φ angle is always close to staggered. Unlike *O*-isomaltose, conformers with Φ = 180° are also populated, showing that *C*-isomaltose is more flexible along this bond. No major differences are observed concerning the flexibility along the Ψ axis.

***C*-gentiobioside** - The points contributing to the final 3D map of the *C*-gentiobioside (β-D-Glc*p*-$CH_2$ (1→6) β–D-Glc*p*-OMe) are shown in Fig. 6. The three contour diagrams in Fig. 7 show the energy surface at the three staggered conformations of angle Ω. Thirteen minima were found within the 8 kcal/mol energy region and their geometric properties are summarized in Table 4. The majority of the structures with Ω = 60° have relative energies lower than 8 kcal/mol. The global minimum as well as the three other low energy minima are located in this Ω region. The two lowest energy conformers have Φ = -60° in conformity with the *exo*-anomeric effect; however, minima with energy less than 2 kcal/mol are located in regions with Φ = 60° and Ψ = 180°. The

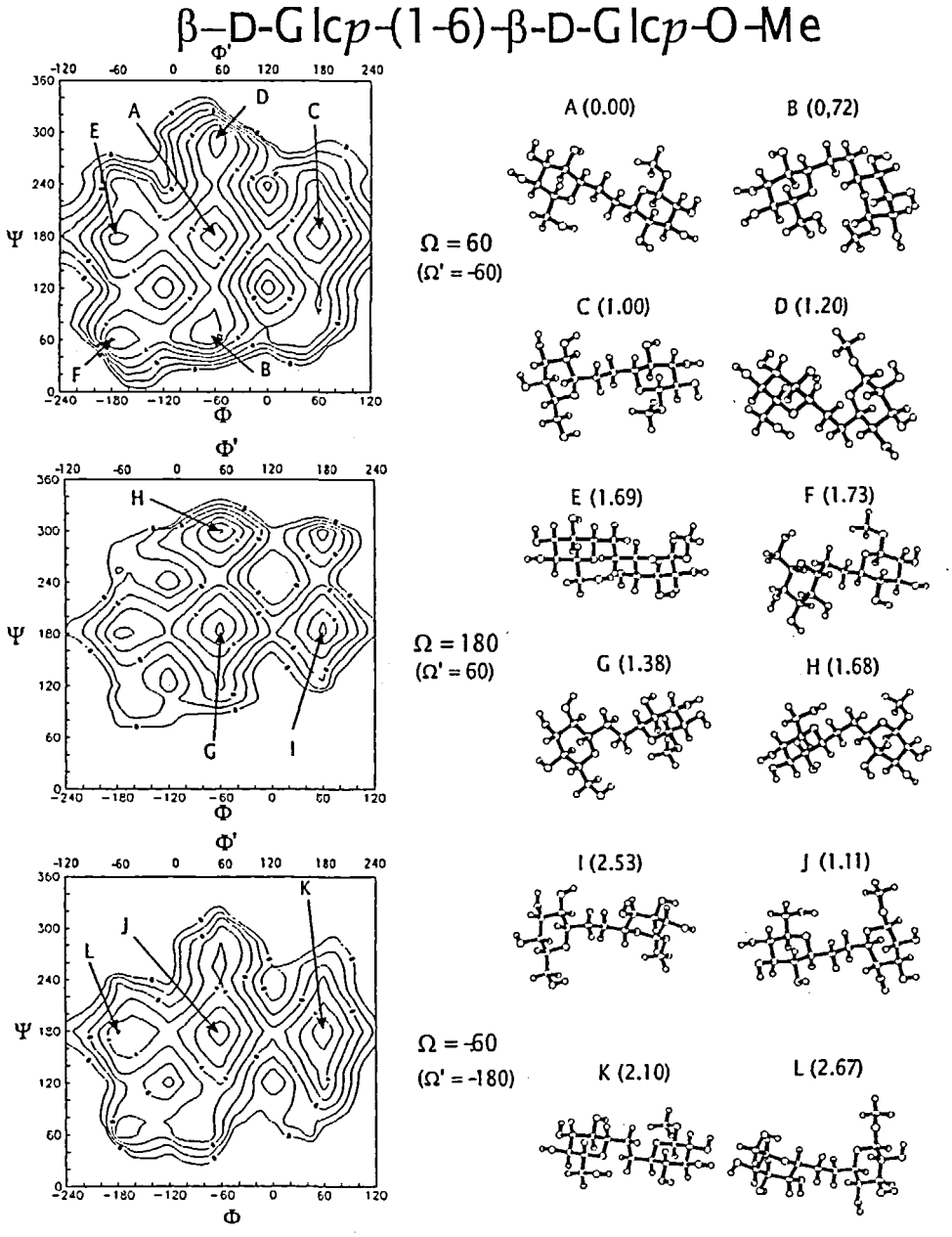
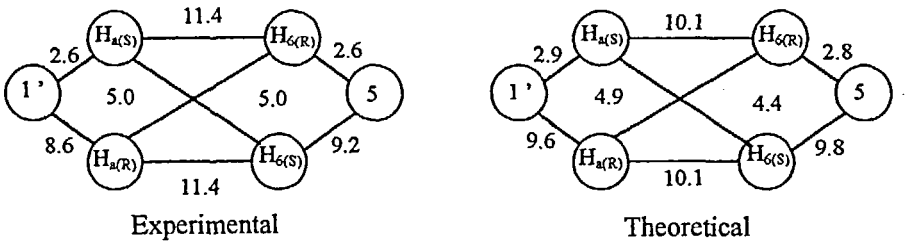



**Figure 7.** Contour maps for *C*-gentiobioside at constant Ω: (a) Ω = 60°, (b) Ω = 180°, (c) Ω = -60°. The Φ, Ψ planes are contoured at 1 kcal/mol increments to 8 kcal/mol above the global minimum. The low energy conformers A-I indicated on the maps are represented by ball and stick models.

**Table 6.** Starting structures contributing to MM3-generated relaxed-residue Φ,Ψ,Ω disaccharide maps.

| Disaccharide | Complete Map | | | | | Lowest 8 kcal/mol | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Hydroxy-Methyl | | Secondary Hydroxyls | | % Area of map | Hydroxy-methyl | | Secondary Hydroxyls | |
| Isomaltose | *gg* | 3.63% | *cc* | 3.02% | 39.1% | *gg* | 3.67% | *Cc* | 3.00% |
| | *gt* | 84.16% | *cr* | 13.49% | | *gt* | 84.12% | *Cr* | 13.52% |
| | *tg* | 12.20% | *rc* | 2.93% | | *tg* | 12.21% | *Rc* | 2.89% |
| | | | *rr* | 80.55% | | | | *Rr* | 80.58% |
| Gentiobiose | *gg* | 4.68% | *cc* | 4.71% | 59.6% | *gg* | 5.27% | *Cc* | 5.10% |
| | *gt* | 63.37% | *cr* | 28.56% | | *gt* | 65.39% | *Cr* | 28.47% |
| | *tg* | 29.50% | *rc* | 5.39% | | *tg* | 29.33% | *Rc* | 5.77% |
| | | | *rr* | 61.34% | | | | *Rr* | 60.66% |




**Scheme 6**

order of increasing energy in the Ω direction is *gt*, *gg* and *tg*. The orientation of the hydroxymethyl group of the non reducing ring that contributes mainly to the final map is the *gt* (Table 6) and the structures having relative energies lower than 8 kcal/mol represent 60% of the whole 3D map.

The energy barriers along the Φ axis are less than 4 kcal/mol, while along the Ω axis are less than 5 kcal/mol (Ω~0°). The eight $^3J$ coupling constants between protons of the three *C*-glycosidic bonds were calculated (Table 5). The theoretical values together with the experimental data given by Kishi *et al.*[5a] are shown in Scheme 6. They are in reasonable agreement showing that the conformational preferences in solution are well described by the calculated Φ/Ψ/Ω map.

The global minimum for *O*-gentiobiose is located in the corresponding region on the *C*-disaccharide hypersurface and the general shape of the *O*-disaccharide maps are similar to those of *C*-gentiobiose. The differences observed are the shifts of the values calculated for the Φ angles of the minima away from the perfectly staggered values present in *O*-gentiobiose, and the existence of minima for the *C*-disaccharide in regions where in the *O*-disaccharide there are only descending plateaus.

Neuman *et al.* have reported a crystal structure for *C*-gentiobioside.[20] Its location Φ, Ψ, Ω = 55.9°, 175.1°,-63.9° agrees with conformer K which is 2.1 kcal/mol higher in energy than the lowest MM3 minimum. In this crystal structure, the hydroxymethyl group of the non-reducing ring is in the *gg* conformation, while in the K conformer the hydroxymethyl group is in the *gt* arrangement. The crystal structure of *C*-gentiobioside involves an extensive network of intermolecular hydrogen bonds resulting in the optimum mode of packing with each molecule surrounded by 10 neighbors. This dense crystal structure may explain the differences with the conformations preferred in solution.

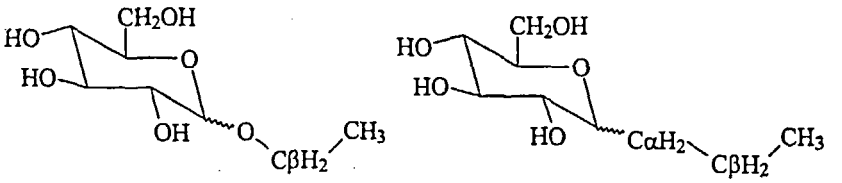


According to the calculations, both *C*-isomaltose and *C*-gentiobiose should have significant populations in at least two of the three Ω potential wells. *C*-isomaltose seems to be more conformationally restricted in Φ and Ψ, as was the case for the axial-equatorial two-bond-linked disaccharides, with a preference for the *exo* anomeric orientation about Φ. On the other hand, *C*-gentiobiose is calculated to be rather flexible in Φ and Ψ. In comparison with the *O*-disaccharides, the same main differences are observed as for the other disaccharides. In the case of the 1→6 linked disaccharides, where the distances between the two rings are bigger and the steric interactions are less severe the differences are more easily understood.

In *O*-glycosides, the minima are found in similar regions but they are shifted from the staggered conformations. In *C*-disaccharides, additional minima occur in regions where only relatively high energy plateaus exist on the *O*-glycoside maps. These additional minima do not conform to the *exo*-anomeric orientation of the Φ bond indicating agreater conformational flexibility for C-C bonds than for C-O bonds in glycosidic sequences.

These differences exist already in the MM3 energy profile of the Φ bond in the case of compounds **9** and **10**. The energy profiles of ethyl glycopyranosides **9** as a function of the Φ angle for α- and β- configurations are represented in Figure 8 and

a)




9 10

b)

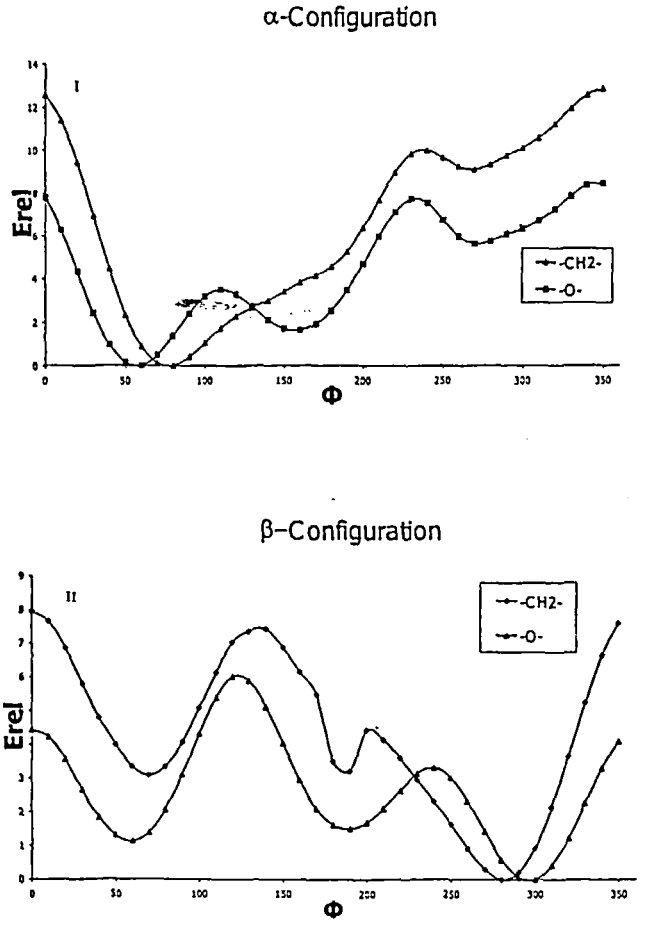



**Figure 8.** a) compounds **9** and **10**. b) Steric energy as a function of the Φ dihedral angle for -OEt glucoside **9** (Φ = O5-C1-O-Cβ) and the -$CH_2$Et analogue **10** (Φ = O5-C1-Cα-Cβ). In (I) and in (II) the profiles of the α and β configurations are compared.

compared with the corresponding profiles for the -$CH_2$- linked analogue **10**. The *O*-linked α- and β- glucosides exhibit global minima at Φ = 80° and -80°, respectively, while the $CH_2$-linked analogues have minima at staggered conformations Φ = 60° and -60°. The energy differences between the low energy conformers are also larger for the *O*-monosacchrides compared to their $CH_2$-linked analogues.

## CONCLUSIONS

MM3-based relaxed maps of *C*-kojibioside, *C*-nigeroside, *C*-maltoside, *C*-sophoroside *C*-laminarabioside, *C*-cellobioside, *C*-isomaltoside and *C*-gentiobioside have been calculated. The axial-equatorial two-bond-linked molecules have similar map characteristics and the same is true for the equatorial-equatorial two-bond linked, molecules which are seen as being rather flexible. The relative positions of the lower energy regions are in agreement with the *exo*-anomeric effect. Although the shapes of the the *C*- and the *O*-disaccharide maps are similar, differences are noted in the exact location, the relative energy and the number of the minima. In the *O*-disaccharides, the parametrization for the O-C-O-C torsion angle for the glycosidic bond results in deviation from staggered orientations and relative rigidity of this bond in the *O*-linked molecules. A more pronounced difference between the *O*- and the *C*- exists in the cellobiose case, extending the difference observed in β(1→4) linked disaccharides by Jimenez-Barbero *et al.* for lactose. [6a] Limitations affected the calculations, in particular, in the maltose case, because of the use of ε = 4, which diminishes the strength of electrostatic interactions such as hydrogen bonding. Nevertheless, MM3 still produces good geometries for the low energy minima, which appears to be important in determining the populations of conformers in solution.



## ACKNOWLEDGMENTS

This work was supported by a scholarship awarded to G. L. by the Greek Ministry of Education (Program EPEAEK)

## REFERENCES AND NOTES

1. Associated with the University Joseph Fourier, Grenoble, France. (mikros@pharm.uoa.gr and Serge.Perez@cermav.cnrs.fr
2. D. Rouzaud and P. Sinaÿ, *J. Chem. Soc., Chem. Commun.*, **104**, 6468 (1982)

3. P. Vogel, R. Ferritto, K. Kraehenbuehl and A. Baudat, *Carbohydrate Mimics*, Y. Chapleur, Ed.; Wiley-VCH, 1998, p 20.
4. N. Ishida, K. Kumagai, T. Niida, T. Tsuruokan H. Yumoto, *J. Antibiot.*, **20**, 66 (1967); b) D. L. Taylor, L. E. Fellows, G. H. Farrar, R. J. Nash, D. Taylor-Robinson, M. A. Mobberley, T. A. Ryder, D. J. Jeffries and A. S. Tyms, *Antiviral Res.*, **10**, 11 (1988); c) Y. Nishimura, T. Satoh, H. Adachi, S. Kondo, T. Takeuchi, M. Azetaka, H. Fukuyasu and Y. Iizuka, *J. Am. Chem. Soc.*, **118**, 3051 (1996); d) P. M. Wassarman, *Science*, **235**, 553 (1987); e) H. Bischoff, *Eur. J. Clin. Invest.*, **3**, 3 (1994); f) P. Layer, G. L. Carlson and E. P. Di-Magno, *Gastroenterology*, **88**, 1895 (1985).
5. a) P. G. Goekjian, T. C. Wu and Y. Kishi, *J. Org. Chem.*, **56**, 6422 (1991) and references cited therein; b) Y. Wang, P. G. Goekijan, D. M. Ryckman, W. H. Miller, S. A. Babirad and Y. Kishi, *J. Org. Chem.*, **57**, 482 (1992) and references cited therein; c) T. Haneda, P. G. Goekjian, S. H. Kim and Y. Kishi, *J. Org. Chem.*, **57**, 490 (1992); d) D. J. O'Leary and Y. Kishi, *J. Org. Chem.*, **58**, 304 (1993); e) A. Wei and Y. Kishi, *J. Org. Chem.*, **59**, 88 (1994); f) A. Wei, A. Haudrechy, C. Audin, H-S. Jun, N. Haudrechy-Bretel and Y. Kishi, *J. Org. Chem.*, **60**, 2161 (1995).
6. a) J. F. Espinosa, M. Martín-Pastor, J. L. Asensio, H. Dietrich, M. Martin-Lomas, R. R. Schmidt and J. Jiménez-Barbero, *Tetrahedron Lett.*, **36**, 6329 (1995); b) J. F. Espinosa, H. Dietrich, M. Martín-Lomas, R. R. Schmidt and J. Jiménez-Barbero, *Tetrahedron Lett.*, **17**, 1467 (1996).
7. a) J. F. Espinosa, F. J. Cañada, J. L. Asensio, M. Martín-Pastor, H. Dietrich, M. Martín-Lomas, R. R. Schmidt and J. Jiménez-Barbero, *J. Am. Chem. Soc.*, **118**, 10862 (1996); b) J. F. Espinosa, E. Montero, A. Vian, J. L. Garcia, H. Dietrich, R. R. Schmidt, M. Martín-Lomas, A. Imberty, F. J. Cañada and J. Jiménez-Barbero, *J. Am. Chem. Soc.*, **118**, 1309 (1998).
8. IUPAC – IUBMB : Nomenclature of Carbohydrates (Recommendations 1996) *Carbohydr. Res.*, **297**, 1 (1997)
9. The corresponding dihedral angles formed between the hydrogen atoms of each bond are also defined as following : Φ(R) = H1-C1'-Cα-HαR, Φ(S) = H1'-C1'-Cα-HαS, Ψ(R)=HαR-Cα-Ci-Hi, Ψ(S)=HαS-Cα-Ci-Hi.
10. The corresponding dihedral angles between the hydrogen atoms of the glycosidic linkage are: Φ(R)=H1'-C1'-Cα–HαR, Φ(S)=H1'-C1'-Cα-HαS, Ψ(RR)=HαR-Cα-C6-H6R,Ψ(RS)=HαR-Cα-C6-H6S,Ψ(SR)=HαS-Cα-C6-H6R, Ψ(SS)=HαS-Cα-C6-H6S, Ω(R)=H5-C5-C6-H6R, Ω(S)=H5-C5-C6-H6S.
11. a) N. L. Allinger, Y. H. Yuh and J. H. Lii, *J. Am. Chem. Soc.*, **111**, 8551 (1989); b) J. H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, **111**, 8566 (1989); c) N.L. Allinger, M. Rahman and J. H. Lii, *J. Am. Chem. Soc.*, **112**, 8293 (1990).
12. S. Pérez, A. Imberty, S. B. Engelsen, J. Gruza, K. Mazeau, J. Jiménez-Barbero, A. Poveda, J. F. Espinoza, B. Van Eyck, G. Johnson, A. D. French, M. L. Kouwijzer, P. D. J. Grootenhuis, A. Bernadi, L. Raimondi, H. Senderowitz,V. Durier, G. Vergoten and K. Rasmussen, *Carbohydr. Res.*, **314**, 141 (1998).
13. a) C. van Alsenoy, A. D. French, M. Cao, S. Q. Newton and L. Schäfer, *J. Am. Chem.*, **116**, 9520 (1994); b) G. P. Rockwell and T. B. Grindley, *J. Am. Chem. Soc.*, **120**, 10953 (1998).
14. A. D. French, R. S. Rowland and N. L. Allinger, *Computer Modeling of Carbohydrate Molecules"*, A.C.S. Symposium Series **430**, A. D. French and J. W. Brady Eds., ACS: Washington, 1990, p 120.

15. M. Karplus, *J. Chem. Phys.*, **30**, 11 (1959).
16. C. A. G. Haasnoot, F. A. A. M. de Leuw and C. Altona, *Tetrahedron*, **36**, 2783 (1980).
17. M. K. Dowd, J. Zeng, A. D. French and P. J. Reilly, *Carbohydr. Res.*, **230**, 223 (1992).
18. M. K. Dowd, A. D. French and P. J. Reilly, *Carbohydr. Res.*, **233**, 15 (1992).
19. M. K. Dowd, P. J. Reilly and A. D. French, *Biopolymers*, **34**, 625 (1994).
20. A. Neuman, F. Longchambon, O. Abbes, H. Gillier-Pandraud, S. Pérez, D. Rouzaud and P. Sinay, *Carbohydr. Res.*, **195**, 187 (1990).