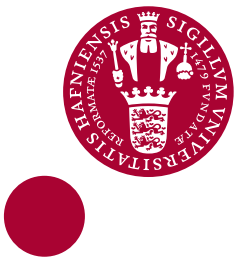
UNIVERSITY OF COPENHAGEN



Conformations of disaccharides by empirical force field calculations.

Engelsen, Søren Balling; Rasmussen, Kjeld

Published in:
International journal of biological macromol

DOI:
0141-8130/93/010056-07

Publication date:
1993

Document version
Early version, also known as pre-print

Citation for published version (APA):
Engelsen, S. B., & Rasmussen, K. (1993). Conformations of disaccharides by empirical force field calculations. *International journal of biological macromol*, *15*(1), 56-62. https://doi.org/0141-8130/93/010056-07

Download date: 23. apr.. 2021

Conformations of disaccharides by empirical force field calculations. Part V: Conformational maps of β-gentiobiose in an optimized consistent force field

Søren Balling Engelsen and Kjeld Rasmussen*

Chemistry Department A, The Technical University of Denmark, DK-2800 Lyngby, Denmark

(Received 3 August 1992; revised 20 October 1992)

A recently optimized set of potential energy functions is used to investigate the conformational flexibility of the β-(1 → 6) glycosidic linkage in β-gentiobiose. Relaxed Ramachandran maps in vacuo *are presented in the torsional angles φ and ω, with torsional angle ψ allowed to relax freely, as are all other internal degrees of freedom. The study reveals two almost iso-energetic low energy domains in (φ,ψ,ω) space, and cross-sections in the low-energy domains at ω = −60° and ω = 60° show that more than 60% and 70% respectively of the area of the conformational maps are accessible within 40 kJ mol⁻¹. The molecular structure in the crystal, including the exoanomeric effect, is well reproduced. The structure belongs to the potential energy well which includes the global potential energy minimum* in vacuo. *The most profound structural difference between the crystal structure and the calculated global minimum* in vacuo *is the 20° deviation of the ψ torsional angle (24° from perfect* trans*). It occurs in the most flexible glucosidic degree of freedom, ψ, and is caused by optimization of the hydrogen bonding network and not by the exoanomeric effect.*

Keywords: Molecular mechanics; optimized force field; β-gentiobiose; Ramachandran maps; conformational analysis; anomeric effect; exo-anomeric effect

Introduction

β-Gentiobiose, or β-D-glucopyranosyl-(1 → 6)-β-D-glucopyranose, is a disaccharide which is found in the roots of gentians and, what is more important, in an evolutionarily conservative kernel in cell walls of a vast number of microorganisms[1].

Using β-gentiobiose as a test case for new potential energy functions has a strategic importance to us since it is our aim to refine our force fields on glucolipids and glucoconjugates in order to apply them to lipid A, important as a fundamental structural part of cell membranes, as an endotoxin and as an evolutionarily conservative anchor of the lipopolysaccharides, and thereby of great importance to cell recognition.

Therefore β-gentiobiose is well suited as a model for studying the large conformational flexibility of the β-(1 → 6) glycosidic linkage which is of importance as a branching point in many polysaccharides, giving the polysaccharides a much more dense structure suitable as compact yet accessible depots for glucose. β-Gentiobiose has received some attention from conformational analysts and structural chemists[2–7].

Force field

This is the very first application of a newly optimized set of potential energy functions PEF91L[8,9] intended for use on glycolipids. This force field uses Morse potentials for modelling covalent bonds, Lennard-Jones 12-6 potentials for modelling non-bonded van der Waals forces, and a simple coulombic term for modelling electrostatic interactions. The latter is effected by placing partial charges in the centres of the nuclei and using a dielectric constant D of 2.0†. Partial charge assignment and neutralization are made by a charge redistribution algorithm which results in the following net charge assignments in disaccharides: hydroxyl oxygen O = −0.2008, ether oxygen O′ = −0.0956, methine carbon C = 0.0055, methylene carbon C = −0.1572, anomeric carbon C′ = −0.1261, neighbouring methine carbon to the anomeric carbon C = 0.0327, hydrogen H(O) = 0.0748 and H(C) = 0.1456. Intramolecular interactions separated by three or more bonds are considered non-bonded.

To complement the too crude model of only bonded and non-bonded interactions, and in order to reproduce physical rigidity and behaviour in normal modes, it is necessary to add more terms. We use a harmonic valence angle potential and a Pitzer torsional angle potential. The torsional potential was parametrized using the bond torsional approach (see *Table 1*).

†The choice of value for the dielectric constant is seldom commented on. When a value of 1 is used, electrostatic interactions usually become completely dominant, which is unacceptable. When a value of 10 is used they are negligible. For many years, values of 3.0 and 3.5 were used. The actual value is unimportant when optimization has been performed as is the case here.

*To whom correspondence should be addressed.

0141-8130/93/010056-07

© 1993 Butterworth-Heinemann Limited

Table 1 Potential energy function parameters

Covalent bonds:	D_e	α	b_0
C-C	88.2	2.2545	1.5185
C-C′	88.2	2.2545	1.5185
C-O	90.0	2.5998	1.4245
C-O′	90.0	2.2431	1.4057
C-H	101.6	1.8139	1.0950
C′-O	90.0	2.6569	1.4061
C′-O′	90.0	2.2431	1.3815
C′-H	101.6	1.8139	1.0950
O-H	100.0	2.3533	0.8636

Angle deformations:	K_θ	θ_0
C-C-C	86.9780	109
C′-C-C	86.9780	109
O-C-C	122.7495	109
O′-C-C	122.7381	109
O-C-C′	122.7404	109
O′-C-C′	122.7381	109
C-C-H	87.1579	109
C′-C-H	87.1579	109
O-C-H	79.4840	109
O′-C-H	79.4446	109
H-C-H	74.0717	109
C-C′-C	86.9780	109
O-C′-C	122.7469	109
O′-C′-C	122.7381	109
C-C′-H	87.1579	109
O-C′-O	142.3717	109
O′-C′-O′	142.3717	109
O-C′-O′	142.3647	109
O-C′-H	79.4840	109
O′-C′-H	98.6649	109
H-C′-H	74.0717	109
C-O′-C	117.1070	1.8659
C-O′-C′	135.7123	1.9221
C′-O′-C′	145.4899	1.8636
C′-O-H	80.1629	1.9861
C-O-H	80.1629	1.9861
H-O-H	55.4181	1.8251

Torsional terms:	K_ϕ	$k = 3$ in all cases
C-C-C-C	1.6067	
C-C-C-H	0.2508	
C-C-C-O	1.1719	
C-C-C-O′	1.2165	
H-C-C-H	0.2216	
O-C-C-H	0.0001	
O′-C-C-H	0.0001	
O-C-C-O	0.7015	
O′-C-C-O′	3.4654	
O-C-C-O′	2.0954	
C′-C-C-H	0.2508	
C′-C-C-O	1.1719	
C′-C-C-O′	1.2165	
C′-C-C-C	1.6067	
C-C′-C-H	0.2508	
H-C′-C-H	0.2216	
O-C′-C-H	0.1100	
O′-C′-C-H	0.0001	
H-C′-C-O	1.0748	
H-C′-C-O′	0.6700	
H-C′-C-C	0.2508	
O-C′-C-O	1.3353	
O′-C′-C-O	1.2000	
O-C′-C-O′	1.3114	
O-C′-C-C	1.1719	
O′-C′-C-C	1.2165	
C-C′-C-C	1.5707	
C-O′-C-H	0.3898	

Table 1 *continued*

Torsional terms:	K_ϕ	$k = 3$ in all cases
C-O′-C-C	0.8070	
C′-O′-C-H	0.3898	
C′-O′-C-C	0.8070	
H-O-C-C′	0.1166	
H-O-C-H	0.1089	
H-O-C-C	0.0927	
C-O′-C′-H	0.4177	
C-O′-C′-C	0.2130	
C-O′-C′-O	1.6612	
C-O′-C′-O′	3.2479	
C′-O′-C′-H	1.0485	
C′-O′-C′-C	0.8582	
C′-O′-C′-O	3.3795	
C′-O′-C′-O′	3.5872	
H-O-C′-H	0.1089	
H-O-C′-C	1.0428	
H-O-C′-O	0.0001	
H-O-C′-O′	0.7116	

Non-bonded interactions:	A	B	e
C--	1100.351	32.1094	0.0934
C′--	1100.351	32.1094	0.0537
O--	654.097	26.5119	−0.1837
O′--	659.733	16.8374	−0.0785
H--	73.3218	2.3730	0.1627

Atomic symbols are: H hydrogen, C aliphatic (sp3) carbon, C′ anomeric (sp3) carbon, O′ ether oxygen and O hydroxyl oxygen.
Units: D_e kcal mol^{-1}, α Å^{-1}, b_0 Å, K_θ kcal mol^{-1} rad^{-2}, θ_0 rad, K_ϕ kcal mol^{-1}, A (kcal mol^{-1} Å^{12})$^{1/2}$, B (kcal mol^{-1} Å^6)$^{1/2}$, e elementary charge. These are the units demanded by the CFF program. When $\theta_0 = 109$, a switch assigns the value in radians of $\arccos(-\frac{1}{3})$. See text for details of the assignment of charges

The total potential energy function is shown in equation (1), and the parameters with their units are given in *Table 1*. The factor $F = 4.184 \times 332.09 \times 100$ converts from elementary charge to kJ mol^{-1} pm.

$$E_{pot} = \sum_{N_b} D_e[e^{-2\alpha(b-b_0)} - 2e^{-\alpha(b-b_0)}] + \sum_{i>j}\left[\frac{A_iA_j}{r^{12}} - \frac{B_iB_j}{r^6} + \frac{Fe_ie_j}{Dr}\right] + \sum_{N_\theta}\frac{1}{2}K_\theta(\theta-\theta_0)^2 + \sum_{N_\phi}\frac{1}{2}K_\phi(1+\cos(k\phi)) \quad (1)$$

The new set of potential energy functions PEF91L was parametrized using the Consistent Force Field (CFF) program[8–12]. The method involves semiautomatic least squares fitting simultaneously on six fundamentally different types of experimental data on functional groups in series of model compounds. The present set of potential energy functions PEF91L has been optimized[8,9,12] on internal coordinates, non-bonded distances, unit cell dimensions, lattice energies, dipole moments and vibrational frequencies of series of alkanes, ethers, alcohols and carbohydrates both in condensed phase and *in vacuo*.

An important and integral part of the consistent force field is optimization of the model on vibrational data available from i.r. and Raman spectroscopy. Only

through repeated and careful identification of vibrational frequencies followed by force field optimization is it possible to obtain the experimental structural rigidity which is of major importance to the calculation of thermodynamic properties. In this respect, carbohydrates constitute a special problem because of the lack of interpretable vibrational data, wherefore a consistent force field for carbohydrates can only be developed by combining optimizations on data from functional groups of small model compounds such as hydrocarbons, ethers and alcohols. So the present PEF91L may not be the ultimate solution, even for the present choice of functional forms. We intend to re-optimize the parameter set when new data become available.

A problem one faces when attempting to parametrize a consistent force field for carbohydrates is that the gas phase structures will almost certainly differ from the condensed phase structures which we want to use for optimization. The difference originates from the intermolecular hydrogen bonding structure within the crystals, opposed to the tendency to create intramolecular hydrogen bonds in the gas phase. For these reasons it is of major importance to simulate the substances in condensed phase, as they were studied experimentally, in order not to obtain a potential of mean force having the condensed phase packing implicit in the parameters. The latter was the case when, around 1980, one of us parametrized force fields for carbohydrates[13,14]. In PEF91L all optimizations on condensed phase data were performed exclusively using a Williams variant of the Ewald lattice summation[11,15].

Two important issues were considered before attempting to parametrize sp^3 oxygen for ethers and alcohols. The first issue is whether or not to include lone pairs explicitly. In order not to introduce difficulties associated with the mass-less lone pairs into normal mode analysis, and thus introduce additional complexity into the force field, we have chosen a compromise. By treating the values of the tetrahedral angles about oxygen (C-O-C and C-O-H) as variable parameters, as opposed to the use of the tetrahedral angle as a fixed parameter around tetrahedral carbons (see *Table 1*), lone pairs are included implicitly in the potential energy functions. This approximation will still include the non-spherical defects of the Lennard-Jones potential; it has previously been shown to give quite reasonable results[11,13].

The next issue is how to deal with stereo-electronic effects. The anomeric effect which was first recognized as something peculiar to the anomeric carbon (C1) of sugars, is a consequence of geminal oxygens, and it is due to resonance interaction between lone pair electrons on closely spaced functional groups. The anomeric effect is thus an electronic effect which we have no hope of reproducing with molecular mechanics methods without special treatment. The rather large deviation in the anomeric C-O bond length in simple force fields using the same parameters for ordinary carbon and anomeric carbon has caused us to experiment with a unique anomeric carbon atom[14,16]. This relatively crude approach to dealing with anomeric effects is carried on in PEF91L, despite the large number of additional parameters and the consequently increased optimization efforts.

It must be emphasized that PEF91L represents our first attempt to create a truly vacuum type force field and that the complexity and computational challenge of using condensed phase simulation in the optimization process are formidable. Non-constrained energy minimization with retention of full crystal symmetry has to be performed on each substance in each optimization iteration, and the computer time increases with approximately the third power of the non-bonded cut-off limit. Although we have a very stable and convergent crystal simulator in CFF even when all unit cell dimensions are allowed to relax, it is a fact that the choice of method to simulate infinity (Ewald summation, periodic boundaries), the choice of method to neutralize molecules and the size of the cut-off sphere as well as the cut-off method all seem to affect the outcoming result. For this reason one should use a molecular mechanics force field which is consistent with the condensed phase simulator to which it has been developed.

Methodology

All force field development was made with the Lyngby version of CFF, conceived by Lifson and Warshel[17] and further developed and documented by Niketić and Rasmussen[10,11]. The final force field was transported to an as yet unpublished PC-based molecular mechanics program named MoleCast (MOLEcular CAlculations and STereo drawings)[18].†

In MoleCast, as in CFF, the molecular topology is generated automatically from a pseudo chemical formula, and rigid as well as relaxed Ramachandran maps of torsional angles can be recorded using a single command line specifying the internal numbers of the torsions in a input stream file. Relaxed Ramachandran maps were recorded in two ways: (1) by the use of rigid geometric rotations followed by conjugate gradients energy minimization, and (2) by using the harmonic drag, from the conjugate gradients energy minimization, to change the glycosidic conformation from one grid point to another. In both cases energy minimization was performed unconstrained except for the three glycosidic torsions, stopping only when the RMS gradient was less than 10^{-3} J mol^{-1} pm^{-1}. The torsions of interest were fixed by a harmonic constraint potential, equation (2), with a force constant of 2000 kJ mol^{-1} rad^{-2}. As initial structure for each map we used the X-ray structure reported by Rohrer *et al.*[5], relaxed in PEF91L.

$$E_{con} = \sum_{N_c} K_{con}(\phi - \phi_{con})^2 \qquad (2)$$

All contour plots were made with a commercial program[19] which allows for drawing of contour maps from irregularly spaced data.

The relaxed Ramachandran-type maps of β-gentiobiose were recorded using a torsional grid increment of 10 degrees, implying that 1369 (37 × 37) different conformations were energy minimized with all internal degrees of freedom allowed to relax in order to create one relaxed map.‡

†MoleCast, written in Pascal, is intended as a teaching as well as an analytical tool, lacking only optimization capabilities. The program was written from scratch, with the benefit of valuable experience from the CFF-coding, but incorporating no CFF-code segments, and MoleCast is now a valuable help in cross-checking results and new codings as well as for use as an analytical tool.

‡On our AST 486/33 PC, energy minimization of one conformer of β-gentiobiose took 5–10 min and accordingly a complete relaxed map required 5–10 days, whereas rigid maps could be calculated in approximately 15 min.

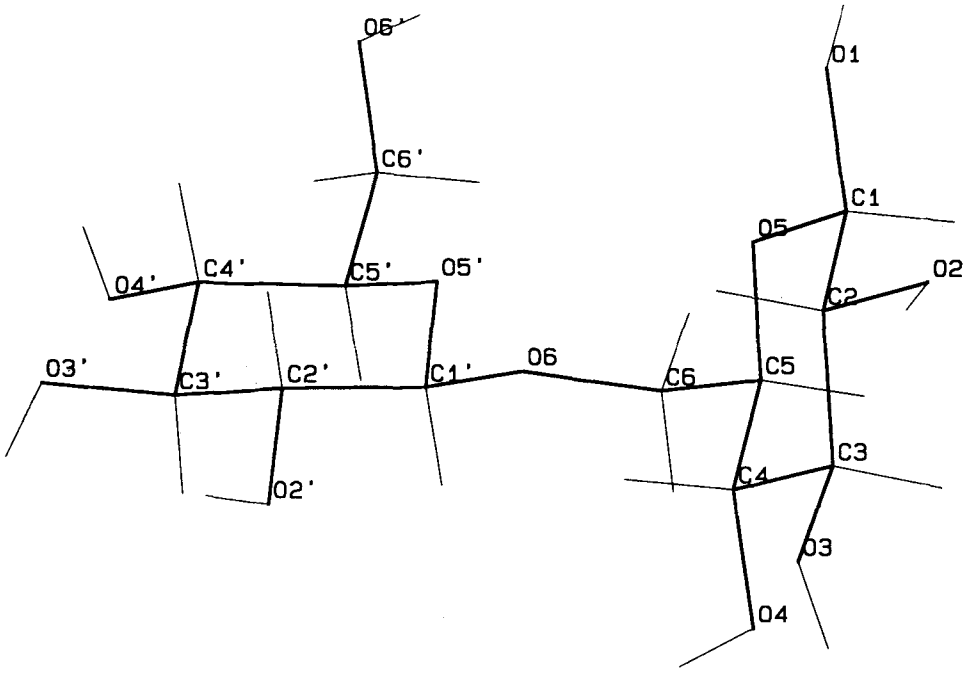


Figure 1 The crystal structure conformation of β-gentiobiose relaxed in PEF91L

When recording Ramachandran maps of disaccharides with the intention of finding the global energy minimum one has to remember that for each combination of ϕ, ψ and ω (see *Figure 1*) β-gentiobiose has nine exo-cyclic torsions giving $3^9 = 19\,683$ possible conformations, not to mention alternative ring conformations. The problem is well known to mathematicians as the multiple minima problem, for which no-one has yet derived a convergent algorithm to find the global minimum of a general function. Ramachandran maps also reflect the fact that we have yet to invent ways to visualize the effects of more than two or three variables. Until the multiple minima problem is solved, relaxed Ramachandran maps probably represent the best tool for investigating the complexity of the potential energy surface of disaccharides. This problem was discussed at some length in earlier publications[20,21].

Two types of relaxed conformational maps were calculated using harmonic constraints on two of the three glucosidic torsions, and one rigid conformational map was calculated for comparison. The three torsions in the glucosidic linkage are defined as: ϕ: O5′-C1′-O6-C6, ψ: C1′-O6-C6-C5 and ω: O6-C6-C5-O5 (see *Figure 1*). These definitions conform to those of Neuman *et al.*[7], but differ from those previously used[2].

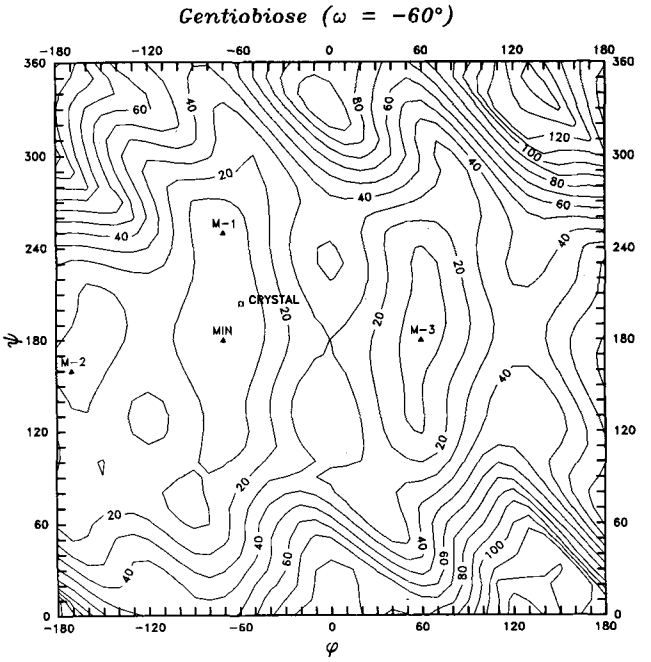


Figure 2 Relaxed conformational map of β-gentiobiose as a function of the torsional angles ϕ and ψ, with the torsional angle ω constrained to $-60°$. Iso-energy contours are drawn in 10 kJ mol^{-1} intervals

Results and discussion

Relaxed (ϕ,ψ) maps ($\omega = -60$, 60 and 180)

Figures 2, 3 and *4* show the conformational (ϕ,ψ) maps with ω constrained in perfect staggered positions. The global and other minima are shown, as well as the (ϕ,ψ) values of the crystal conformation. When this is minimized it converges towards the global minimum MIN.

These maps were recorded for comparison with a similar study by Neuman *et al.*[7] who used a force field which includes hydrogen bonds explicitly and gives special treatment to the exo-anomeric effect. In this case we use a different procedure to record the relaxed maps. The minimized coordinates from one grid point were used as the initial set for the following grid point and no rigid rotations were performed when going from one grid point to another letting the harmonic drag change the glycosidic conformation.

When comparing the three maps we find that they all have three low energy domains centred about $\phi = -170°$, $\phi = -70°$ and $\phi = 60°$. Of the three low energy domains the domain centred about $\phi = -70°$ spans more than half of the ψ cycle, and the $\omega = 60°$ map shows almost perfect threefold symmetry. The lowest energy minima

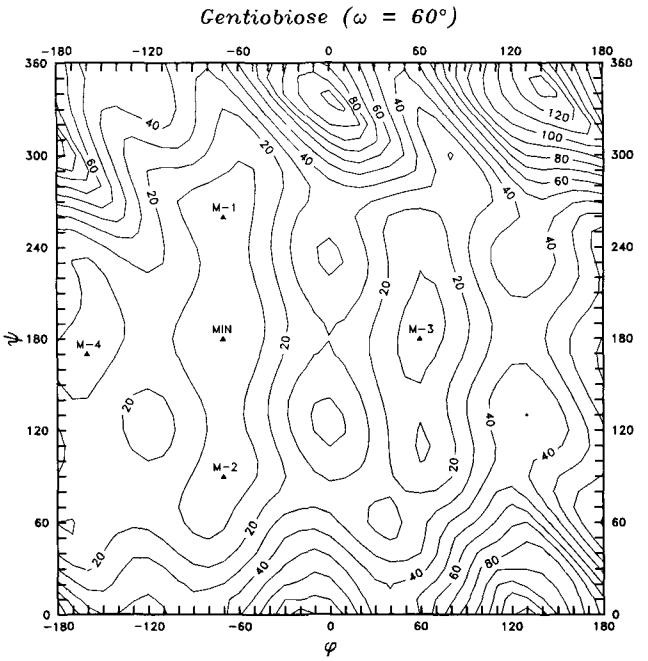


Figure 3 Relaxed conformational map of β-gentiobiose as a function of the torsional angles ϕ and ψ, with the torsional angle ω constrained to 60°. Iso-energy contours are drawn in 10 kJ mol^{-1} intervals

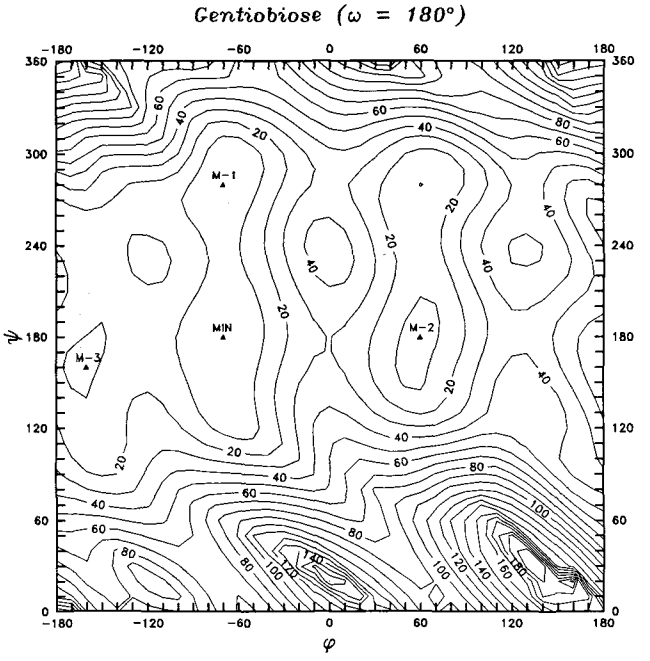


Figure 4 Relaxed conformational map of β-gentiobiose as a function of the torsional angles ϕ and ψ, with the torsional angle ω constrained to 180°. Iso-energy contours are drawn in 10 kJ mol^{-1} intervals

in (ϕ,ψ) space are, independent of ω, placed close to the point $(\phi,\psi) = (-60°, 180°)$ with the minimum $(\phi,\psi,\omega) = (-70°, 180°, -60°)$ being the global minimum, only 0.4 kJ mol^{-1} lower than the corresponding $\omega = 180°$ minimum and 1.2 kJ mol^{-1} lower than the corresponding $\omega = 60°$ minimum.

In the $\omega = 60°$ and $\omega = -60°$ maps more than 70% and 60% of the (ϕ,ψ) space, respectively, is accessible within 40 kJ mol^{-1} of the global minimum, demonstrating the flexible nature of the glycosidic linkage. The $\omega = 180°$ map is more restricted with only little more than 50% of the (ϕ,ψ) space accessible within 40 kJ mol^{-1}.

On the $\omega = -60°$ map the barrier heights when going from the global minimum MIN to the M-1 minimum and the M-3 minimum respectively (*Figure 2*) are calculated as 4.3 kJ mol^{-1} and 30.0 kJ mol^{-1}. On the $\omega = 60°$ map the barrier heights when going from the global minimum MIN to the M-1 minimum and the M-3 minimum respectively (*Figure 3*) are calculated as 5.7 kJ mol^{-1} and 29.8 kJ mol^{-1}. The corresponding barrier heights on the $\omega = 180°$ map (*Figure 4*) are 7.9 kJ mol^{-1} (MIN to M-1) and 30.3 kJ mol^{-1} (MIN to M-2).

Relaxed (ϕ,ω) map $(\psi \sim 180)$

In order to study a cross-section of the previously recorded (ϕ,ψ) maps (*Figures 2–4*), we found it interesting to study the variation of ϕ and ω, with ψ left completely relaxed in the *trans* conformation. The resulting map (*Figure 5*) reveals highly symmetric iso-energy contours with the three lowest energy domains all centred about $\phi = -70°$ in a perfect threefold symmetry.

The global minimum is again found at $\omega = -60°$ in full agreement with the previously recorded (ϕ,ψ) maps but differs by 1.5 kJ mol^{-1} in energy, a difference which can easily be attributed to the now 'released' ψ torsional angle and to differences in the recording method. The three lowest lying energy minima at $\phi = -70°$ are reproduced in the same order as in the (ϕ,ψ) maps, having energy differences of less than 0.4 kJ mol^{-1}. We can only conclude that β-gentiobiose *in vacuo* has three almost iso-energetic low-energy domains centred around the glycosidic points $(\phi,\omega) = (-70°, -60°)$, $(\phi,\omega) = (-70°, 60°)$ and $(\phi,\omega) = (-70°, 180°)$ with the minimum at $\omega = -60°$ being the global minimum. The glycosidic torsional angle ψ is the most flexible having a flat minimum around the unstrained *trans* conformation.

The barrier height between the global minimum MIN and the other *gauche* conformation M-2 (*Figure 5*) is

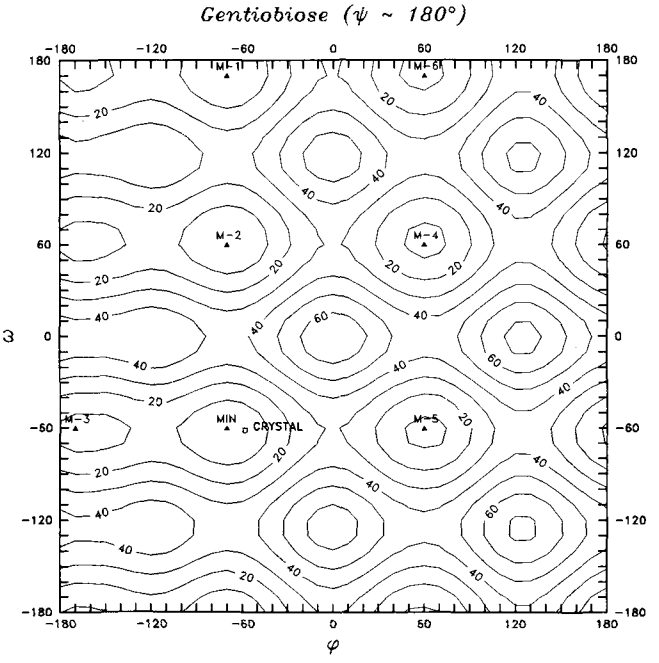


Figure 5 Relaxed conformational map of β-gentiobiose as a function of the torsional angles ϕ and ω, with the torsional angle ψ fully relaxed in the *trans* conformation. Iso-energy contours are drawn in 10 kJ mol^{-1} intervals

Table 2 Crystal structure conformation of β-gentiobiose calculated with different potential energy functions

	ϕ	ψ	ω	O1C1O5	C1O1	C1'O6	C1'O6C6	O6C1'O5'	H1'---H_R	H1'---H_S
Crystal										
Aréne *et al.*[4]	−59	−155	−60	107	140	139	113	104		
Rohrer *et al.*[5]	−58	−156	−62	108	139	139	113	107	240	312
PEF91L (this work)	−61	−160	−65	109	141	140	116	110	235	325
'Gas'										
PEF91L (this work)	−69	−176	−62	109	141	140	117	110	238	300
PEFAC1[16]	−60	−178	−67	109	141	141	114	110	242	308
PEF300[11]	−73	−177	−62	109	142	143	113	109	228	350
PEF400[2]	−35	−155	−88	109	142	143	115	111	266	356
Rohrer *et al.*[5]	−140	−156	−62							
Bock and Vignon[6]	−60	−170	−60							
Neuman *et al.*[7]	−75	−175	−60						232	308

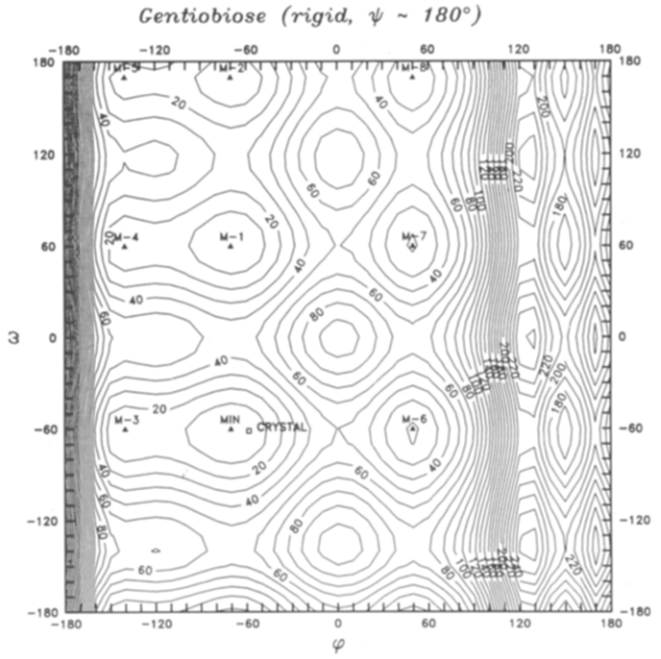


Figure 6 Rigid conformational map of β-gentiobiose as a function of the torsional angles ϕ and ω, with the torsional angle ψ in the *trans* conformation. Iso-energy contours are drawn in 10 kJ mol^{-1} intervals

calculated as 37 kJ mol^{-1}, whereas the barrier to the *trans* conformation M-1 is calculated as 35 kJ mol^{-1}, facilitating almost equally frequent MIN to M-2 transformations as MIN to M-1 transformations. The third barrier between the M-1 minimum and the M-2 minimum is calculated as 26 kJ mol^{-1}. Barriers from the $\phi = -70°$ minima to the $\phi = 60°$ minima are all calculated to be close to 30 kJ mol^{-1}, whereas the barriers from the $\phi = -70°$ minima to the $\phi = -170°$ minima are all calculated to be close to 14 kJ mol^{-1}. Thus the conformational freedom of the ϕ torsional angle is larger than that of the ω torsional angle.

Rigid (ϕ,ω) map ($\psi \sim 180$)

The rigid (ϕ,ω) map of β-gentiobiose (*Figure 6*) was calculated as a comparative study and was not expected to differ fundamentally from the relaxed map (*Figure 5*) due to the free and undisturbed rotation about ϕ and ω when keeping ψ in the *trans* conformation close to 180°.

A comparison with the relaxed map reveals that the three lowest energy minima at $\phi = -70°$ are reproduced, whereas the minima previously found at $\phi = -170°$ have moved to $\phi = -140°$ due to the unreleased strain in the structures. The energies of the three lowest energy minima are very well reproduced (within 2 kJ mol^{-1}), only the $\omega = 60°$ and the $\omega = 180°$ minima have swapped order. The global minimum at $\omega = -60°$ is reproduced within 0.1 kJ mol^{-1} when compared to the relaxed map, an extraordinary fine result due to the absence of steric barriers. Only about 30% of the map area lies within 40 kJ mol^{-1} of the global minimum compared to about 70% for the relaxed map. Even for non-steric rotations, as presented here, the rigid map exposes the deficiencies connected with the recording of rigid maps, and accordingly rigid maps should be used only for preliminary investigations.

The crystal structure

When the crystal structure conformation of β-gentiobiose[5], (ϕ,ψ,ω) = (−58.3°, −156.3°, −61.5°), is allowed to relax in PEF91L *in vacuo*, the glycosidic conformation changes to (ϕ,ψ,ω) = (−68.6°, −175.7°, −61.9°), belonging to the potential energy well around the global minimum. However, the glycosidic conformation changes only slightly when the crystal structure is allowed to relax in PEF91L retaining full crystal symmetry. In this case the glucosidic conformation changes to (ϕ,ψ,ω) = (−61.4°, −160.4°, −64.8°), and the orthorhombic unit cell dimensions are altered from (a,b,c) = (886.9 pm, 2284.6 pm, 720.1 pm) to (a,b,c) = (910.0 pm, 2350.4 pm, 748.5 pm), expanding the unit cell volume by 9.7%.

The error is most probably due to deficiencies in the parameters, particularly in respect to reproducing the hydrogen bonding network. For alkanes, the corresponding errors are about the same size, but with reversed sign. Therefore we aim at performing a re-optimization when supplementary data become available, probably using a slightly different function for non-bonded interactions.

The results of minimizations of the energy of the crystal structure conformation emphasize the importance of the use of an optimized set of energy function parameters as well as convergent energy minimization in the crystal phase[15]. *Table 2* shows a comparison of calculations* made with different potential energy functions. PEF400

*The crystal calculation using PEF91L was performed on an IBM RS 6000 320 and took about 1 h.

and PEF300 were 'handmade', with parameters adapted by trial and error. PEFAC1 was partly optimized; it was a step towards our new PEF91L using harmonic bond functions and special treatment of the anomeric carbon atom, but not discriminating between the two types of oxygen. The force fields of Rohrer *et al.*[5], only non-bonded interactions, and of Bock and Vignon[6], stiff residues and hard spheres, are far more primitive.

The anomeric and the exo-anomeric effects

The calculated anomeric C1-O distances are in good agreement with experimental data. The anomeric C1-OH bond is calculated as 141.1 pm compared to the measured 139.3 pm and the C1′-O6 of the glucosidic linkage is calculated as 139.8 pm compared to the measured 139.0 pm. The ψ torsional angle of the glucosidic linkage is observed to deviate 24° from the *trans* conformation. This has been attributed to the exo-anomeric effect caused by non-bonded interactions between the atoms O5′ and one of the H(C6)[5]. The torsion angles C6-O6-C1′-O5′ and C6-O6-C1′-C2′ defining the rotational position of the glycosidic linkage are calculated as $-68.6°$ $(-58.3°)_{obs}$ and $170.1°$ $(-176.4°)_{obs}$ respectively, a conformation considered preferable when the exo-anomeric effect is present. Likewise, the torsional angles O5-C1-O1-H(O1) and C2-C1-O1-H(O1) defining the rotational position of the free glucosidic hydroxyl C1-O1 are calculated as $-63.2°$ $(-56.1°)_{obs}$ and $175.4°$ $(-173.4°)_{obs}$ respectively. These results indicate to us that the anomeric as well as the exo-anomeric effects are reasonably well reproduced by PEF91L.

The rather large geometric but small energetic deviation from perfect *trans* conformation of ψ as measured in the crystal is found to be caused by the crystal packing and not by the exo-anomeric effect. This conclusion is strongly supported by the crystal simulation where ψ is only slightly altered from the crystal conformation. The very flexible nature of the glycosidic linkage especially in the ($\phi = -70°$) plane and the fact that the crystal packing involves an extensive intermolecular hydrogen bonding network (only the OH(2′) hydroxyl, the bridge and the ring oxygens do not form hydrogen bonds in the crystal) further supports the conclusion. Neuman *et al.*[7] reached a similar conclusion when comparing calculated structures of β-gentiobiose and the related 'methyl C-gentiobioside' which lacks the exo-anomeric effect.

Previous work[2,3,6,7,11] has not investigated the total potential energy surface in such detail as to allow us to conclude whether or not the picture described above represents a change from earlier descriptions using non-optimized energy parameters.

Conclusion

Development of a consistent force field is a continuing process: when new data appear or small deficiencies are revealed the force field has to be revised. A step-by-step process is necessary to keep integrity with experimental data, and molecular dynamics simulation of β-gentiobiose in water seems to us to be a natural continuation of this study. PEF91L with its Morse functions appears to be a relative 'flat' force field which holds some promise for future attempts to relate the MD simulations to n.m.r. data, since results have indicated that currently existing force fields are energetically too restricted.

Acknowledgements

The work reported in this paper was part of the PhD study of S. Balling Engelsen. The following foundations and organizations contributed towards financing it: The Technical University of Denmark, The Danish Research Academy, The Carlsberg Foundation, The Jorck Foundation, Thomas B. Thriges Foundation and The COWI Foundation. All contributions were essential and are gratefully acknowledged. We would also like to thank Dr Al French, Southern Regional Research Center, New Orleans (Louisiana, USA), for valuable discussions and hints on how to record Ramachandran maps.

References

1 Lüderitz, O., Freudenberg, M. A., Galanos, C., Lehmann, V., Rietschel, E. T. and Shaw, D. H. *Current Topics in Membranes and Transport* 1982, **17**, 79
2 Melberg, S. and Rasmussen, Kj. *Carbohydr. Res.* 1980, **78**, 215
3 Rasmussen, Kj. in 'Molecular Structure and Dynamics', (Ed. M. Balaban), Balaban, Jerusalem, 1980, p 171
4 Arène, F., Neuman, A. and Longchambon, F. *Compt. rend.* 1979, **C288**, 331
5 Rohrer, D. C., Sarko, A., Bluhm, T. L. and Lee, Y. N. *Acta Crystallogr.* 1980, **B36**, 650
6 Bock, K. and Vignon, M. *Nouv. J. Chim.* 1982, **6**, 301
7 Neuman, A., Longchambon, F., Abbes, O., Gillier-Pandraud, H., Pérez, S., Rouzaud, D. and Sinaÿ, P. *Carbohydr. Res.* 1990, **195**, 187
8 Engelsen, S. B. PhD Thesis, Technical University of Denmark, 1991
9 Engelsen, S. B., Fabricius, J. and Rasmussen, Kj. (in preparation)
10 Niketić, S. R. and Rasmussen, Kj. 'The Consistent Force Field: A Documentation', Lecture Notes in Chemistry, Vol. 3, Springer-Verlag, Heidelberg, 1977
11 Rasmussen, Kj. 'Potential Energy Functions in Conformational Analysis', Lecture Notes in Chemistry, Vol. 37, Springer-Verlag, Heidelberg, 1985
12 Fabricius, J. PhD Thesis, Technical University of Denmark, 1991
13 Melberg, S. and Rasmussen, Kj. *J. Mol. Struct.* 1979, **57**, 215
14 Rasmussen, Kj. *Acta Chem. Scand.* 1982, **A36**, 323
15 Pietilä, L.-O. and Rasmussen, Kj. *J. Comput. Chem.* 1984, **5**, 252
16 Rasmussen, Kj. and Fabricius, J. in 'Computer Modeling of Carbohydrate Molecules' (Eds A. D. French and J. W. Brady), ACS Symposium Series No. 430, American Chemical Society, 1990, p 177
17 Lifson, S. and Warshel, A. *J. Chem. Phys.* 1968, **49**, 5116
18 Engelsen, S. B. and Rasmussen, Kj. (submitted)
19 *Surfer*, V.4, Golden Software Inc.
20 Kildeby, K., Melberg, S. and Rasmussen, Kj. *Acta Chem. Scand.* 1977, **A31**, 1
21 Melberg, S. and Rasmussen, Kj. *Carbohydr. Res.* 1979, **69**, 27