Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 340 (2005) 1019–1024
Carbohydrate RESEARCH

Conformation of the exopolysaccharide of *Burkholderia cepacia* predicted with molecular mechanics (MM3) using genetic algorithm search

Francesco Strino,[a,†] Abraham Nahmany,[a,†] Jimmy Rosen,[a] Graham J. L. Kemp,[b] Isabel Sá-correia[c] and Per-Georg Nyholm[a,*]

[a]*Department of Medical Biochemistry/Centre for Structural Biology, Göteborg University, Medicinaregatan 7B, 405 30 Göteborg, Sweden*
[b]*Department of Computing Science, Chalmers University of Technology, SE-412 96 Göteborg, Sweden*
[c]*Centro de Engenharia Biológica e Química, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisboa Codex, Portugal*

Received 13 September 2004; accepted 16 December 2004

Dedicated to Professor David A. Brant

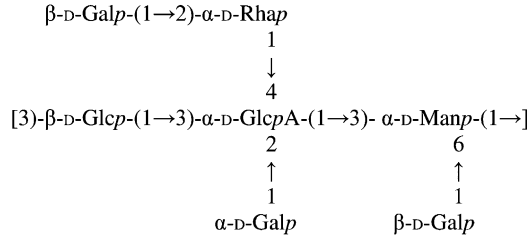
Abstract—We present a computational conformational analysis of the exopolysaccharide of *Burkholderia cepacia*, which is believed to play a role in colonization and persistence of *B. cepacia* in the lungs of cystic fibrosis patients. The repeating unit of the exopolysaccharide is a heptasaccharide with three branches, which cause significant steric restraints. Conformational searches using GLYGAL, an in-house developed software using genetic algorithm search methods, were performed on fragments as well as on the complete repeating unit with wrap-over residues. The force field used for the calculations was MM3(96). The search showed four favored conformations for an isolated repeating unit. However, for a sequence of several repeating units, the calculations indicate a single, well-defined linear conformation.

Keywords: *Burkholderia cepacia*; Polysaccharide; Conformation; Molecular mechanics; Genetic algorithms

1. Introduction

In the last two decades *Burkholderia* (previously known as *Pseudomonas*) *cepacia* has appeared as a human pathogen, causing several outbreaks, mostly among cystic fibrosis (CF) patients.[1,2]

Cescutti et al. recently reported the primary structure of the exopolysaccharide produced by a clinical isolate of the bacterium *Burkholderia cepacia*.[3] In their study Cescutti et al. determined that the repeating unit of the exopolysaccharide is a highly branched heptasaccharide:



It was apparent that the elucidation of the 3D structure of this saccharide would be of interest to allow a rational design of glycoconjugate vaccines. Furthermore this system was a challenge for the available software for conformational analysis of oligosaccharides. The present paper presents a prediction of the three-dimensional structure of the repeating unit, as well as a model of several repeating units of the exopolysaccharide.

* Corresponding author. Tel.: +46 (0)31 773 34 54; fax: +46 (0)31 823758; e-mail: nyholm@medkem.gu.se
† Contributed equally to this paper.

0008-6215/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2004.12.031

2. Methods

The notation used for torsion angles is the *heavy atom* definition with Φ defined as O-5–C-1–O–C-X and Ψ as C-1–O–C-X–C-(X+1). In the (1→6) linkage, the extra angle ω exists and is defined as O–C-6–C-5–O-5 and the Φ and Ψ are then defined as O-5–C-1–O–C-6 and C-1–O–C-6–C-5, respectively. The conformational search was primarily performed in the Φ/Ψ (or $\Phi/\Psi/\omega$ if required) conformational space of the glycosidic linkages. The conformational space was investigated both using systematic search and a genetic algorithm search. In both cases the energy calculations were carried out with the molecular mechanics MM3(96) program[4–6] using a dielectric constant (ε) of 80.

Systematic search using MM3(96) calculations was performed on all the disaccharide substructures contained in the polysaccharide. The systematic searches were performed with 15° step size of the Φ, Ψ, and in the case of the (1→6) linkage also ω, resulting in $24 \times 24 = 576$ sampled points and $24 \times 24 \times 24 = 13{,}824$ calculations in the case of the (1→6) linkage. In order to account for the possible conformations of the primary hydroxyls, the two favored conformations of the C-5–C-6 bond were selected for each residue in the systematic search. The sampled points were then used as the basis for the adiabatic energy maps.[7]

The genetic algorithm search was carried out using the GLYGAL[8] program, which is an in-house developed software implementing several genetic algorithm (GA) search methods. GLYGAL was used to carry out extensive calculations on the heptasaccharide. The searches were first performed on different fragments of the molecule, ranging in size from trisaccharide to hexasaccharide. A complete study of the whole repeating unit was then performed.

In order to study the connection of two repeating units, and especially the interactions of the (1→6)-linked Gal, another complete search was performed for the octasaccharide formed by one repeating unit without the (1→6)-linked Gal but including the upstream Man and (1→6)-linked Gal.

All the GA searches were performed using a parallel Lamarckian GA [cf. Ref. 9] implemented in GLYGAL. The investigations started from completely randomized structures and minimized the main glycosidic linkages and the pendant groups, namely C-5–C-6 linkages and hydroxyl groups. Some of the secondary minima needed independent minimization in order to obtain a complete minimization of the hydroxyl groups. Finally the conformations of structures with two, three, four, and eight repeating units were investigated.

In total approximately 50,000 structures were sampled in the different searches using the GLYGAL program. All calculations were performed on five computers with dual 2200+ MHz AMD processors using a Linux cluster (CSOL HOBORG). During the searches the α-D-Glc*p*A carboxyl group was protonated, but preliminary studies with MM3 showed no significant conformational differences between the charged and the protonated forms.

The molecular graphics was produced using Sybyl 6.8 (Tripos Inc.). Solvent accessible surfaces were created using the MOLCAD module of Sybyl.

3. Results and discussion

3.1. Systematic search

The MM3 adiabatic energy Φ/Ψ maps for all the disaccharide structures, occurring in the repeating unit (including one wrap-over residue) were the starting point for predicting the 3D structure of the molecule.

The linkage α-D-Glc*p*A-(1→3)-α-D-Man*p* (Fig. 1a) has one elongated energy minimum, showing flexibility in respect to Ψ (between 90°–180°). The map for the linkage α-D-Gal*p*-(1→2)-α-D-Glc*p*A (Fig. 1b) shows the global minimum at $\Phi/\Psi \approx 90°/180°$. The map for the linkage β-D-Glc*p*-(1→3)-α-D-Glc*p*A (Fig. 1c) has one elongated energy minimum with flexibility with respect to Ψ (between 60°–170°). The map for the linkage α-D-Rha*p*-(1→4)-α-D-Glc*p*A (Fig. 1d) shows a single constrained global minimum at $\Phi/\Psi \approx 80°/-170°$ and a secondary minimum region at $\Phi/\Psi \approx 100°/75°$ with a 3 kcal difference from the global minimum. The barrier between the two minima is due to a steric clash between H-1 of the Rha*p* and H-3 of the Glc*p*A.

The map for the linkage β-D-Gal*p*-(1→2)-α-D-Rha*p* (Fig. 1e) shows the global minimum at $\Phi/\Psi \approx -80°/160°$. Summing up the information given in the maps for the linkage β-D-Gal*p*-(1→6)-α-D-Man*p* (Fig. 1$f_{1,2,3}$) we see that the two main minima are found at $\Phi/\Psi/\omega \approx -80°/170°/-60°$ and at $\Phi/\Psi/\omega \approx -80°/170°/60°$. The linkage α-D-Man*p*-(1→3)-β-D-Glc*p* (Fig. 1g) has one elongated energy minimum, showing flexibility with respect to Ψ (between 90°–180°).

3.2. Genetic algorithm search

The searches performed using the GLYGAL program enabled us to define the conformational preferences of larger structural elements of the polysaccharide. The investigation of the whole repeating unit showed a fairly constrained structure, with the exception of the Gal-(1→6)-Man linkage. This flexibility allows the structure to assume two favored conformations (Fig. 2 and Table 1) with an energy difference of only about 0.1 kcal. Another alternative conformation was found for the linkage α-D-Rha*p*-(1→4)-α-D-Glc*p*A, which can assume a second minimum around 100°/68°. The energy value for this conformation in the isolated repeating unit is

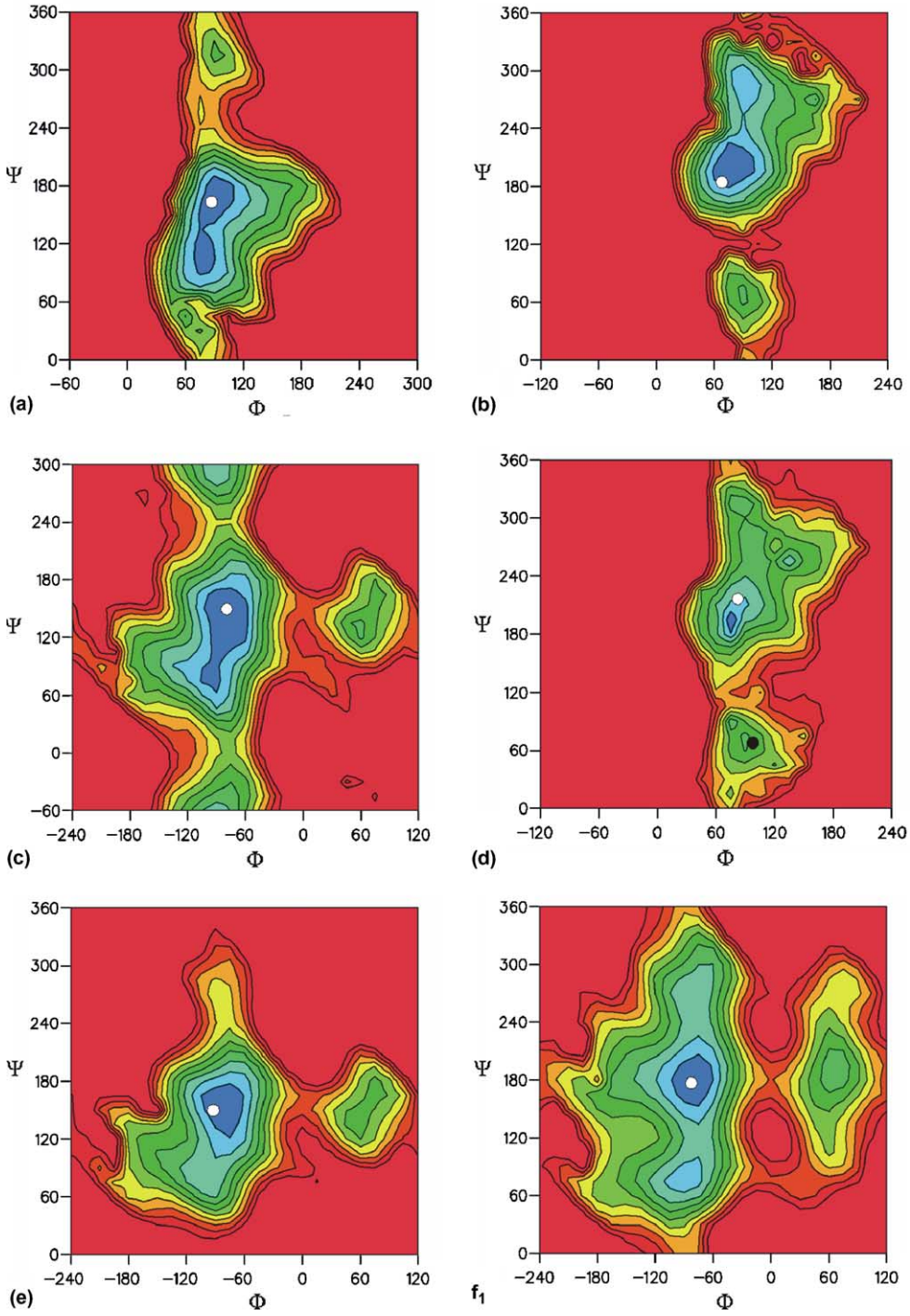


Figure 1. Adiabatic Φ/Ψ energy maps for the glycosidic linkages for the different disaccharide substructures of the repeating unit (plus one downstream residue) of *B. cepacia*. Contour levels are shown at every kcal/mol with blue for low-energy regions and red for high-energy levels. (a) The α-D-Glc*p*A-(1→3)-α-D-Man*p* linkage; (b) α-D-Gal*p*-(1→2)-α-D-Glc*p*A; (c) β-D-Glc*p*-(1→3)-α-D-Glc*p*A; (d) α-D-Rha*p*-(1→4)-α-D-Glc*p*A; (e) β-D-Gal*p*-(1→2)-α-D-Rha*p*; (f) the β-D-Gal*p*-(1→6)-α-D-Man*p*. f_1 for the Φ/Ψ, f_2 for the ω/Ψ, and f_3 for Φ/ω; (g) α-D-Man*p*-(1→3)-β-D-Glc*p*. The white dots show the torsion angle values of the global minimum for the repeating unit according to GLYGAL, the black dots shows the position of the major secondary minima of the repeating unit.

about 1 kcal higher than the global minimum. Other possible minima were explored in the elongated low energy Ψ valleys of the mainchain linkages β-D-Glc*p*-(1→3)-α-D-Glc*p*A, α-D-Glc*p*A-(1→3)-α-D-Man*p*, and α-D-Man*p*-(1→3)-β-D-Glc*p*. We observed that minimizations started from different parts of the low-energy regions in the disaccharide maps converged smoothly to the global minimum.

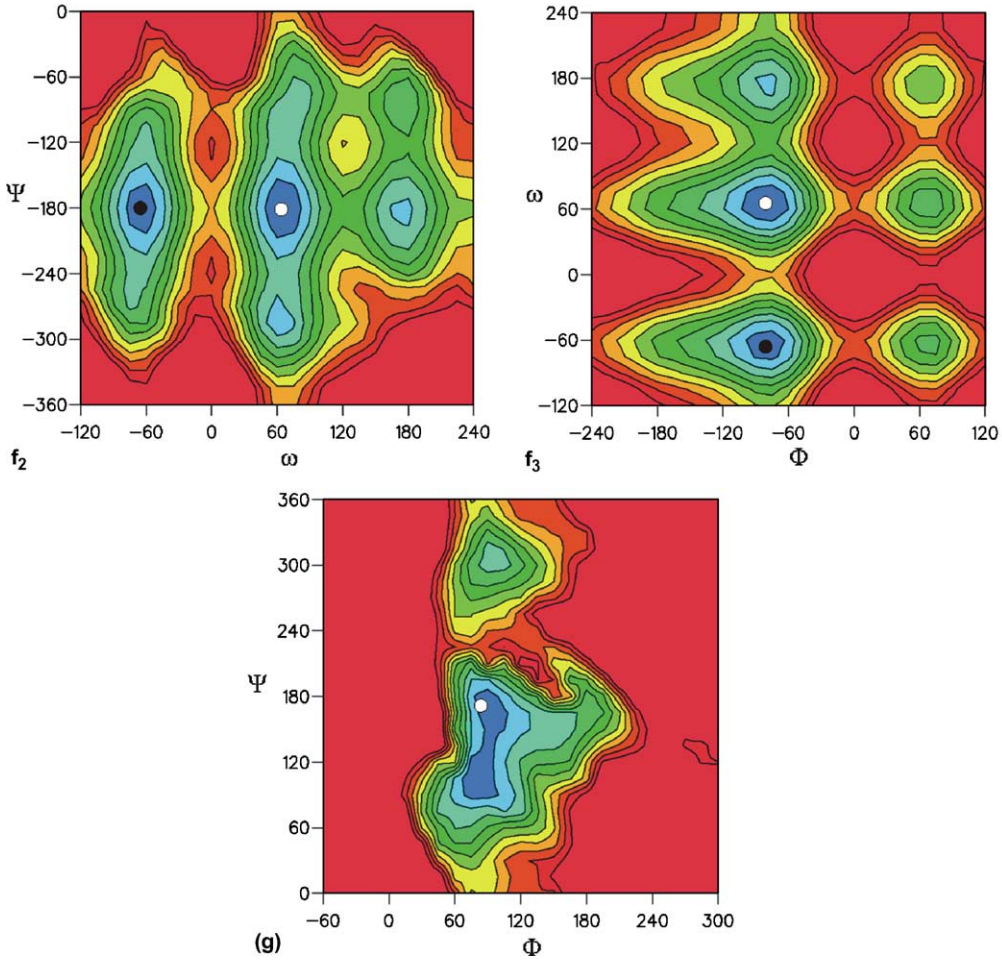


Figure 1 (*continued*)

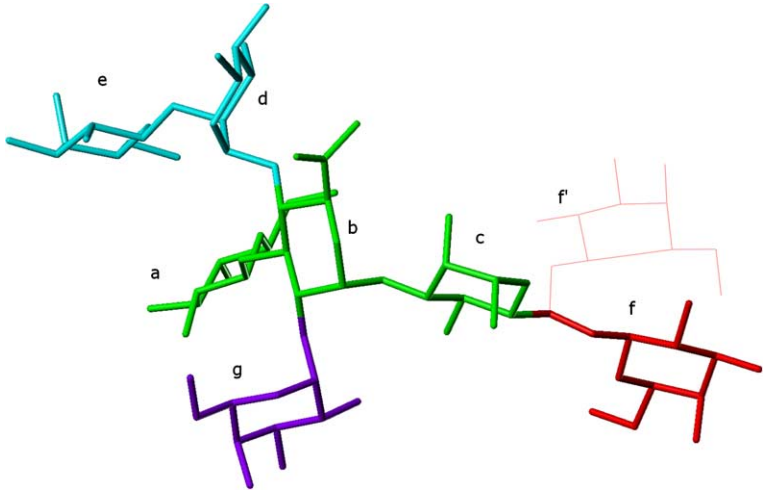


Figure 2. The two best conformations obtained with GLYGAL for one repeating unit of the heptasaccharide of *B. cepacia*. Letters a, b, and c denote the mainchain residues β-D-Glc*p*, α-D-Glc*p*A, and α-D-Man*p*, respectively, while d, e, f, and g denote the residues α-D-Rha*p*, β-D-Gal*p*, β-D-Gal*p*, and α-D-Gal*p*, respectively, existing in the branches. The f' residue in red thin lines indicates the conformation of the (1→6)-linked Gal of the secondary minimum. Hydrogen atoms are not shown for clarity.

The GA search on the octasaccharide (the most favorable conformation is shown in Fig. 3) showed a fairly constrained minimum energy conformation, which matched completely the global minimum energy struc-

Table 1. Favorable $\Phi/\Psi/\omega$ conformations of one and two repeating units of *B. cepacia*

Linkage	One repeating unit $\Phi°/\Psi°(/\omega°)$	Two repeating units $\Phi°/\Psi°(/\omega°)$
α-D-Glc*p*A-(1→3)-α-D-Man*p*	90/165	90/165
α-D-Gal*p*-(1→2)-α-D-Glc*p*A	63/−177	63/178
β-D-Glc*p*-(1→3)-α-D-Glc*p*A	−75/155	−75/155
α-D-Rha*p*-(1→4)-α-D-Glc*p*A	82/−142[a] or 100/68	82/−142
β-D-Gal*p*-(1→2)-α-D-Rha*p*	−88/150	−88/150
β-D-Gal*p*-(1→6)-α-D-Man*p*	−79/175/65[a] or −79/−178/−64	−75/177/65
α-D-Man*p*-(1→3)-β-D-Glc*p*	—	84/169

[a] Global minimum.

ture of the repeating unit. The (1→6) Gal branch is not allowed to assume its secondary minimum due to strong interactions with the residues of the downstream unit. According to the GA runs the octasaccharide conformation is very stable. With respect to the primary and secondary hydroxyls, the calculations on the repeating unit and the octasaccharide converged on the *gtR* conformation (for definition see Ref. 7) for all residues with a primary hydroxyl group. The rhamnose also had an *R* arrangement of the secondary hydroxyls and the ω torsion of the GlcA converged to −55° (124°). In order to investigate the effect of a lower dielectric constant, comparative GA calculations were carried out on the octasaccharide at $\varepsilon = 4$ and $\varepsilon = 15$. The result showed no significant conformational differences as compared to the calculations at $\varepsilon = 80$.

Further evaluations with $\varepsilon = 80$ were done for two, three, four, and eight repeating units. The results closely matched the values obtained in the GA search on the octasaccharide. The final Φ, Ψ, and ω torsion angle values obtained with GLYGAL for two repeating units are shown in Table 1. In particular it is noted that the (1→6)-linked Gal is not free to rotate in structures containing two or more repeating units. These calculations confirmed the previous results and indicate strongly that longer chains of repeating units assume a linear, fairly stiff conformation (Fig. 4).

3.3. Implications regarding molecular properties

In a recent study by Sist et al. calculated the stiffness parameter based on intrinsic viscosity data[10] for the exopolysaccharide of *B. cepacia* to be 0.059, which is intermediate between alginate and carboxymethyl cellulose, which are considered to be semi-flexible polymers.[11] The results of the present conformational analysis do not allow a quantitative prediction of the chain stiffness, since the dynamics, especially the interactions with the surrounding water, is not addressed. However, the results of the present conformational analysis can explain the results by Sist et al. in a qualitative manner. Moreover, mapping of the lipophilicity on two repeating units shows a hydrophobic region on the Rha residue and the crevice downstream of the (1→2) Gal of the same repeating unit. However, this hydrophobic patch is less pronounced in the deprotonated form (Fig. 5).

It is expected that the present results will be of interest in the development of glycoconjugate based vaccines against *B. cepacia*. The repeating unit has four different favored conformations and appears less suitable as a

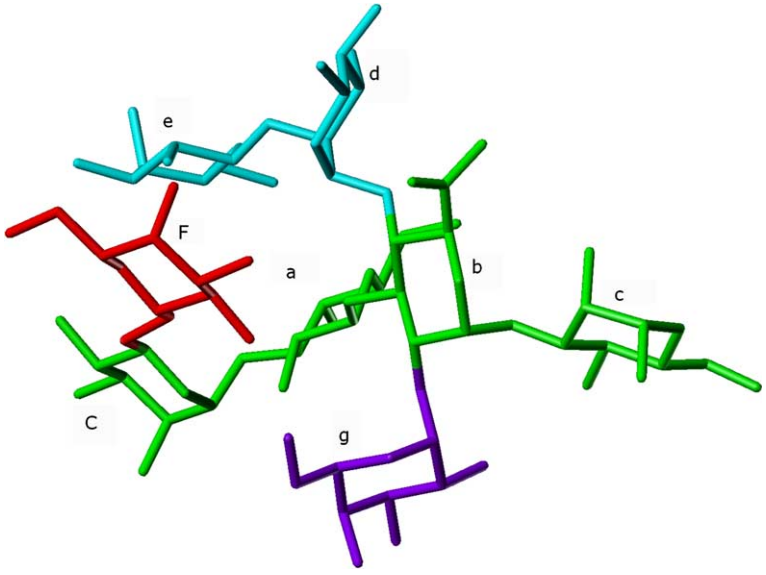


Figure 3. The global minimum conformation obtained with GLYGAL for an octasaccharide of the *B. cepacia* exopolysaccharide frame-shifted with respect to Figure 2 (see text). Letters a, b, and c denote the mainchain residues β-D-Glc*p*, α-D-Glc*p*A, and α-D-Man*p*, respectively, while d, e, f, and g denote the residues α-D-Rha*p*, β-D-Gal*p*, β-D-Gal*p*, and α-D-Gal*p*. Capital letters indicate the upstream residues: F for β-D-Gal*p* in the branch and C for α-D-Man*p*. Hydrogen atoms have been omitted.

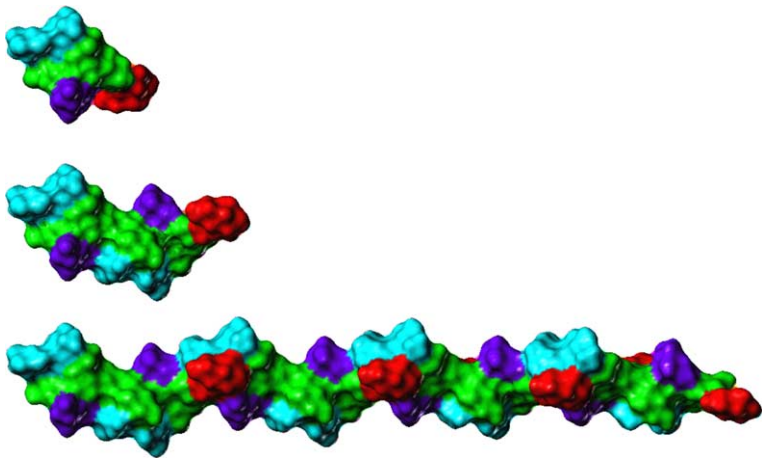


Figure 4. Global minimum conformations of one, two, and eight repeating units of the *B. cepacia* exopolysaccharide. Mainchain is shown in green, Gal-Rha branch in cyan, Gal (1→6) branch in red and Gal (1→2) branch in violet.

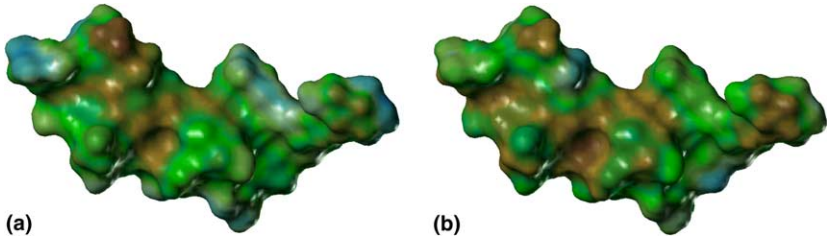


Figure 5. Solvent-accessible MOLCAD surface of two repeating units of the exopolysaccharide of *B. cepacia* color coded according to lipophilicity (blue hydrophilic, brown hydrophobic, and green intermediate). (a) Protonated form, (b) deprotonated.

hapten. However, as shown in these studies, an octasaccharide fragment (Fig. 3) has only one favored conformation and appears interesting as a hapten for vaccine development.

Acknowledgements

Financial support from the Swedish Medical Research Council is acknowledged (K2000-03x-00006-36A).

References

1. Coenye, T.; LiPuma, J. J. *Front Biosci.* **2003**, *8*, 55–67.
2. Speert, D. P. *Paediatr. Respir. Rev.* **2002**, *3*, 230–235.
3. Cescutti, P.; Bosco, M.; Picotti, F.; Impallomeni, G.; Leitao, J. H.; Richau, J. A.; Sa-Correia, I. *Biochem. Biophys. Res. Commun.* **2000**, *273*, 1088–1094.
4. Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
5. Allinger, N. L.; Rahman, M.; Lii, J. H. *J. Am. Chem. Soc.* **1990**, *112*, 8293–8307.
6. Lii, J. H.; Allinger, N. L. *J. Comput. Chem.* **1991**, *12*, 186–199.
7. French, A. D.; Tran, V. H.; Perez, S. In *Computer Modelling of Carbohydrate Molecules.* ACS Symp. Ser. 430; American Chemical Society: Washington, DC, 1990; pp 191–212.
8. Nahmany, A.; Strino, F.; Rosen, J.; Kemp, G. J. L.; Nyholm, P.-G. *Carbohydr. Res.* **2005**, *340*, see doi:10.1016/j.carres.2004.12.037.
9. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *14*, 1639–1662.
10. Smidsrød, O.; Haug, A. *Biopolymers* **1971**, *10*, 1213–1227.
11. Sist, P.; Cescutti, P.; Skerlavaj, S.; Urbani, R.; Leitao, J. H.; Sa-Correia, I.; Rizzo, R. *Carbohydr. Res.* **2003**, *338*, 1861–1867.