Cyclic Dodecasaccharide Structure

Crystal Structure of a Cyclic Enterobacterial Common Antigen**

*Magnus Färnbäck, Lars Eriksson, Sof'ya Senchenkova, Krystyna Zych, Yuriy A. Knirel, Zygmunt Sidorczyk, and Göran Widmalm**

Carbohydrates have several functions in nature and, among other things, they are involved in biochemical regulation and trafficking processes.[1] Likewise, bacteria have different types

[*] Prof. Dr. G. Widmalm, M. Färnbäck
Department of Organic Chemistry
Arrhenius Laboratory, Stockholm University
106 91 Stockholm (Sweden)
Fax: (+46) 8-154-908
E-mail: gw@organ.su.se

M. Färnbäck, Dr. L. Eriksson
Division of Structural Chemistry
Arrhenius Laboratory, Stockholm University
106 91 Stockholm (Sweden)

S. Senchenkova, Dr. Y. A. Knirel
N. D. Zelinsky Institute of Organic Chemistry
Russian Academy of Sciences
119991 Moscow (Russia)

K. Zych, Prof. Dr. Z. Sidorczyk
Department of General Microbiology
Institute of Microbiology and Immunology
University of Lodz, Banacha 12/16
90-237 Lodz (Poland)

[**] This work was supported by a grant from the Swedish Research Council.

 DOI: 10.1002/ange.200351250

of polysaccharides on their surfaces and in their outer membranes. A certain species of bacteria usually carries both a specific and a common type of polysaccharide. All members of the enterobacteriaceae family share an antigen known as the enterobacterial common antigen (ECA), which normally is found as a phosphoglycolipid conjugate in the outer membrane of the bacteria.[2] However, the biological importance of ECA is so far not known. In some cases cyclic ECA structures are also produced,[3] but the most common type of cyclic carbohydrate oligomers are the cyclodextrins that are composed of α-(1→4)-linked glucosyl residues ranging from hexamers to hexaicosamers.[4] Crystal structures of other cyclic glycans have also been reported.[5,6] In the present communication, we describe the crystal structure of a cyclic dodecasaccharide of ECA, in which the repeating unit is composed of three amino sugars, namely, →3)-α-D-Fuc*p*4NAc-(1→4)-β-D-Man*p*NAcA-(1→4)-α-D-Glc*p*NAc-(1→ (Figure 1).

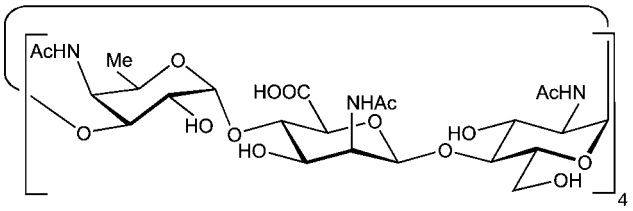

Figure 1. Structure of the cyclic enterobacterial common antigen dodecasaccharide (cECA$_{12}$) having a repeating unit of →3)-α-D-Fuc*p*4-NAc-(1→4)-β-D-Man*p*NAcA-(1→4)-α-D-Glc*p*NAc-(1→.

An ECA preparation from *Proteus penneri* strain 17 was shown to exist as a cyclic dodecasaccharide (cECA$_{12}$), since the FAB mass spectrum revealed the pseudomolecular ion at *m*/*z* 2429.89, which corresponds to $[M+H]^+$[7] and showed good agreement with previously published ^{1}H and ^{13}C NMR data of a cyclic ECA dodecasaccharide from a *Plesiomonas shigelloides* strain.[8] It was noted by serendipity that crystals of cECA$_{12}$ had formed in the NMR tube. Subsequently, a preliminary investigation of the brittle crystals revealed that the unit cell consisted of four molecules (data not shown).

A systematic investigation was now performed where cECA$_{12}$ was crystallized employing the hanging-drop technique, by which prism-like crystals (monoclinic space group *C*2) were formed. Data were collected with a rotating anode and synchrotron radiation, and the structure was solved by direct methods, which reveal that the asymmetric unit is composed of two symmetry-independent parts, each containing half of the cyclic dodecasaccharide.[9] The two different cECA$_{12}$ molecules are denoted **A** and **B**. The former adapts the form of a slightly tilted rhomb, while the latter adapts an almost perfect square. One half of each molecule is related to the other half by crystallographic symmetry, thus forming a complete cECA$_{12}$ unit (Figure 2). Each side is composed of a repeating unit that is followed by a sharp turn at the (1→3)-linkage. A similar structure was also found in a prior molecular dynamics simulation.[8] The sugar rings all adopt the $^4C_1$ chair conformation and the glycosidic torsion angles

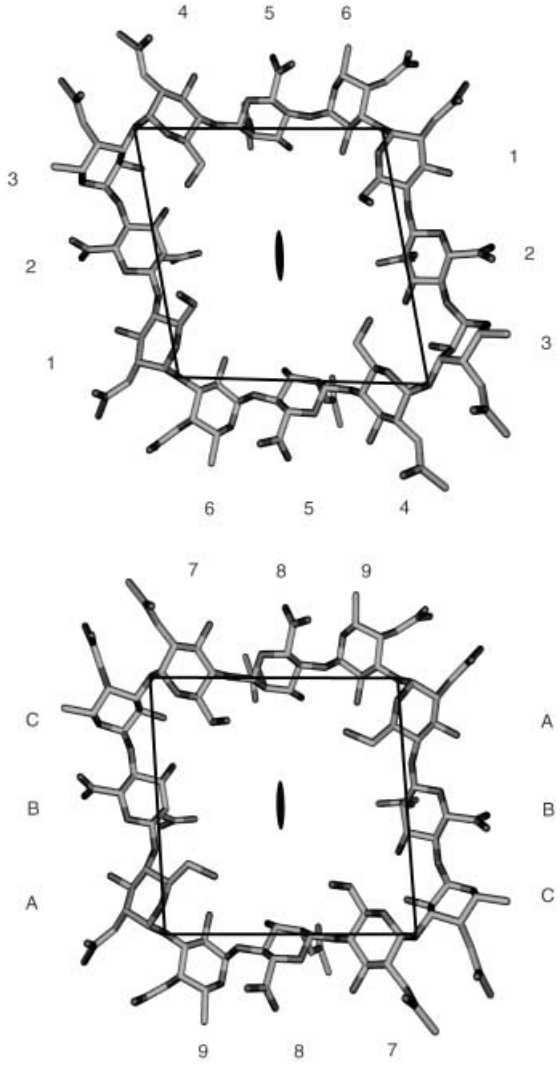

Figure 2. Stick representation of the two crystal forms of cECA$_{12}$ including the one-character sugar residue labeling 1–6 in **A** and 7–C in **B**. Hydrogen atoms are removed for clarity.

lead to a *syn* relationship[10] between adjacent residues. The ϕ versus ψ scatter plot (Figure 3) reveals a grouping of torsional angles[11] according to each glycosidic linkage, with the largest range observed for the fucosyl residue. A summary of the glycosidic torsion angles is compiled in Table 1.

The crystal packing is shown in Figure 4. The molecules are arranged in layers in the *ac* plane and stack in the *b* direction. There are different orientations observed in the *ac* plane, with two symmetry independent cECA$_{12}$ environments exhibited. Molecule **A** is surrounded by four **B** molecules and two **A** molecules in the *ac* plane, while molecule **B** is surrounded solely by four **A** molecules. Two distinct layers are present throughout the structure. They are identical and related to each other through *c* centering. The columns consist solely of either **A** or **B** molecules.

A noticeable feature of cECA$_{12}$ is the positioning of the *N*-acetyl groups. On the mannuronosyl residues the *N*-acetyl groups protrude out from the plane of cECA$_{12}$, that is, all in the same direction (see Figure 2). The next molecule along

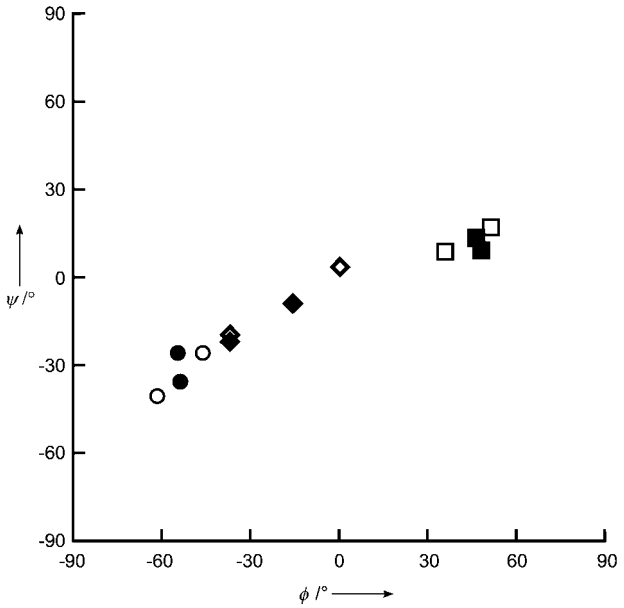

Figure 3. Scatter plot of the glycosidic torsion angles of the crystal structure of $cECA_{12}$. Glucosyl residues are denoted by circles, fucosyl residues by diamonds, and mannuronosyl residues by squares. Open and closed symbols refer to molecules **A** and **B**, respectively.

Table 1: Glycosidic torsion angles in the two crystal forms of $cECA_{12}$.

Residue	A			B		
	Label	ϕ [°]	ψ [°]	Label	ϕ [°]	ψ [°]
→4)-α-D-Glc*p*NAc-(1→	1	−46	−26	7	−55	−26
→4)-β-D-Man*p*NAcA-(1→	2	36	9	8	48	9
→3)-α-D-Fuc*p*4NAc-(1→	3	0	4	9	−37	−22
→4)-α-D-Glc*p*NAc-(1→	4	−62	−41	A	−54	−36
→4)-β-D-Man*p*NAcA-(1→	5	52	17	B	46	14
→3)-α-D-Fuc*p*4NAc-(1→	6	−37	−20	C	−16	−9

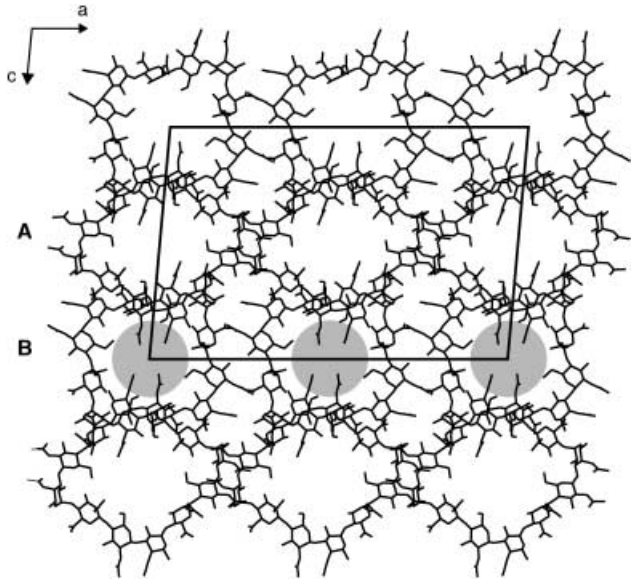

Figure 4. Crystal packing including the unit cell of $cECA_{12}$ viewed along the *b* axis. The shaded regions highlight the channels closed by *N*-acetyl groups on glucosyl and fucosyl residues. Water molecules are not displayed.

the *b* axis is turned upside down and translated by half a molecule, and its *N*-acetyl groups are in close vicinity of two of the *N*-acetyl groups in the former molecule. The *N*-acetyl groups on adjacent glucosyl and fucosyl residues extend in a parallel manner (see shaded regions in Figure 4) at the corners of $cECA_{12}$.

Channels are formed along the *b* axis, mainly through molecule **A**. The channels through the **B** molecules are closed by *N*-acetyl groups on glucosyl and fucosyl residues protruding from adjacent molecules (solely of type **A**) above and below molecule **B**. This feature is not observed for molecule **A**. Instead, the channel is filled to a greater extent with partially disordered water molecules. Of those that could be located, all are geometrically arranged to form hydrogen bonds to either the sugars or to other water molecules.

A few intramolecular hydrogen bonds are indicated between residues. Both molecules **A** and **B** show a similar hydrogen-bond network with hydrogen bonds between O2 in the fucosyl residue and O3 in the mannuronosyl residue, as well as between O3 in the glucosyl residue and O5 in the mannuronosyl residue, based on the distance between donor and acceptor atoms in the range 2.7–3.3 Å. Molecule **B** also displays an additional hydrogen bond between O6 in the glucosyl residue and O2 in the fucosyl residue when the hydroxymethyl group of the donor has the *gauche*/*trans* (*gt*) conformation, but not when it has the *gauche*/*gauche* (*gg*) conformation. Although both the *gg* and *gt* conformations are present for the glucosyl residues in **A**, the latter hydrogen bond is not formed.

In larger cyclic crystal structures, solvent molecules make up a significant part of the volume of the unit cell.[12] Likewise, the $cECA_{12}$ crystal contains a large amount of water (~34% by weight or ~6 water molecules per sugar residue in the unit cell, if a volume of 29.7 $Å^3$ is assumed).[13] In the analysis of hydrogen-bonding patterns, 30 water molecules were localized per $cECA_{12}$ unit, all of which participate in interactions to either the sugar residues or to other water molecules. The identified water molecules amount to approximately half of those present per $cECA_{12}$ unit.

In conclusion, the crystal structure of a cyclic oligosaccharide has been determined. Its characterization is of special interest since only amino sugars are found in the repeating unit and in another form they are the constituents of a polysaccharide anchored in outer bacterial membranes. Future studies should include analysis of NMR relaxation-time data to investigate whether the ϕ versus ψ variation observed in the crystal structure can be reconciled with our previous NMR and simulation studies of $cECA_{12}$.

Experimental Section

Proteus penneri strain 17 (1410-75) was cultivated using a procedure previously described.[14] A lipopolysaccharide (LPS) preparation was obtained by extraction of dry bacterial mass with a hot phenol/water mixture and purified by treatment with a cold aqueous solution of CCl_3CO_2H (50%) followed by dialysis of the supernatant. Degradation of LPS was performed with an aqueous acetic acid solution (1%) at 100 °C for 2 h and the product was fractionated by gel permeation chromatography and lyophilized. The resulting polysaccharide preparation (700 mg) was suspended in water (10.5 mL) and centrifuged

(13000 *g*, 20 min). The precipitate was washed with water (7 mL), centrifuged as above, and the combined supernatant was concentrated in vacuum and fractionated on a column (70 × 2.5 cm) of Sephadex G-50 (Pharmacia) in 0.05 M pyridinium acetate buffer (pH 4.5). Fractions containing ECA were collected (90 mg). An NMR sample of ECA in D_2O was prepared (~100 mg mL^{-1}), from which prism-like crystals with dimensions up to 0.3 × 0.3 × 0.1 mm formed. These crystals were investigated with X-ray diffraction methods using an IPDS system.

For further investigations, larger crystals of a higher quality were desired. ECA was crystallized employing the hanging-drop technique from an aqueous solution (~100 mg mL^{-1}) containing PEG-400 (40%). Crystals formed as slightly irregular prisms, up to 0.5 mm in length. The structure was solved from data obtained from the STOE IPDS data set with direct methods using the density modification technique available in SHELXD.[15] Attempts to use the direct methods available in SHELXS97[16] failed. Both halves of the $cECA_{12}$ structure in the asymmetric unit, with the exception of some side chains, were visible in the initial electron-density maps. Nonexchangeable hydrogen atoms were added at calculated geometrical positions. Several bond length restraints were applied using standard bond lengths[17] and some conformational features, such as *N*-acetyl group planarity, with the DFIX and FLAT command available in SHELXH.

Water molecules were located from the electron-density difference map and kept if they were located at a reasonable distance (2.6–3.2 Å) from suitable hydrogen acceptors or donors, together with the observation that they did not diverge in the least-squares refinement. Not all water molecules in the cavities could be located, probably as a result of high disorder. The remaining solvent-accessible volume was calculated with PLATON.[18]

Received: February 24, 2003 [Z51250]

Keywords: carbohydrates · polysaccharides · structure elucidation · X-ray diffraction

[1] *Essentials of Glycobiology* (Eds.: A. Varki, R. Cummings, J. Esko, H. Freeze, G. Hart, J. Marth), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **1999**.
[2] H.-M. Kuhn, U. Meier-Dieter, H. Mayer, *FEMS Microbiol. Rev.* **1988**, *54*, 195–222.
[3] A. Dell, J. Oates, C. Lugowski, E. Romanowska, L. Kenne, B. Lindberg, *Carbohydr. Res.* **1984**, *133*, 95–114.
[4] K. Gessler, I. Usón, T. Takaha, N. Krauss, S. M. Smith, S. Okada, G. M. Sheldrick, W. Saenger, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4246–4251.
[5] M. Sawada, T. Tanaka, Y. Takai, T. Hanafusa, T. Taniguchi, M. Kawamura, T. Uchiyama, *Carbohydr. Res.* **1991**, *217*, 7–17.
[6] G. M. Bradbrook, K. Gessler, G. L. Côté, F. Momany, P. Biely, P. Bordet, S. Pérez, A. Imberty, *Carbohydr. Res.* **2000**, *329*, 655–665.
[7] E. V. Vinogradov, Y. A. Knirel, J. E. Thomas-Oates, A. S. Shashkov, V. L. L'vov, *Carbohydr. Res.* **1994**, *258*, 223–232.
[8] M. Staaf, C. Höög, B. Stevensson, A. Maliniak, G. Widmalm, *Biochemistry* **2001**, *40*, 3623–3628.
[9] $T = 293$ K, crystal dimensions: 0.08 × 0.10 × 0.15 mm, monoclinic, space group C_2, $Z = 4$, $a = 46.42(2)$, $b = 13.542(2)$, $c = 30.15(1)$ Å, $\beta = 95.15(5)°$, $V = 18875(12)$ Å^3, $\rho_{calcd} = 1.023$ g cm^{-3}. Data measured with a STOE IPDS (Stoe & Cie, Darmstadt, Germany) on a rotating anode equipped with a graphite monochromator set to diffract $Mo_{K\alpha}$ ($\lambda = 0.71073$ Å) and a BRUKER SMART 1k (Bruker AXS GmbH, Karlsruhe, Germany) diffraction system using synchrotron radiation, ($\lambda \approx 0.856$ Å) at beam line I711, MAXLAB, Lund, Sweden. The crystal was mounted in a glass capillary with some residual mother liquor. Scan mode: area detector φ-scans (STOE IPDS) and ω-scans (BRUKER SMART). Three data sets, one from the in-house STOE IPDS and two from the BRUKER SMART system were scaled and merged to a common data set. A total of 52476 reflections were measured of which 17543 were independent with $R_{int} = 0.143$. The structure was solved by direct methods (SHELXD). The non-hydrogen carbohydrate skeleton was refined with anisotropic displacement parameters and water oxygen atoms were refined with isotropic displacement parameters. Hydrogen atoms were not placed on water oxygen atoms. All 17543 independent reflections were used in the full-matrix least-squares calculation (SHELXH97) against F^2. Number of refined parameters = 1633, $2\theta_{max} = 64.0°$ ($\lambda \approx 0.856$ Å), $R1 = 0.129$, $wR2 = 0.316$ for 5409 reflections with $F^2 \geq 2\sigma(F^2)$, max shift/esd = 0.12 in the final model. Lorentz and polarization corrections were applied during data reduction with STOE diffractometer control software for the STOE IPDS data set and SAINT 5.5 for the BRUKER data. Absorption correction with XRED 1.09 (STOE), $\mu = 0.92$ cm^{-1}, transmission factors were min: 0.98, max: 0.99. Residual electron density: min: 0.37 e Å^{-3}; max: 0.50 e Å^{-3}. CCDC-202591 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/conts/retrieving.html (or from the Cambridge Crystallographic Data Centre, 12, Union Road, Cambridge CB2 1EZ, UK; fax: (+44) 1223-336-033; or deposit@ccdc.cam.ac.uk).
[10] C. Höög, C. Landersjö, G. Widmalm, *Chem. Eur. J.* **2001**, *7*, 3069–3077.
[11] Definition of ϕ and ψ torsion angles: H1(m)–C1(m)–On($m+1$)–Cn($m+1$) and C1(m)–On($m+1$)-Cn($m+1$)–Hn($m+1$), respectively, where m denotes a sugar residue and n denotes a linkage position.
[12] J. R. Dias, R. A. Pascal, Jr., J. Morrill, A. J. Holder, H. Gao, C. Barnes, *J. Am. Chem. Soc.* **2002**, *124*, 4647–4652.
[13] M. Gerstein, C. Chothia, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10167–10172.
[14] K. Kotelko, W. Gromska, M. Papierz, Z. Sidorczyk, D. Krajewska, K. Szer, *J. Hyg. Epidemiol. Microbiol. Immunol.* **1977**, *21*, 271–284.
[15] G. M. Sheldrick, T. R. Schneider in *Methods in Macromolecular Crystallography* (Eds.: D. Turk, L. Johnson), IOS Press, Amsterdam, **2001**, pp. 72–81.
[16] G. M. Sheldrick, SHELX-98, Program for the Solution of Crystal Structures, University of Göttingen, Göttingen (Germany), **1998**
[17] *International Tables for Crystallography, Vol. C* (Ed.: A. J. C. Wilson), Kluwer, Dordrecht, **1995**.
[18] A. L. Spek, PLATON, University of Utrecht, The Netherlands.

