Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 4246–4251, April 1999
Chemistry

V-Amylose at atomic resolution: X-ray structure of a cycloamylose with 26 glucose residues (cyclomaltohexaicosaose)

KATRIN GESSLER*, ISABEL USÓN†, TAKESHI TAKAHA‡, NORBERT KRAUSS*, STEVEN M. SMITH§, SHIGETAKA OKADA‡, GEORGE M. SHELDRICK†, AND WOLFRAM SAENGER*¶

*Institut für Kristallographie, Freie Universität Berlin, Takustrasse 6, D-14195 Berlin, Germany; †Institut für Anorganische Chemie, Universität Göttingen, Tammannstrasse 4, D-37077 Göttingen, Germany; ‡Biochemical Research Laboratory, Ezaki Glico Co., Ltd., 4-6-5 Utajima, Nishiyodogawa, Osaka 555, Japan; and §Institute of Cell and Molecular Biology, University of Edinburgh, The King's Buildings, Mayfield Road, Edinburgh EH9 3JH, United Kingdom

Edited by Isabella L. Karle, Naval Research Laboratory, Washington, DC, and approved February 10, 1999 (received for review September 10, 1998)

ABSTRACT **The amylose fraction of starch occurs in double-helical A- and B-amyloses and the single-helical V-amylose. The latter contains a channel-like central cavity that is able to include molecules, "iodine's blue" being the best-known representative. Molecular models of these amylose forms have been deduced by solid state ^{13}C cross-polarization/magic angle spinning NMR and by x-ray fiber and electron diffraction combined with computer-aided modeling. They remain uncertain, however, as no structure at atomic resolution is available. We report here the crystal structure of a hydrated cycloamylose containing 26 glucose residues (cyclomaltohexaicosaose, CA26), which has been determined by real/reciprocal space recycling starting from randomly positioned atoms or from an oriented diglucose fragment. This structure provides conclusive evidence for the structure of V-amylose, as the macrocycle of CA26 is folded into two short left-handed V-amylose helices in antiparallel arrangement and related by twofold rotational pseudosymmetry. In the V-helices, all glucose residues are in *syn* orientation, forming systematic interglucose $O(3)_n \cdots O(2)_{n+1}$ and $O(6)_n \cdots O(2)_{n+6}$/$O(3)_{n+6}$ hydrogen bonds; the central cavities of the V-helices are filled by disordered water molecules. The folding of the CA26 macrocycle is characterized by typical "band-flips" in which diametrically opposed glucose residues are in *anti* rather than in the common *syn* orientation, this conformation being stabilized by interglucose three-center hydrogen bonds with $O(3)_n$ as donor and $O(5)_{n+1}$, $O(6)_{n+1}$ as acceptors. The structure of CA26 permitted construction of an idealized V-amylose helix, and the band-flip motif explains why V-amylose crystallizes readily and may be packed tightly in seeds.**

Starch is composed of two fractions, the linear amylose consisting exclusively of α(1–4)-linked glucose residues in $^4C_1$-chair conformation, and the branched amylopectin, which also contains α(1–6) links at characteristic intervals. The polysaccharide chain of amylose may be folded into three different structures denoted A, B, and V (1–3). Since crystallization of amylose fragments with defined chain lengths has remained elusive, structural information relies on x-ray fiber diffraction, electron diffraction on tiny single crystals, and solid-state ^{13}C cross-polarization/magic angle spinning (CP/MAS) NMR spectroscopy combined with computer-aided modeling (3–8). The only available single crystal x-ray study of an amylose-type oligosaccharide in the complex (*p*-nitrophenyl α-maltohexaoside)$_2$·Ba(I$_3$)$_2$·27H$_2$O features an antiparallel left-handed double helix (9, 10) that has no resemblance to A-, B-, or V-amylose.

The structures of A- and B-amylose are similar and differ only in packing arrangement and water content, the A-form occurring preferentially in cereals and the B-form in tubers (3). They both form double helices with parallel strands of 6 × 2 glucoses per turn and right-handed (3–8) or left-handed (11) twist, this ambiguity illustrating the weakness of the above methods, which do not provide structural information at atomic resolution.

The polysaccharide chain of V-amylose found naturally in non-A and non-B segments of amylose is folded into a left-handed single helix; it contains 6 glucoses per turn with 7.91- to 8.17-Å pitch height (3–5) and forms a central channel-like cavity. V-Amylose may be obtained synthetically when amylose is precipitated from aqueous solution by addition of alcohols, ketones, fatty acids, iodine, or salts that form inclusion complexes; "iodine's blue" is the most familiar one. These properties are similar to those found in the related cyclodextrins, α-cyclodextrin or cyclohexaamylose (CA6) being a mimic of one turn of V-amylose (12, 13). Because the glucoses in V-amylose (as in the cyclodextrins) are all in *syn* orientation, the secondary hydroxyl groups are hydrogen bonded $O(3)_n \cdots O(2)_{n+1}$, and the V-amylose helix is additionally stabilized by hydrogen bonds $O(6)_n \cdots O(2)_{n+6}$ between the turns (3–5).

Whereas treatment of amylose with bacterial cyclodextrin glucanotransferases produces the common cyclodextrins with 6 to 8 (CA6 to CA8) glucose residues and small amounts of the higher homologues (14, 15), disproportionating enzyme yields cycloamyloses (CAs) containing from 17 to several hundred glucose residues (16, 17). Because the two ends of the amylose chain are linked together in CAs, the conformational space is substantially reduced; CA10, CA14, and CA26 could be crystallized, and the structures of the first two molecules have been described (18–20). They are not annular in shape as are the smaller cyclodextrins CA6 to CA8 (12, 13) and the elliptically distorted CA9 (21), in which all glucoses are in *syn*, this conformation being stabilized by $O(3)_n \cdots O(2)_{n+1}$ hydrogen bonds. In CA10 and CA14, excessive steric strain is relieved by flips (18) at two diametrically opposed sites in the macrocycles where one glucose each is rotated into the *anti* orientation (like cutting a band, rotating one end by 180°, and then gluing the ends together). At this well-characterized motif, termed "band-flip" (18), the common $O(3)_n \cdots O(2)_{n+1}$ hydrogen bond is broken; a new three-center hydrogen bond is formed instead between the two *anti*-oriented glucoses with $O(3)_n$-H as donor and $O(5)_{n+1}$, $O(6)_{n+1}$ as acceptors; see Fig. 1*A*. To balance residual strain, each band-flip is associated with a "kink" where adjacent glucoses in *syn* orientation are rotated about the common glycosidic link so that the $O(3)_n \cdots O(2)_{n+1}$ hydrogen bond is lengthened (as in CA14) or even broken (as in CA10) (18).

Here we describe the structure of CA26 (cyclomaltohexaicosaose), which also features band-flips and (minor) kinks. It is not folded in the double-helical form predicted by us on the basis of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office.

Abbreviations: CP/MAS, cross-polarization/magic angle spinning; CA*n*, cycloamylose with *n* glucose residues.

Data deposition: The atomic coordinates have been deposited in the Cambridge Crystallographic Data Centre database. ID code CCDC 115146.

¶To whom reprint requests should be addressed. e-mail: saenger@chemie.fu-berlin.de.

the CA10 and CA14 structures (18); by contrast, it contains two short V-helices, showing the structure of this form of amylose at atomic resolution.

EXPERIMENTAL

Preparation of CA26. Cyclic α-1,4-glucan mixtures were produced by the action of recombinant potato D-enzyme (4-α-glucanotransferase, EC 2.4.1.25) on synthetic amylose AS-320 (Nakano Vinegar, Aichi, Japan) as described previously (16, 17). About 500 mg of the cyclic α-1,4-glucan mixture obtained was size-fractionated by gel-filtration chromatography through a 26 × 500 mm Superdex 30 column (Pharmacia). The fractions containing cyclic glucans with degree of polymerization from 17 to 30 were pooled, and the glucans were precipitated by adding 10 vol of ethanol. The size-fractionated CA mixture (25 mg) was dissolved in water, loaded onto a 20 × 250 mm ODS column (Daisopak SP-120–5-ODS-BP; Daiso, Osaka, Japan), and eluted with 6% (vol/vol) methanol. The glucan in the eluate was monitored with a refractive index detector (RID-6A; Shimadzu, Kyoto, Japan), and the peak for CA26 was collected, precipitated with 10 vol of ethanol, and lyophilized. The purity of CA26 was confirmed by high-performance anion-exchange chromatography and time-of-flight mass spectrometry.

Crystals of CA26 were grown from an aqueous solution containing 37.5% (vol/vol) PEG 400 (Merck) by using the vapor diffusion method. X-ray diffraction data were collected at synchrotron sources and processed by using the HKL package (22): 19,970 independent reflections to 1.1-Å resolution at beamline X31, European Molecular Biology Laboratory outstation [Deutsches Elektronen Synchrotron (DESY), Hamburg] and 25,416 reflections to 0.99-Å resolution at beamline DW32 [Laboratoire pour l'Utilisation du Rayonnement Electromagnétique (LURE), Orsay] were merged, yielding a data set containing 26,766 unique reflections (R_{merge} 0.037): space group triclinic, $P1$, a = 21.844 Å, b = 22.922 Å, c = 29.050 Å, α = 87.36°, β = 89.51°, γ = 61.98°. The asymmetric unit contains two CA26 molecules (312 C and 260 O atoms) and 76.75 water molecules distributed over 147 positions, of which 26 are ordered and fully occupied; total molecular mass of 2CA26·76.75H_2O is 8.8 kDa. Attempts to solve the structure by conventional direct methods were fruitless. Because the success rate of these methods falls off rapidly above about 150 unique atoms, and the largest structures so far determined by direct methods are about half the size of CA26, this was not unexpected. The structure was eventually solved by a real/reciprocal space recycling procedure (23) inspired by, but different in detail, from the "Shake and Bake" method of Weeks, Miller, Hauptmann, and their colleagues (24, 25). Instead of using randomly distributed atoms to generate starting phases, a diglucose fragment was employed along the lines suggested by Sheldrick and Gould (26). The phase sets were refined alternately in reciprocal space, optimizing the agreement with the predicted cosine invariants, and in real space, eliminating atoms to maximize 229 E_{calc}^2 (E_{obs}^2 − 1), in both cases using only E >1.4. Finally, the best solutions were refined further against all E values by iterative E-Fourier synthesis and elimination of peaks to maximize the correlation coefficient between E_{obs} and E_{calc} (26). The best solution had a correlation coefficient (27) of 81.8 and contained 595 sites of carbon and oxygen atoms. Subsequent fine-tuning of the algorithm enabled us to solve the structure by real/reciprocal space recycling starting from random atoms rather than from an oriented fragment.

Remaining atoms were located from electron and difference electron density maps (program O) and by using the SHELXWAT (28) procedure. The occupation factors of water oxygen atoms were assessed from electron densities and assigned as 0.25, 0.5, or 1.0, respectively, depending on maxima and their distances. They were held constant during refinement and add up to a total of 76.75 water molecules. For refinement of the crystal structure, the number of parameters was reduced by restraining chemically similar 1–2 and 1–3 interatomic distances (equivalent to bond distances and angles) in the 52 glucoses according to the protocol in the program SHELX 97. All non-hydrogen atoms were refined anisotropically; C—H hydrogens were placed in their calculated positions (C—H at 0.97 Å) and treated according to the "riding model," and O—H hydrogens identified from difference electron density maps were put in positions with normalized O—H distances, at 0.82 Å.

In the refinement of twofold disordered O(6) atoms, occupation factors were restrained to sum up to 1.0 and antibumping restraints were applied. To avoid "unstable behavior" of disordered water oxygen atoms, their occupation factors were held fixed and only positional and thermal parameters were refined; the minimal distance between disordered water molecules is 1.01 Å.

In the last cycle of refinement, the number of parameters was 6,617, the goodness-of-fit was 1.932, and the crystallographic R factor converged at 0.082 for 23,172 reflections with $F_{obs} \geq 4\sigma F_{obs}$ and at 0.089 for all the 25,473 reflections, based on F^2. To avoid overfitting of the model to the data, 5% of the reflections were randomly flagged for calculation of the free R factor before starting refinement (29), which converged at 0.10 for 1,224 reflections with $F_{obs} \geq 4\sigma F_{obs}$ and at 0.11 for all the 1,329 flagged data. The quality of the refinement and data may be assessed from electron densities shown in Fig. 1.

RESULTS

Overall View of the Crystal Structure. The triclinic unit cell of the 2CA26·76.75H_2O crystals (22–27) contains a total of 648.75 non-hydrogen atoms belonging to two symmetry-independent molecules A and B and 76.75 water molecules. The 76.75 water molecules are distributed over 147 sites, of which only 26 are fully occupied. The two CA26 molecules have comparable structures; they are clearly divided into two halves, glucose residues G1–G13 and G14–G26 (Figs. 2 and 3). In these segments, all glucoses are oriented *syn*; by contrast, the relative orientations of glucoses in the pairs G13–G14 and G26–G1 are *anti*, forming band-flip

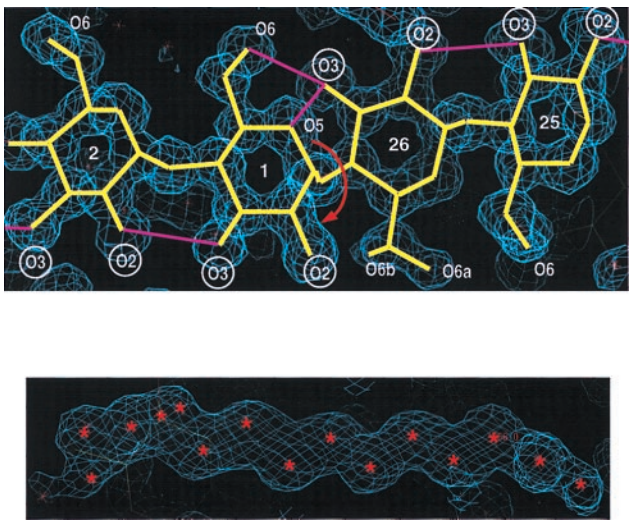

FIG. 1. Two sections of electron density at 1.0-Å resolution. (*A*) Segment of four glucoses, G25 to G2, showing one of the two band-flip sites in CA26 defined by G26-G1; these two glucoses are stabilized in *anti* orientation by the three-center hydrogen bond $O(3)_{26}$—H···($O(5)_1$,$O(6)_1$. The adjacent glucoses on both sides of the flip are oriented *syn* as usually found in amylose chains and hydrogen bonded $O(3)_n$···$O(2)_{n+1}$. Because the flip at G26-G1 involves not just a single glucose but the whole appended amylose chain, it was called "band-flip" (18). Labels of O(2) and O(3) hydroxyl groups are circled to emphasize the abrupt structural change at the band-flip site. (*B*) Disordered water molecules located in the channel-like cavity of the V-amylose helix. Because the distances between their positions (marked ∗) are shorter than the minimum hydrogen bonding distance of 2.5 Å (30), the occupations are around 0.5. Drawn with O (31).

motifs (ref. 18; see below and Fig. 1*A*). The two halves of CA26 are each folded into nearly two turns of single-stranded, left-handed V-amylose helix with six glucoses per turn. The helices are virtually antiparallel to each other and almost regular between glucoses G3–G13 and G16–G26, the remaining 2 × 2 glucoses forming interhelix links containing the band-flips. Along the axes of the V-helices are channel-like cavities with van der Waals diameters of 5.0–5.5 Å filled by statistically disordered water molecules (Fig. 1*B*). The CA26 molecule has pseudo-twofold rotational symmetry with the dyad axis located between the two V-helices and oriented perpendicular to their axes (Figs. 2 and 3). The rms deviations between the two halves, glucose residues G1–G13 and G14–G26, are only 0.33 Å and 0.35 Å for molecules A and B, respectively. If molecules A and B are superimposed, the rms deviation increases to 0.834 Å because the axes of helical segments are not strictly antiparallel in the two molecules but differ within 10°.

Conformational Features. The glucose residues in the two CA26 molecules are in the common $^4C_1$-chair form as observed thus far in all unsubstituted cyclodextrins.[‖] Their conformations are unstrained as shown by the Cremer–Pople parameters (35) (Table 1). Other parameters to define the $^4C_1$ glucose conformation are the virtual $O(4)_n$···$O(4)_{n+1}$ distances, with mean value of 4.39 Å in the helical parts and 4.65 Å at the band-flip sites (G13, G14; G26, G1), indicating minor strain. The mean angles at the glycosidic oxygens, $C(1)_{n+1}$–$O(4)_n$–$C(4)_n$, are 118.3° in the two CA26 molecules. These parameters are in agreement with β-cyclodextrin** and probably representative of an unstrained amylose chain.

Because the glucose residues G1–G13 and G14–G26 in the two halves of CA26 are oriented *syn*, associated ϕ and ψ torsion angles[††] are in the usual range: ϕ, 91° to 115°; and ψ, 97° to 131°. The torsion angles χ about the C(5)–C(6) bonds defined by O(5)–C(5)–C(6)–O(6) are predominantly in the preferred (−)*gauche* range, some are disordered in (−)*gauche* and (+)*gauche*, and those at G1, G7, G14, and G20 are systematically (+)*gauche* in molecules A and B (see below); *anti* is not observed, again in agreement with the cyclodextrins (12, 13).

Geometry of the Band-Flips and Kinks. The band-flip is a structural feature of the amylose chain that has not been observed before for the smaller cyclodextrins, CA6 to CA8. In CA9, elliptical distortion of the macrocycle (21) indicates severe steric strain, supported by potential energy calculations (18). In CA10 and the larger CA14, this strain would be even more increased were it not relieved by two band-flips occurring at diametrically opposed sites in the macrocycles, where one glucose each is flipped from the common *syn* orientation into *anti* (18). In 2CA26·76.75H_2O, the same band-flip motifs are also observed at diametrically opposed glucose pairs G13–G14 and G26–G1. In both CA26 molecules the band-flips are stabilized by three-center hydrogen bonds (Fig. 1*A* and Table 1), with $O(3)_n$—H donating minor (long; mean 3.07 Å) and major (short; mean 2.73 Å) components to the acceptors $O(5)_{n+1}$ and $O(6)_{n+1}$, respectively, and characterized by mean torsion angles ϕ at 88° and ψ at −48° for molecules A and B.

The band-flips are associated with "kinks" where adjacent glucoses are rotated relative to each other, but still in *syn* orientation, leading to widening of the $O(3)_n$···$O(2)_{n+1}$ hydrogen bonds (18). In CA26, the observed kinks are minor, widening the $O(3)_{12}$···$O(2)_{13}$ and $O(3)_{25}$···$O(2)_{26}$ distances to 2.98 Å and 3.28 Å, respectively, in molecule A and to 3.15 Å and 3.11 Å in molecule B. This finding confirms the expectation (18) that kinks will become smaller with increasing chain length of a cycloamylose because conformational strain can dissipate along the oligosaccharide.

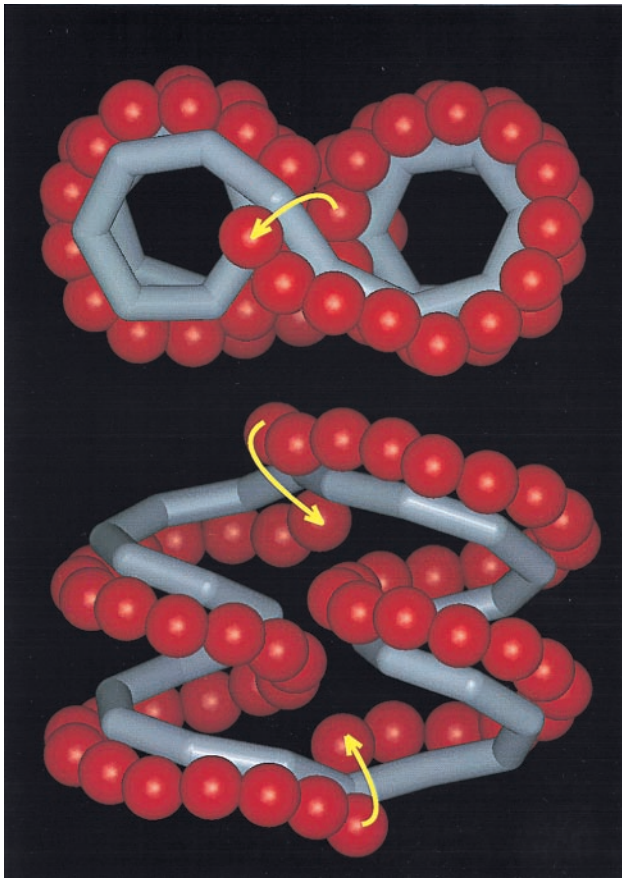

FIG. 2. Top and side views of CA26, showing schematically the folding of the macrocycle into two left-handed single helices connected in the form of a "figure eight." O(2) and O(3) oxygen atoms are drawn red and the other atoms are symbolized by a gray tube passing through O(4) atoms; all other atoms are omitted. Band-flips occur at the two positions in the macrocycle indicated by yellow arrows. Overall size of CA26: 27 Å × 15 Å × 21 Å (height). Drawn with INSIGHT II (32).

Intramolecular Direct and Water-Mediated Hydrogen Bonds Stabilize the Folding of CA26. Except for the band-flip sites where adjacent glucoses are oriented *anti*, all other glucoses in 2CA26·76.75H_2O are *syn* and hydrogen bonded $O(3)_n$···$O(2)_{n-1}$ at a mean of 2.86 Å (Table 1), the same distance as found in β-cyclodextrin.** In addition, the V-helical turns described by chain segments G3–G13 and G16–G26 in molecules A and B are stabilized by five $O(6)_n$···$O(2)_{n+6}$ and/or $O(6)_n$···$O(3)_{n+6}$ hydrogen bonds each. The hydrogen bonds differ in length depending on the orientation of the C(6)–O(6) groups defined by torsion angles χ; they are (−)*gauche* in 11, disordered (−)*gauche* (+)*gauche* in 5, and (+)*gauche* in 4 glucoses. For C(6)–O(6) in (−)*gauche*, hydrogen bonds $O(6)_n$···$O(3)_{n+6}$ are shorter than $O(6)_n$···$O(2)_{n+6,}$ whereas the reverse is found for the (+)*gauche* form (Table 1).

A notable hydrogen bonding scheme connects glucoses G1, G7, G14, and G20, whose torsion angles χ are all (+)*gauche*; G1 and G14 are at band-flip sites and are hydrogen bonded to G7 and G20 and two water molecules, W3 and W4 (see below; Fig. 3*B*). Besides being acceptors of the major (short) components of the three-center hydrogen bonds donated by $O(3)_{26}$—H and $O(3)_{13}$, respectively, the O(6) hydroxyls of G1 and G14 are engaged in hydrogen bonds $O(6)_1$···$O(3)_7$ and $O(6)_{14}$···$O(3)_{20}$, respectively. This scheme adds to the stability of the band-flip motifs, and it

[‖]In two complexes of the fully methylated heptakis-2,3,6-trimethyl-β-cyclodextrin, one glucose each was found in $^1C_4$-chair (33) and ^{3,0}B-boat (34) forms.

**Some mean geometrical data for α-, β-, and γ-cyclodextrin taken from ref. 12: glycosidic bonds $C(1)_{n+1}$-$O(4)_n$-$C(4)_n$: α 119.0°, β 117.7°, γ 112.6°; virtual distances $O(4)_n$··· $O(4)_{n-1}$: α 4.23 Å, β 4.36 Å, γ 4.48 Å; hydrogen bonds $O(2)_n$··· $O(3)_{n-1}$: α 3.00 Å, β 2.86 Å, γ 2.81 Å.

[††]Definition of ϕ, ψ torsion angles: ϕ, $O(5)_n$-$C(1)_n$-$O(4)_{n-1}$-$C(4)_{n-1}$ and ψ, $C(1)_n$-$O(4)_{n-1}$-$C(4)_{n-1}$-$C(3)_{n-1}$; see International Union of Pure and Applied Chemistry recommendations (36).

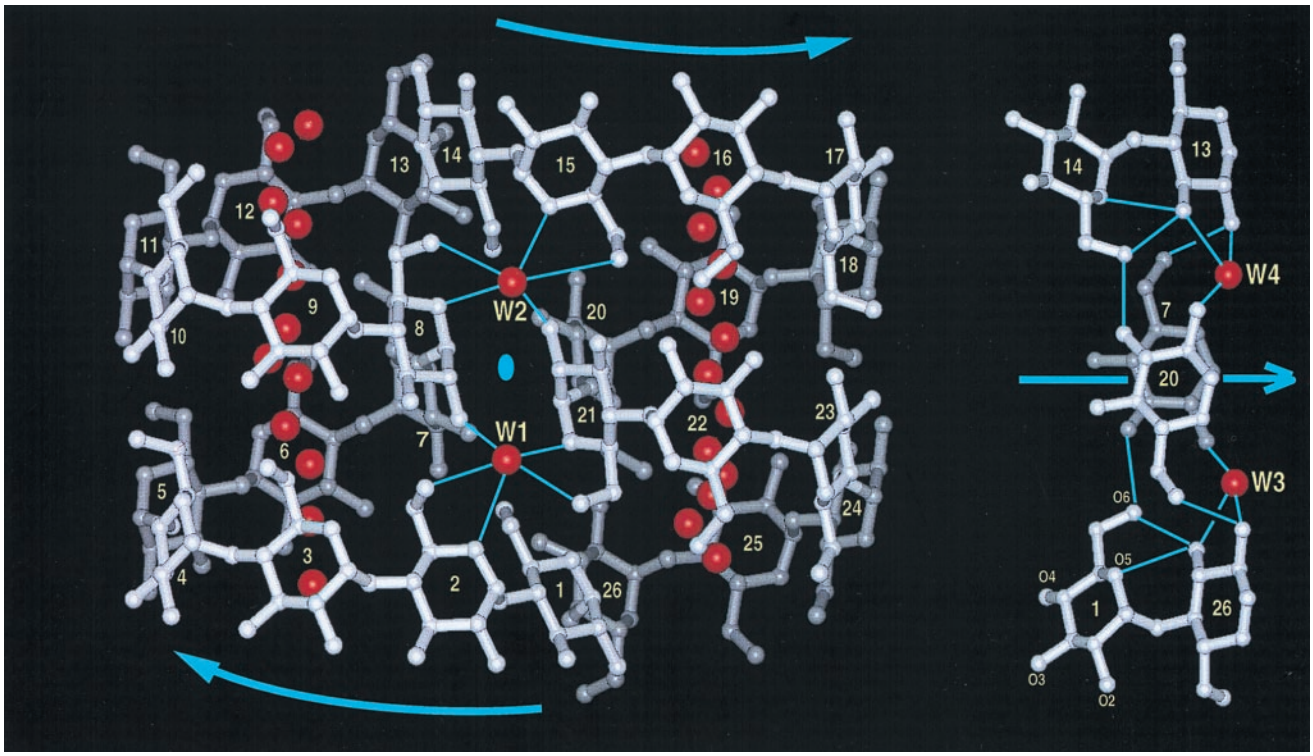

FIG. 3. *(Left)* View of CA26, molecule A. Those parts of the molecule located at the front are light gray, those at the back dark; glucose residues G are numbered 1–26. Curved arrows indicate the "directions" of the two V-helices. The location of the pseudo-twofold rotation axis (●) is between water molecules W1 and W2 and vertical to the plane of the paper. Disordered water molecules filling the channel-like cavities in the two V-helical segments are shown in red, as are two five-coordinated water molecules, W1 and W2, in strategic positions conferring stability to the folding of CA26. They are hydrogen bonded in bidentate mode to O5/O6 atoms; W1 to G21 and G2; W2 to G8 and G15, with additional single hydrogen bonds W1···O(2)$_8$ and W2···O(2)$_{21}$. Water molecules W3 and W4 are not shown, as they are at the "back" of the molecule and nearly overlap with W1 and W2. *(Right)* View as in *Left*, but rotated 90° so that the pseudo-twofold axis (blue arrow) is now horizontal. Shown is the interface between the two V-helices where the band-flip segments, glucoses G26–G1 and G13–G14, are connected by direct and water (W3 and W4)-mediated hydrogen bonds to glucoses G7 and G20. Note superposition of G7 on G20, forming part of the interface between the V-helices. Oxygen atoms of glucose 1 are labeled. Drawn with INSIGHT II (32).

superimposes G7 and G20 at the interface between the two V-helices.

It is remarkable that there are no other direct hydrogen bonds between glucoses in this tightly coiled molecule. This lack also holds for intermolecular interactions between the V-helices, which are all mediated by water molecules.

The coordination of the water molecules to CA26 follows patterns that have been observed previously (9, 10, 12, 18). The most prominent motifs are five-membered rings formed by water molecules coordinated in bidentate mode to O(2) and O(3) and/or to O(5) and O(6) belonging to the *same* glucose. A number of such interactions are found in CA26, the most

Table 1. Selection of geometrical data in CA26, molecules A and B

Parameter		In helical part			At band-flip sites		
		Max	Min	Mean	Max	Min	Mean
Glycosidic angle, °[a]	C(1)$_{n+1}$–O(4)$_n$–C(4)$_n$	120.7	114.6	118.3			
Virtual distance, Å[b]	O(4)$_n$···O(4)$_{n-1}$	4.56	4.11	4.39	4.79	4.48	4.65
Virtual angles, °[c]	O(4)$_n$···O(4)$_{n+1}$···O(4)$_{n+2}$	136.7	117.9	126.4	149.8	138.7	143.6
Hydrogen bonds, Å	O(3)$_n$···O(2)$_{n+1}$ (interglucose)	3.10	2.65	2.86			
	O(3)$_n$···O(6)$_{n+1}$ (at band-flips)				2.78	2.66	2.73
	O(3)$_n$···O(5)$_{n+1}$				3.12	3.03	3.07
	(−)*gauche*-O(6)$_n$···O(2)$_{n+6}$[d]	3.42	3.09	3.27			
	(−)*gauche*-O(6)$_n$···O(3)$_{n+6}$[e] (between turns)	3.27	2.65	2.85			
	(+)*gauche*-O(6)$_n$···O(2)$_{n+6}$[f]	2.97	2.66	2.75			
	(+)*gauche*-O(6)$_n$···O(3)$_{n+6}$[g]	3.28	3.09	3.23			
Torsion angles, °[h] ϕ,	O(5)$_n$–C(1)$_n$–O(4)$_{n-1}$–C(4)$_{n-1}$	115.3	91.1	103.6	90.2	86.4	88.1
Torsion angles, °[h] ψ,	C(1)$_n$–O(4)$_{n-1}$–C(4)$_{n-1}$–C(3)$_{n-1}$	131.4	97.4	115.3	−51.7	−45.8	−48.4
Cremer–Pople glucose puckering amplitude Q, Å		0.58	0.52	0.54	0.58[i]	0.53[i]	0.54[i]
angle θ, °		9.8	1.0	4.6	14.7[i]	1.1[i]	8.3[i]

[a]For all glucoses because these angles are not different at band-flips.
[b]Helical parts: n from 3 to 13 and 16 to 26.
[c]Helical parts: n from 2 to 11 and 15 to 24.
[d–g]Helical parts as in b; averaged over different number of glucoses: d, 11; e, 10; f, 9; and g, 2.
[h]Defined as in footnote ††.
[i]For glucoses G26–G1 and G13–G14.

interesting involving water molecules W1 to W4 located *between* the V-helices (Fig. 3). W1 is hydrogen bonded simultaneously to O(5) and O(6) of glucose residues G21 and G2 and additionally to O(2) of G8; W2 is related to W1 through the pseudo-twofold symmetry and in the same configuration with G8 and G15 and forms an additional bond to O(2) of G21. Another pair of water molecules related by the pseudosymmetry, W3 and W4, mediates hydrogen bonds between O(2) and O(3) hydroxyls of glucoses G13 and G26 at the band-flip sites and O(2) of G7 and G20, respectively, namely, $O(2)_7$···W3···$O(2)_{26}$, $O(3)_{26}$ and $O(2)_{20}$···W4···$O(2)_{13}$, $O(3)_{13}$ (Fig. 3 *Right*). All these interactions contribute to the stabilization of the band-flip sites and of the structure of the linker segment between the V-helices formed by glucoses G13, G14, G15 and G26, G1, G2.

Disordered Water Molecules Fill the Channel-Like Cavities in the V-Helix. Because the central channels formed by the V-helices in CA26 are coated by C(3)—H, C(5)—H, C(6)—H_2 and ether-like O(4) groups, they are hydrophobic in character so that only van der Waals interactions can form with molecules that fit into them (Fig. 3 *Left*). They are disordered with occupancies around 0.5, and distances between the water sites are in the range 1.01–1.94 Å (Figs. 3 *Left* and 1*B*)—i.e., water molecules in adjacent sites cannot be present simultaneously because their distances are shorter than the minimum hydrogen-bonding distance of about 2.5 Å (30). This restriction does not hold for the water sites at the "ends" of the channels, where direct hydrogen bonding to glucose O(2), O(3), and O(6) hydroxyls may stabilize full occupation.

NMR Studies on CA26. ^{13}C NMR studies on CAs dissolved at 50°C in ^2H_2O and containing 6–26 glucoses in the macrocycle (18, 37) showed a distinct shift of the signals for the ^{13}C atoms at the glycosidic link, $C(1)_{n+1}$–$O(4)_n$–$C(4)_n$. For CA6 to CA8, signals occur at 102.4 ppm for ^{13}C(1) and 81.8 ppm for ^{13}C(4), but for CA10 and higher homologues, they are at 100.2 ppm for ^{13}C(1) and 78.3 ppm for ^{13}C(4); the corresponding ^{13}C signals for CA9 are at intermediate positions. These shifts are interpreted as the band-flips observed for CA10 and their higher homologues but not for CA6 to CA8 (18). The positions of ^{13}C signals for C(2), C(3), C(5), and C(6) are as expected and do not show a defined trend. Since only six sharp signals are observed for each CA, the molecules experience rapid conformational changes in solution; the band-flips and V-helices in CA26 are not static but migrate through the macrocycle so that all glucoses are equivalent on the NMR time-scale.

A Model for V-Amylose. An "infinite" V-amylose helix has been constructed by superposition of the central 6 glucose residues of the four independent short helices in the CA26 crystal; with rms deviation of only 0.20 Å [omitting O(6) atoms in (+)*gauche*]. The resulting structure agrees well with that proposed for V-amylose on the basis of x-ray fiber diffraction, electron diffraction, and ^{13}C CP/MAS NMR (refs. 1–8; Fig. 4). There are exactly 6 glucose residues per turn with the following mean geometrical data: pitch, 7.9 Å; virtual bond $O(4)_n$···$O(4)_{n+1}$, 4.26 Å; glycosidic angle $C(1)_{n+1}$–$O(4)_n$–$C(4)_n$, 118.7°; all C(6)–O(6) groups are oriented (−)*gauche*. The V-helix is stabilized by hydrogen bonds $O(3)_n$···$O(2)_{n+1}$, 2.89 Å, between adjacent glucoses and, between the turns, by $O(6)_n$···$O(3)_{n+6}$, 2.86 Å; the hydrogen bonds $O(6)_n$···$O(2)_{n+6}$, 3.16 Å, are somewhat longer. If the V-helix is modeled with C(6)–O(6) groups oriented (+)*gauche*, hydrogen bonds $O(6)_n$···$O(2)_{n+6}$ are shorter, 2.75 Å, than $O(6)_n$···$O(3)_{n+6}$, 3.23 Å, as discussed in the section *Intramolecular Direct and Water-Mediated Hydrogen Bonds*. . . This alternative possibility for hydrogen bond formation will add to the flexibility of V-amylose which is, in fact, observed—the V-helix might accommodate 6 to 8 glucoses per turn, depending on the size of the molecules that are enclosed in the channel-like cavity (3).

According to x-ray fiber and single-crystal electron diffraction data from fibrous and microcrystalline V-amylose preparations, the helices are oriented antiparallel and preferentially arranged in pseudohexagonal lattices with the orthorhombic space group $P2_12_12_1$ and typical unit cell constants a = 13.6 Å, b = 23 Å, and c = 8.0 Å (3). In CA26, the reduced unit cell containing only one V-helix in the asymmetric unit is pseudohexagonal with $a \approx b$ = 13.8 Å and $\gamma \approx 60°$. If this is transformed into a C-centered pseudo-orthorhombic lattice, the unit cell constants become a = 13.8 Å and b = 24.5 Å, in close agreement with above data from x-ray fiber diffraction. This comparison suggests that the packing arrangements and intra- and intermolecular interactions found between the short V-helices in CA26 correspond to those derived from fibrous and microcrystalline V-amylose.

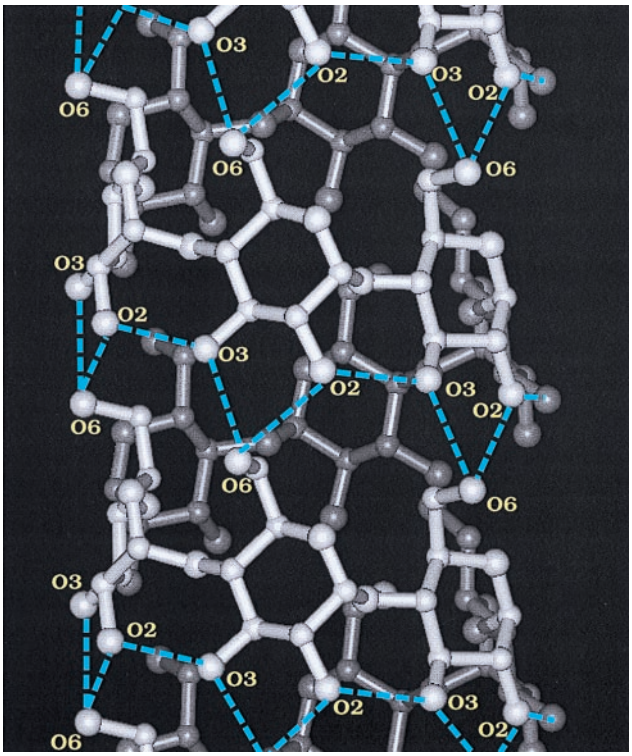

FIG. 4. View of the ideal left-handed, single-stranded V-amylose helix constructed on the basis of the central 6 glucoses in the four helical parts of the two symmetry-independent CA26 molecules in the triclinic unit cell. All C(6)–O(6) groups are oriented (−)*gauche*, hydrogen bonds $O(3)_n$···$O(2)_{n+1}$ and $O(6)_n$···$O(2)_{n+6}$; $O(6)_n$···$O(3)_{n+6}$ drawn in blue; geometrical data are given in the text. Drawn with INSIGHT II (32).

Comparing the Models: Single-Crystal Versus Fiber X-Ray Diffraction. The available structural data on V-amylose obtained in the present study and by x-ray fiber diffraction analyses of V-amylose complexed with iodine (38), water (39), or DMSO (40) indicate that the overall structures are identical. The fiber work identified 6 glucoses per turn with pitch in the range 7.91–8.17 Å, the V-helix being left-handed and stabilized by hydrogen bonding $O(3)_n$···$O(2)_{n+1}$ and $O(6)_n$···$O(2)_{n+6}$, with C(6)–O(6) groups being oriented (−)*gauche* or (+)*gauche*, depending on model. The alternate hydrogen bond $O(6)_n$···$O(3)_{n+6}$, found to be 0.18 Å shorter than $O(6)_n$···$O(2)_{n+6}$ in our study (see above), was not modeled in the fiber diffraction analyses. The overall geometries are comparable, but they are more narrowly confined in the x-ray fiber diffraction work to those known for α-cyclodextrin than, as found in the present study, to the more relaxed β-cyclodextrin.** This might be the reason for the discrepancies in hydrogen bonding interactions involving $O(6)_n$ and $O(2)_{n+6}$ or $O(3)_{n+6}$.

CONCLUSIONS

This study reports the three-dimensional structure of the largest oligosaccharide obtained so far in single-crystal form suitable for x-ray diffraction studies. With 648.75 non-hydrogen atoms it is

also the largest "equal-atom" structure for which the crystallographic phase problem has been solved by real/reciprocal space recycling methods. The CA26 macrocycle is tightly folded into a compact structure with pseudo-twofold symmetry featuring two short, two-turn V-amylose helices in antiparallel orientation. The helices are connected by three-glucose linkers with *trans*-oriented glucoses forming band-flips, as in CA10 and CA14. This structural motif serves to relieve steric strain in amylose chains (18), permitting all glucoses to adopt the preferred $^4C_1$-chair form and to be oriented *syn* with formation of interglucose $O(3)_n \cdots O(2)_{n+1}$ hydrogen bonds.

The band-flips explain readily why amylose precipitates as plate-like, flaky crystals with molecules forming inclusion complexes. The V-amylose chain folds back and forth in these crystals to yield antiparallel packing of the helices in pseudohexagonal form, the folding being facilitated or even made possible by the band-flip motif. This also appears to be of importance for the packing of amylose in seeds and tubers. As observed with x-ray methods, 20–40% of the amylose in starch granules adopts A and B forms (41, 42), but there is no indication of the V-form. According to ^{13}C CP/MAS NMR, however, the V-form is present as well (7, 8), and probably forms inclusion complexes with small molecules such as fatty acids and alkanes occurring in substantial amounts in seeds and tubers. It appears that the segments of the amylose chain adopting the V-form are so short that they escape detection by x-rays, which require more extended repetitive structures (helical turns) to yield measurable diffraction intensities. The same might be true for amylopectin, where short, unbranched segments could occur in V-form associated with band-flips.

Of great interest is the formation of inclusion complexes because of the channel-like cavity in the V-amylose helix. In the crystal structure of $2CA26 \cdot 76.75H_2O$ the channels are occupied by disordered water molecules. The availability of CA26 and of other CAs with a wide range of chain lengths will permit investigation of a variety of inclusion complexes of V-amylose. We expect a correlation between the size and nature of the enclosed guest molecules with the geometries of the V-helices defined by the number of glucoses per turn and by the hydrogen bonds formed between O(6) and O(2), O(3) hydroxyl groups. This information will also provide an answer to the long-standing question about the structural characteristics of "iodine's blue," which was discovered 185 years ago (43).

In the series of CAs from CA10 to CA32, crystallization succeeded so far only with CA10, CA14, and CA26. The CAs show pronounced differences in solubilities which we correlate with differences in structural flexibility. The least-soluble CAs were those obtained in crystalline forms. This behavior is probably similar to that of the cyclodextrins, where CA7 is about 1/10 as soluble in water as CA6 and CA8 (11, 44) and crystallizes more readily.

Crystallization of CA10, CA14, and CA26 may occur because (*i*) due to cyclization, they are conformationally restricted; (*ii*) they have inherent molecular twofold symmetry so that the band-flips are diametrically opposed; (*iii*) they are internally stabilized by tight folding and intramolecular hydrogen bonding; and (*iv*) in CA26, water molecules W1 to W4 are in strategic positions. For these reasons, it is questionable whether CA24 or CA28 would crystallize—one is too short, the other too long to provide structural features as found for CA26; the best candidate may be CA38, with V-helices three turns long.

^{13}C NMR studies and model building used to interpret small-angle x-ray data for CA21 in aqueous solution (45) indicate that the larger CAs have enormous conformational flexibility. Of the many models fitted against the data for CA21, the best one was shaped like a threefold propeller and had no band-flip and no V-type helix. We also tried to model CA26 before the x-ray structure determination. Although our models contained the band-flips, all were wrong, showing how elusive model building may be.

Because CA26 is available on a preparative scale and has been characterized structurally, it may now serve as a prototype of the V-amylose helix. Of interest will be investigations in solution where spectroscopic methods can be used to study the structure of CA26 per se and its kinetics and affinities of complex formation with a large variety of guest molecules. Crystallographic studies will also continue, as we have obtained an orthorhombic form of CA26 hydrate and two inclusion complexes with triiodide, one light and the other dark brown.

K.G. acknowledges discussions with Prof. H. Chanzy and his group on amylose folding and crystallization, and thanks J. Shimada and D. Hoffmann for performing molecular modeling on models of larger CA. We are grateful to the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie for financial support of this work and for synchrotron beamtime at EMBL/DESY, Hamburg, and LURE, Orsay. This work was supported in part by a grant for the development of the next generation of bioreactor systems from the Society for Techno-Innovation of Agriculture, Forestry and Fisheries (STAFF), Portugal.

1. Galliard, T., ed. (1987) *Starch Properties and Potential* (Wiley, New York).
2. French, D. (1984) *Starch: Chemistry and Technology* (Academic, New York), 2nd Ed.
3. Sarko, A. & Zugenmaier, P. (1980) in *Fiber Diffraction Methods,* eds. French, A. D. & Gardner, K. C. H., ACS Symposium Series (Am. Chem. Soc., Washington, DC), No. 141, pp. 459–482.
4. Rappenecker, G. & Zugenmaier, P. (1981) *Carbohydr. Res.* **89,** 11–19.
5. Murphy, V. G., Zaslow, B. & French, A. D. (1975) *Biopolymers* **14,** 1487–1501.
6. Brisson, J., Chanzy, H. & Winter, W. T. (1991) *Int. J. Biol. Macromol.* **13,** 31–39.
7. Veregin, R. P., Fyfe, C. A. & Marchessault, R. H. (1987) *Macromolecules* **20,** 3007–3012.
8. Gidley, M. J. & Bociek, S. M. (1988) *J. Am. Chem. Soc.* **110,** 3820–3829.
9. Hinrichs, W., Büttner, G., Steifa, M., Betzel, C., Zabel, V., Pfannemüller, B. & Saenger, W. (1987) *Science* **238,** 205–208.
10. Hinrichs, W. & Saenger, W. (1990) J. Am. Chem. Soc. **112**, 2789–2796.
11. Imberty, A., Chanzy, H., Pérez, S., Buléon, A. & Tran, V. (1988) *J. Mol. Biol.* **201**, 365–378.
12. Saenger, W. (1984) in *Inclusion Compounds,* eds. Atwood, J. L., Davies, J. E. D. & MacNicol, D. D. (Academic, London), Vol. 2, pp. 231–259.
13. Harata, K. (1996) in *Comprehensive Supramolecular Chemistry,* eds. Atwood, J. L., Davies, J. E. D., MacNicol, D. D. & Vögtle, F. (Pergamon, Oxford), Vol. 3, pp. 279–304.
14. French, D., Pulley, A. D., Effenberger, J. A., Ronghvie, V. A. & Abdullah., M. (1965) *Arch. Biochem. Biophys.* **111,** 153–160.
15. Huber, O. & Szejtli, J., eds. (1988) *Proceedings of the IV. International Symposium on Cyclodextrins* (Kluwer, Dordrecht, the Netherlands).
16. Takaha, T., Yanase, M., Takata, H., Okada, S. & Smith, S. M. (1996) *J. Biol. Chem.* **271,** 2902–2908.
17. Terada, Y., Yanase, M., Takata, H., Takaha, T. & Okada, S. (1997) *J. Biol. Chem.* **272**, 15729–15733.
18. Jacob, J., Gessler, K., Hoffmann, D., Saube, H., Koizumi, K., Smith, S. M., Takaha, T. & Saenger, W. (1998) *Angew. Chem.* **110,** 606–609.
19. Ueda, H., Endo, T., Nagase, H., Kobayashi, S. & Nagai, T. (1996) *J. Inclusion Phenom. Mol. Recognit. Chem.* **25,** 17–20.
20. Harata, K., Endo, T, Ueda, H. & Nagai, T. (1998) *Supramol. Chem.* **9,** 143–149.
21. Fujiwara, T., Tanaka, N. & Kobayashi, S. (1990) *Chem. Lett.*, 739–742.
22. Otwinowski, Z. & Minor, W. (1996) *Methods Enzymol.* **276,** 307–326.
23. Sheldrick, G. M. (1997) in *CCP4 Meeting* York, January, pp. 147–157.
24. Miller, R., DeTitta, G. T., Jones, R., Langs, D. A., Weeks, C. M. & Hauptmann, H. A. (1993) *Science* **259,** 1430–1433.
25. Miller, R., Gallo, S. M., Khalak, H. G. & Weeks, C. M. (1994) *J. Appl. Crystallogr.* **27,** 613–621.
26. Sheldrick, G. M. & Gould, R. O. (1995) *Acta Crystallogr. B* **51,** 423–431.
27. Fujinaga, M. & Read, R. J. (1987) *J. Appl. Crystallogr.* **20,** 517–521.
28. Sheldrick, G. M. & Schneider, T. R. (1997) *Methods Enzymol.* **277,** 339–343.
29. Brünger, A. T. (1992) *Nature (London)* **355,** 472–475.
30. Jeffrey, G. A. & Saenger, W. (1991) *Hydrogen Bonding in Biological Structures* (Springer, Heidelberg).
31. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47,** 110–119.
32. Molecular Simulations (1997) INSIGHT II 97.0, (Molecular Simulations, San Diego).
33. Caira, M. R., Griffith, V. J., Nassimbeni, S. R. & van Oudtshoorn, B. (1994) *J. Chem. Soc. Perkin Trans.* **2,** 2071–2072.
34. Harata, K. (1988) *J. Chem. Soc. Chem. Commun.,* 928–929.
35. Cremer, D. & Pople, J. A. (1975) *J. Am. Chem. Soc.* **97,** 1354–1358.
36. IUPAC–IUB Joint Commission on Biochemical Nomenclature (1983) *Eur. J. Biochem.* **131,** 5–7.
37. Endo, T., Nagase, H., Ueda, H., Kobayashi, S. & Nagai, T. (1997) *Chem. Pharm. Bull.* **45,** 532–536.
38. Bluhm, T. L. & Zugenmaier, P. (1981) *Carbohydr. Res.* **89,** 1–10.
39. Zaslow, B., Murphy, V. G. & French, A. D. (1974) *Biopolymers* **13,** 779–790.
40. Winter, W. T. & Sarko, A. (1974) *Biopolymers* **13,** 1461–1482.
41. Morsi, M. K. S. & Sterling, C. (1966) *Carbohydr. Res.* **3,** 97–105.
42. Nara, S. (1978) *Stärke* **30,** 183–186.
43. Colin, J. J. & de Claubry, H. G. (1814) *Ann. Chim.* **90,** 87–92.
44. Szejtli, J. (1988) *Cyclodextrins and Their Inclusion Complexes* (Académiai Kiado, Budapest), p. 34.
45. Kitamura, S., Isuda, H., Shimada, J., Takada, T., Takaha, T., Okada, S., Mimura, M. & Kajiwara, K. (1997) *Carbohydr. Res.* **304,** 303–314.