Computation of Low-Energy Crystalline Arrangements of Cellulose Triacetate

Romain M. Wolf* and Eric Francotte

Central Research Laboratories, CIBA-GEIGY Ltd., 4002 Basel, Switzerland

Leslie Glasser,† Istvan Simon,‡ and Harold A. Scheraga*

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853-1301

Received April 19, 1991

ABSTRACT: Low-energy crystal packings of cellulose triacetate (CTA) have been computed, based on a "two-dimers" unit cell as suggested for CTA I. In addition to the different packing arrangements, parallel versus antiparallel, the effect of changes in the main-chain conformation on the crystal packing was investigated. Minor alterations of the main-chain torsional angles were sometimes found to induce significant changes of unit-cell parameters, although yielding structures with comparable energies. Since small changes of the main-chain conformation do not involve large energy barriers, it is therefore assumed that, in crystalline CTA, local transitions between unit cells with slightly different parameters are possible. Even highly ordered CTA samples might thus consist of a distribution of marginally differing crystalline domains. The resulting imperfections at the interfaces of slightly different unit cells, conceivable on the basis of the computed structures, might explain the problems encountered in the experimental elucidation of the crystal structures in both CTA morphologies, CTA I and CTA II.

Introduction

Cellulose triacetate (CTA) is the most common derivative of cellulose. It is readily obtained from cellulose by homogeneous or heterogeneous acetylation. Depending on the crystal structure of the starting cellulose and on the acetylation conditions, two distinct crystal structures of CTA can be obtained. By analogy with cellulose, these have been called CTA I and CTA II. As mentioned below, the supramolecular structure of CTA can strongly influence its behavior under certain conditions, and it is highly desirable to obtain information about the structure of CTA. Detailed crystal structures of the two CTA morphologies have been proposed,[1,2] but some questions raised by more recent findings[3-5] are as yet unanswered.

The relation between the major cellulose structures (I and II) and CTA I and CTA II has often been represented by the so-called Sprague diagram (shown in expanded form in Figure 1).[2,6] The original Sprague diagram suggests that CTA I can be obtained only from cellulose I (i.e., "native" cellulose), and only if the acetylation is carried out heterogeneously, i.e., in such a way that the polymer is never dissolved or strongly swollen. The proposed parallel chain arrangement of cellulose I[7-9] would thus basically be conserved in CTA I.[1]

Dissolution of cellulose triacetate during the acetylation of cellulose I is thought to result in CTA II. Acetylations of cellulose II always yield CTA II under either heterogeneous or homogeneous conditions. Similarly, dissolution of CTA I leads irreversibly to CTA II. As for cellulose II,[10-13] CTA II is suggested as having an antiparallel chain arrangement.[2] This interpretation was thought to be confirmed by the fact that the back-reaction (i.e., saponification) normally yields the corresponding cellulose.[6] It was supposed for a long time that CTA II cannot be reconverted to CTA I, implying that CTA II is the thermodynamically favored structure. This would again be in accord with results obtained on cellulose itself.

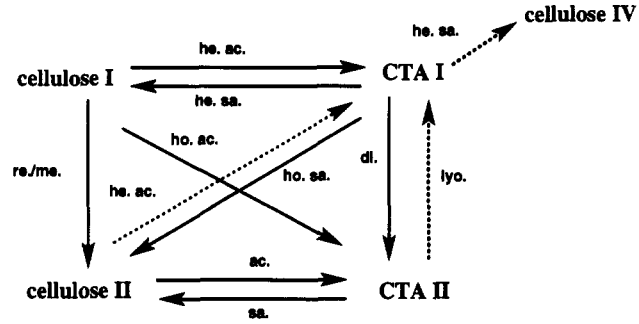

Figure 1. Relations among cellulose and cellulose triacetate (CTA) morphologies. Solid arrows represent the connections as suggested by Sprague et al.,[6] and dotted arrows illustrate more recent findings,[3-5] partially contradicting the earlier data. The diagram does not differentiate between cellulose IV_I and cellulose IV_{II} (see also the text). he. ac. = heterogeneous acetylation, he. sa. = heterogeneous saponification, ho. ac. = homogeneous acetylation, ho. sa. = homogeneous saponification, ac. = any acetylation conditions (heterogeneous or homogeneous), sa. = any saponification conditions (heterogeneous or homogeneous), di. = dissolution and recrystallization, re./me. = regeneration or mercerization, lyo. = lyotropic (liquid–crystalline) solution.

A number of contradicting experimental findings have contested the simple relations presented in the original Sprague diagram. (These more recent results are indicated schematically in Figure 1 by the dotted arrows.) In 1968, Watanabe et al.[14] presented evidence that CTA I might be obtained from cellulose II. These data were confirmed by more recent work from the same group.[4] Furthermore, CTA samples regained from lyotropic (i.e., liquid–crystalline) solutions of CTA II were found to display the X-ray diffraction pattern of CTA I.[5] Heterogeneous saponification of such samples leads preferentially to cellulose IV, another polymorph of cellulose, supposed to exist in two modifications, namely cellulose IV_I and cellulose IV_{II}.[15] Prior to the work on lyotropic solutions,[5] which does not specify the form of the cellulose IV actually obtained, it had already been reported[3] that heterogeneous saponification of CTA I yields cellulose IV_{II} (contrary to older data, shown in the original Sprague diagram, suggesting that this material is cellulose I). Conflicting

* To whom all correspondence should be addressed.

† On leave from the Department of Chemistry, University of Witwatersrand, Wits 2050, South Africa.

‡ On leave from the Institute of Enzymology, Hungarian Academy of Sciences, P.O. Box 7, Budapest 1518, Hungary.

0024-9297/92/2225-0709$03.00/0

experimental data of this kind indicate that, despite numerous investigations, knowledge about the actual structures of CTA I and CTA II is rather sparse.

Besides its numerous applications (fibers for the textile industry, membranes, etc.), cellulose triacetate has been recognized in the last decade as an extremely powerful chiral polymeric sorbent for the chromatographic separation of enantiomers.[16-18] It has been shown that the supramolecular structure of CTA definitely influences the resolving power and the overall interaction mechanisms.[19] Amorphous CTA, as well as CTA II, and highly annealed CTA I all exhibit very poor resolving power. Only CTA I, as obtained directly from the heterogeneous acetylation of microcrystalline cellulose I,[20] has optimal properties to serve as a versatile chiral sorbent. Interactions with small molecules are supposed to occur by inclusion of the low molecular mass compounds between the (ordered) chains of the polysaccharide.[16,19,21] Recent investigations imply that this type of CTA I must offer a multitude of structurally different interaction sites for low molecular mass compounds.[18] Although numerous different sites can be envisaged even in a single unique crystal structure, one can imagine that CTA I is actually an assembly of slightly different polymorphs and that the various crystal packings (and especially their interfaces) within the same CTA I sample may be responsible for the broad applicability of the polymer. The loss of chromatographic retention (and thereby of resolution power) on annealing optimal CTA I sorbents[19] might be explained by the fact that the annealing process favors one morphology and removes the imperfections introduced by interfaces between different unit cells that do not match perfectly. Precisely such interfaces could be locations where inclusion of low molecular mass compounds is preferred, in contrast to regions of high long-range order. The interest in such interactions between small molecules and CTA I[18,21] and an earlier successful application of the combination of the crystal packing program WMIN[22] and the force fields ECEPP[23-25] and UNICEPP[26] to the structures of cellulose[27,28] have motivated this effort to compute low-energy crystalline arrangements of CTA.

Determination of the crystal structures of CTA I and CTA II has already involved some computations of intra- and interchain interaction energies.[1,2] The major goal of those studies was to fit experimental X-ray diffraction data to calculated data (i.e., minimizing the R factor) and simultaneously exclude structures in which excessively close van der Waals contacts occur. In the present work, we have analyzed different modes of packing CTA chains in a crystalline array, and have investigated the influence of parameters such as main-chain and side-chain torsional angles on the packing. For this purpose, unit cells with only two dimers of two separate chains (i.e., four monomer units per cell) have been considered. This is the minimum requirement to model the proposed unit cell of CTA I.[1] It is also the minimum unit cell size to examine antiparallel chain packings. Since our present interest concentrates on the supramolecular structure of CTA I, larger unit cells, such as the one suggested for CTA II,[2] are not treated in this work. Even so, a number of our findings also extend to proposals regarding such larger unit cells.

Methods

Atom types referred to in the text and tables are defined in Figure 2, which depicts the dimer unit of CTA. In order to keep the problem within reasonable computer limits, the following approximations have been made: Nonpolar hydrogens are not explicitly included. $-CH_3$, $-CH_2-$, and $-CH<$ are treated as united atoms. This approach has been justified previously.[26] The

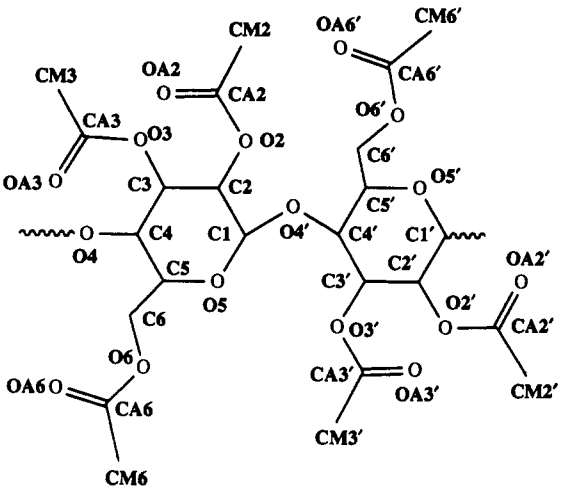

Figure 2. Definition of atom types in the dimer of CTA, used as the basic unit in the computed unit cells. Pyranose ring atoms, attached oxygens, and the exocyclic C atom in position 6 are named as in cellulose; acetate group atoms are called OA (for carbonyl oxygen), CA (for carbonyl carbon), and CM (for the methyl carbon).

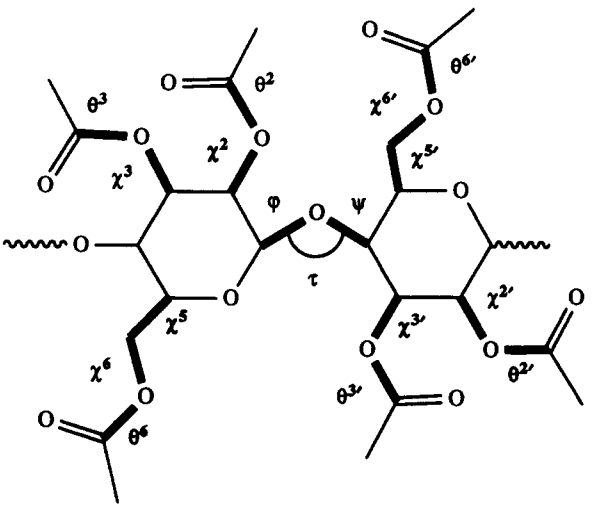

Figure 3. Definition of side-chain torsional angles and the torsional angles at the glycosidic bridge. See also Table I.

geometry of the pyranose ring as well as all bond lengths and bond angles in the side chains are kept fixed (see below). Only the side chains are free to rotate (subject to empirical potentials). The two main-chain torsional angles φ and ψ are in principle free to vary, but they are subject to special constraints to ensure certain crystal properties, as discussed in detail below. Bonds about which rotation can occur, and the nomenclature used for torsional angles, are shown in Figure 3; the torsional angles are defined in Table I. It should be noted that some of the definitions of torsional angles are different from those in the standard literature on polysaccharides; since we do not include hydrogen atoms explicitly, the torsional angles referring to hydrogens must be redefined. This is of particular relevance for the two main-chain torsional angles φ and ψ.

A. Geometry of the Monomer. Crystal structures of both cellobiose octaacetate[29] and cellotriose undecaacetate[30] have been determined. R factors for the crystal structures of these dimer and trimer units are 0.075 and 0.091, respectively. Following the reasoning of Stipanovic and Sarko[1] and of Roche et al.,[2] we have chosen the geometry of the central pyranose ring of cellotriose undecaacetate for the monomer structure of the polymer. The ring geometry is shown in Figure 4, and has been kept rigid in all subsequent calculations, as mentioned above. For the side-chain geometries, however, we have opted for the structural parameters shown in Figure 5, considering that side chains must be affected more strongly by neighboring groups in the crystal lattice than is the pyranose ring geometry. The structural features

Table I
Four Atoms Used To Define Dihedral Angles in the Dimer of CTA

dihedral angle[a]	atoms[b]				
φ	O5	C1	O4′	C4′	
ψ	C1	O4′	C4′	C5′	
χ^2	C1	C2	O2	CA2	
χ^3	C4	C3	O3	CA3	
χ^5	C4	C5	C6	O6	
χ^6	C5	C6	O6	CA6	first monomer
θ^2	C2	O2	CA2	OA2	
θ^3	C3	O3	CA3	OA3	
θ^6	C6	O6	CA6	OA6	
$\chi^{2'}$	C1′	C2′	O2′	CA2′	
$\chi^{3'}$	C4′	C3′	O3′	CA3′	
$\chi^{5'}$	C4′	C5′	C6′	O6′	
$\chi^{6'}$	C5′	C6′	O6′	CA6′	second monomer
$\theta^{2'}$	C2′	O2′	CA2′	OA2′	
$\theta^{3'}$	C3′	O3′	CA3′	OA3′	
$\theta^{6'}$	C6′	O6′	CA6′	OA6′	

[a] See also Figure 3. [b] Definitions of atom types correspond to those given in Figure 2.

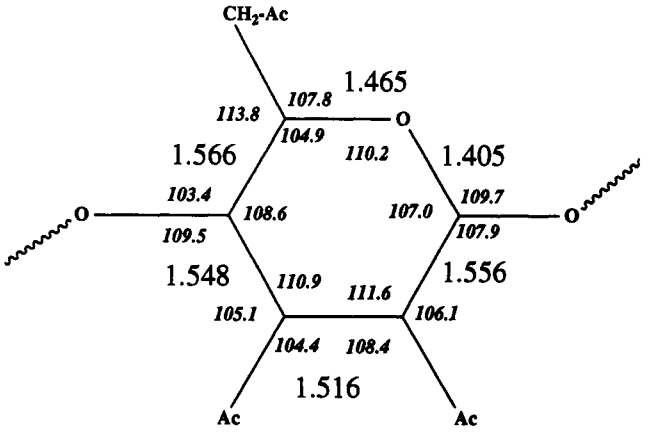

Figure 4. Geometry of the pyranose ring of the monomer unit (taken from the X-ray structure of the central unit of cellotriose undecaacetate).[30] Bond lengths (large numerals) are in angstroms; bond angles (smaller italic numerals) are in degrees.

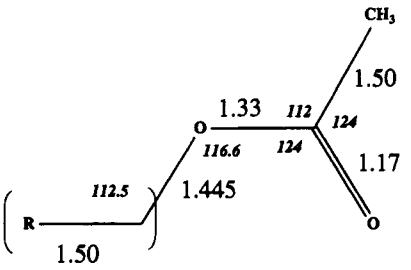

Figure 5. Geometry of the acetate side chains in CTA, obtained by averaging the bond lengths and bond angles from all acetate groups in cellobiose octaacetate and cellotriose undecaacetate,[29,30] including terminal acetate groups (cf. text for details). Bond lengths (large numerals) are in angstroms; bond angles (smaller italic numerals) are in degrees.

in Figure 5 were obtained by averaging the bond lengths and bond angles of all acetates (including the terminal ones) in the dimer and the trimer model compounds. The chosen values for the bond lengths in the acetate match well with average values determined from a large number of X-ray data on esters, compiled by Allen et al.[31] (For better agreement with the compilation of Allen et al.,[31] and because of uncertainties in the X-ray data,[29,30] the bond length C5–C6 was rounded from 1.49 to 1.5 Å.) Initial torsional angles θ^i around the O—(C=O) bond in esters were all set to zero, where $\theta^i = 0°$ for C and OA being eclipsed in structures C—O—CA=OA (see also Figure 2).

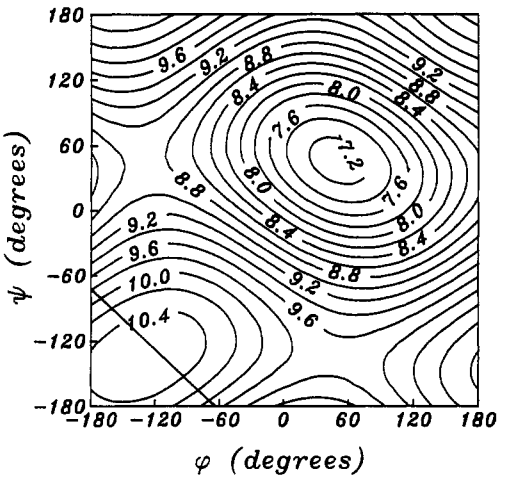

Figure 6. Dimer length as a function of the glycosidic torsional angles φ and ψ for an *isolated* disaccharide. Contours are labeled in angstroms and are spaced by 0.2 Å. The angles are defined in Figure 3 and Table I. The diagonal line at the lower left in the figure represents eq 2, which defines φ,ψ pairs yielding zero offset (i.e., a 2-fold helix) in the repeating chain.

B. Geometry of the Glycosidic Bridge. The bond lengths C1–O4′ and O4′–C4′ and the bond angle C1–O4′–C4′ (denoted τ in Figure 3) differ in the structures of the dimer and the trimer. This is understandable since every pyranose unit is in a different situation. In the (infinite) polymer there is a priori no reason for the glycosidic bridges to differ between different monomers (although "complicated" helical arrangements might, in principle, involve a regular sequence of structurally differing glycosidic bridges). The geometry used here has been obtained by averaging the geometry of the three glycosidic bridges in the dimer and the trimer. Bond lengths (rounded to 0.01 Å) are 1.38 Å for C1–O4′ and 1.44 Å for O4′–C4′; the mean value for the bond angle C1–O4′–C4′ (i.e., τ) is 116.5°. In this work, we did not vary τ since, in the case of cellulose, variations of a few degrees in τ had no sizeable influence on the results.[27,28]

C. Construction of the Asymmetric Unit. The experimentally determined fiber-repeat distances in both CTA I (c = 10.43 Å)[1] and CTA II (c = 10.54 Å)[2] correspond to the length of a disaccharide unit. The experimental unit cell[1] of CTA I contains two parallel dimer units which differ only by a slight rotation around the fiber axis and a marginal translation of 0.4 Å with respect to one another in the fiber direction. The proposed unit cell[2] for CTA II contains four dimer units and is more complicated; it consists of two antiparallel pairs of parallel chains.

On the basis of these experimental data, we have chosen the dimer of CTA as the smallest possible structural unit. The dimer was constructed by linking two of the monomer units by the glycosidic bridge with the bond lengths and the bond angle τ mentioned above.

Three major constraints were imposed on the dimer unit in the crystal: (i) the length of the dimer (i.e., the fiber-repeat distance) must fall within the range of experimental values given above (ca. 10.2–10.8 Å); (ii) in order for the polymer to form a perfect two-fold helix, the pair of main-chain torsional angles (φ,ψ) must be chosen in such a way that the third monomer unit is in exactly the same alignment as the first, shifted along the helix axis by the dimer length corresponding to the given values of φ,ψ; (iii) the conformational energy for the chosen φ,ψ pairs should be reasonably low (acknowledging that, in crystal packings, attractive *inter*chain interactions may partially compensate unfavorable *intra*chain repulsions).

In order to find the two-dimensional φ,ψ region within which constraint (i) on the fiber-repeat distance c is fulfilled, a contour map $c = f(\varphi,\psi)$ was generated (Figure 6). It is seen that only the region in the lower left corner of the map, delimited roughly by the range $-180° < \varphi < -70°$, $-180° < \psi < -70°$, yields values of c close to the experimental ones.

There are, in principle, straightforward (although cumbersome) mathematical solutions to constraint (ii).[32-36] We have chosen an empirical approach, shown schematically in Figure 7. Placing

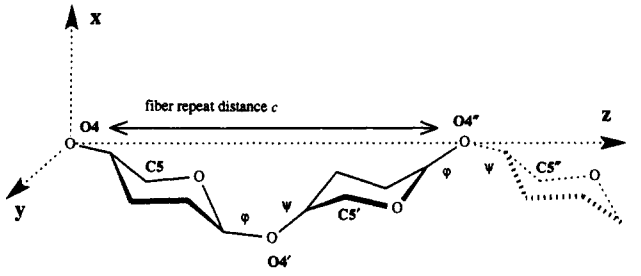

Figure 7. Schematic representation of the alignment of a dimer unit to evaluate φ,ψ pairs yielding a 2-fold helix. O4 is at the origin, C5 in the x–z plane. The third monomer unit (dashed) is used to place the hypothetical O4″ and C5″ atoms, required to be in the same alignment as O4 and C5, shifted by the fiber-repeat distance (i.e., the length of a dimer).

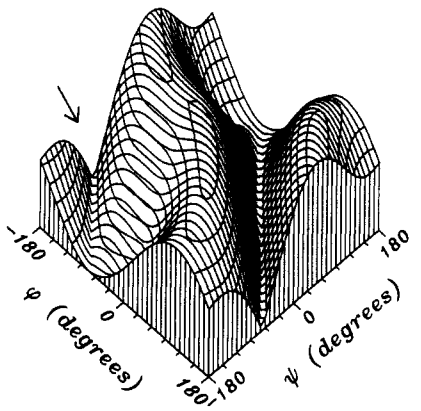

Figure 8. *Offset* evaluated by eq 1, as a function of the glycosidic torsional angles φ and ψ (as defined in Figure 3 and Table I). The valley trough marked by the arrow corresponds to the region where perfect 2-fold helices with acceptable fiber-repeat distances can be found (cf. also Figure 6).

O4 at the origin of a Cartesian coordinate system, atom O4″ on the z axis and atom C5 in the x–z plane, we evaluate the *offset* of a hypothetical atom C5″ (i.e., belonging to a double-primed residue connected to the primed one by the glycosidic bridge geometry defined above) as a function of φ and ψ, defined by

$$offset(\varphi,\psi) = \sqrt{d_x^2 + d_y^2 + (d_z - c)^2} \quad (1)$$

where d_x, d_y, and d_z are the Cartesian components of the vector connecting atoms C5 and C5″ and c is the fiber-repeat distance (i.e., the distance O4–O4″) corresponding to the given φ,ψ pair. For a perfect two-fold helix $d_x = d_y = 0$, $d_z = c$, i.e., the *offset* must be zero. In Figure 8, the value of the *offset* is depicted as a function of φ and ψ. The lowest points in the two valleys (which is actually a single, continuous valley because of the cyclic nature of φ and ψ) correspond to φ,ψ pairs yielding perfect 2-fold helices. Since only the region with acceptable fiber-repeat distances is of interest [as described above in constraint (i) and denoted by an arrow in Figure 8], we have determined five values of φ and ψ for which the *offset* is zero in that region. The five points were determined interactively by fine-tuning φ and ψ until the *offset* was <0.001 Å. These points are located on a straight line (correlation coefficient $R > 0.999$) which relates ψ to φ (in degrees) by

$$\psi = -(1.0683)\varphi - 257.71 \quad (2)$$

This line is drawn as the diagonal line in the lower left corner of Figure 6. Equation 2 will be used to calculate the corresponding allowed values of ψ for a given value of φ. It should be noted that eq 2 depends on the actual geometry of the pyranose rings (the "length" of the monomer, i.e., the distance C1–C4, representing a *virtual* bond) and of the glycosidic bridge. Therefore, the parameters are similar but not identical to the values found for the same linear relation by Simon et al. for cellulose.[28]

The last constraint (iii) has not been explored in detail during the course of this work. Pérez and Brisse[37] have carried out conformational energy calculations on the dimer of CTA. They confirmed that the actual crystal-structure conformation of the dimer model compound[29] corresponds closely to one of the calculated energy minima. The crystallographic φ,ψ pair falls into the range specified above, supporting the argument that constraint (iii) is also obeyed.

D. Generation of the Crystal Structure. As mentioned above, the disaccharide unit has been used as the smallest asymmetric unit. Starting unit-cell structures for subsequent energy minimizations were prepared with a program allowing the manipulation of dimer units in space as well as changes in side-chain conformations. The main-chain conformation was always specified by φ, the value of ψ then being automatically set by the program, using the linear relation of eq 2. The values of φ and ψ were constrained to be the same in all dimer units. Since fixing φ and ψ determines the length of a dimer unit, the fiber axis of the unit cell (c, by convention in polymers) is thus automatically given.

Starting arrangements were created by positioning one dimer with O4 at the origin of a Cartesian coordinate system, the hypothetical O4″ atom (belonging to the first monomer of the next dimer unit) on the z axis, and atom C5 in the x–z plane. A second identical dimer was then translated (and sometimes rotated around the fiber axis) to yield different initial arrangements. Antiparallel starting configurations were obtained by rotating the second unit by 180° around the y axis. In all cases, the vectors connecting atoms O4 and O4″ in the two dimer units were either parallel or antiparallel. Since the density derived from the experimental X-ray work[1] agrees reasonably well with directly measured densities,[1] we also used the experimental unit cell parameters a and b of CTA I[1] for setting up starting configurations (c being fixed by the 2-fold helix constraint mentioned above). The z axis of the Cartesian coordinate system thus coincides with the fiber axis which, in accordance with the literature, becomes the crystallographic c axis.

The data preparation program allowed a rough check of steric interactions in side chains so that only "reasonable" starting conformations were set up. Within the same dimer unit, the corresponding side chains in both monomer units may have identical (A–A) or different (A–B) torsional angles; i.e., both A–A and A–B structures were generated initially.

It should be noted that, in all cases, we started with the same side-chain conformation sets in both dimer units. There is no a priori reason for this choice, except that, with rotation about four bonds per monomer (not counting the ester bond torsions around O–CA bonds), and assuming as few as three discrete rotational isomeric states per bond, a total of 4.3×10^7 starting combinations would result. Although many of these combinations might be ruled out because of severe steric overlaps, the remaining number would still exceed present computational resources. However, although identical at the start, all of the side chains in both asymmetric units were allowed to rotate independently during energy minimizations (see below).

E. Conformational and Packing Energy Calculations. Intra- and intermolecular energies were calculated as the sum of pairwise nonbonded and electrostatic interaction energies and intrinsic torsional energies for each single bond about which rotation can occur, using the FORTRAN program WMIN,[22] which contains procedures for evaluating electrostatic and van der Waals interactions by an accelerated convergence procedure,[38] and which was adapted earlier for calculation of ECEPP/2 interaction terms.[39] Since CTA has no hydrogen-bond-donor groups, hydrogen bonding did not need to be included, by contrast with the previous work on cellulose,[27,28] where hydrogen bonding contributes predominantly to the overall energy.

Electrostatic interactions were evaluated as Coulomb terms by eq 3,

$$U_e = \frac{332 \cdot q_i \cdot q_j}{D \cdot r_{ij}} \quad (3)$$

where q_i and q_j are partial charges in electronic units on atoms i and j, separated by a distance r_{ij}, and D is the dielectric constant. The partial charges were derived from CNDO/2,[40] and a dielectric constant of 2.0 was used. These CNDO/2 charges correspond to an effective dielectric constant of ca. 4.[23,41]

Table II
Intrinsic Torsional Potential Parameters

bond type	U_0[a]	n[b]	k
C(sp^3)–C(sp^3)	2.8	3	1[c]
C(sp^3)–O(sp^3)	0.6	3	1[c]
O—C=O	8.0[d]	2	−1[e]

[a] Barrier height in kcal/mol. [b] n-fold barrier. [c] $k = 1$ for minima at $\theta = +60°$, 180°, and −60°. [d] Value chosen in this work (see also Yan et al.[54]). [e] $k = -1$ for minima at $\theta = 0°$ and 180° (here, θ is defined as 0° when the atoms Ci and Ol in Ci—Oj—Ck=Ol are eclipsed).

Nonbonded energy terms were computed by a Lennard-Jones 6–12 potential, expressed as

$$U_{nb} = \epsilon_{ij}[F(r_{ij}^0/r_{ij})^{12} - 2(r_{ij}^0/r_{ij})^6] \quad (4)$$

where ϵ_{ij} and r_{ij}^0 are the potential well depth and the atom separation at the minimum of the interaction energy between atoms i and j, respectively, and r_{ij} is the actual distance separating them. The same parameters were used as previously.[27,28] For 1–4 interactions, i.e., atoms separated by three covalent bonds, F was 0.5, otherwise F was 1.0 for 1–5 and higher interactions. For this rigid geometry, 1–3 interactions are fixed and hence omitted from our calculations.

Intrinsic torsional energies were calculated by

$$U_t = (U_0/2)[1 + k \cos(n\theta)] \quad (5)$$

where U_0 is the barrier height for variation of the torsional angle θ for an n-fold torsional barrier. Values for U_0, k, and n are given in Table II and have been used previously.[27,28,41]

The total applied force field was the sum of all terms U_e, U_{nb}, and U_t. As in the work on cellulose,[28] the computed total energy per asymmetric unit consists of three terms: an "internal" (here "intramolecular") energy term accounting for interactions between atoms whose distance with respect to one another can vary, plus all variable torsional energies, including those of the glycosidic bond connecting adjacent unit cells; a term generated by an accelerated convergence procedure to evaluate electrostatic interactions between atoms in the structure;[38] an "external" (here "intermolecular") energy term accounting for all nonbonded and electrostatic interactions between the atoms of the asymmetric unit and those within a prescribed radius from each of these atoms; this radius was eventually set to 12 Å, based on convergence checks confirming that larger values did not yield different structures at the corresponding energy minimum.

The program WMIN assumes that asymmetric units contained in different unit cells are not covalently bonded. Therefore, interactions between atoms belonging to different unit cells would normally be evaluated as 1–5 interactions. Because of the glycosidic bonds connecting adjacent unit cells in the fiber direction, interactions between atoms closer than actual 1–5-positions have had to be treated individually; i.e., interaction energies between 1–2 and 1–3 pairs were set to zero, while interactions between 1–4 bonded atoms were treated explicitly as 1–4 interactions.

F. Energy Minimizations. In all our computations, starting arrangements always consisted of one unit cell containing two identical dimer units (i.e., with identical conformations for both the main chain and the side chains). WMIN uses periodic boundary conditions to optimize these starting structures to local-energy-minimum configurations.

Besides the side-chain torsional angles χ^i and θ^i already defined as variables, the two dimers in the unit cell were individually free to translate along x, y, and z, and rotate around the axis c (i.e., the Cartesian z axis), implying that the crystal unit-cell parameters a, b, α, β, and γ were free to vary during energy minimizations. The main-chain torsional angles φ and ψ and the fiber-repeat distance c are not independent, because of the constraints of the 2-fold helical arrangement (see above). If one is defined as variable, the other two are automatically determined. However, having one of these three parameters variable during minimizations might lead to a disruption of the covalent bonds between adjacent unit cells. Therefore φ, ψ, and c were kept constant in WMIN, and a grid search was used to determine the triplet φ,ψ,c which leads to the lowest energy. After roughly determining the φ domain of interest, φ was varied in steps of 1°

Table III
CNDO/2 Partial Charges[a] of Atoms and United Atoms of the Neutral Monomer Unit of CTA

main-chain atoms		side-chain atoms[b]	
O4	−0.2394	O(i)	−0.2476
C4	+0.1328	CA(i)	+0.4243
C5	+0.1128	OA(i)	−0.3214
C6	+0.1279	CM(i)	+0.0104
O5	−0.2461		
C1	+0.2555		
C2	+0.1252		
C3	+0.1342		

[a] Charges are given in electronic charge units. [b] Atom types are defined in Figure 2; i stands for 2, 3, or 6.

in successive WMIN energy minimizations. It should be mentioned that, in some cases with very unfavorable starting arrangements, WMIN became stuck in high local energy minima. Such structures were discarded and will not be discussed here. Initial values for a and b were taken from the experimental X-ray work on CTA I.[1] The unit cell angles α, β, and γ were always assigned initial values of 90°.

Minimizations generally converged with all variables included. In some cases, initial steps were carried out with a reduced number of variables (mostly by fixing the angles α, β, and γ). Generally the POWELL minimizer was used.[39,42] In cases where POWELL failed, a ROSENBROCK minimization[43] was first carried out and then continued by POWELL. Convergence criteria used in POWELL corresponded to energy convergence to ca. 0.01 kcal/mol.

Results and Discussion

A. Partial Charge Distribution in the Monomer Unit. CNDO/2 calculations were carried out on a monomer unit of CTA with the fixed geometry of the pyranose ring described above, and terminated by O–CH_3 in positions O4 and O4′ (actually O1 in the monomer nomenclature). The methoxy groups were attached with the geometry defined for the glycosidic bonds. Calculations were carried out with all hydrogens present. The CH, CH_2, and CH_3 united atoms were then assigned the sum of the partial charges on the corresponding C and H atoms.

It is obvious that, in such a monomer unit, O4 and O4′ are not in equivalent positions and hence may have different partial charges. In the polymer, however, O4 and O4′ are indistinguishable. Furthermore, since the polymer is neutrally charged, the sum of the partial charges in the monomer unit (after removing one methyl and one methoxy group) must be zero. The partial charges listed in Table III were obtained after assigning to O4 the sum of all partial charges in both terminal methoxy groups, thus yielding a neutral monomer.

Partial charges in side chains are slightly sensitive to the actual conformations, so that CNDO/2 calculations yielded somewhat different charge distributions for the three acetate groups. Since the differences were marginal and since, in any event, the side chains are expected to change conformation during the subsequent energy minimizations, partial charges on atoms belonging to the acetates (i.e., O, CA, OA, and CM atoms) were averaged, so that all three acetate groups were electrostatically equivalent.

B. Computed Unit Cells. In Table IV are summarized all unit-cell parameters for structures which WMIN succeeded in optimizing to relatively low (local) energy minima. Since the main-chain torsion angle φ and the chain orientation (parallel versus antiparallel) uniquely define the structures presented here, the following notation will be used from now on to refer to a given structure: p{φ} stands for a parallel chain structure with a main-chain conformation fixed by φ, whereas a{φ} designates a cor-

Table IV
Unit Cell Parameters for CTA[a] Obtained from Energy Minimizations by WMIN

fixed									
φ, deg	ψ, deg	energy,[b] kcal/mol	a, Å	b, Å	c, Å	α, deg	β, deg	γ, deg	ρ,[c] g/cm^3
				Parallel Chain Arrangements					
-90	-161.6	-83.9	23.50	5.48	10.31	90.1	89.9	83.0	1.45
-95	-156.2	-101.4	23.25	5.47	10.39	90.5	90.4	83.9	1.46
-100	-150.9	-111.5	23.24	5.45	10.45	90.1	90.1	84.4	1.45
-105	-145.5	-116.0	23.49	5.40	10.51	90.2	90.1	84.3	1.44
-106	-144.5	-116.7	24.16	5.21	10.50	89.6	89.7	84.3	1.45
-107	-143.4	-117.0	24.28	5.17	10.52	90.1	90.1	84.1	1.46
-108	-142.3	-117.2	24.31	5.15	10.53	90.1	90.0	83.9	1.46
-109	-141.3	-117.0	24.31	5.13	10.54	90.1	90.0	83.7	1.46
-110	-140.2	-116.6	24.32	5.11	10.54	90.1	90.0	83.5	1.47
-115	-134.8	-109.5	24.20	5.06	10.57	90.1	90.0	82.6	1.49
-120	-129.5	-86.2	24.01	5.01	10.58	90.1	90.0	81.3	1.52
				Antiparallel Chain Arrangements					
-90	-161.6	-83.1	25.46	5.34	10.31	89.5	72.8	85.2	1.43
-95	-156.2	-101.8	26.76	5.21	10.39	90.9	67.7	79.1	1.46
-100	-150.9	-104.5	24.83	5.42	10.45	90.3	83.7	89.5	1.37
-105	-145.5	-109.3	25.57	5.20	10.50	90.0	81.2	85.4	1.39
-106	-144.5	-111.7	24.21	5.33	10.51	91.0	90.0	90.1	1.41
-107	-143.4	-111.8	24.29	5.29	10.52	89.2	90.0	89.9	1.42
-108	-142.3	-111.7	24.30	5.28	10.53	89.1	90.0	89.9	1.42
-109	-141.3	-111.3	24.15	5.32	10.54	91.4	90.0	90.1	1.41
-110	-140.2	-110.8	24.32	5.24	10.54	89.3	90.0	89.9	1.42
-115	-134.8	-100.0	25.44	5.11	10.57	90.0	90.0	90.0	1.39
-120	-129.5	-84.8	23.78	5.00	10.58	90.0	90.0	90.0	1.52

[a] The experimental unit-cell parameters for CTA I are[1] a = 23.63 Å, b = 6.27 Å, c = 10.43 Å, $\alpha = \beta = \gamma = 90°$. For the larger unit cell of CTA II, containing the dimers of four chains,[2] the experimental parameters are a = 24.68 Å, b = 11.52 Å, c = 10.54 Å, $\alpha = \beta = \gamma = 90°$.[b] Energy refers to the asymmetric unit (i.e., four monomer units). [c] Calculated from the unit cell volume V (in Å^3) by the relation ρ (g/cm^3) = 1913.23/V, obtained by using as molar mass for one monomer the value of 288.08.

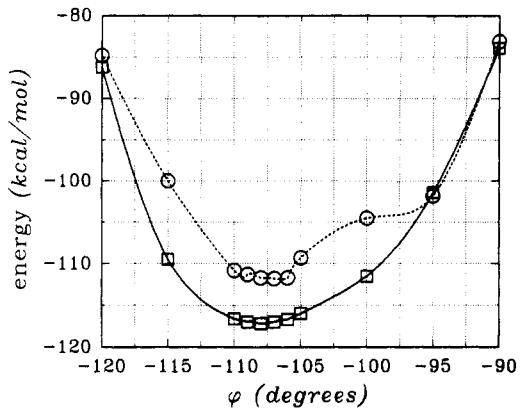

Figure 9. WMIN (ECEPP/2, UNICEPP) energy in kilocalories/mole of asymmetric unit as a function of the main-chain torsional angle, φ. Squares represent parallel chain packings (p{φ}) and circles antiparallel packings (a{φ}).

responding antiparallel chain structure with the main-chain conformation fixed by φ.

In Figure 9 is shown the final WMIN energy in kilocalories/mole per asymmetric unit for parallel and antiparallel chain packings as a function of φ. In both cases, the minimum falls within the range of $-105° \leq \varphi \leq -110°$.

The structure for the lowest energy parallel chain arrangement, p{-108} (shown in Figure 10), is 5.4 kcal/mol per unit cell lower than that for the lowest energy antiparallel structure, a{-107} (shown in Figure 11). With interpretation of this difference as a difference in cohesive energy, the main-chain and side-chain conformations in the two structures being virtually identical (cf. the next section), the parallel structure would be favored thermodynamically by 1.35 kcal/mol of monomer units over the ("simple") antiparallel arrangement shown in Figure 11. While CTA II is known to be more stable than CTA I,[44] and the former is claimed[2] to be antiparallel, there is a discrepancy between this conclusion and our finding that

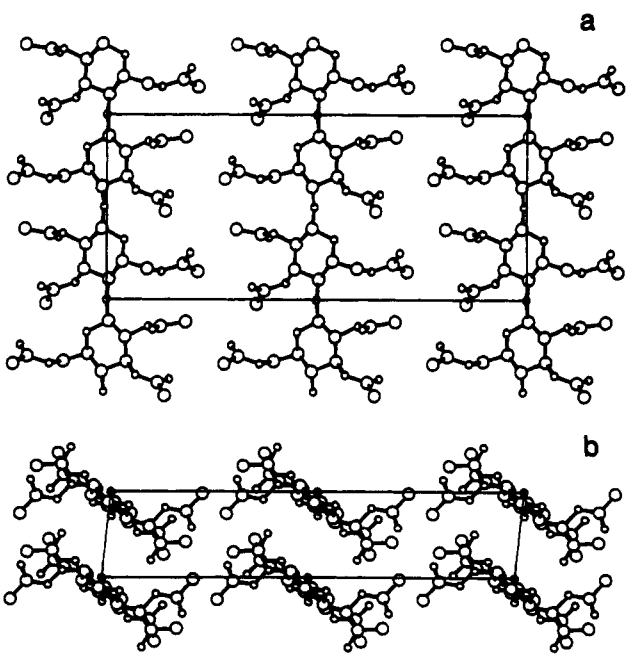

Figure 10. Unit cell of the lowest energy parallel chain packing p{-108}: (a) view projected onto the a–c plane; (b) view projected onto the a–b plane, with all chains pointing toward the viewer (defining the positive chain vector as the O4–O4″ direction in Figure 7).

the parallel structure is more stable. It should be noted that, while p{-108} closely resembles the proposed packing of CTA I,[1] the antiparallel arrangement a{-107} does *not* correspond to the more complicated packing in the larger proposed unit cell for CTA II, where pairs of parallel chains are arranged in an antiparallel array.[2] Indeed, the two energy vs φ curves in Figure 9 show that parallel structures such as p{-100} and p{-115} have WMIN energies in the same range as the minimum for the antiparallel packings. It is therefore not possible to deduce the favored packing from

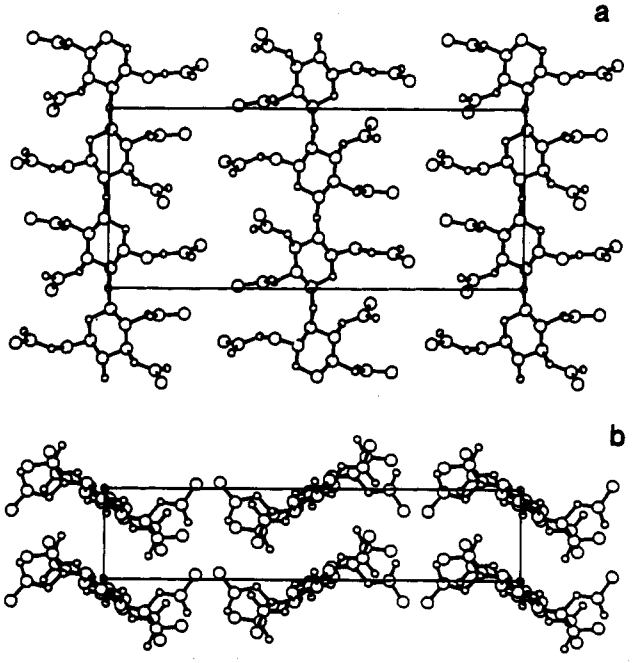

Figure 11. Unit cell of the lowest energy antiparallel chain packing a{−107}: (a) view projected onto the a–c plane; (b) view projected onto the a–b plane, with the corner chains pointing toward the viewer and the center chains pointing in the opposite direction (defining the positive chain vector as the O4–O4″ direction in Figure 7).

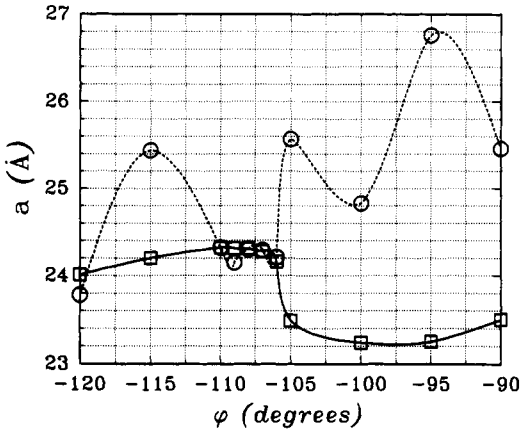

Figure 12. Unit-cell length a (in Å) as a function of the main-chain torsional angle, φ. Squares represent parallel chain packings (p{φ}) and circles antiparallel packings (a{φ}).

these calculations alone.

Small variations in φ around the values for the minimum-energy structures do not affect the energy very much for either parallel or antiparallel structures. It turns out, however, that the unit-cell parameters a and b can vary quite strongly with small alterations in φ. Figures 12 and 13 depict a and b as functions of φ. For parallel packings, the functions have a "jump" in both a and b between p{−106} and p{−105}. The density ρ (last column in Table IV) is altered only marginally, since the two jumps are in opposite directions and partially compensate one another as far as the unit-cell volume is concerned. It is notable that the two structures differ by only 0.18 kcal/mol per monomer unit. Similar observations are made for antiparallel structures, but in this case, the functions are less "smooth". The passage from a{−106} to a{−105} leads to quite a strong jump in both a and b while the unit-cell angle β falls from 90° to 81.2°. In this case, the energy difference is 0.6 kcal/mol per monomer, which is still small, being in the range of kT at 25 °C. These results indicate

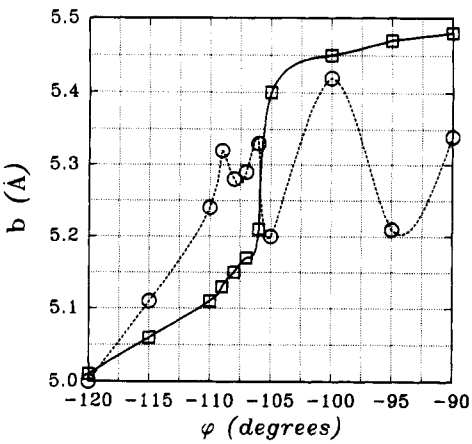

Figure 13. Unit-cell length b (in Å) as a function of the main-chain torsional angle, φ. Squares represent parallel chain packings (p{φ}) and circles antiparallel packings (a{φ}).

that small conformational changes in the main chain can result in rather different packing arrangements, but with similar energies. When such energetically similar unit cells do not match perfectly in at least two dimensions, the actual (experimentally characterized) crystalline domains must be riddled with imperfections, which may be fundamentally responsible for the problems encountered in the elucidation of the crystal structures of CTA.

On a molecular level, all parallel chain packings closely resemble the structure p{−108} represented in Figure 10. The chain in the center of the unit cell is not tilted around the fiber axis with respect to a corner chain and it is not shifted along the fiber axis. Since the side-chain conformations of the center chain also closely match those of a corner chain (cf. following section), all computed parallel chain unit cells could in principle be reduced to "one chain" unit cells by dividing the a length by 2. For antiparallel chain arrangements, the lowest energy packings resemble a{−107} represented in Figure 11. The shift along the fiber axis is ca. −1 Å for all four structures a{−106}, a{−107}, a{−108}, and a{−109}. In higher energy unit cells, the relative translations of corner and center chains differ quite strongly. In antiparallel chain structures there exist many distinct possibilities to pack CTA chains in a two-chain unit cell, depending basically on the main-chain conformation adopted. In parallel chain orientations, the number of possible low-energy packings is far more limited.

Besides the general packing schemes which resulted in two-chain unit cells as shown in Figures 10 and 11, a number of other starting configurations were generated. WMIN either converted these structures to those shown here or their energies were far above the values given in Table IV.

C. Side-Chain Conformations. In a cellulose monomer, only one heavy atom (O6) out of 11 in all is connected to the pyranose ring by a bond about which rotation can take place. In CTA, however, the number of heavy atoms in rotatable side chains increases to 10 out of a total of 20, i.e., to 50%. In other words, half of the atoms contributing to the X-ray intensities in CTA, in principle, can participate in more or less free movement. On the other hand, the larger side chains must be accommodated in a densely packed crystalline array and the number of combinations of side-chain conformations leading to optimum packings can be expected to be quite small.

Tables V and VI compile all values of side-chain torsional angles obtained at the respective local energy minima listed in Table IV. As mentioned earlier, the data preparation program allowed for a rough check of van der Waals

Table V
Side-Chain Torsional Angles[a] in the First Basic Dimer Unit (Chains at Unit Cell Corners)

fixed															
φ	ψ	χ^2	$\chi^{2'}$	χ^3	$\chi^{3'}$	χ^5	$\chi^{5'}$	χ^6	$\chi^{6'}$	θ^2	$\theta^{2'}$	θ^3	$\theta^{3'}$	θ^6	$\theta^{6'}$
Parallel Chain Arrangements															
-90	-161.6	118	118	154	154	61	60	175	175	11	10	-23	-24	-3	-4
-95	-156.2	126	126	147	147	64	63	168	169	8	7	-20	-21	1	-1
-100	-150.9	133	132	140	140	65	65	162	163	2	2	-16	-16	2	1
-105	-145.5	139	139	134	134	65	65	155	156	-7	-7	-10	-10	0	0
-106	-144.5	139	139	137	137	61	62	144	143	-11	-12	-5	-5	-9	-9
-107	-143.4	140	140	136	136	60	61	142	142	-14	-15	-3	-3	-9	-10
-108	-142.3	142	141	135	135	61	61	140	140	-16	-17	-2	-2	-10	-11
-109	-141.3	143	143	134	134	61	61	139	139	-18	-19	-0	-1	-10	-11
-110	-140.2	145	144	132	133	61	61	138	138	-20	-20	1	1	-10	-11
-115	-134.8	152	152	126	126	61	62	130	130	-27	-27	6	6	-9	-10
-120	-129.5	160	160	119	119	62	62	121	121	-33	-33	11	11	-8	-8
Antiparallel Chain Arrangements															
-90	-161.6	114	114	155	156	69	66	183	194	-22	-20	-21	-23	9	0
-95	-156.2	121	120	151	152	58	58	155	161	-13	-10	-14	-14	-10	-13
-100	-150.9	135	132	141	142	69	73	150	161	-9	-5	-14	-15	1	10
-105	-145.5	138	136	138	137	78	57	166	135	-19	-27	-6	-5	10	-1
-106	-144.5	141	139	135	136	64	78	140	159	-20	-15	-7	-7	1	7
-107	-143.4	141	141	135	134	78	64	159	138	-16	-23	-5	-5	8	0
-108	-142.3	142	142	134	132	78	64	158	137	-18	-25	-3	-4	8	0
-109	-141.3	144	143	131	133	64	78	136	154	-23	-19	-4	-3	2	7
-110	-140.2	144	145	132	130	79	63	156	134	-21	-27	-1	-1	8	-1
-115	-134.8	152	152	125	125	59	60	124	124	-30	-30	5	5	-1	-2
-120	-129.5	160	160	117	117	61	61	121	121	-35	-35	13	13	-5	-5

[a] All torsional angles are given in degrees. For the definition of torsional angles, cf. Figure 3 and Table I.

Table VI
Side-Chain Torsional Angles[a] in the Second Basic Dimer Unit (Chains in the Center of the Unit Cell)

fixed															
φ	ψ	χ^2	$\chi^{2'}$	χ^3	$\chi^{3'}$	χ^5	$\chi^{5'}$	χ^6	$\chi^{6'}$	θ^2	$\theta^{2'}$	θ^3	$\theta^{3'}$	θ^6	$\theta^{6'}$
Parallel Chain Arrangements															
-90	-161.6	119	117	154	154	61	60	176	175	11	8	-23	-25	-3	-4
-95	-156.2	126	126	147	147	64	63	168	170	7	7	-20	-21	0	-2
-100	-150.9	133	132	140	140	65	65	162	163	2	2	-16	-16	2	1
-105	-145.5	139	139	134	134	65	65	155	156	-7	-7	-10	-10	0	0
-106	-144.5	139	139	137	137	60	62	144	143	-11	-12	-4	-5	-9	-8
-107	-143.4	140	140	136	136	61	61	142	142	-15	-15	-3	-3	-9	-10
-108	-142.3	142	141	135	135	61	61	140	140	-17	-17	-2	-2	-10	-11
-109	-141.3	143	143	134	134	61	61	139	139	-18	-18	0	-1	-10	-11
-110	-140.2	145	144	132	133	61	61	138	138	-20	-20	1	1	-10	-11
-115	-134.8	152	152	126	126	61	62	130	130	-27	-27	6	6	-9	-10
-120	-129.5	160	160	119	119	62	62	121	121	-33	-33	11	11	-8	-8
Antiparallel Chain Arrangements															
-90	-161.6	115	114	155	156	62	65	157	196	-23	-4	-21	-22	-7	-1
-95	-156.2	121	120	151	152	57	59	157	156	-11	-11	-14	-15	-11	-10
-100	-150.9	135	133	141	142	75	73	158	160	-8	-8	-14	-15	5	11
-105	-145.5	137	134	138	139	62	58	141	135	-20	-20	-6	-6	-6	-2
-106	-144.5	140	140	136	135	77	65	158	139	-13	-21	-6	-7	7	1
-107	-143.4	142	140	134	136	63	78	138	159	-23	-17	-5	-5	0	8
-108	-142.3	143	142	132	134	63	79	137	158	-24	-19	-4	-4	0	8
109	-141.3	144	144	132	131	77	64	154	135	-18	-24	-3	-4	7	2
-110	-140.2	145	144	130	132	62	80	134	156	-27	-22	-1	-1	0	8
-115	-134.8	152	152	125	125	59	60	124	124	-30	-30	5	-5	-1	-2
-120	-129.5	160	160	117	117	61	61	121	121	-35	-35	13	13	-5	-5

[a] All torsional angles are given in degrees. For the definition of torsional angles, cf. Figure 3 and Table I.

repulsions, and only starting side-chain torsions without close contacts were subjected to energy minimizations by WMIN. The torsional angles in Tables V and VI always represent the best choice of the minimizer. WMIN proved to be efficient in locating the best conformations of side chains; despite different combinations of starting torsional angles, torsional movements of more than 180° were sometimes encountered during minimizations to yield the listed values. We therefore assume that the displayed side-chain torsional-angle combinations represent "true" local minima.

The comparison of side-chain torsional angles can be divided into two parts: within one chain (i.e., a dimer in the unit cells here), the respective unprimed and primed side-chain torsional angles can be compared; because the pyranose ring geometries are held constant in all monomer units, (nearly) identical values for unprimed and primed χ^i (and θ^i) values correspond to A–A structures; stronger disparities would yield A–B structures; within one unit cell, the corner and the center chain can be compared with respect to their side-chain torsional angles.

Let us first compare unprimed and primed torsional angles within a single chain. In parallel chain structures, χ^i,θ^i and corresponding $\chi^{i'}$,$\theta^{i'}$ values are almost identical, meaning that only A–A structures were generated by WMIN. Although various A–B starting structures were tested, WMIN

either brought them back to A–A structures or the resulting A–B structures were of much higher energy. In antiparallel chain arrangements, the unprimed and primed monomer units have the same values for the "short" side chains (i.e., $\chi^2 \approx \chi^{2'}$ and $\chi^3 \approx \chi^{3'}$). The larger side chains (with χ^5 and χ^6, or $\chi^{5'}$ and $\chi^{6'}$) display some differences, but they still fall in the same conformational range, so that it is not really justified to label such structures A–B. As for the parallel structures, definite A–B starting structures were reconverted by WMIN to the A–A arrangements reported in Tables V and VI.

Comparing corner chains to center chains in parallel chain structures reveals that the corresponding side-chain torsional angles in both dimers of the unit cell are nearly identical. Combined with the equivalent orientation of the corner and center chains mentioned in the previous section, this finding suggests that the unit cell dimension a can in principle be reduced by one-half. The comparison is less important in antiparallel chains where the center chain is different anyway. Still, even in this case, the respective torsional angles are in the same conformational range, even when they differ more strongly than in the parallel case.

The calculations indicate that the side chains have quite restricted conformational domains, irrespective of the actual main-chain packing scheme. Consider the effect on the side-chain torsional angles of the principal parameter varied in our computations, namely the main-chain torsional angle φ. The torsional angles χ^2 and $\chi^{2'}$ increase whereas χ^3 and $\chi^{3'}$ decrease with more negative values of φ, in both parallel and antiparallel packings. Values of χ^5 and $\chi^{5'}$ are much less affected by different main-chain conformations, but χ^6 and $\chi^{6'}$ are strongly influenced. Interestingly, the largest jumps in χ^6 and $\chi^{6'}$ are observed at the transitions from $\varphi = -106°$ to $\varphi = -105°$ in both parallel and antiparallel structures. The corresponding observation of the strongest jumps in the unit-cell parameters at these same transitions shows that the C5–C6 side-chain conformations are affected by the packing, and vice versa. An indirect confirmation of the influence of the packing on χ^6 and $\chi^{6'}$ also comes from the obvious fact that the side-chain torsional angles in parallel and antiparallel structures should be identical if the side chains were merely to adjust to the conformation of the backbone. In this context, it is understandable that the short side chains at C2- and C3-positions are less influenced by packing effects than the longer side chains at C5-positions. Indeed, values of χ^2 and χ^3 in parallel and antiparallel arrangements are very similar. Only values of χ^5 and (especially) χ^6 differ strongly and clearly reveal "packing effects".

The ester distortions (around the O–CA bonds) never exceed 35° and mostly are much smaller. However, preliminary runs with fixed ester geometries (i.e., all $\theta^i = 0°$) led to higher WMIN energies, indicating that even moderate distortions of the (initially) flat ester groups can appreciably reduce the strain arising from van der Waals repulsion. This is understandable since the methyl groups in the acetates (here the combined atom types CM) are the bulkiest groups in the structure and are in closest contact simultaneously with adjacent chains in the crystallographic a direction. Figures 10a and 11a clearly show that the acetate groups of neighboring chains are intercalated in that the longer side chain (at C5) of one polysaccharide chain extends between the two shorter side chains (at C2 and C3) of the nearby chain.

D. Comparisons with Experimental Data. Besides the detailed structures reported for CTA I[1] and CTA II[2], other (sometimes quite differing) X-ray data have been published for the different CTA morphologies.[45-51]

Table VII
Side-Chain Torsional Angles in Peracetylated Poly(β-D-glucose) X-ray Structures

structure	χ^2	χ^3	χ^5	χ^6	ref[a]
first (unprimed) residue in peracetylated cellobiose	114	107	49	−141	29
second (primed) residue in peracetylated cellobiose	111	98	156	77	29
first (unprimed) residue in peracetylated cellotriose	141	123	64	160	30
second (primed) residue in peracetylated cellotriose	130	138	52	168	30
third (double primed) residue in peracetylated cellotriose	116	110	176	88	30
CTA I, all monomers, both chains	142	148	68	138	1
CTA II, first monomer, all chains	130	123	42	135	2
CTA II, second monomer, all chains	129	124	32	65	2

[a] All torsional angles (in deg) are expressed according to our definitions (cf. Figure 3 and Table I) and were evaluated from coordinates listed in the respective references.

Sprague et al.[6] have summarized the data to 1958, adding their own findings for unit cells of both CTA I and CTA II. For CTA I, they report a number of reflections in disagreement with early data by Hess and Trogus.[46,47] Stipanovic and Sarko[1] revised the results of Sprague et al., rejecting their unit cell on the basis of some medium-intensity reflections which it fails to reproduce. Also, the density of 1.39 g/cm^3 resulting from the Sprague unit cell dimensions was judged to be too high by Stipanovic and Sarko, who found an experimental density of 1.29 g/cm^3. The density derived from their refined unit cell is 1.24 g/cm^3.

The unit cell of CTA I of Stipanovic and Sarko (S & S from now on) encompasses the dimers of two distinct chains. The chain in the center of the unit cell is translated by 0.4 Å and tilted by ca. 13° with respect to the corner chains. Otherwise the two chains in a unit cell are virtually identical. With a space group of $P2_1$ and the screw axis in the fiber direction c, the second (primed) monomer in each chain has the same geometry (including side-chain conformations) as the first.

The computed lowest energy structures with parallel chains (p{−106} to p{−110}) in this work reproduce several, but not all, of the features of the experimental CTA I structures. The a and c dimensions are somewhat different while the b dimension is ca. 1 Å shorter than in the S & S unit cell (6.27 Å). Also, their unit cell is orthorhombic whereas, in our computed unit cells, while α and β are practically 90°, $\gamma = 83.9°$.

The main-chain conformation angle φ in S & S (based on hydrogen-atom positions) is reported to be close to the value of φ in the glycosidic bridge connecting the nonreducing (i.e., unprimed) and the middle (i.e., primed) unit of the peracetylated cellotriose structures (−98° in our definition).[30] According to eq 2, this corresponds to $\psi = -153°$ for a 2-fold helix with the geometry used in this work. Given the lack of precision with which hydrogens can be located, these values must be regarded as poorly determined. Side-chain conformations in the S & S structure (evaluated from the published coordinates) and in p{−108} match exceptionally well (cf. Tables V–VII).

A fundamental difference between the detailed structure proposed for the S & S unit cell and all of our computed p{φ} structures is the equivalence of the center and the corner chain in all p{φ} (i.e., no translation along the fiber

axis, no tilt with respect to the corner chain, and almost identical side-chain conformations). In other words, a reduced unit cell with half the a dimension could, in principle, be accepted. It should be noted, however, that the chain in the center of the S & S unit cell differs only slightly from the corner chain. The need for a two-chain unit cell was justified by Stipanovic and Sarko, arguing that equatorial X-ray reflections are consistent with a tilt of the chain in the middle with respect to the corner chain.

The differences between the unit-cell parameters of S & S and the lowest energy p{φ} structures result in an 18% higher density in our calculated packings, i.e., the computed packing density is appreciably higher than the experimentally derived one. On the other hand, the calculated density is only 5% higher than the value derived from the Sprague unit cell.[6]

Although Stipanovic and Sarko rejected an antiparallel chain arrangement, it seems worthwhile to compare the lowest energy antiparallel structures in this work with their unit-cell parameters. Our a{−106}, a{−107}, a{−108}, a{−109} are all very similar and our unit cell dimensions a, b, and c are in as good accord with the S & S unit cell as those of the lowest energy parallel structures. Furthermore, like the S & S unit cell, all low-energy antiparallel structures have essentially orthorhombic unit cells. The antiparallel arrangement clearly leads to an at least two-chain unit cell. The chain in the center of the low-energy antiparallel structures listed above is always shifted with respect to the corner chains by ca. 1 Å in the fiber direction.

We must conclude from the computations carried out here that we cannot deduce, even tentatively, the actual chain orientation in CTA I. The energy difference between parallel and antiparallel structures is not large enough to confirm that the most stable arrangements are parallel, so that an antiparallel chain structure cannot be ruled out on the basis of our computed results.

As mentioned earlier, we have not carried out calculations for the larger unit cell proposed for CTA II by Roche et al.[2] Yet, it is useful to compare certain features of our calculated structures with structural parameters in CTA II as suggested by experimental results. Besides the detailed analysis of Roche et al., both Dulmage[50,51] and Sprague et al.[6] have proposed unit cells for CTA II. Whereas the Sprague unit cell for CTA II is not confirmed by Roche et al., the Dulmage unit-cell dimensions are basically the same. Both the Roche and the Dulmage unit cells are orthorhombic with very similar axial lengths. Dulmage proposed a $P2_1$ space group, although his suggested chain arrangement is very close to $P2_12_12_1$. Roche et al. opted for the $P2_12_12_1$ space group, relying not only on X-ray but also on electron-diffraction data obtained from single crystals of CTA II. A density of 1.28–1.30 g/cm^3 can be deduced from the unit-cell volumes. Dulmage had determined an experimental density of 1.29 g/cm^3 for well-crystallized CTA II and 1.28 g/cm^3 for amorphous CTA. Again, these values are smaller than any computed densities in this work (cf. Table IV).

A detailed comparison of the main-chain and side-chain conformations between our calculated structures and the refined CTA II structure of Roche et al. reveals a close agreement, suggesting again that the overall conformational freedom in two-fold helical CTA structures is quite restricted. The main-chain conformational torsional angles φ and ψ in the Roche structure are, in our definition, −106° and −145°. This corresponds almost perfectly to the main-chain conformations in the computed minimum-energy structures (p{φ} as well as a{φ}) in Table IV. The values of the side-chain torsional angles listed for CTA II in Table VII and for the calculated structures in Tables V and VI are in the same conformational range for the first monomer. In the second monomer, the suggested experimental value of $\chi^{6'}$ falls into another range. Roche et al. have discussed in detail the problems of assigning the conformation of the C6 acetyl group, arguing that this group might have a statistical distribution, since a Fourier refinement in this region of the molecule fails.

The side-chain torsional angles in all calculated structures close to the energy minima represented by p{−108} and a{−107} also agree well with the respective values of the unprimed and primed residues in the peracetylated trimer structure.[30] The main-chain torsional angles φ and ψ between the unprimed and primed pyranose ring in the trimer are −98° and −143°, respectively, in our notation. These values diverge only slightly from those in the low-energy structures computed here, although there is no 2-fold helix constraint in the low molecular mass analogues of CTA. In other words, the 2-fold helix constraint does not seem to enforce a main-chain conformation diverging strongly from a (local) energy minimum. This again justifies the choice of the geometry of the central unit of the trimer as a monomer for extended cellulose triacetate chains.

Conclusion

The present work has explored possible structures of cellulose triacetate, by computation of intra- and intermolecular energies by the force field approach. We have been especially interested in the structure of CTA I because of its capabilities as a chiral sorbent for the chromatographic resolution of racemates, hoping that these calculations would be a first step in modeling interactions between CTA I and low molecular mass compounds.

When comparing the computed structures in this work with experimental results, several features are similar, especially on a molecular level, which are useful for further investigations.

Despite the differences mentioned in unit-cell parameters, the overall packing scheme in low-energy parallel chain arrangements closely resembles the parallel chain structure proposed by Stipanovic and Sarko for CTA I. Any strong deviations from this scheme yield unacceptably high energies. However, our antiparallel packings, although somewhat higher in energy, cannot be excluded on the basis of our calculations. Only a detailed comparison of computed X-ray intensities with very carefully obtained experimental diffraction intensities might definitely rule out this type of packing for CTA I.

The "jumps" observed in unit-cell dimensions for small alterations in the main-chain conformation demonstrate that the packing of CTA is quite sensitive to changes that would be insignificant for isolated chains (e.g., a rotation of one degree in φ). If the main-chain conformation has a strong influence on the packing, it is obvious that the multidimensional potential surface of CTA crystals could be riddled with local minima which must eventually lead to extended polymorphism, i.e., imperfections in actual crystalline CTA samples. These imperfections would lower the measured density and could be responsible for the difference between computed and directly measured densities. Obviously, the resulting perturbation of X-ray diffraction patterns would limit the precision with which detailed crystal structures can be obtained. The inconsistent literature data might be a direct consequence of that phenomenon. Contrary to the case of cellulose, there are no clearly defined and strictly localized interactions like hydrogen bonds in CTA. Precisely the absence of

such determining interactions for a defined supramolecular structure of polymers (as, for example, appear in all kinds of polyamides, including polypeptides, and in unsubstituted polysaccharides like cellulose itself) might lead to the suggested imperfections between CTA crystalline domains.

The side-chain conformations have turned out to be restricted to quite narrow conformational domains. Packing effects have been observed mainly for the longest side chains, i.e., at the C5-position. Strong deviations from the side-chain torsional angles reported here (and also found experimentally) would inevitably lead to much higher energies.

There are some discrepancies between the lowest energy structures computed here and the proposed structure for CTA I. The deviations result in an appreciable difference in density. All calculated structures have higher density than the density derived from the proposed unit cell. This difference cannot be explained by the imperfections between crystalline domains because the X-ray density is in principle obtained from a perfectly ordered material, even considering that the R factor obtained for CTA I is 0.24, indicating only a moderate agreement between the reflections for the proposed structure and the observed diffraction pattern. There are also several possible explanations related to the approximations inherent in force field calculations as performed here.

Since we use fixed bond lengths and bond angles, the selected geometry of the monomer units might not be adequate. However, the monomer structure used in the computations is derived from reliable X-ray data obtained on the trimer of CTA.[30] Only in the worst case should small deviations in bond lengths and bond angles add up instead of partially compensating each other. Another problem with the fixed geometry arises from the restricted number of pathways for the energy minimizer. Even then, the important deviation of the crystallographic b axis of 15–18% is hard to explain. Also, the excellent agreement of the side-chain conformations between the calculated structures and the proposed X-ray data indicates that the selected geometry is acceptable and that the side-chain relaxation process behaves reasonably.

There might be deficiencies in the force field parameters. Because of the fixed bond lengths and bond angles, errors might arise mainly from nonbonded interactions, especially from too soft Lennard–Jones interactions. The Lennard–Jones parameters used in eq 3 have been confirmed in numerous applications.[52] The choice of combined atoms (especially for CH_3) might influence the packing somewhat, but it should affect both the a and the b lengths in the same direction, which is not the case. Some tests, using reasonable variations of the Lennard–Jones parameters, showed that the packing was not sensitive to the values chosen.

The fact that the puzzle of CTA supramolecular structures has not been solved satisfactorily may also have a number of experimental origins. Cellulose triacetate samples are not necessarily fully acetylated. Careful analysis in our laboratories (Wolf, R. M.; Francotte, E., unpublished results) of the degree of acetylation by different methods has shown that values of 2.8–2.9 (compared with a theoretical value of 3) are always obtained under the heterogenous acetylation reaction conditions which yield CTA I. However, a quite strong degradation of the cellulose backbone is sometimes observed under such conditions, also yielding a distribution of low molecular mass oligomers which may perturb regular crystallization. A strong influence of the nature of the starting cellulose material and the precise acetylation conditions has also been reported elsewhere.[4]

It should be noted that highly crystalline CTA I samples cannot be obtained by recrystallization, for the obvious reason that dissolving and recrystallizing cellulose triacetate produces CTA II. Annealing of CTA I samples definitely improves their X-ray diffraction patterns, but it does not necessarily remove all imperfections. Domains of identical structure may grow or fuse together, hence forming larger crystallites and yielding better resolved diffraction patterns. Energetically degenerate states with different structures may, however, persist and the sample may remain polymorphous, i.e., still giving rise to imperfections at interfaces of domains with unit cells of different parameters. In other words, only a very careful preparation of experimental samples may eventually result in solving the dilemna. The problem is less critical with CTA II. Solvent-free single crystals of CTA II can be obtained.[53] Roche et al., for example, used such samples for electron diffraction.[2] Their problem with the conformation of the long side chain at the C5-position, however, shows that a statistical distribution of distinct structures may coexist even in highly crystalline material.

Acknowledgment. We thank Dr. W. R. Busing, Oak Ridge National Laboratory, Oak Ridge, TN 37830, for providing us with WMIN, his FORTRAN program for crystal-packing calculations. This work was supported at Cornell University by research grants from the National Institute of General Medical Sciences of the National Institutes of Health (GM-14312) and from the National Science Foundation (DMB84-01811). Most of the computations were carried out at the Cornell National Supercomputer Facility, a resource of the Cornell Center for Theory and Simulation in Science and Engineering, which receives major funding from the National Science Foundation and IBM Corp. with additional support from New York State and members of its Corporate Research Institute. L.G. acknowledges support for his sabbatical leave at Cornell University by the University of the Witwatersrand, Wits 2050, South Africa. I.S. acknowledges support from the Hungarian Academy of Sciences (OTKA 318 and 1361).

Addendum: In response to a referee's comment that the bond angles should have been allowed to vary, we first of all cite our comment about τ at the end of section B of Methods. In addition, as pointed out by Roterman et al.,[55,56] artifacts can be introduced if bond-angle bending is treated only by a harmonic approximation. Therefore, it is preferable to keep the bond angles fixed, as is done here.

Supplementary Material Available: Tables of computed fractional coordinates in the lowest energy structures for parallel (p{-108}) and antiparallel (a{-107}) chain arrangements after minimization (4 pages). Ordering information is given on any current masthead page.

References and Notes

(1) Stipanovic, A. J.; Sarko, A. *Polymer* **1978**, *19*, 3–8.
(2) Roche, E.; Chanzy, H.; Boudeulle, M.; Marchessault, R. H.; Sundararajan, P. *Macromolecules* **1978**, *11*, 86–94.
(3) Venkataraman, A.; Subramanian, D. R.; Soosamma, P. C. *Text. Res. J.* **1982**, *52*, 506–509.
(4) Takai, M.; Yaginuma, Y.; Kon, H.; Murata, M.; Hayashi, J. *Sen-i Gakkaishi* **1984**, *40*, T415–T424.
(5) Roche, E. J.; O'Brien, J. P.; Allen, S. R. *Polym. Commun.* **1986**, *27*, 138–140.
(6) Sprague, B. S.; Riley, J. L.; Noether, H. D. *Text. Res. J.* **1958**, *28*, 275–287.

(7) Gardner, K. H.; Blackwell, J. *Biopolymers* **1974**, *13*, 1975–2001.
(8) Gardner, K. H.; Blackwell, J. *Biochim. Biophys. Acta* **1974**, *343*, 232–237.
(9) Sarko, A.; Muggli, R. *Macromolecules* **1974**, *7*, 486–494.
(10) Stipanovic, A. J.; Sarko, A. *Macromolecules* **1976**, *9*, 851–857.
(11) Kolpak, F. J.; Blackwell, J. *Macromolecules* **1976**, *9*, 273–278.
(12) Kolpak, F. J.; Weih, M.; Blackwell, J. *Polymer* **1978**, *19*, 123–131.
(13) Kolpak, F. J.; Blackwell, J. *Polymer* **1978**, *19*, 132–135.
(14) Watanabe, S.; Takai, M.; Hayashi, J. *J. Polym. Sci. Part C* **1968**, *23*, 825–835.
(15) Hayashi, J.; Sufoka, A.; Ohkita, J.; Watanabe, S. *J. Polym. Sci., Polym. Lett. Ed.* **1975**, *13*, 23–27. Sarko, A. *Tappi* **1978**, *61*, 59–61. Marchessault, R. H.; Sundararajan, P. R., In *The Polysaccharides*; Aspinall, G. O., Ed.; Academic: New York, 1983; Vol. 2, pp 11–95.
(16) Hesse, G.; Hagel, R. *Liebigs Ann. Chem.* **1976**, 996–1008.
(17) Shibata, T.; Mori, K.; Okamoto, Y. In: *Chiral Separations by HPLC*; Krstulovic, A. M., Ed.; John Wiley & Sons: New York, 1989; p 336–398.
(18) Francotte, E.; Wolf, R. M. *Chirality* **1990**, *2*, 16–31.
(19) Francotte, E.; Wolf, R. M.; Lohmann, D.; Mueller, R. *J. Chromatogr.* **1985**, *347*, 25–37.
(20) Husemann, E.; Werner, R. *Methoden der organischen Chemie (Houben Weyl)*, 4th ed.; Georg Thieme Verlag: Stuttgart, 1963; Vol. 14/2, p 875.
(21) Wolf, R. M.; Francotte, E.; Lohmann, D. *J. Chem. Soc., Perkin Trans. 2* **1988**, 893–901.
(22) Busing, W. R. *ORNL-5747*; Oak Ridge National Laboratory: Oak Ridge, TN, 1981 (revised March 1984).
(23) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361–2381.
(24) Némethy, G.; Pottle, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1983**, *87*, 1883–1887.
(25) Sippl, M. J.; Némethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1984**, *88*, 6231–6233.
(26) Dunfield, L. G.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1978**, *82*, 2609–2616.
(27) Simon, I.; Scheraga, H. A.; Manley, R. St. J. *Macromolecules* **1988**, *21*, 983–990.
(28) Simon, I.; Glasser, L.; Scheraga, H. A.; Manley, R. St. J. *Macromolecules* **1988**, *21*, 990–998.
(29) Leung, F.; Chanzy, H. D.; Pérez, S.; Marchessault, R. H. *Can. J. Chem.* **1976**, *54*, 1365–1371.
(30) Pérez, S.; Brisse, F. *Acta Crystallogr.* **1977**, *B33*, 2578–2584.
(31) Allen, F. H.; Kennard, O.; Watson, D. G.; Brammer, L.; Orpen, A. G.; Taylor, R. *J. Chem. Soc., Perkin Trans. 2* **1987**, S1–S19.
(32) Shimanouchi, T.; Mizushima, S. *J. Chem. Phys.* **1955**, *23*, 707–711.
(33) Miyazawa, T. *J. Polym. Sci.* **1961**, *55*, 215–231.
(34) Sugeta, H.; Miyazawa, T. *Biopolymers* **1967**, *5*, 673–679.
(35) Yokouchi, M.; Tadokoro, H.; Chatani, Y. *Macromolecules* **1974**, *7*, 769–776.
(36) Sundararajan, P. R.; Marchessault, R. H. *Can. J. Chem.* **1975**, *53*, 3563–3566.
(37) Pérez, S.; Brisse, F. *Biopolymers* **1978**, *17*, 2083–2096.
(38) Williams, D. E. *Topics Curr. Phys.* **1981**, *26*, 3–40.
(39) Glasser, L.; Scheraga, H. A. *J. Mol. Biol.* **1988**, *199*, 513–524.
(40) Pople, J. A.; Beveridge, D. L. *Approximate Molecular Orbital Theory*; McGraw-Hill: New York, 1970; pp 57–153.
(41) Momany, F. A.; Carruthers, L. M.; McGuire, R. F.; Scheraga, H. A. *J. Phys. Chem.* **1974**, *78*, 1595–1620.
(42) Powell, M. J. D. *Comput. J.* **1964**, *7*, 155–162.
(43) Rosenbrock, H. H. *Comput. J.* **1960**, *3*, 175–184.
(44) Chanzy, H. D. *Colston Pap. (London)* **1975**, *26*, 417–434.
(45) Naray-Szabo, S.; Susich, G. *Z. Phys. Chem.* **1928**, *B4*, 264–270.
(46) Hess, K.; Trogus, C. *Z. Phys. Chem.* **1929**, *B5*, 161–176.
(47) Hess, K.; Trogus, C. *Z. Phys. Chem.* **1930**, *B9*, 160–168.
(48) Baker, W. O.; Fuller, C. S.; Paper, N. R. *J. Am. Chem. Soc.* **1942**, *64*, 776–782.
(49) Happey, F. *Text. Inst.* **1950**, *41*, T381–T403.
(50) Dulmage, W. J. *Nature* **1953**, *172*, 1053.
(51) Dulmage, W. J. *J. Polym. Sci.* **1957**, *26*, 277–288.
(52) Scheraga, H. A. *Chem. Scr.* **1989**, *29A*, 3–13.
(53) Chanzy, H. D.; Roche, E. J. *J. Polym. Sci., Polym. Ed.* **1974**, *12*, 1117–1126.
(54) Yan, J. F.; Vanderkooi, G.; Scheraga, H. A. *J. Chem. Phys.* **1968**, *49*, 2713–2726.
(55) Roterman, I. K.; Lambert, M. H.; Gibson, K. D.; Scheraga, H. A. *J. Biomol. Struct. Dyn.* **1989**, *7*, 421–453.
(56) Gibson, K. D.; Scheraga, H. A. *J. Biomol. Struct. Dyn.* **1991**, *8*, 1109–1111.

Registry No. CTA, 9012-09-3.