Polysaccharide Conformation. Part IX.[1] Monte Carlo Calculation of Conformational Energies for Disaccharides and Comparison with Experiment

By **David A. Rees** * and **Peter J. C. Smith,** Unilever Research Colworth/Welwyn Laboratory, Colworth House, Sharnbrook, Bedford MK44 1LQ

The treatment of conformational energies that has already been tested for monosaccharides is now applied to disaccharides and tested against available evidence for the crystalline state and solutions. van der Waals, polar, and torsional terms are included as before, with thermodynamic averaging over all possible conformations of side groups. Inter-residue hydrogen bonding is also included, in a way that allows for non-linear hydrogen bonds and makes no assumptions about the particular oxygen atoms likely to participate. For the maltose residue, better agreement is obtained with experiment than has been possible with any cruder approaches used hitherto. It is possible for the first time to account simultaneously for the seven different crystal conformations and two different solution conformations. For the cellobiose residue the agreement is at least as good as any obtained before. For both systems it is possible to define the favourable regions of conformational space more closely than by any previous approach. A major disadvantage of the method is that it involves extended computation.

In the preceding paper[1] we described a method for the estimation of conformational energies of carbohydrates which appeared promising after tests against experimental evidence for monosaccharides in aqueous solution. We used functions to take into account van der Waals, electrostatic, and torsion energies, with thermodynamic averaging of side-group conformations by a Monte Carlo Markov chain approach. We have now attempted to predict linkage conformations for certain disaccharides and to make comparison with the experimental evidence for the crystalline state and in solution. This has required that the potential function be extended to include hydrogen bonding terms in a way that is realistic about the effects of departure from collinearity of the O–H · · · O system, and which makes no assumptions about the particular oxygen atoms that are likely to be involved.

METHOD

The van der Waals coulombic and torsional terms were calculated as previously described.[1] The electronic partial charges calculated for a 1→4 linked disaccharide are given in Table 1. Torsional energies were calculated for rotation about the bonds indicated in Figure 1, and the amplitudes for these energy barriers were estimated as before from literature data[2] and are given in Table 2.

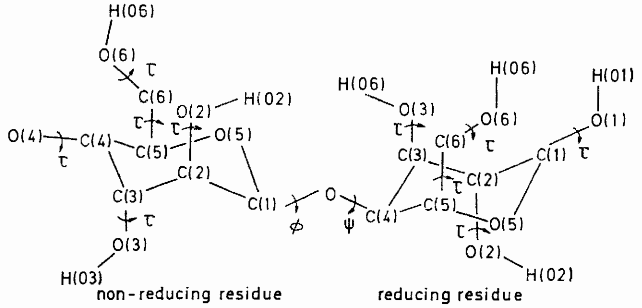

FIGURE 1 Stereochemistry of a 1→4 linked disaccharide (certain hydrogen atoms omitted). The statistically varied torsion angles are shown by τ. The $\phi = 0$, $\psi = 0°$ position is defined as that in which C(4) of both residues are all coplanar, with the remote C(4) and C(1) closest. The sense of ϕ and ψ is positive for anticlockwise rotation of each residue, viewed from the glycosidic oxygen

The hydrogen bond energies were calculated using a modification of the empirical method of Poland and Scheraga.[3] This assumes, for a linear O–H · · · O′ hydrogen bond, (a) that there is a known, constant O–O′ distance, r_0, which corresponds to the minimum energy, (b) that this minimum energy is of known and constant magnitude, ε, relative to the energy at infinite O–O′ separation, and (c) that at O–O′ separations $>r_0$, there exists an attractive potential which is proportional to r^{-6}, where r is the O–O′

TABLE 1

Calculated partial electronic charges for a 1→4 linked dihexopyranose

Atom	Partial charge: Non-reducing	Partial charge: Reducing
C(1)	0·177	0·190
C(2)	0·113	0·114
C(3)	0·106	0·105
C(4)	0·104	0·092
C(5)	0·095	0·094
C(6)	0·045	0·045
O(1)	−0·264	−0·446
O(2)	−0·457	−0·457
O(3)	−0·458	−0·458
O(4)	−0·458	−0·264
O(5)	−0·264	−0·262
O(6)	−0·457	−0·457
H(1)	0·060	0·061
H(2)	0·053	0·053
H(3)	0·052	0·052
H(4)	0·052	0·050
H(5)	0·051	0·050
H(6)	0·053 [a]	0·053 [a]
H(O1)		0·303
H(O2)	0·302	0·302
H(O3)	0·301	0·301
H(O4)	0·301	
H(O6)	0·302	0·302

[a] Two such atoms exist.

TABLE 2

Experimentally determined barriers to single bond rotation[2] and contributions from the postulated van der Waals potential (kcal mol^{-1})

Model compound	Barrier height	van der Waals contribution	Torsion amplitude
CH_3–O–CH_3	2·72	0·32	2·40
CH_3–O–H	1·07	0·11	0·96
CH_3–CH_2–H	2·90	0·34	2·56

---

[1] Part VIII, D. A. Rees and P. J. C. Smith, preceding paper.

[2] D. J. Millen, *Progr. Stereochem.*, 1962, **3**, 138.

[3] D. Poland and H. A. Scheraga, *Biochemistry*, 1967, **6**, 3791.

distance. All three requirements of the postulated function are shown in Figure 2.

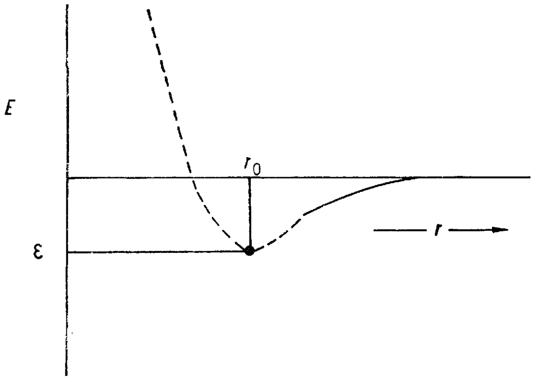

FIGURE 2 Known features of the hydrogen bond. The energy minimum at O · · · O′ separation $r = r_0$ of energy ε, and the longer range dependence on r^{-6} are shown by the dot and by the solid line respectively

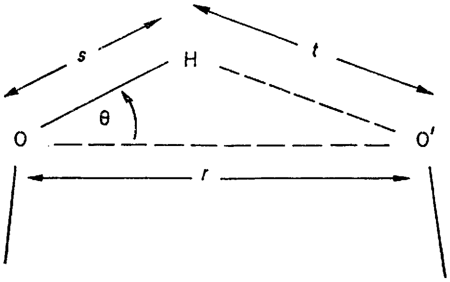

FIGURE 3 The O–H · · · O′ hydrogen bond

Using the nomenclature of Figure 3, consider the linear case, $\theta = 0°$. Suppose the total conformational energy of the –O–H · · · O– system is given by equation (1) where a

$$V = at^{-9} - bt^{-6} + ct^{-1} + S(r) \qquad (1)$$

and b are to be determined, c is the usual polar attraction (see ref. 1) between O′ and H, and $S(r)$ comprises all the other energy terms in the molecule. When $r = r_0$, $V = \varepsilon$, and $\partial V/\partial t = 0$.

Let $T(r)$ be the 'normal' interaction between O and O′, *i.e.* equation (2) applies, then approximate

$$T(r) = dr^{-9} - er^{-6} + fr^{-1} \qquad (2)$$

$\partial S/\partial t \approx \partial T/\partial t = \partial T/\partial r$; then by differentiation, substitution, and elimination equations (3a) and (b) are obtained where

$$a = 1/3(5ct_0^{\,8} + 6(\alpha - \varepsilon)t_0^{\,9} + \beta t_0^{\,10}) \qquad (3a)$$

$$b = 1/3(8ct_0^{\,5} + 9(\alpha - \varepsilon)t_0^{\,6} + \beta t_0^{\,7}) \qquad (3b)$$

$t_0 = t$ when $r = r_0$ and $\theta = 0°$, $\alpha = S(r_0)$ and $\beta = (\partial T/\partial r)$ at $r = r_0$.

Although this appears complex, it is computationally efficient, since once ε and r_0 have been decided upon, a and b are fixed. For the present work, r_0 was assigned a value of 2·7 Å and ε −4·0 kcal mol^{-1}.

For the usual case where $\theta \neq 0°$ the same function (1) is used. Although little is known experimentally concerning the energies of bent hydrogen bonds, the present function embodies the qualitative observation that as θ increases the stabilisation declines fairly rapidly and the potential merges with that for the normal non-bonded interaction. Figure 4 shows these characteristics as exhibited by the present function.

The ten conformation angles treated statistically are shown in Figure 1. Conformational energies were calculated as a function of the linkage conformation angles ϕ and ψ (Figure 1), each of which was stepped by intervals of 15° from 0 to 345°. Some regions in which the potential changed rapidly were reinvestigated with smaller step sizes.

The residue co-ordinates were derived from a survey of crystal structures.[4] The glycosidic bond angle was 117°. To economise in the calculation, it was necessary to calculate various types of interactions separately. Defining 'floating' atoms as those whose positions relative to the ring atoms are dependent on the statistically varied conformation angles, and 'fixed' atoms conversely; and 'self' interactions to be those between the atoms on one residue with each other, as distinct from 'mutual' interactions, the following distinctions can be made. (i) The self fixed–fixed interactions are independent of linkage conformation and can be ignored, since we are only interested in the relative energies. (ii) The mutual fixed–fixed interactions (V_1) are independent of the statistically varied angles but depend on linkage conformation. (iii) The self fixed–floating, mutual fixed–floating, self floating–floating, and mutual floating–floating interactions (V_2) are all dependent on the statistically varied angles and must therefore be averaged.

It was found that the calculation of V_2, the most time-consuming part, could be omitted (*i.e.* the conformation rejected) if V_1 either exceeded 500 kcal mol^{-1}, or exceeded the lowest $V_1 + V_2$ already found by more than 50 kcal mol^{-1}. These simple tests eliminated rather more than half the 576 possible ψ–ϕ combinations for maltose.

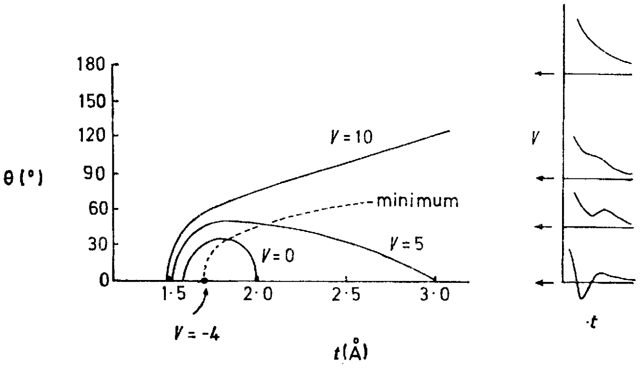

FIGURE 4 Postulated hydrogen bond potential for straight and bent hydrogen bonds. Left: The variation parameters defined in Figure 3, shown as a series of contours. Right: the variation of hydrogen bond energy with O · · · O distance, shown for a series of values of θ. The variation with θ of the distance corresponding to minimum energy is shown by the broken line

For similar reasons, fewer cycles than in the monosaccharide work could be used for the Monte Carlo averaging; 600 cycles were found to be sufficient, the first 100 of which were rejected. Calculated energies were found to be reproducible to within *ca.* 0·5 kcal mol^{-1}.

RESULTS AND DISCUSSION

A major limitation in our approach, apart from those that must be inherent in any use of atom–atom functions, is that intermolecular interactions are not considered even though these must be important, particularly in the crystalline state but also in solution (see below).

---

[4] S. Arnott and W. E. Scott, *J.C.S. Perkin II*, 1972, 324.

However, we shall show that it is useful to predict the limits outside which the intramolecular energy is so high that intermolecular terms could not reasonably be expected to outweigh them. Setting this limit at 5 kcal mol^{-1} of disaccharide, we find that only 3·5% of the total conformational space remains accessible for the maltose residue. Since seven different crystal conformations have been characterised, and two solution conformations, we can carry out a searching test of our energy map. It is of course true that the method we use of thermodynamically averaging the energies of side-group conformations, is more obviously appropriate to solution conformations rather than crystal conformations. In the latter, the thermal motions implicit in our treatment would have some relation to crystallographic thermal parameters but it is outside our scope to consider this here. However, our approach is not expected to introduce any error that is important compared with the neglect of intermolecular terms in the crystal.

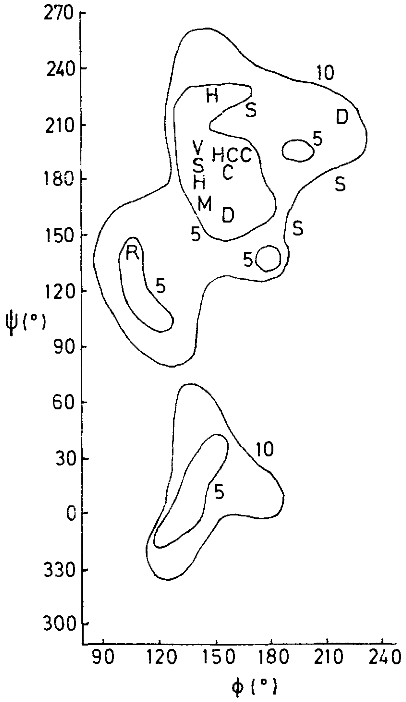

FIGURE 5 Calculated energy surface for maltose and comparison with solid state conformations. For symbols, see text.

The conformational energy surfaces for maltose are shown in Figure 5, with contours at 5 and 10 kcal mol^{-1} above the energy minimum shown as M. About 95% of the Boltzmann probability distribution is inside the 5 kcal mol^{-1} contour around $\phi = 150$, $\psi = 180°$. The three points marked C correspond to disaccharide crystal structures, being β-maltose[5] (upper left), cyclohexa-amylose[6] (upper right), and methyl β-maltoside[7] (lower). The points marked H are the linkages in the cyclohexa-amylose–water complex,[8] showing (close to the crystal values) the mean of the four linkages that are similar and (separately) the two linkages to the twisted sixth residue.

Also plotted are the linkage conformations that correspond[9] to the parameters derived by X-ray diffraction for known crystal forms of amylose; V-amylose[10] ($n = -6$, $h = 1·32$ Å) * is shown by V. For B-amylose,[10] X-ray diffraction shows $|n| = 6$ and models can be proposed of the double helix type ($h = 3·46$ Å) or the single helix type ($h = 1·73$ Å). The four points marked S show the possible single helical conformations, the upper and furthest to the right being right- and the others left-handed. The double helical model is in fact geometrically impossible using the residue geometry that we have assumed since h cannot exceed *ca.* 3·0 Å for a six-fold helix. However, a reduction of the glycosidic angle to *ca.* 113° would permit the structure, and this much distortion could be distributed around the sugar ring without exceeding the uncertainty limits of the co-ordinates.[4] It has been pointed out before[10,12] that some changes in geometry must occur within the sugar residues to account for the wide variation of amylose conformations. The closest approach using the standard geometry is shown on Figure 5 as D, the lower value being for a left-handed helix.

For conformations in solution, the calculations should give better predictions for non-aqueous rather than aqueous solvents because the latter might diminish the effect of inter-residue hydrogen bonding by competition for the sites. 'Special' hydration could stabilise particular conformations as we have discussed for monosaccharides[1] and in other ways (see below) but these would not be expected to take the conformation outside the 5 kcal mol^{-1} contour. It turns out that the optical rotation of methyl β-maltoside in dioxan and in dimethyl sulphoxide, corresponds[13] to an average conformation for which ϕ and ψ are both within 5° of the overall energy minimum, M (Figure 5), and within 10° of the conformation angles that are predicted by thermodynamic averaging over[13] the map. In aqueous solution the optical rotation is most reasonably interpreted[13,14] in terms of an important contribution from the conformation shown as R (Figure 5).

Although the intramolecular energy is <5 kcal mol^{-1} above the minimum, this aqueous conformation is within a subsidiary contour of much lower Boltzmann

---

* Following our usual convention,[11] n is negative for a left-handed helix, $|n| \geqslant 2$, and h always positive.

[5] G. J. Quigley, A. Sarko, and R. H. Marchessault, *J. Amer. Chem. Soc.*, 1970, **92**, 5834.

[6] S. S. C. Chu and G. A. Jeffrey, *Acta Cryst.*, 1967, **23**, 1038.

[7] A. Hybl, R. E. Rundle, and D. G. Williams, *J. Amer. Chem. Soc.*, 1965, **87**, 2779.

[8] P. C. Manor and W. Saenger, *Nature*, 1972, **237**, 392; W. Saenger, personal communication.

[9] T. Miyazawa, *J. Polymer Sci.*, 1961, **55**, 215.

[10] D. A. Rees, *MTP Internat. Rev. Sci.*, *Org. Chem.*, *Series One*, 1973, **7**, 251.

[11] D. A. Rees and R. J. Skerrett, *Carbohydrate Res.*, 1968, **7**, 334.

[12] A. D. French and V. Murphy, *Carbohydrate Res.*, 1973, **27**, 391.

[13] D. A. Rees and D. Thom, in preparation; D. Thom, Ph.D. Thesis, University of Edinburgh, 1973.

[14] D. A. Rees, *J. Chem. Soc.* (*B*), 1970, 877.

probability than the main one and is indeed the only conformation that does not fall within the main contour. Two reasons can be suggested for this. First, it is easily seen from the calculations (or from molecular models) that the main area derives much stability from the internal O(2) · · · O(3) hydrogen bond that could be sequestered in aqueous solvent. Secondly, as pointed out by Neal and Goring and shown by a photograph of a molecular model in their paper,[15] conformations in the region of R might derive special stabilization from internal hydrophobic bonding which has not been included in our calculations.

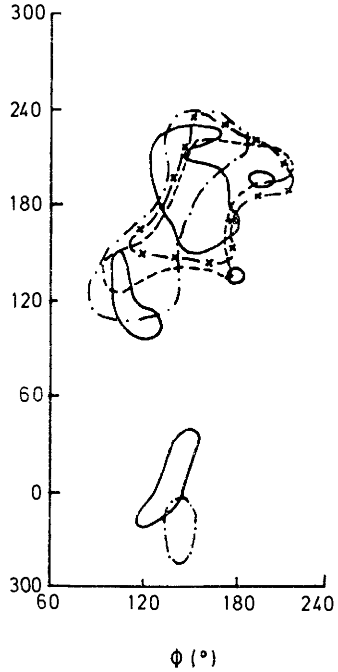

FIGURE 6 Comparison of potential energy calculations for maltose. Contours are 5 kcal mol^{-1} above the minimum: —— present calculation; - - - - - Rao *et al.*[a] using a Flory potential (6-exp) with no polar or torsion terms; Brant and Dimpfl[b] using a Lennard-Jones potential (6-12) and some polar but no torsion terms, is very similar; — × — × — Rao *et al.*[c] using a Kitaigorodsky potential with no polar or torsion terms; Blackwell *et al.*[d] using a similar function to Rao (Flory) with the inclusion of a hydrogen bonding term, is similar; — · — · — Pullman *et al.*[e] using a simple quantum mechanical treatment

[a] V. S. R. Rao, P. R. Sundararajan, G. N. Ramachandran, and C. Ramakrishnan in ' Conformation of Biopolymers,' ed. G. N. Ramachandran, Academic Press, London, 1967, vol. 2, p. 271. [b] D. A. Brant and W. L. Dimpfl, *Macromolecules*, 1970, **3**, 655. [c] V. S. R. Rao, N. Yathindra, and P. R. Sundararajan, *Biopolymers*, 1969, **8**, 325. [d] J. Blackwell, A. Sarko, and R. H. Marchessault, *J. Mol. Biol.*, 1969, **42**, 379. [e] M. Giacomini, B. Pullman, and B. Maigret, *Theor. Chim. Acta*, 1970, **19**, 347.

A comparison of our energy map with several reported in the recent literature is made in terms of the 5 kcal mol^{-1} contours (Figure 6). Not all workers report the depth of the minimum, and some define ϕ and ψ in ways which cannot be exactly related to the present definition, and therefore the comparison is not exact. Although it can be seen that all methods give results that are broadly similar, no treatment other than ours can account for all the conformations known. By studying the components of the present energies it can be seen that the inclusion of a full polar and hydrogen bond treatment substantially modifies the map, giving sharper minima and favouring areas of lower ϕ.

The energy map for cellobiose (Figure 7) can also be tested by comparison with experiment. The crystal conformations shown by C are (upper) β-cellobiose,[16] (middle) α-lactose,[17] and (lower) methyl β-cellobioside.[18] All these, and the generally accepted [10] bent-chain conformation of cellulose (shown as B; $n = 2$, $h = 5{\cdot}15$ Å) fall close to the overall minimum, M, and within the 5 kcal mol^{-1} contour. Optical rotation studies [13,14] suggest that the linkage conformations in aqueous solution and in dimethyl sulphoxide are similar and close to that of the crystal, and correspond almost exactly to prediction made by thermodynamic averaging of optical rotation over the map.[13] The Boltzmann probability of the region around $\phi = 210$, $\psi = 150°$ is *ca.* 77% and that around $\phi = 210$, $\psi = 0°$ is *ca.* 20%.

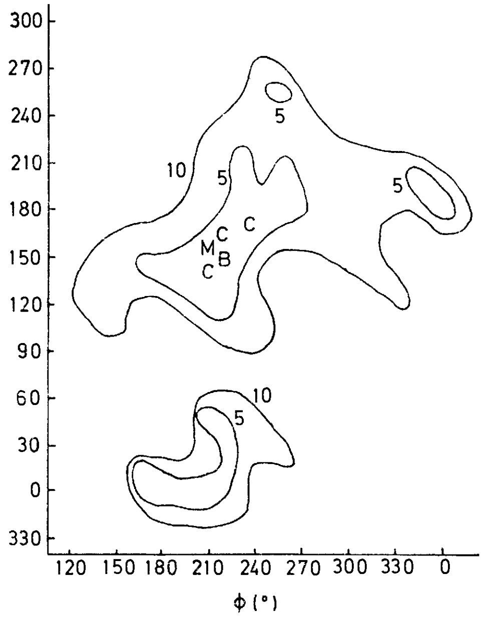

FIGURE 7 Calculated energy surface for cellobiose. For explanation of symbols, see text

The conformations close to the minimum derive much of their stability from a hydrogen bond between O(5) on the nonreducing residue and O(3) on the reducing residue and those around $\phi = 210$, $\psi = 0°$ are stabilized by a hydrogen bond between the two O(6) atoms. Although the latter group would seem to be a favourable range of conformations, they have not been found experimentally for cellobiose derivatives. If such conformations were adopted in cellulose, they would show a

---

[15] J. L. Neal and D. A. I. Goring, *Canad. J. Chem.*, 1970, **48**, 3745.

[16] S. S. C. Chu and G. A. Jeffrey, *Acta Cryst.*, 1968, **B24**, 830.

[17] D. C. Fries, S. T. Rao, and M. Sundaralingam, *Acta Cryst.*, 1971, **B27**, 994.

[18] J. T. Ham and D. G. Williams, *Acta Cryst.*, 1970, **B26**, 1373.

projected residue height close to zero and this would not fit any known cellulose structure except perhaps at fold regions.[19]

In maltose, the general shape of the accessible region of the map is dictated largely by van der Waals interactions between fixed atoms. Because the β-linkage is less constricted than the α, however, these terms allow a much wider region for cellobiose and the polar and torsional terms are correspondingly more important. The correspondence between minima and particular hydrogen bonds is also much more marked for cellobiose. However, the minima for maltose and cellobiose are of almost the same energy (maltose is *ca.* 0·8 kcal mol^{-1} lower).

A comparison with other recent studies is shown in Figure 8, and similar comments apply here as in the maltose comparison, except that the variation is much greater. This is presumably a consequence of the flatter energy surface for cellobiose.

Conclusions.—Clearly, the potential function matches well the observed properties of the linkages studied. The hydrogen bond potential reproduces the observed tendencies to form such bonds, without, contrary to many other published calculations, making any prior assumptions as to which oxygen atoms are involved.

The known conformations of the cellobiose and maltose residues in the solid state and in aqueous and non-aqueous solutions, do seem to cluster around the forms (Figure 9) that we predict to correspond to the minimum energies.

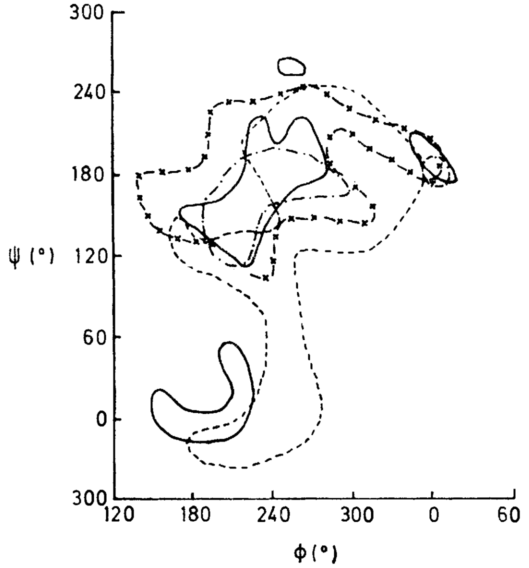

FIGURE 8 Comparison of potential energy calculations for cellobiose. Contours are 5 kcal mol^{-1} above the minimum: —— present calculation; – – – – – Pullman *et al.*[a] using a simple quantum mechanical treatment; — × — × — Yathindra and Rao[b] using a Kitaigorodsky function and glycosidic bond torsion, but no polar terms; — · — · — Sathranayarana and Rao[c] using a similar function to the preceding, with the inclusion of hydrogen bonding

[a] Figure 6, note *e*. [b] N. Yathindra and V. S. R. Rao, *Biopolymers*, 1970, **9**, 783. [c] B. K. Sathyanarayana and V. S. R. Rao, *Biopolymers*, 1971, **10**, 1605.

The chief disadvantage of our method is the amount of computation required for the statistical averaging: the maltose calculation, for example, requires 12 processor hours on an IBM 360/50 machine. Given present resources, therefore, recourse will have to be made to simpler descriptions for higher polymers. Nevertheless this work does show the kind of limitations that enter into energy calculations as a result of such approximations as the neglect of side group rotations, and the improvement that can result when a more realistic model is taken.

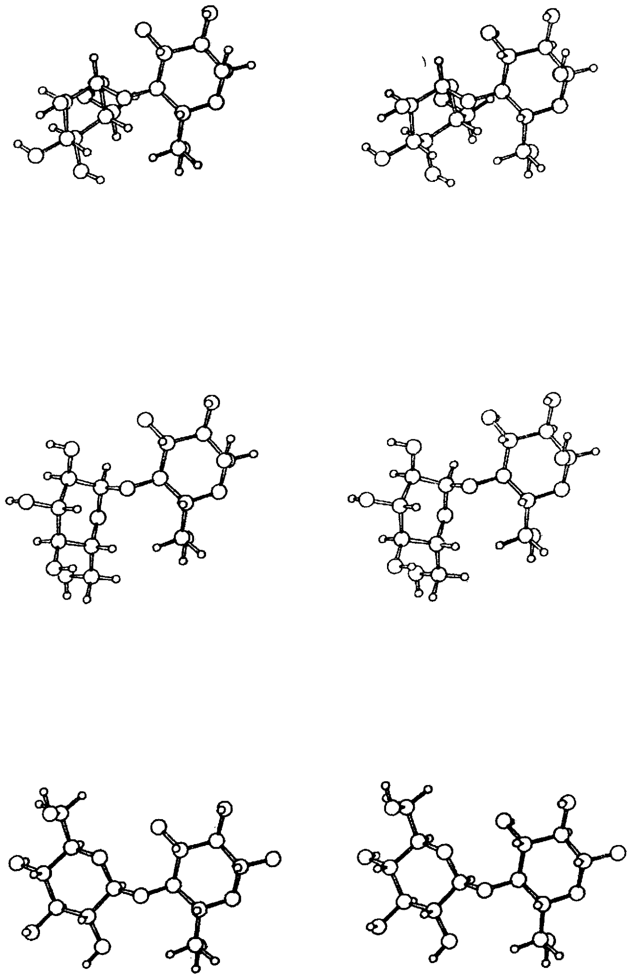

FIGURE 9 Stereoviews of minimum energy conformations for (lower) cellobiose (conformation M in Figure 7; $\phi = 210$, $\psi = 155°$); (middle) maltose (conformation M in Figure 5; $\phi = 145$, $\psi = 165°$); (upper) maltose with inter-residue hydrogen bonding quenched (conformation R in Figure 5; $\theta = 110$, $\psi = 140°$). These are drawn by the computer program ORTEP,[20] with the left-hand residue in the same orientation in each structure

P. J. C. S. thanks Unilever Research for a research studentship. We are grateful to Dr. W. Saenger for crystal structure information.

[4/1691 *Received*, 12*th August*, 1974]

---

[19] D. A. Rees and R. J. Skerrett, *J. Chem. Soc.* (*B*), 1970, 189.
[20] C. K. Johnson, Report ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1965.