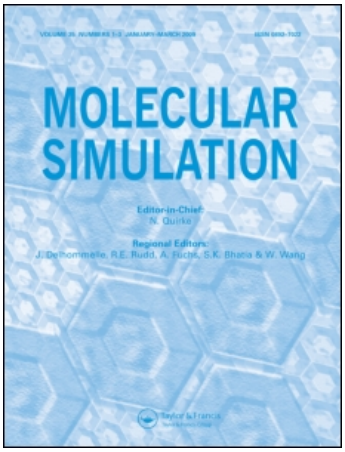

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Roles of starting geometries in quantum mechanics studies of cellobiose

A. D. French[a]; G. P. Johnson[a]
[a] Southern Regional Research Center, US Department of Agricuture, New Orleans, LA, USA

To cite this Article French, A. D. and Johnson, G. P.(2008) 'Roles of starting geometries in quantum mechanics studies of cellobiose', Molecular Simulation, 34: 4, 365 — 372
To link to this Article: DOI: 10.1080/08927020701663347
URL: http://dx.doi.org/10.1080/08927020701663347

Molecular Simulation
Vol. 34, No. 4, April 2008, 365–372

Taylor & Francis
Taylor & Francis Group

Roles of starting geometries in quantum mechanics studies of cellobiose†

A.D. French* and G.P. Johnson

Southern Regional Research Center, US Department of Agricuture, New Orleans, LA, USA

(*Received 30 June 2007; final version received 4 September 2007*)

In earlier work, a relaxed HF/6-31G(d) energy surface was constructed for the fraction of ϕ, ψ space that contains most geometries from crystals of molecules similar to cellobiose. Two regions around other minima were examined with unconstrained B3LYP/6-31 + G(d) minimisations, as were two sub-regions covered by the map. More than 100 different exo-cyclic group arrangements ("starting geometries"), selected for stability and low energies, were tested at each ϕ, ψ point and in the four sets of unconstrained minimisations. The influence of these starting geometries was studied in the present work. Twenty-four unique structures gave the lowest energy at one or more of the 81 ϕ, ψ grid points. Structures from the unconstrained minimisations covered wide ranges of ϕ, ψ space. Also, the 11–18 kcal/mol ranges of relative energies for the unconstrained minimisations resulting from the different starting geometries were comparable to the 20 kcal/mol range of energies on the relaxed map where some conformations were held in high-energy ϕ, ψ shapes by the mapping procedure. Thus, exo-cyclic group orientations and the ϕ and ψ torsion angles are both important factors in determining the likely structures and energies of a disaccharide.

Keywords: cellulose; carbohydrate; modelling; electronic structure theory

1. Introduction

Computerised conformational analysis can be carried out in several ways, even after the type of energy calculation is chosen. The present work applies adiabatic mapping to β-cellobiose (Figure 1), the shortest of the β-(1,4)-linked chains that correspond to cellulose, the most prevalent biopolymer on earth. Such maps depict the conformational landscape for a molecule. Two limitations of this method are seemingly substantial: the use of potential energy does not take into account entropy that could give free energy results, and it is difficult to explicitly consider solvent molecules. Despite these limitations, studies have shown that the experimentally determined shapes of crystalline disaccharides can be predicted from the lowest energy regions of their adiabatic maps [1]. Figure 2 shows a HF/6-31G(d) energy surface that we recently published for cellobiose [2], along with the locations of the observed crystal structures of cellobiose and its close relatives.

The axes of Ramachandran maps such as shown in Figure 2 are the torsion angles, ϕ = O5′—C1′—O4—C4 and ψ = C1′—O4—C4—C5 (see Figure 1 for atom labels). Figure 2 covers only part of conformation space because of the large amount of computer time required when maps are calculated with *ab initio* quantum mechanics. By the definition of an adiabatic surface, the calculated potential energy is at a minimum at each increment of ϕ and ψ [3]. This has two practical implications. First, the geometry must be optimised through energy minimisation. Before minimisation, the particular coordinates of the structure will depend on the source of the structure. One might use a structure from a crystallographic study, or some sort of computer model. That structure's internal geometry might not have conflicts at its original ϕ and ψ, but it will not be optimal at other values of ϕ and ψ. Minimisation results in structural adjustments that relieve the artifactual deficiencies that would otherwise penalise the structures with other ϕ and ψ values. Maps based on minimisation at each increment are called "relaxed".

Secondly, the changes of energy that result from variation of the exo-cyclic group orientations must be considered as well, again at each increment of ϕ and ψ. While minimisation should orient the hydroxyl and hydroxymethyl groups to local minima, carbohydrates will always have multiple stable minima with various energies. The magnitude of this problem is substantial [4,5], as indicated in Figure 1: if all placements of each group in the three staggered orientations are considered, there would be $3^{10} = 59{,}049$ ways that the groups could be arranged. Not all will be stable, and experience has shown that arrangements that allow formation of intra-residue hydrogen bonds will have lower energy than those that do not. To create an adiabatic map, a straight-forward approach is to prepare a *sufficient* number of individual maps based on structures that each have a different, low-energy arrangement of the exo-cyclic groups. After

†Presented in part at the 231st National Meeting of the American Chemical Society, Atlanta, Georgia.
*Corresponding author. Email: afrench@srrc.ars.usda.gov

ISSN 0892-7022 print/ISSN 1029-0435 online

DOI: 10.1080/08927020701663347
http://www.informaworld.com

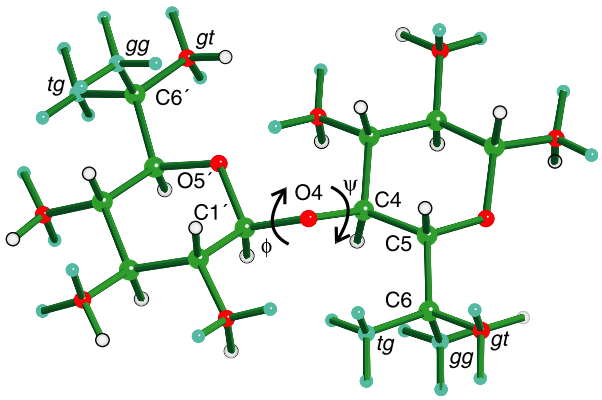

Figure 1. β-Cellobiose, showing all of the staggered orientations of the exo-cyclic groups and ϕ and ψ. The atoms that define ϕ and ψ (see text) are labelled, as are C6 and C6′, as well as all three staggered orientations of O6 and O6′ (*tg*, *gg* and *gt*). Blue atoms are in alternative positions.

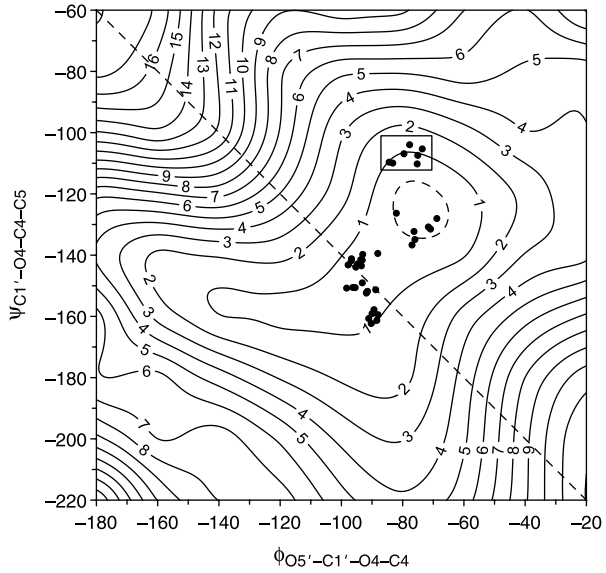

Figure 2. HF/6-31G(d) energy surface for cellobiose. The dashed contour line is at 0.25 kcal/mol above the global minimum, the other contour lines are at 1 kcal/mol intervals. Crystal structures from appropriate small molecules with β-(1,4)-linkages are indicated by dots. The diagonal line indicates that a polymer with these values of ϕ and ψ would conform to 2-fold screw symmetry. Crystal structures in the rectangular box cannot have a typical O3···O5′ hydrogen bond because of their linkage geometry.

considering all of these maps, the lowest energy at each ϕ, ψ point is then selected for the adiabatic map. The role of the structures with different exo-cyclic group orientations, or "starting geometries," is the subject of this paper.

Besides the mapping in Figure 2, we also studied cellobiose by starting energy minimisations at four specific values of ϕ and ψ. These locations (A–D) are shown in Figure 3, which depicts the entire ϕ, ψ conformational space. This procedure was used in part because there is, so far, only one crystallographically observed point in the D region; none are in the C region. This is despite the fact that C corresponds to the calculated global minimum [2,6] and D is the location of a substantial secondary minimum. That global minimum structure has a highly cooperative hydrogen bonding system that accounts for its stability [7]. It is arguably superior to hydrogen bonding systems in the crystallographic region because its two inter-residue hydrogen bonds are hydroxyl-to-hydroxyl. In the crystallographic region, only one hydroxyl–hydroxyl bond can occur, a bond that requires O6′ to be in the rare *tg* orientation, plus the usually-observed O3—H···O5′ bond. Bonds to the ring oxygen are weaker than interactions between two hydroxyl groups.

Explanations for the lack of correspondence between the global minimum and the experimental structures include the absence of solvation in the model [8] (which would disrupt the hydrogen bonding system). Also, there is higher entropy in the crystallographic region, inferred from low-temperature spectroscopy [9–11]. Our unpublished maps based on several empirical force fields also favor minimum C unless the dielectric constant is elevated. Also, models of cellobiose analogues that lack hydroxyl groups favor the same conformational region as cellobiose crystal structures [12]. This indicates that the preference for the C minimum is based on relatively

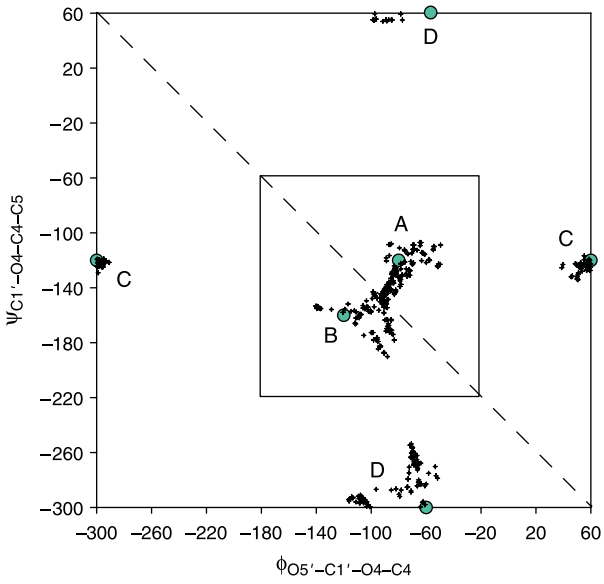

Figure 3. ϕ and ψ space, showing the subset (inset square) that was subjected to mapping (Figure 2) and the A–D locations for the start of unconstrained minimisation. Plus signs indicate the locations of the structures resulting from unconstrained minimisation. The axes are approximately equivalent to ϕ_H (H1′—C1′—O4—C4) and ψ_H (C1′—O4—C4—H4) taking values from −180° to +180°.

simple principles, as suggested by the above explanations. In any case, it will be useful to have adiabatic maps of continually improving quality so that we can better understand the roles of all factors in determining the conformational behavior of cellobiose.

2. Methods

Basically, our initial cellobiose structures (starting geometries) came from three sources. A Monte Carlo conformational search used the AMBER* force field as implemented in Schrödinger MacroModel [13]. The conformations that resulted were then refined with XCluster [13], a geometric cluster analysis. The conformation generator routine as implemented in Chem-X (no longer distributed) furnished a second set. The set of 26 structures from Strati *et al.* [6] were also used, kindly sent by Dr Momany. Initially we had sketched those structures (with Maestro [13]) based on their published parameters but made some mistakes due to confusion regarding their published torsion angles. We decided to keep the incorrectly sketched structures as another set of starting conformations because of the differences, both from the incorrect hydroxyl orientations as well as the small differences in bond lengths, etc. that resulted from sketching instead of using the coordinates from their B3LYP/6-311++G(d,p) minimisations. For the mapping studies, all structures were modified by increasing their glycosidic angles from their minimised value to 150°, except that we neglected to do this for the complete set of the correct Strati *et al.* [6] structures. We repeated the minimisations of the Strati *et al.* geometries with the increased glycosidic angles and reached reasonable energy values at more places on the map. The increase in the glycosidic angle is important in avoiding interpenetration of the two glucose rings prior to minimisation. In total, we used 181 starting geometries for each incremental point on the maps, although from a practical viewpoint there were some redundancies. To clarify, we use the phrase "starting geometry" to indicate the exo-cyclic group orientations and other intramolecular coordinate variations, regardless of the ϕ and ψ values.

We consider that 113 of these starting geometries are grossly different from each other by the criterion that one or more of the exo-cyclic group orientations are not in the same staggered orientation, $\pm 30°$. Other differences, besides the differences in glycosidic bond angles for one of the sets of structures from Strati *et al.* are more subtle. Still, it is fair to say that none of the three primary methods (the conformational search with MacroModel, the Chem-X conformation generator, or the search by Strati *et al.*) gave any structures that are duplicates in a practical sense. This happened even though each method was intended to produce the structures that would correspond to the important low-energy forms.

Maps were constructed at nine 20° increments each of ϕ and ψ, giving 81 grid points. The ϕ and ψ values were adjusted for each grid point on the map, always starting from the same initial structure for a given starting geometry. Regarding the unconstrained minimisations, two starting points were in sub-regions of the central region, called A and B (Figure 3). For the A region, initial ϕ and ψ values were $-80°$, $-120°$ and for the B region, $-120°$, $-160°$. The other two starting points (C and D) correspond to the global minimum region of Strati *et al.* [6] and another well-known secondary minimum. They had initial ϕ and ψ values of 60°, $-120°$ (left and right edges) and $-60°$, 60° (top and bottom edges). These unconstrained minimisations were carried out without increased glycosidic bond angles, so 155 starting geometries were used.

We used the Jaguar software [13] on a Linux cluster. All structures were minimised with HF/6-31G(d) theory, followed by single point calculations at the HF/6-311+G(d) level. The unconstrained minimizations were also calculated at the B3LYP/6-31+G(d) level, followed by 6-311+G(d) single point calculations. For the maps, only the HF/6-31G(d) results are discussed herein and only the B3LYP/6-31+G(d) results are discussed for the unconstrained minimisations. We are aware that these levels of theory have inadequacies but still they are widely used on problems far smaller than the present project. We look forward to more elaborate calculations, but even with faster computers, we will need to reduce the number of starting geometries. We expect that the present work will be helpful in making this reduction.

Analysis of the energy map consisted of three main operations. (1) The 20 lowest-energy results (of the 181 energies) for each of the 81 ϕ, ψ points were listed in a spreadsheet to see how they varied. (2) We visually inspected all 81 lowest-energy structures, measured the exo-cyclic group torsion angles and classified them for comparison with the starting structures. (3) All 81 structures were inspected for the particular final geometry that had given the overall lowest map energy (structure B1 in Figure 6, see below). Such detailed analyses are not normally done for mapping studies, but because an extreme amount of computational effort had been expended, each individual result was more valuable than otherwise might have been the case. Details of the important minimum-energy structures were presented previously [2].

3. Results and discussion

3.1 Unconstrained minimisations

Both the geometries and the energies from the unconstrained minimisations are of interest. Their distributions are shown in Figures 3 and 4 (geometries) and five (energies). The Figures clarify that there are many

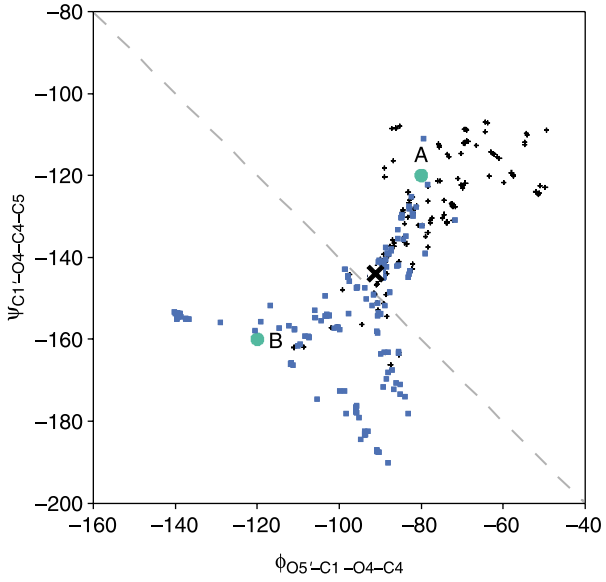

Figure 4. The A/B region, showing how many of the original A structures changed to B, and vice versa, during minimisation. Black + signs indicate the structures started at A and blue squares indicate the structures started at B. The large round dots show the original ϕ and ψ values, and the large × indicates the structure of the overall minimum for this region. The final structure is almost on the 2-fold screw axis, having started from both A and B.

different geometries and corresponding energies. Figure 3 has remarkably different distributions depending on which of the locations, A–D, was used as the starting point for the minimisations. The C geometries were confined to a very small range of ϕ and ψ whereas the structures started from the A, B and D regions had larger ranges of ϕ and ψ. The HF results (not shown) were similar, so the restricted range of the C structures supports the idea of reduced entropy in this region. Figure 4 shows a close-up view of the A and B regions. Here, the minimisations often took the original structures across the 2-fold screw axis line into the other region. In at least one case, the same starting geometry, started from both A and B, resulted in identical final structures. It came to our attention because that final structure corresponds to the overall minimum for the A/B region.

The distributions of energies shown in Figure 5 are remarkable for several reasons, especially the ranges of relative energy values. Bearing in mind that these all correspond to stable structures that were obtained from procedures that were intended to locate the lowest-energy conformers, we consider the ranges of 11 to nearly 18 kcal/mol to be large. While the energy map in Figure 2 depicts a roughly equivalent 20 kcal range of energies, and the maximum for the entire ϕ, ψ surface might be as much as 30 kcal/mol, those high map energies are for constrained structures. It is widely held that the ϕ and ψ torsions are the most important structural variables for disaccharides. However, these ranges of energy for the different exo-cyclic group orientations in fully minimised, uncon-strained structures show that exo-cyclic group orientations are equally important in terms of finding the lowest energy structures. Another important aspect of Figure 5 is the continuum of energy values, although the C well minimisations had a couple of quantum jumps of about 1 kcal/mol. Five of the 155 structures in well C minimised to the same final (global minimum) structure. Finally, the distributions of conformations in the A, B and D regions indicate the many different locations for the minima that would be found, if energy maps were made with each of these different starting geometries.

3.2 *Analyses of the relaxed map*

Our first question to be answered by this analysis was the basis for the low energy of the observed crystal structures that have ϕ and ψ values that would not be compatible with

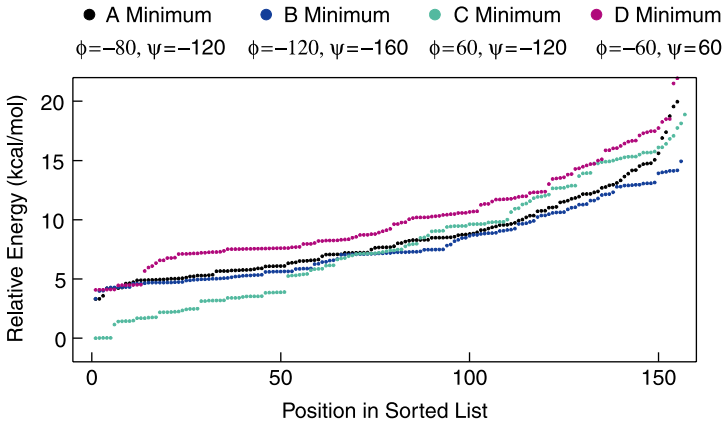

Figure 5. Distribution of B3LYP/6-31 + G(d) energies for unconstrained minimisations starting from the four regions depicted in Figure 3. Energies for each region were sorted independently.

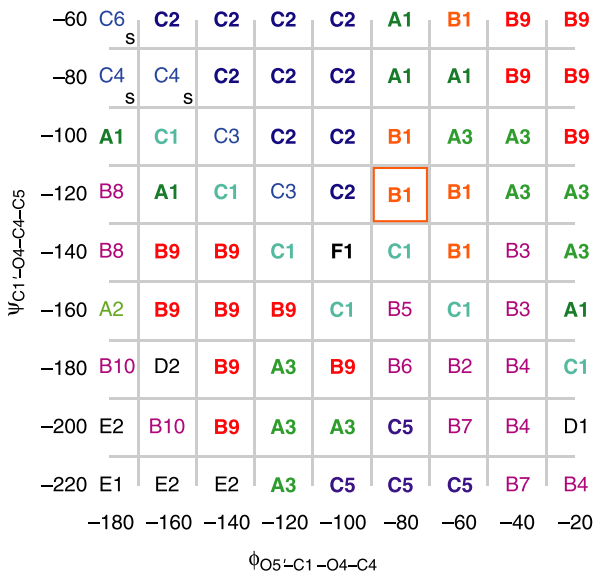

Figure 6. Lowest-energy structure at each ϕ, ψ grid point for the map in Figure 2. Structures were classified with letters to represent the orientations of the O6 and O6′ groups, respectively: A, *gg*, *gg*; B, *gg*, *gt*; C, *gg*, *tg*; D, *gt*, *gt*; E, *tg*, *gt*; and F, *tg*, *tg*. The s characters denote skew form rings. The number following the letter indicates a particular arrangement of the hydroxyl groups (described in Table 1).

normal O3—H···O5′ hydrogen bonds [14,15]. Because a hydrogen bond provides several kcal/mol of stabilisation, it was perplexing that the boxed structures in Figure 2 had only 1 kcal/mol or so higher relative energy. We found that the lowest-energy structure at the $\phi = -80°$, $\psi = -100°$ grid point had the same exo-cyclic group orientations as the overall map minimum but the O3—H···O5′ hydrogen bond was lengthened, to 2.77 Å, well beyond what would normally be considered a hydrogen bond. Thus, a distorted version of the overall map minimum has the lowest energy at that point, rather than a basically different structure. This similarity of resulting geometries is indicated in Figure 6, which shows that the lowest overall energy was attained for structure B1, indicated in the small box. Just above it is the $\phi = -80°$, $\psi = -100°$ grid point and Figure 6 indicates that the B1 structure gave the lowest energy there, also.

Examination of all 81 minimum energy structures revealed that 24 had unique orientations of the exo-cyclic groups. Figure 6 shows which O6 group orientations gave the minimum energies in each of the 81 locations, and Table 1 provides the approximate orientations of hydroxyl groups for the 24 unique models. There were as many as 10 unique arrangements of the hydroxyl groups for a given pair of O6 orientations (the B1–B10 structures with O6 and O6′ *gg* and *gt*, respectively). The most frequently observed structure, B9, accounted for 13 of the 81 lowest energies. Of the nine possible combinations of staggered O6 and O6′ orientations, the *gt*–*gg*, *gt*–*tg* and *tg*–*gg* combinations were not observed.

Only five of the structures at the 81 points had any hydroxyl groups that were positioned more than ± 30° from one of the three staggered orientations (see structures B6, B7 and B8 in Table 1). These outliers were attached to either O3 or O2′ and were participating in inter-ring hydrogen bonds. Of the remaining hydroxyl groups, eight O2′—H groups had H—C—O—H torsions near 180°, all others were ± 60°.

The 20 lowest energy structures at each of the 81 ϕ, ψ points had energy ranges from 0.16 to 3.20 kcal/mol, not counting locations where skew-form rings resulted from the minimisations. In the case, when the range was only 0.16 kcal/mol, the 20 structures included examples in which O6′ was in *gg*, *gt* and *tg* orientations.

A number of structural transformations occurred during the energy minimisations, including transitions of the reducing ring to a $^{2}S_{O}$ form at three ϕ, ψ points. These skews had energies 3.5–6.0 kcal/mol less than the $^{4}C_{1}$ form at those points. An approach to dealing with various ring forms is discussed in an accompanying paper in this special issue [16]. Figure 7 shows the pre-minimisation structure at $\phi = -180°$, $\psi = -80°$. The short contact between C6 and O2′ at that conformation is the primary cause of the conversion to the skew form ring. In the starting model, the glycosidic angle is 150° and the distance is 2.46 Å. This is already short, but if the glycosidic angle had been 115°, the distance would have been only 1.16 Å. After the minimisation that gave the skew form, the glycosidic angle was large, at 126° and the contact between C6 and O2′ was increased to 3.09 Å. Also, the skew form allowed formation of an excellent O6—H···O1 hydrogen bond. Still, the final relative energy at $\phi = -180°$, $\psi = -80°$ was high: 17.8 kcal/mol.

Another transition was the reorientation of the hydroxyl groups on the reducing ring at $\phi = -120°$, $\psi = -200°$ (or $+160°$). At that point, the trigger for the transition was a short contact between O3—H and H1′. With the glycosidic angle of 150°, the distance was 1.77 Å but it would have been only 0.56 Å at 115°. The reducing ring of the starting geometry had a ring of hydroxyl groups that were pointing clockwise. Rotation of O3—H to relieve the contact with H1′ resulted in subsequent rotations of O2—H and O1—H, to make a ring of hydroxyl groups pointed counter-clockwise. We visualised this rotation by making a movie from the coordinates at each step of minimisation. There was a "falling domino" sequence of these rotations.

A highly unwelcome structural change was the formation of an anhydro ring with loss of a water molecule when $\phi = -180°$, $\psi = -80°$, the same highly strained grid point that led to one of the skew form rings when a different starting geometry was used. In this case,

Table 1. Exo-cyclic group orientations for the 24 unique structures in Figure 6.

	Exo-cyclic group[a,b,c]									
Structure	O6	O6′	O1H	O2H	O3H	O6H	O2′H	O3′H	O4′H	O6′H
A1	*gg*	*gg*	m	p	m	p	m	p	m	p
A2	*gg*	*gg*	p	m	p	p	t	p	m	p
A3	*gg*	*gg*	p	m	p	p	m	p	m	p
B1	*gg*	*gt*	m	p	m	p	m	p	m	m
B2	*gg*	*gt*	m	p	m	p	m	p	m	t
B3	*gg*	*gt*	m	p	m	p	m	p	m	m
B4	*gg*	*gt*	m	p	m	p	m	p	m	t
B5	*gg*	*gt*	m	p	m	p	m	p	p	p
B6	*gg*	*gt*	m	p	−26.5	p	m	p	p	t
B7	*gg*	*gt*	m	p	−17.6	p	m	p	m	t
B8	*gg*	*gt*	p	m	p	p	−26.6	p	m	m
B9	*gg*	*gt*	p	m	p	p	m	p	m	m
B10	*gg*	*gt*	p	m	p	p	t	p	m	m
C1	*gg*	*tg*	m	p	m	p	m	p	m	p
C2	*gg*	*tg*	m	p	m	p	p	m	p	t
C3	*gg*	*tg*	m	p	m	p	t	p	m	p
C4(s)	*gg*	*tg*	p	m	p	p	p	m	p	t
C5	*gg*	*tg*	p	m	p	p	m	p	m	p
C6(s)	*gg*	*tg*	p	p	m	p	p	m	p	t
D1	*gt*	*gt*	m	p	m	m	m	p	m	t
D2	*gt*	*gt*	p	m	p	m	t	p	m	p
E1	*tg*	*gt*	p	m	p	p	m	p	m	m
E2	*tg*	*gt*	p	m	p	p	t	p	m	m
F1	*tg*	*tg*	m	p	m	m	m	p	m	p

[a] O6 and O6′ orientations are described with the usual *gauche,gauche (gg)*, *gauche-trans (gt)* and *trans-gauche (tg)* nomenclature (see Figure 1). [b] Secondary hydroxyl group orientations are described with "p" (+60°), "m" (−60°) and "t" (180°) for the torsion angles HOn—On—Cn—H, where *n* is the position of the carbon atom around the ring. A numeric value indicates an actual value in degrees for the torsion angle, which was more than 30° from any of the staggered values. Values for B7 and B8 are averages of two similar values. [c] Primary hydroxyl orientations are also described with "p", "m" and "t", as above but referencing the torsion angle HO6—O6—C6—C5.

where the chemical structures changed, the glycosidic angle had not been increased prior to minimisation and the O6 orientation was *tg* instead of *gt*. The resulting structures (water and 6,2′-anhydrocellobiose) had lower electronic energy than any of the cellobiose structures at that ϕ, ψ location. This result was removed from the mapping, but it was only detected because of the unusual, post-minimisation examination of the structures.

Despite the frequent findings that most cellulose crystal structures have $P2_1$ symmetry, there has been substantial controversy as to whether such 2-fold screw axis structures are at a minimum in energy for isolated cellulose molecules. The F1 structure, which lies on the diagonal line that denotes 2-fold screw axis symmetry, is not the structure with the minimum energy in the region of the crystal structures, at least not at the HF/6-31G(d) level. However, F1 is very similar to the lowest energy structure in this region that was obtained by unconstrained B3LYP/6-31 + G(d) minimisation from both the A and B starting points. It also falls very near the 2-fold screw axis line (Figure 3). Both the HF and B3LYP structures have O6 and O6′ in *tg* positions and O3—H···O5′ and O2′—H···O6 hydrogen bonds, similar to proposed structures of cellulose I [17,18].

Although previous work [4,5] pointed to the need to consider the various exo-cyclic group orientations, several fairly recent studies have ignored it. Figure 8 is presented

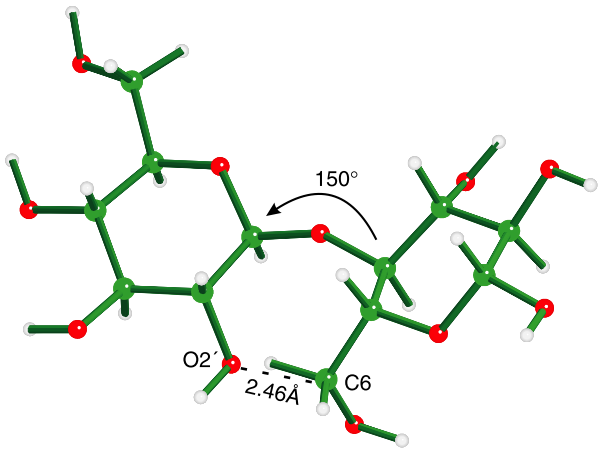

Figure 7. Starting geometry that led to a skew form (^2S_O) reducing ring at $\phi = -180°$, $\psi = -80°$. The C1′—O4—C4 glycosidic bond angle is 150°. If it had a conventional value of 115°, the C6···O2′ distance would be 1.16 Å instead of the indicated 2.46 Å.

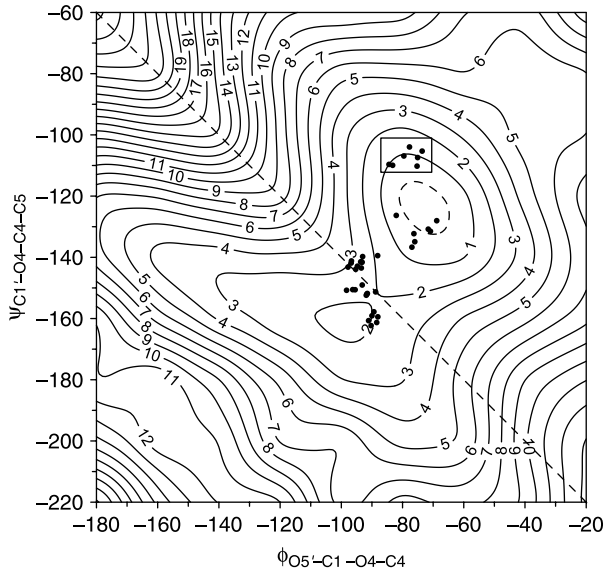

Figure 8. Map made with only one starting structure (the structure that gave the lowest overall energy in the adiabatic mapping.) This map should be compared with the adiabatic map in Figure 2.

to underscore this need. This map was constructed with the B1 structure, which gave the lowest overall HF/6-31G(d) energy. Energy differences with Figure 2 are as large as 7 kcal/mol in the high-energy, upper left region where the skews occurred for some starting geometries. However, the differences of about two kcal/mol around the crystal structures, near the diagonal 2-fold axis line are more important.

4. Conclusions

This paper provides a more detailed analysis of the results than is typically done in a conformational analysis paper. In this case, the availability of so many structures determined by *ab initio* quantum mechanics provided an unusual opportunity. In particular, this deep analysis was the first step in the selection of a smaller number of structures for higher-level calculations. Although we explored the roles of exo-cyclic group orientations in adiabatic mapping, the same consideration is also needed in molecular dynamics simulations. While adiabatic mapping requires multiple starting geometries, molecular dynamics studies will require scrutiny to assure adequate sampling of the various exo-cyclic group orientations.

The importance of a thorough study of these orientations was underscored initially when we realised that our three procedures for obtaining starting geometries each gave unique sets of structures. After producing the adiabatic map, the importance became more apparent because 24 of the different starting geometries were required to obtain the lowest energies at the 81 grid points. Earlier [19], we had published a preliminary HF/6-31G* map that was based on "only" 61 of these starting geometries, and it was clearly different and less predictive for the crystal structures, especially in the area of the 2-fold screw axis. The definition of "sufficient" for the number of starting geometries is larger than we had originally anticipated.

A close examination of the underlying data revealed a new pitfall for mapping studies when electronic structure theory is used. Although, it could have been avoided by explicitly specifying all atom connectivity when setting up the calculations, the formation of an anhydro sugar was unexpected. Another surprise was the complete conversion of a clockwise sequence of intra-residue hydrogen bonds to a counter-clockwise sequence. The finding that a skew form arose from the minimisation was less of a surprise, based on our studies of fluorinated disaccharide analogues [20].

References

[1] A.D. French, A.-M. Kelterer, G.P. Johnson, M.K. Dowd, and C.J. Cramer, *Constructing and evaluating energy surfaces of crystalline disaccharides*, J. Mol. Graph. Model. 18 (2000), p. 95.
[2] A.D. French and G.P. Johnson, *Quantum mechanics studies of cellobiose conformations*, Can. J. Chem. 84 (2006), p. 603.
[3] V.H. Tran and J.W. Brady, *Conformational flexibility of sucrose*, ACS Symp. Ser. 430 (1990), p. 213.
[4] A.D. French, V.H. Tran, and S. Pérez, *Conformational analysis of a disaccharide (cellobiose) with the molecular mechanics program (MM2)*, ACS Symp. Ser. 430 (1990), p. 191.
[5] C.A. Stortz, *Disaccharide conformational maps: how adiabatic is an adiabatic map?* Carbohydr. Res. 322 (1999), p. 77.
[6] G.L. Strati, J.L. Willett, and F.A. Momany, Ab initio *computational study of β-cellobiose conformers using B3LYP/6-311++G***, Carbohydr. Res. 337 (2002), p. 1833.
[7] G.L. Strati, J.L. Willett, and F.A. Momany, *A DFT/*ab initio *study of hydrogen bonding and conformational preference in model cellobiose analogs using B3LYP/6-311++G***, Carbohydr. Res. 337 (2002), p. 1851.
[8] W. Bosma, U. Schnupf, J.L. Willett, and F.A. Momany, *Computer modeling of β-cellobiose in a solvent environment: DFT calculations using implicit and explicit solvent models*, Abstracts, 231st National Meeting, American Chemical Society, March, 2006, CELL016.
[9] R.A. Jockusch, R.T. Kroemer, F.O. Talbot, L.C. Snoek, P. Çarçabal, J.P. Simons, M. Havenith, J.M. Bakker, I. Compagnon, G. Meijer, and G. von Helden, *Probing the glycosidic linkage: UV and IR ion-dip spectroscopy of a lactoside*, J. Am. Chem. Soc. 126 (2004), p. 5709.
[10] M. Hineno and H. Yoshinaga, *Far-infrared specta of glucose, sorbose, sucrose and cellobiose at liquid helium temperatures*, Spectrochim. Acta 28A (1972), p. 2263.
[11] M. Hineno and H. Yoshinaga, *Far-infrared spectra of β-D-glucose, cellobiose, galactose, lactose, α-D-glucose and saccharose in 50–30 cm^{-1} at liquid helium temperatures*, Spectrochim. Acta 29A (1973), p. 1575.
[12] A.D. French, A.-M. Kelterer, G.P. Johnson, M.K. Dowd, and C.J. Cramer, *HF/6-31G* energy surfaces for disaccharide analogs*, J. Comput. Chem. 22 (2001), p. 65.
[13] Macromodel, Xcluster, Maestro and Jaguar are all developed and marketed by Schrödinger, 101 SW Main Street, Suite 1300, Portland, Oregon 97204. USA. www.schrodinger.com.
[14] Z. Peralta-Inga, G.P. Johnson, M.K. Dowd, J.A. Rendleman, E.D. Stevens, and A.D. French, *The crystal structure of the α-cellobiose·2 NaI·2 H_2O complex in the context of related structures and conformational analysis*, Carbohydr. Res. 337 (2002), p. 851.

[15] A.D. French and G.P. Johnson, *What crystals of small analogs are trying to tell us about cellulose structure*, Cellulose 11 (2004), p. 5.
[16] C.A. Stortz, A.D. French. *Disaccharide conformational maps: adiabaticity in analogs with variable ring shapes*. Mol. Simul., 34 (2008), p. 373.
[17] Y. Nishiyama, J. Sugiyama, H. Chanzy, and P. Langan, *Crystal structure and hydrogen bonding system in cellulose Iα from synchrotron X-ray and neutron fiber diffraction*, J. Am. Chem. Soc. 125 (2003), p. 14300.
[18] Y. Nishiyama, P. Langan, and H. Chanzy, *Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron X-ray and neutron fiber diffraction*, J. Am. Chem. Soc. 124 (2002), p. 9074.
[19] A.D. French and G.P. Johnson, *Advanced conformational energy surfaces for cellobiose*, Cellulose 11 (2004), p. 449.
[20] A.D. French, G.P. Johnson, A.-M. Kelterer, and G.I. Csonka, *Fluorinated disacccharides as computational models*, Tetrahedron: Asymmetry 16 (2005), p. 577.