

# Torsional Anharmonicity in the Conformational Analysis of $\beta$ -D-Galactose<sup>†</sup>

Yvette K. Sturdy, Chris-Kriton Skylaris, and David C. Clary\*

Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom

Received: June 15, 2005; In Final Form: July 29, 2005

Schemes to include a treatment of torsional anharmonicity in the conformational analysis of biological molecules are introduced. The approaches combine ab initio electronic energies and harmonic frequencies with anharmonic torsional partition functions calculated using the torsional path integral Monte Carlo method on affordable potential energy surfaces. The schemes are applied to the conformational study of the monosaccharide  $\beta$ -D-galactose in the gas phase. The global minimum structure is almost exclusively populated at 100 K, but a large number of conformers are present at ambient and higher temperatures. Both quantum mechanical and anharmonic effects in the torsional modes have little effect on the populations at all temperatures considered, and it is, therefore, expected that standard harmonic treatments are satisfactory for the conformational study of monosaccharides.

## 1. Introduction

Carbohydrates are an important component of all biological systems, whether performing structural, metabolic, or recognition functions. They can be found as free molecules, bound to proteins or lipids, and within nucleotide units.<sup>1</sup> The three-dimensional shape of a carbohydrate is central to its particular function. For example, recognition processes show very high specificity to the extent that one monosaccharide can be selected over another. Two of the most common monosaccharides, glucose and galactose, can be distinguished between despite differing only in the chirality of one of their carbons. For example, peanut lectin is exclusively specific for galactose,<sup>2</sup> while the glucose/galactose-binding protein (GBP) of Gram-negative bacteria (important in chemotaxis and transport) shows a higher affinity for glucose, due to a precise steric fit.<sup>3</sup> Oligosaccharides are very flexible and exhibit a large number of possible conformations. A knowledge of the relative thermodynamic stability of these different conformers is an important step toward understanding the function of carbohydrates.

We discuss here two schemes for calculating the conformational preferences of biological molecules, based on quantum mechanical harmonic partition functions with corrections to include the torsional anharmonicity. These schemes combine standard electronic structure harmonic frequency calculations with torsional path integral Monte Carlo (TPIMC) simulations.<sup>4–8</sup> The TPIMC method gives a fully quantum mechanical and anharmonic description of the torsional coordinates of a molecule and was initially introduced for the calculation of internal and free energies of flexible molecules.<sup>4,5</sup> The technique was then applied to the conformational analysis of small biological molecules.<sup>6,7</sup> Calculations on the smallest amino acid, glycine,<sup>7</sup> using an ab initio potential energy surface (PES), predicted conformer populations in good agreement with experimental results and demonstrated the importance of the inclusion of a full quantum mechanical description of the

torsional modes. Incorporation of a harmonic treatment of the non-torsional modes into the analysis had little effect on the predicted populations, illustrating the importance of the torsional coordinates in determining the conformation of the molecule. Similar studies on the nor-adrenaline analogue, APE,<sup>6</sup> using a molecular mechanics PES also proved successful and showed that an anharmonic description of the torsions was necessary to give an accurate prediction of the relative conformer populations of the molecule. More recently, the TPIMC method has been combined with ab initio harmonic frequency calculations to give relative conformer partition functions with an anharmonic treatment of the torsional modes.<sup>8</sup> Such anharmonic corrections were found to have significant effects on the predicted conformer populations of the adrenaline molecule, with some relative populations being altered by over 20%.

The two schemes, which combine TPIMC simulations for the torsional modes with electronic structure energies and harmonic frequencies, are used here to predict the equilibrium conformer populations of the simple monosaccharide  $\beta$ -D-galactose in the gas phase.  $\beta$ -D-galactose consists of a pyranose ring with pendant hydroxyl and hydroxymethyl groups and has six torsional internal coordinates, which are treated here anharmonically. A harmonic treatment is used for the remaining coordinates, namely, the stretches and bends. Galactose, along with its structural analogues, has been the subject of many experimental and theoretical studies to date. Rahal-Sekkal et al.<sup>9</sup> carried out calculations on the isolated  $\beta$ -D-galactose molecule using both the HF/6-31G\* ab initio level of theory and the AM1 semiempirical method and concluded the “T” conformation of the C–C hydroxymethyl torsion to be the most stable (for conformer nomenclature, see Section 3). However, NMR studies by Tvaroška et al.<sup>10</sup> based on vicinal proton–proton coupling constants on methyl  $\beta$ -D-galactopyranoside showed dominance of the “G+” conformation in water and, to a lesser extent, of the “G–” conformation in methanol. These findings were supported by their density functional calculations at the B3LYP/6-311++G\*\* level of theory. Tvaroška et al. also reported predicted gas-phase populations from free energy estimations at the same level of theory, in which the G+/G–/T

<sup>†</sup> Part of the special issue “Michael L. Klein Festschrift”.

\* Author to whom correspondence should be addressed. E-mail: david.clary@chem.ox.ac.uk.

ratios were 30.4:29.8:39.8. Jockusch et al.,<sup>11</sup> in their gas-phase UV and IR ion-dip experiments, concluded that the G+(tttG+g-) conformer dominated at 90% and was found to be the global minimum at the MP2/6-311++G(d,p)//B3LYP/6-31+G(d) level of theory.

It is clear, then, that the conformation of galactose is influenced strongly by the environment, and that the predicted global minimum is dependent on the level of theory used. We focus here on the isolated  $\beta$ -D-galactose molecule, for which, to our knowledge, there is only one source of experimental data.<sup>11</sup>

In Section 2, we present the two methods for introducing the anharmonic treatment of the torsional internal coordinates, while in Section 3, the  $\beta$ -D-galactose molecule is introduced and calculation details are given. Section 4 presents results from the application of both schemes to the conformational analysis of  $\beta$ -D-galactose. A discussion of the conformational preference of the molecule and a comparison of the two schemes are given. Our conclusions are drawn in Section 5.

## 2. Theory: Torsional Anharmonic Correction

We present schemes that enable the calculation of molecular vibrational-electronic quantum partition functions based on an anharmonic treatment of the torsional internal coordinates and a harmonic treatment of the remaining modes. In particular, we consider the application of these schemes to the calculation of relative conformer populations of flexible molecules. Small differences in low-frequency modes (such as torsions) between conformers contribute much more to relative conformer free energies than do those for the higher frequency modes,<sup>7</sup> meaning that an accurate treatment of the former is essential to a good description of the conformational thermodynamics of a molecule.

The TPIMC method<sup>4,5</sup> provides a fully anharmonic, quantum mechanical treatment of torsional modes and has previously been successfully applied to the calculation of relative conformer populations of small flexible biological molecules.<sup>6,7</sup> The TPIMC method applies the path integral formalism<sup>12</sup> to molecular torsions. The torsional Hamiltonian is expressed as

$$H = - \sum_{j=1}^M \frac{\hbar^2}{2I_j} \frac{\partial^2}{\partial \theta_j^2} + V(\theta) \quad (1)$$

where  $M$  is the number of torsional degrees of freedom,  $I_j$  is the moment of inertia for the  $j$ th torsion, and  $V(\theta)$  is the intramolecular potential energy as a function of the torsional coordinates, calculated here using a molecular mechanics potential. The torsional quantum canonical partition function for a given conformer  $i$ ,  $q_i^{\text{TPI}}$ , relative to the total torsional quantum canonical partition function,  $q^{\text{TPI}}$ , is then equal to the population of conformer  $i$ :<sup>8</sup>

$$\rho_i^{\text{TPI}} = \frac{q_i^{\text{TPI}}}{q^{\text{TPI}}} = \lim_{P \rightarrow \infty} \frac{\prod_{j=1}^M \left[ \int_{D_i} d\theta_j \right] e^{-\beta(V^{\text{int}} + V^{\text{ext}})}}{\prod_{j=1}^M \left[ \int_0^{2\pi} d\theta_j \right] e^{-\beta(V^{\text{int}} + V^{\text{ext}})}} \quad (2)$$

$$V^{\text{int}} = \frac{P}{2\beta^2 \hbar^2} \sum_{j=1}^M \left[ \sum_{t=1}^P (\theta_{j,t} - \theta_{j,t+1})^2 \right] \quad (3)$$

$$V^{\text{ext}} = \frac{1}{P} \sum_{t=1}^P V(\theta_t) \quad I = \prod_{k=1}^N I_k \quad \theta_{j,P+1} = \theta_{j,1} \quad (4)$$

$P$  is the number of Trotter beads,<sup>13</sup>  $\theta_{j,t}$  is the value of the  $j$ th torsional angle for the  $t$ th Trotter bead, and  $\int_{D_i}$  denotes integration over the portion of the torsional coordinates that define conformer  $i$ . The classical result is recovered in the case where  $P$  is set to unity, while quantum values correspond to the limit as  $P \rightarrow \infty$ . For low frequency vibrations such as the torsional modes of a biological molecule, only three to five Trotter beads are usually required to obtain converged quantum values.<sup>6,7</sup>

High-quality ab initio PESs can only be used for systems of a limited size.<sup>7</sup> For larger systems, more affordable methods of calculating the intramolecular potential must be used. As in previous studies, we aim to describe the torsional PES using a force field, taking advantage of the relative computational affordability of evaluating molecular mechanics energies, while combining this with a more accurate, but computationally expensive, representation of the conformer electronic energies calculated using a higher level of theory. In this study, the higher-level energies and frequencies are calculated using density functional theory (DFT), although the procedure can, in principle, use any electronic structure method, and for this reason, all our formulas will refer to the high-level methods using the term “ab initio”. The more expensive ab initio methods will only be required to calculate minimum-energy conformer energies and their corresponding harmonic frequencies, rather than a whole PES. One way to combine a TPIMC calculation on an affordable potential with ab initio conformer electronic energies and harmonic frequency calculations for the non-torsional modes has recently been presented by Miller and Clary (which we will term the “ratio scheme”).<sup>8</sup> A “torsional anharmonic correction” to standard harmonic calculations was proposed, and the scheme yielded encouraging results when applied to the conformational study of adrenaline.<sup>8</sup> We will outline this scheme below and then develop an alternative approach.

The relative population of two conformers, a and b, of a molecule that is assumed to rotate rigidly can be expressed as

$$K_b^a = \frac{(I_x^a I_y^a I_z^a)^{1/2} q_{\text{ve}}^a}{(I_x^b I_y^b I_z^b)^{1/2} q_{\text{ve}}^b} \quad (5)$$

where  $q_{\text{ve}}^i$  is the vibrational-electronic partition function for conformer  $i$  and  $I_x^i$ ,  $I_y^i$ , and  $I_z^i$  are the corresponding principal axis rotational moments of inertia.

Within the harmonic approximation, the ab initio vibrational-electronic partition function for conformer  $i$ ,  $q_{\text{ve}}^{i,\text{AIHO}}$ , can be evaluated as

$$q_{\text{ve}}^{i,\text{AIHO}} = e^{-\beta E_{0,\text{AI}}^i} q_{\text{vib}}^{i,\text{AIHO}} \quad q_{\text{vib}}^{i,\text{AIHO}} = \prod_{j=1}^{3N-6} \frac{e^{-\beta \hbar \omega_j^i / 2}}{1 - e^{-\beta \hbar \omega_j^i}} \quad (6)$$

where  $\beta = 1/k_B T$ ,  $N$  is the number of atoms in the molecule,  $E_{0,\text{AI}}^i$  is the ab initio electronic energy at the equilibrium geometry of conformer  $i$ , and  $\omega_j^i$  are the corresponding harmonic frequencies. To include an anharmonic treatment of the torsional modes, the partition function composed of the torsional harmonic frequencies from a molecular mechanics potential,  $q_{\text{tor}}^{i,\text{MMHO}}$ , is factored out from  $q_{\text{ve}}^{i,\text{AIHO}}$  and the anharmonic partition function

based on the TPIMC calculation,  $q_{\text{tor}}^{i,\text{TPI}}$ , is inserted in its place:

$$q_{\text{ve}}^i = q_{\text{ve}}^{i,\text{AIHO}} \xi_{\text{tor}}^i \quad \xi_{\text{tor}}^i = \frac{q_{\text{tor}}^{i,\text{TPI}}}{q_{\text{tor}}^{i,\text{MMHO}}} \quad (7)$$

where  $\xi_{\text{tor}}^i$  is termed the anharmonic correction factor for conformer  $i$ .

Thus, the ratio of partition functions  $q_{\text{ve}}^a/q_{\text{ve}}^b$  in eq 5 is replaced by

$$\frac{\tilde{q}_{\text{ve}}^a}{\tilde{q}_{\text{ve}}^b} = \frac{q_{\text{ve}}^{a,\text{AIHO}} q_{\text{tor}}^{a,\text{TPI}} q_{\text{tor}}^{b,\text{MMHO}}}{q_{\text{ve}}^{b,\text{AIHO}} q_{\text{tor}}^{b,\text{TPI}} q_{\text{tor}}^{a,\text{MMHO}}} \quad (8)$$

where  $q_{\text{ve}}^{i,\text{AIHO}}$  and  $q_{\text{tor}}^{i,\text{MMHO}}$  are calculated trivially from the ab initio Hessian and molecular mechanics torsional Hessian, respectively.  $q_{\text{tor}}^{i,\text{TPI}}$  is evaluated according to

$$q_{\text{tor}}^{i,\text{TPI}} = e^{+\beta E_{0,\text{MM}}^i} \rho_i^{\text{TPI}} q^{\text{TPI}} \quad (9)$$

where  $E_{0,\text{MM}}^i$  is the molecular mechanics energy of conformer  $i$  and  $\rho_i^{\text{TPI}}$  is calculated using Boltzmann Monte Carlo sampling to evaluate eq 2. The total torsional partition function,  $q^{\text{TPI}}$ , cancels in the evaluation of relative conformer populations. The anharmonically corrected ratio of vibrational-electronic partition functions (eq 8) is thus

$$\frac{\tilde{q}_{\text{ve}}^a}{\tilde{q}_{\text{ve}}^b} = \frac{q_{\text{ve}}^{a,\text{AIHO}} \xi_{\text{tor}}^a}{q_{\text{ve}}^{b,\text{AIHO}} \xi_{\text{tor}}^b} = \frac{e^{-\beta(E_{\text{AI}}^a - E_{\text{AI}}^b)} \rho_a^{\text{TPI}} q_{\text{vib}}^{a,\text{AIHO}} q_{\text{tor}}^{b,\text{MMHO}}}{e^{-\beta(E_{\text{MM}}^a - E_{\text{MM}}^b)} \rho_b^{\text{TPI}} q_{\text{vib}}^{b,\text{AIHO}} q_{\text{tor}}^{a,\text{MMHO}}} \quad (10)$$

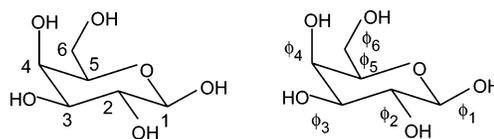
Here, we also develop an alternative anharmonic correction scheme, in which the ratio of partition functions in eq 5 is replaced by

$$\frac{\tilde{q}_{\text{ve}}^{a,\text{Proj}}}{\tilde{q}_{\text{ve}}^{b,\text{Proj}}} = \frac{q_{\text{ve}}^{a,\text{AIHOP}} q_{\text{tor}}^{a,\text{TPI}}}{q_{\text{ve}}^{b,\text{AIHOP}} q_{\text{tor}}^{b,\text{TPI}}} = \frac{e^{-\beta(E_{\text{AI}}^a - E_{\text{AI}}^b)} \rho_a^{\text{TPI}} q_{\text{vib}}^{a,\text{AIHOP}}}{e^{-\beta(E_{\text{MM}}^a - E_{\text{MM}}^b)} \rho_b^{\text{TPI}} q_{\text{vib}}^{b,\text{AIHOP}}} \quad (11)$$

$q_{\text{vib}}^{i,\text{AIHOP}}$  is the harmonic vibrational partition function calculated from an ab initio PES according to eq 6 after the torsional internal coordinates have been projected out from the Hessian using the projection operator method.<sup>14,15</sup> The torsional motions are projected out from the Hessian in Cartesian coordinates,  $\mathbf{H}$ , to give a new Hessian,  $\mathbf{H}^{\text{Proj}}$ :

$$\mathbf{H}^{\text{Proj}} = (1 - \mathbf{P})(1 - \mathbf{P}^{\text{RT}})\mathbf{H}(1 - \mathbf{P}^{\text{RT}})(1 - \mathbf{P}) \quad (12)$$

where  $\mathbf{P} = \sum_{i=1}^k \mathbf{e}_i \mathbf{e}_i^{\dagger}$  and  $\mathbf{P}^{\text{RT}}$  is the projection operator for the rotational and translational motions of the molecule as a whole.  $k$  is the number of “constraints” or internal motions to be projected out, and the  $\mathbf{e}_i$  are a set of orthonormal vectors of dimension  $3N$  formed by orthogonalizing a set of unit vectors,  $\mathbf{e}_i^0$ , one for each torsional coordinate to be projected out. Each  $\mathbf{e}_i^0$  has 12 nonzero elements, the form of which are as given by Wilson et al.<sup>16</sup>  $\mathbf{H}^{\text{Proj}}$  is then diagonalized to yield the harmonic frequencies used to calculate  $q_{\text{vib}}^{i,\text{AIHOP}}$ . In a similar way to the ratio method, we can define an anharmonic correction factor for the projection method for a given conformer  $i$ . The correction factor,  $\xi_{\text{tor}}^{i,\text{Proj}}$ , is given as the anharmonically corrected vibra-



**Figure 1.**  $\beta$ -D-galactose. Schematic representation depicting carbon atom and torsion labeling.

tional-electronic partition function divided by the ab initio harmonic partition function:

$$\xi_{\text{tor}}^{i,\text{Proj}} = \frac{\tilde{q}_{\text{ve}}^{i,\text{Proj}}}{q_{\text{ve}}^{i,\text{AIHO}}} = \frac{q_{\text{ve}}^{i,\text{AIHOP}} q_{\text{tor}}^{i,\text{TPI}}}{q_{\text{ve}}^{i,\text{AIHO}}} \quad (13)$$

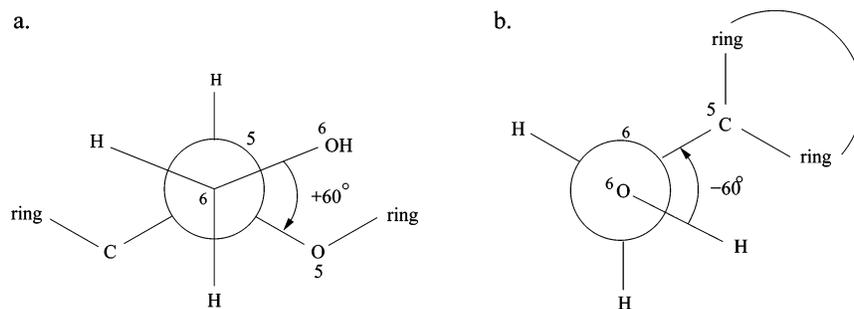
The ratio scheme makes the approximation that the normal modes do not involve any coupling between the torsions and the other internal coordinates. It is assumed that there are modes composed of torsions and modes composed of stretches and bends, but none that combine torsions with the other internal coordinates. The second scheme (which we will term the “projection method”) does not make this assumption but exactly removes the torsional modes irrespective of mixing with stretching or bending motions.

### 3. Application to $\beta$ -D-Galactose: Calculation Details

**Nomenclature.** A schematic representation of  $\beta$ -D-galactose is shown in Figure 1, with the carbon atom numbering as standard for monosaccharides and the torsional angles labeled correspondingly. Conformers are named using six letters (one for each torsion), for example, ttttG+g-. The first four letters refer to the hydroxyl torsions ( $\text{H}_n\text{-O}_n\text{-C}_n\text{-C}_{n+1}$ ,  $n = 1, 4$ ), while the final two letters refer to the two torsions of the hydroxymethyl group. The first of these two, a capital letter, denotes the  $\text{O}_6\text{-C}_6\text{-C}_5\text{-O}_5$  torsion and the second, lower case letter the  $\text{H}_6\text{-O}_6\text{-C}_6\text{-C}_5$  torsion. A “t” (or T) refers to a torsional angle of between  $120^\circ$  and  $-120^\circ$ , g+ (or G+) an angle between  $0^\circ$  and  $120^\circ$ , and g- (or G-) an angle between  $-120^\circ$  and  $0^\circ$  (see Figure 2).

It has previously been observed that the hydroxyl torsions tend to be coupled in the sense that they are more likely to all point in the same direction, for example, tttt, to form a ring of weak hydrogen bonds around the sugar,<sup>17</sup> see Table 1.

**TPIMC Calculations.** During the TPIMC simulations, all six pendant torsions of the galactose molecule (labeled in Figure 1) were active. The restricted torsions of the pyranose ring were held fixed along with the remaining non-torsional internal coordinates at the Boltzmann weighted average values for the different ab initio optimized conformer geometries.<sup>8</sup> We should note here that, while it is well-known that the bond lengths and angles around the anomeric center are coupled to the glycosidic torsion,<sup>18</sup>  $\phi_1$ , our fixed coordinate constraint will only affect the anharmonic correction term, and as we see in the results section, the assumption of separability of torsions from bends and stretches holds very well for the anharmonic corrections. The coupling is taken into full account in the harmonic part of our results, which are generated from DFT calculations. The AMBER force field<sup>19</sup> specifically parametrized for oligosaccharides by Homans<sup>20</sup> was used to describe the intramolecular potential. Simulations were performed using the parallel tempering algorithm<sup>21</sup> to ensure sampling of all of the configuration space despite possible large barriers to interconversion between competing conformers. Ensembles were run at 100, 150, 200, 300, 430 (for comparison with gas-phase experiments by Jockusch et al.<sup>11</sup>), 550, 700, 900, and 1100 K. Three types of



**Figure 2.** Newman projections depicting torsional nomenclature for the hydroxymethyl group. The “G+” (a) and “g-” (b) conformations of the hydroxymethyl group are shown.

trial moves were used:<sup>8</sup> “chain moves” involve movement of all Trotter beads by the same amount, “bead moves” allow movement of each Trotter bead by a different amount, and “swap moves” allow configurations to be swapped between ensembles at adjacent temperatures (which means high-temperature coordinates can be supplied to the low-temperature ensembles that may not otherwise be able to overcome barrier regions in the PES). All moves were accepted according to the Metropolis algorithm.<sup>21</sup> Step size parameters were updated during equilibration to ensure approximately 50% acceptance for both types of displacement moves. Swap steps were attempted for approximately 25% of the Monte Carlo steps. The simulations were run for  $10^8$  Monte Carlo steps. For the calculation of the conformer populations, conformers were defined by the coordinate ranges as described above [composed of  $120^\circ$  ranges centered at the ideal geometries for gauche +(-) and trans conformations, i.e.,  $60^\circ(-60^\circ)$  and  $180^\circ$ ]. Classical TPIMC results reported correspond to simulations with one Trotter bead, while the converged quantum results are obtained using five beads. It should be noted that both give a fully anharmonic description of the torsional modes.

The TPIMC code we developed uses several routines from the TINKER<sup>22</sup> molecular mechanics package adapted to implement the Homans force field parameters. The force field torsional Hessian was calculated numerically, and the full Hessian was calculated analytically using routines from the TINKER package.

**Density Functional Calculations.** Conformer structures were optimized with density functional calculations using the B3LYP hybrid functional,<sup>23,24</sup> whose good accuracy has been demonstrated for a large variety of organic compounds.<sup>25</sup> The 6-31+G\*<sup>26,27</sup> Gaussian basis set was used for all the B3LYP calculations we report here. In organic molecules with inter- and intramolecular hydrogen bonds, the combination of B3LYP with the 6-31+G\* basis set is known to provide structures and energetics of accuracy comparable to higher-level explicitly correlated ab initio methods with much larger basis sets.<sup>28,29</sup> Initial guesses for the structure optimizations were provided by scanning the torsional PES of galactose with the molecular mechanics force field for local minima. No symmetry constraints were applied (point group C1). Geometry optimizations were followed by harmonic vibrational frequency calculations. We verified that all the vibrational frequencies of each equilibrium structure were real and only one of the vibrational frequencies of each transition state was imaginary. All our density functional calculations were carried out with the NWChem 4.5 computational chemistry package.<sup>30</sup>

#### 4. Results and Discussion

The equilibrium geometries of the various conformers of  $\beta$ -D-galactose optimized at the B3LYP/6-31+G\* level of theory are

**TABLE 1: Equilibrium Geometries Calculated at the B3LYP/6-31+G\* Level of Theory<sup>a</sup>**

a : ttttG+g- (0.00)	b : g-tttG+g- (0.16)	c : g-tttTt (2.11)
d : ttttTt (1.83)	e : ttttTg+ (0.41)	f : g-tttTg+ (0.78)
g : g-g-g-g+G-g+ (0.49)	h : g-tttG-g- (0.72)	i : ttttG-g- (0.61)

<sup>a</sup> The relative electronic energies in kcal/mol are given in parentheses.

displayed in Table 1, along with their electronic energies. The conformers can be considered as pairs of closely related structures differing only in one hydroxyl torsion, namely, ttttXX and g-tttXX, which are close in energy. The global minimum conformer **a** corresponds to the ttttG+g- configuration, with its structural partner **b**, the g-tttG+g- conformer, lying only 0.16 kcal/mol higher in energy.

Equilibrium conformer populations calculated under the harmonic approximation from DFT electronic energies and harmonic frequencies according to eqs 5 and 6 are given in Table 2, along with the relative anharmonic correction factors from the ratio method,  $\xi_{\text{tor}}^i/\xi_{\text{tor}}^a$  as defined in eq 7, and the corresponding anharmonically corrected conformer populations. The relative anharmonic correction factors from the projection method,  $\xi_{\text{tor}}^{i,\text{Proj}}/\xi_{\text{tor}}^{a,\text{Proj}}$  as given in eq 13, and the corresponding conformer populations are also listed. The estimated errors in the corrected percentages based on those of the TPIMC simulations are, in no case, greater than 1.08%.

The rigid molecule rotational contributions to the relative conformer populations are included as in eq 5 but, in all cases, are close to unity (the largest rotational contribution relative to that for conformer **a** is 1.01 for conformer **j**, while the smallest, corresponding to conformer **k**, is 0.93). The conformer populations are, therefore, largely determined by the ratio of vibrational-electronic partition functions.

TABLE 2: Percentages of Conformers Present under Various Treatments

conformer	AIHO <sup>a</sup>	$\xi_{\text{tor}}^x / \xi_{\text{tor}}^a$ <sup>b</sup>	% <sup>c</sup>	$\xi_{\text{tor}}^{x, \text{Proj}} / \xi_{\text{tor}}^{a, \text{Proj}}$ <sup>d</sup>	% <sup>e</sup>
100 K					
ttttG+g- (a)	64.5	1.00	72.7	1.00	68.6
g-ttttG+g- (b)	26.3	0.88	26.1	1.00	27.8
g-tttTt (c)	0.01	0.32	0.00	0.12	0.00
ttttTt (d)	0.05	0.32	0.02	0.09	0.00
ttttTg+ (e)	3.0	0.28	0.9	0.98	3.1
g-tttTg+ (f)	0.5	0.19	0.1	0.74	0.4
200 K					
ttttG+g- (a)	40.0	1.00	46.1	1.00	38.0
g-ttttG+g- (b)	25.2	1.00	29.1	1.08	25.8
g-tttTt (c)	0.6	0.91	0.7	0.43	0.3
ttttTt (d)	1.6	0.92	1.7	0.34	0.5
ttttTg+ (e)	9.0	0.90	9.4	1.63	14.0
g-tttTg+ (f)	3.7	0.85	3.7	1.60	5.6
g-g-g-g+G-g+ (g)	2.9	0.56	1.9	1.10	3.1
g-tttG-g- (h)	3.1	0.52	1.9	1.00	2.9
ttttG-g- (i)	4.5	0.38	2.0	0.74	3.2
ttttG+t (j)	0.5	0.41	0.2	0.18	0.1
g+g-g-g+G-g+ (k)	6.8	0.33	2.6	0.81	5.2
tg-g-g+G-g+ (l)	1.5	0.47	0.8	0.86	1.3
430 K					
ttttG+g- (a)	20.8	1.00	21.2	1.00	18.1
g-ttttG+g- (b)	16.2	1.00	16.4	1.07	15.1
g-tttTt (c)	4.6	1.07	5.0	0.52	2.1
ttttTt (d)	8.1	1.08	8.9	0.42	3.0
ttttTg+ (e)	10.2	1.05	10.9	1.45	12.9
g-tttTg+ (f)	6.8	1.03	7.2	1.47	8.7
g-g-g-g+G-g+ (g)	4.1	0.99	4.2	1.76	6.3
g-tttG-g- (h)	4.4	0.94	4.3	1.64	6.3
ttttG-g- (i)	5.4	0.90	5.0	1.59	7.5
ttttG+t (j)	3.8	0.89	3.5	0.46	1.5
g+g-g-g+G-g+ (k)	5.4	0.90	4.9	2.00	9.4
tg-g-g+G-g+ (l)	2.9	0.84	2.5	1.48	3.8

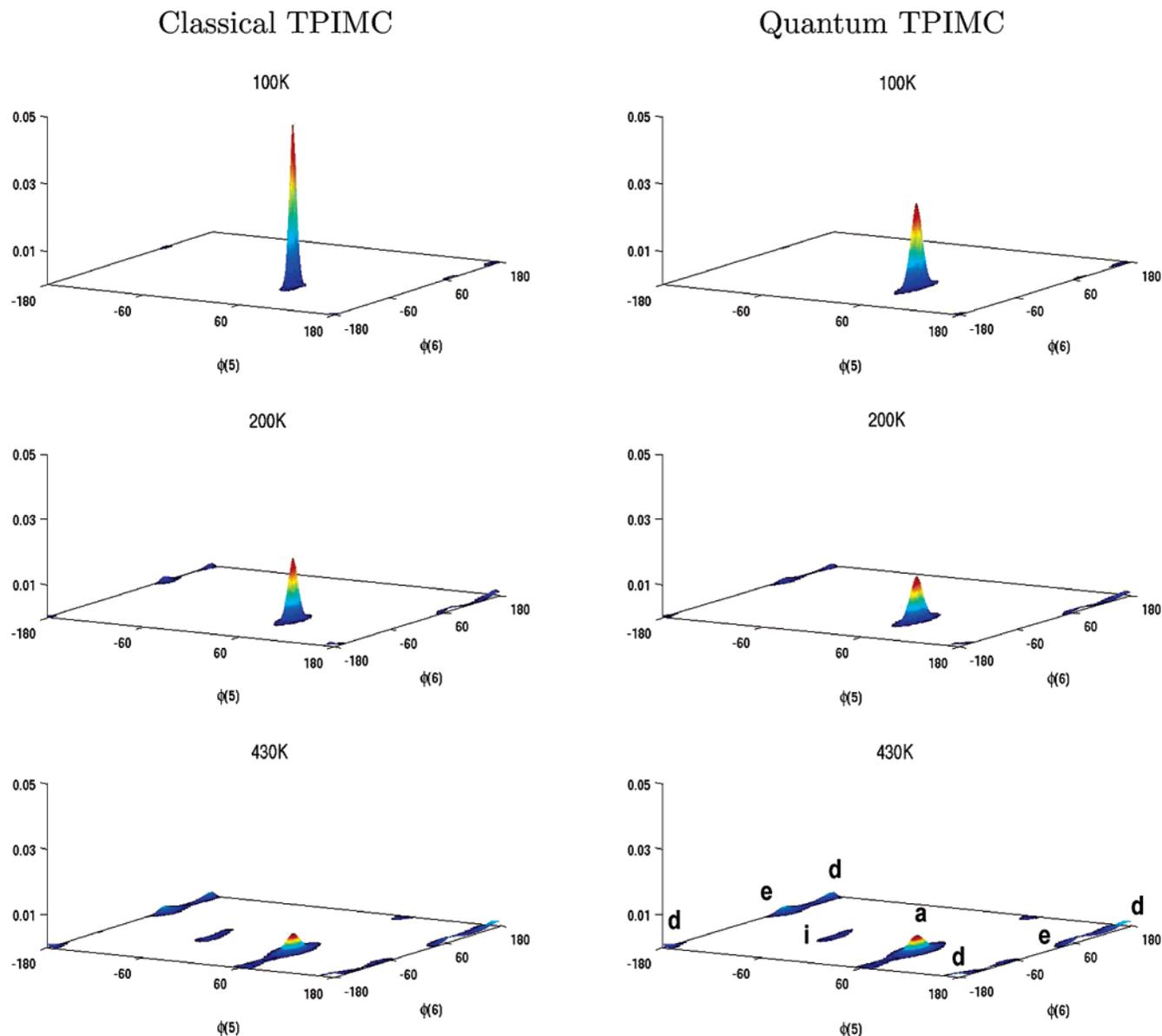
<sup>a</sup> The harmonic approximation using DFT energies and frequencies. <sup>b</sup> Anharmonic correction factors given relative to that for conformer **a** under the ratio scheme. <sup>c</sup> Anharmonically corrected populations using the ratio scheme. <sup>d</sup> Anharmonic correction factors given relative to that for conformer **a** under the projection scheme. <sup>e</sup> Anharmonically corrected populations using the projection scheme.

At 100 K, the global minimum conformer ttttG+g- (**a**) dominates the equilibrium distribution along with its closely related partner g-ttttG+g- (**b**). As the temperature is increased, other conformers become populated and the dominance of the **a** and **b** conformers is gradually diminished.

Jockusch et al. conducted UV and IR ion-dip spectroscopy experiments on phenyl  $\beta$ -D-galactoside in the gas phase at 430 K. The phenyl group in place of the hydroxyl hydrogen at C-1 (see Figure 1) acts as a chromophore enabling UV detection and is expected to have little effect on the geometry of the rest of the molecule. They found 90% of the conformer distribution to consist of the tttG+g- (**a** + **b**) conformer, with the remaining 10% assigned to either the tttTg+ (**e** + **f**) or the tttG-g- (**h** + **i**) conformers. Our anharmonically corrected values predict an equilibrium distribution of (**a** + **b**):(**e** + **f**) = 99:1 at 100 K (see Table 2, column c), but the dominance of the tttG+g- is less pronounced at 430 K with populations of (**a** + **b**):(**e** + **f**):(**a** + **b**):(**h** + **i**) = 38:18:38:9. Both the DFT harmonic frequency calculations and those corrected as a result of torsional anharmonicity underestimate the proportion of (**a** + **b**) present at 430 K compared to the experimental results. This could be due to the limitations of the DFT calculations to accurately represent the relative conformer electronic energies. Another reason for the discrepancy could be that conformer interconversion occurs during the cooling phase of the experiment. The sample is heated in an oven to provide gas-phase  $\beta$ -D-galactose molecules and, then, cooled before the spectrum is recorded. Collisions with the inert carrier gas during cooling quench the vibrational and rotational motions of the molecules. The molecules can either remain in the high-temperature equilibrium

distribution of conformers, which is the ideal scenario, providing spectroscopic data corresponding to the equilibrium distribution at the oven temperature, or interconversion from higher-energy conformers to the lower-energy ones can occur, giving a misleading view of the conformer populations.<sup>31</sup> In the latter case, the experimental data would appear to suggest a higher proportion of low-energy conformers than is actually present at equilibrium. Of course, conformer interconversion will only occur if the barriers to the torsional rotation are low enough. We have calculated the barrier heights at the B3LYP/6-31+G\* level of theory for conversion of conformers **d**, **e**, and **i** to the global minimum to be 2.18, 4.53, and 5.19 kcal/mol, respectively, suggesting that interconversion is possible (although a full dynamical study of the interconversion process<sup>31</sup> would need to be conducted to give conclusive results). The barrier to interconversion for the **d** to **a** process is particularly low and could explain the absence of signals corresponding to conformer **d** in the gas-phase spectra.

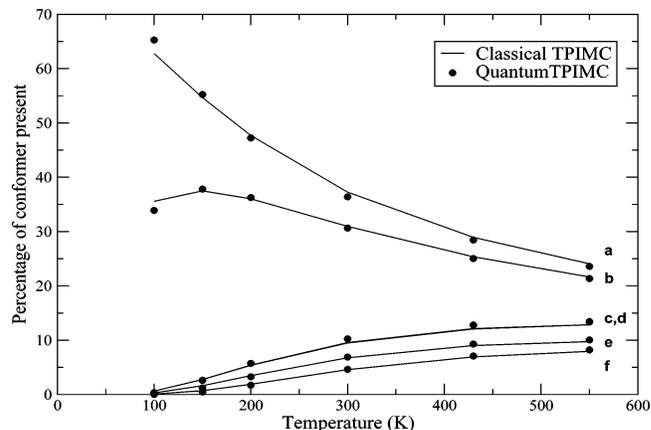
In all cases, the anharmonically corrected relative conformer populations differ little from the values based on the DFT harmonic treatment. The largest difference occurs for conformer **a** at 100 K under the ratio scheme where the population is increased by approximately 8% (see Table 2). Any anharmonicity in the torsional modes can, therefore, be said to have a negligible effect on the relative conformer populations, suggesting that, for this kind of conformational analysis on monosaccharides, in which relative conformer populations are considered, standard harmonic-frequency-based calculations are satisfactory. For some conformers, the individual anharmonic correction factor ratios, particularly at 100 K, do deviate



**Figure 3.** TPIMC torsional probability distribution plots as a function of the  $\phi(5)$  ( $\text{O}_6\text{-C}_6\text{-C}_5\text{-O}_5$ ) and  $\phi(6)$  ( $\text{H}_6\text{-O}_6\text{-C}_6\text{-C}_5$ ) torsions for the tttt hydroxyl torsion configuration.

significantly from unity (when the correction factor is equal to 1, there is no anharmonic correction and the corrected population reduces to the harmonic DFT treatment). However, in these cases, the populations of the conformers predicted by the harmonic treatment are very small and the large correction factors, therefore, have little influence on the final anharmonically corrected ratios (for example, see Table 2, conformers **c–f** at 100 K). The small effect of torsional anharmonicity on the relative conformer populations of galactose does not necessarily imply that the torsional modes themselves behave harmonically. There is cancellation of the anharmonic effects in the cases where the relative anharmonic correction factors are close to unity, as a result of the structural similarity of the conformers considered.

Figure 3 shows torsional probability distribution functions as a function of the two torsions of the hydroxymethyl group with all other torsions held in the “ttt” configuration. The left-hand column displays results from the classical TPIMC simulations ( $P = 1$ ) at various temperatures, while the right-hand column displays the corresponding quantum TPIMC plots ( $P = 5$ ). Although quantum effects in the torsional modes on the



**Figure 4.** Percentages of various conformers present as a function of temperature under both classical and quantum TPIMC treatments.

relative conformer populations are found to be negligible (see Figure 4), we can see from Figure 3 that, for a given conformer, there is tunneling into the neighboring configuration space. Peaks corresponding to the quantum TPIMC are lower and

broader, demonstrating the population of a wider range of torsional angles than for the classical TPIMC. As expected, this effect is most pronounced at the lower temperatures.

The fact that quantum effects in the torsional modes have little effect on the relative conformer populations can be understood in terms of the torsional moments of inertia and the structural similarity of most of the conformers. The torsion that prompts the biggest conformational change is that about the C–C bond of the hydroxymethyl group [ $\theta(5)$ , see Figure 1], which has a torsional moment of inertia of approximately 211620 a.u. This torsion is much too heavy for quantum mechanical behavior, such as tunneling from one conformer to another, to be observed. The hydroxyl torsions all have moments of inertia in the range 5630–5770 au. All those apart from the  $\theta(6)$  hydroxyl torsion are largely unchanged between the different conformers, and even when they are, the structures do not differ significantly. We would, therefore, expect differences in zero-point energies for different conformers to be small for the hydroxyl torsions.

Let us now turn to a comparison of the anharmonic correction terms from the two schemes for a given conformer  $i$ .

The correction from the ratio scheme,  $\xi_{\text{tor}}^i$ , is given by

$$\xi_{\text{tor}}^i = \frac{q_{\text{tor}}^{i,\text{TPI}}}{q_{\text{tor}}^{i,\text{MMHO}}} \quad (14)$$

while that from the “projection method” is

$$\xi_{\text{tor}}^{i,\text{Proj}} = \frac{q_{\text{ve}}^{i,\text{AIHOP}} q_{\text{tor}}^{i,\text{TPI}}}{q_{\text{ve}}^{i,\text{AIHO}}} \quad (15)$$

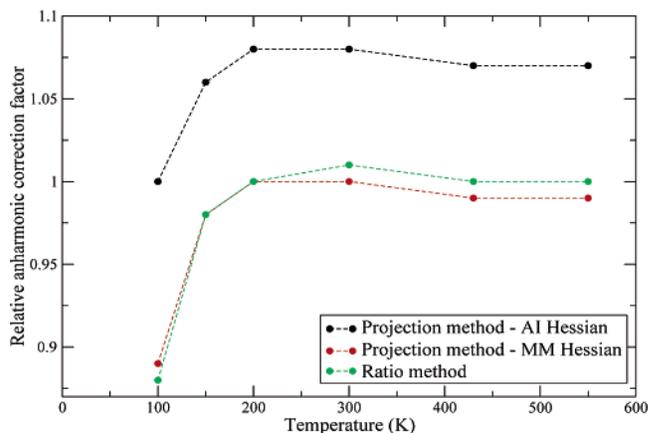
In both cases, the value of  $q_{\text{tor}}^{i,\text{TPI}}$  is the same, so the difference between the two correction factors is dependent on the difference between the molecular mechanics and the DFT PESs and on the degree to which the torsional internal coordinates are mixed with others in the normal modes. In the limit that the force field energies and frequencies are equal to those calculated using DFT and that the normal modes correspond directly to internal coordinates without mixing of the torsions with the stretches or bends, the two correction terms will be equal. Consider applying the projection method to the molecular mechanics Cartesian Hessian to give an anharmonic correction term  $\xi_{\text{tor}}^{i,\text{MMProj}}$ .

$$\xi_{\text{tor}}^{i,\text{MMProj}} = \frac{q_{\text{ve}}^{i,\text{MMHOP}} q_{\text{tor}}^{i,\text{TPI}}}{q_{\text{ve}}^{i,\text{MMHO}}} \quad (16)$$

When this value is compared to  $\xi_{\text{tor}}^i$ , the difference will depend only on the extent of mixing of the torsional internal coordinates with others in the normal modes.

Figure 5 shows the anharmonic correction terms for conformer **b** relative to conformer **a** calculated in three different ways: using the projection scheme with the Hessian from the DFT calculations, the projection scheme with the molecular mechanics Hessian, and the ratio scheme. The plots for other conformers are found to be similar and are not included.

We observe that  $\xi_{\text{tor}}^{b,\text{a,MMProj}}$  (projection method – MM Hessian) and  $\xi_{\text{tor}}^{b,\text{a}}$  almost coincide, suggesting a very low degree of mixing of the torsional internal coordinates with stretches or bends in the composition of the normal modes. This would suggest that the “separability” assumption in the conformational analysis is valid and that not only is the TPIMC Hamiltonian in eq 1 (based on the separability of the torsional



**Figure 5.** Plot of the torsional anharmonic correction factors as a function of temperature.

modes) applicable in this case but, for our treatment of  $\beta$ -D-galactose, the ratio scheme is more appropriate. The largest difference is observed between  $\xi_{\text{tor}}^{b,\text{a,Proj}}$  (projection method – AI Hessian) and  $\xi_{\text{tor}}^{b,\text{a}}$  (ratio method) and can be almost exclusively accounted for by the difference between the DFT and molecular mechanics conformer energies and torsional harmonic frequencies. Only in systems where differences between the anharmonic correction terms arising from discrepancies between the higher-level theory and the molecular mechanics energies and frequencies are expected to be smaller than those arising because of the non-separability of the torsional modes would the projection scheme be expected to yield better results than the ratio method. However, in this regime, the applicability of the TPIMC method based on the Hamiltonian in eq 1 would be questionable. At the very least, the projection scheme can be used to check the validity of the separability assumption when calculating conformer populations.

It can also be seen from Figure 5 that the relative anharmonic correction factors increase with increasing temperature (this is the case for the other conformers also). This can be understood if we consider the correction factor for a conformer  $i$  according to the ratio scheme (eq 14). This will increase with increasing temperature since the anharmonic partition function (in the numerator) will increase more rapidly with temperature than the harmonic partition function (in the denominator) due to the fact that the energy levels become closer together with increasing energy for the anharmonic case, while under the harmonic approximation, they remain equally spaced. Torsional anharmonicity is expected to be more pronounced for conformer **b** (and the other higher-energy conformers) than for the global minimum since it is usual for the lower-energy wells in a PES to be narrower and deeper than the higher-energy wells.  $\xi_{\text{tor}}^b$  will, therefore, increase more rapidly with temperature than  $\xi_{\text{tor}}^a$ , giving rise to the overall increase observed for  $\xi_{\text{tor}}^b/\xi_{\text{tor}}^a$ .

## 5. Conclusions

We have described a new scheme for the conformational analysis of flexible molecules, which combines ab initio or density functional theory electronic energies and harmonic frequencies with an anharmonic description of the torsional motions using the TPIMC method with an inexpensive PES. The recent approach of Miller and Clary is also considered. These schemes are advantageous since they require high-level calculations only to locate conformational minima and evaluate the corresponding harmonic frequencies. The anharmonic treatment of the low-frequency torsional modes is conducted on an

affordable molecular mechanics PES and is introduced to the higher-level calculations as a torsional anharmonicity correction factor that is independent of the molecular mechanics relative conformer energies.

We applied the schemes to the study of the conformational behavior of  $\beta$ -D-galactose and found that both quantum mechanical and anharmonic effects in the pendant torsional modes have little effect on the predicted conformer populations. In this case, the new scheme was found to give anharmonic correction factors similar to those of the scheme of Miller and Clary. From a discussion of the theoretical basis of the two correction schemes used, we conclude that the new scheme will be most useful in tests for the applicability of the TPIMC method to particular molecular systems. The correction factor from the new scheme can be used as a measure of the effects of non-separability of the torsional coordinates from the other modes on the predicted conformer populations.

Both schemes are expected to be useful tools in future conformational studies of flexible biomolecules, with further possible applications to the estimation of binding energies and calculation of transition state prefactors and barrier heights.

**Acknowledgment.** Y.K.S thanks the Engineering and Physical Sciences Research Council for a Ph.D. studentship, and C.-K. S. thanks the Royal Society for a University Research Fellowship. We are grateful to Christofer S. Tautermann and Thomas F. Miller for insightful discussions, and we thank Rebecca Jockusch for useful discussions regarding the galactose molecule.

## References and Notes

- (1) Matthews, C.; van Holde, K. *Biochemistry*, 2nd edition; The Benjamin/Cummings Publishing Company Inc.: San Francisco, CA, 1996.
- (2) Ravishankar, R.; Suguna, K.; Surolia, A.; Vijayan, M. *Acta Crystallogr. Sect. D* **1999**, *55*, 1375.
- (3) Åqvist, J.; Mowbray, S. L. *J. Biol. Chem.* **1995**, *270* (17), 9978.
- (4) Miller, T. F.; Clary, D. C. *J. Chem. Phys.* **2002**, *116*, 8262.
- (5) Miller, T. F.; Clary, D. C. *J. Chem. Phys.* **2003**, *119*, 68.
- (6) Miller, T. F.; Clary, D. C. *J. Phys. Chem. B* **2004**, *108*, 2484.
- (7) Miller, T. F.; Clary, D. C. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2563.
- (8) Miller, T. F.; Clary, D. C. *Mol. Phys.* **2005**, *103*, 1573.
- (9) Rahal-Sekhal, M.; Sekkal, N.; Kleb, D. C.; Bleckmann, P. J. *Comput. Chem.* **2003**, *24*, 806.
- (10) Tvaroška, I.; Taravel, F. R.; Utille, J. P.; Carver, J. P. *Carbohydr. Res.* **2002**, *337*, 353.
- (11) Jockusch, R. A.; Talbot, F. O.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1502.
- (12) Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill, Inc.: New York, 1965.
- (13) Chandler, D.; Wolynes, P. G. *J. Chem. Phys.* **1981**, *74*, 4078.
- (14) Lu, D.; Truhlar, D. G. *J. Chem. Phys.* **1993**, *99*, 2723.
- (15) Nguyen, K. A.; Jackels, C. F.; Truhlar, D. G. *J. Chem. Phys.* **1996**, *104*, 6491.
- (16) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations, The Theory of Infrared and Raman Vibrational Spectra*; Dover Publications Inc.: New York, 1980.
- (17) Carçabal, P.; Jockusch, R. A.; Huenig, I.; Snoek, L. C.; Kroemer, R. T.; Davis, B. G.; Gamblin, D. P.; Compagnon, I.; Oomens, J.; Simons, J. P. *J. Am. Chem. Soc.* **2005**, *127* (32), 11414–11425.
- (18) Tvaroška, I.; Bleha, T. *Adv. Carbohydr. Chem. Biochem.* **1989**, *47*, 45.
- (19) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- (20) Homans, S. W. *Biochemistry* **1990**, *29*, 9110.
- (21) Newman, M. E. J.; Barkema, G. T. *Monte Carlo Methods in Statistical Physics*; Clarendon Press: Oxford, 1999.
- (22) Ponder, J. W. *TINKER Software Tools for Molecular Design*, version 4.1; 2003.
- (23) Becke, A. D. *J. Chem. Phys.* **1993**, *98* (7), 5648.
- (24) Stephen, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Chem. Phys.* **1994**, *98* (45), 11623.
- (25) Baker, J.; Pulay, P. *J. Chem. Phys.* **2002**, *117* (4), 1441.
- (26) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56* (5), 2257.
- (27) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.
- (28) Tautermann, C. S.; Loferer, M. J.; Voegelé, A. F.; Liedl, K. R. *J. Chem. Phys.* **2004**, *120* (24), 11650.
- (29) Tautermann, C. S.; Voegelé, A. F.; Liedl, K. R. *J. Chem. Phys.* **2004**, *120* (2), 631.
- (30) Straatsma, T. P.; Apra, E.; Windus, T. L.; Dupuis, M.; Bylaska, E. J.; de Jong, W.; Hirata, S.; Smith, D. M. A.; Hackler, M. T.; Pollack, L.; Harrison, R. J.; Nieplocha, J.; Tipparaju, V.; Krishnan, M.; Brown, E.; Cisneros, G.; Fann, G. I.; Fruchtl, H.; Garza, J.; Hirao, K.; Kendall, R.; Nichols, J. A.; Tsemekhman, K.; Valiev, M.; Wolinski, K.; Anchell, J.; Bernholdt, D.; Borowski, P.; Clark, D.; Clerc, T.; Dachsels, H.; Deegan, M.; Dyall, K.; Elwood, D.; Glendening, E.; Gutowski, M.; Hess, A.; Jaffe, J.; Johnson, B.; Ju, J.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.; Long, X.; Meng, B.; Nakajima, T.; Niu, S.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wong, A.; Zhang, Z. *NWChem, A Computational Chemistry Package for Parallel Computers*, version 4.5; 2003.
- (31) Miller, T. M.; Meijer, A. J. H. M.; Clary, D. C. *J. Chem. Phys.* **2005**, *122*, 244323.