

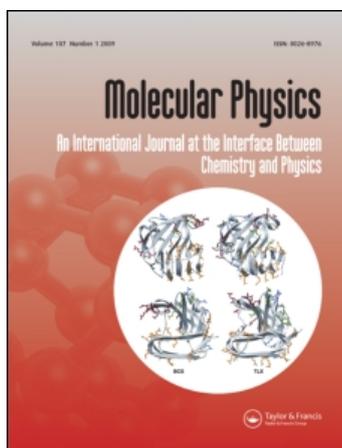
This article was downloaded by: [Naidoo, Kevin J.]

On: 2 June 2009

Access details: Access Details: [subscription number 911909794]

Publisher Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Molecular Physics

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t713395160>

### Free Energies from Adaptive Reaction Coordinate Forces (FEARCF): an application to ring puckering

Christopher B. Barnett <sup>a</sup>; Kevin J. Naidoo <sup>a</sup>

<sup>a</sup> Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa

Online Publication Date: 01 January 2009

**To cite this Article** Barnett, Christopher B. and Naidoo, Kevin J. (2009) 'Free Energies from Adaptive Reaction Coordinate Forces (FEARCF): an application to ring puckering', *Molecular Physics*, 107:8, 1243 — 1250

**To link to this Article:** DOI: 10.1080/00268970902852608

**URL:** <http://dx.doi.org/10.1080/00268970902852608>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## INVITED ARTICLE

# Free Energies from Adaptive Reaction Coordinate Forces (FEARCF): an application to ring puckering

Christopher B. Barnett and Kevin J. Naidoo\*

Department of Chemistry, University of Cape Town, Rondebosch 7701, South Africa

(Received 4 November 2008; final version received 24 February 2009)

Previously an adaptive reaction coordinate force biasing method based on probability distributions has been used to calculate the free energy of conformation, configuration and chemical reactions. This method has recently been generalised to perform calculations on multidimensional reaction coordinates. This paper presents details of this method, termed Free Energies from Adaptive Reaction Coordinate Forces (FEARCF). The efficiency of sampling of this method is demonstrated by applying it to the problem of sampling the many characteristic pucker conformations of a  $\beta$ -D-glucose pyranose using a semi-empirical PM3 Hamiltonian. The sampling ratio of the global minimum conformer ( ${}^4C_1$ ) to the highest energy conformer (a planar hexapyranose ring) is 1.7:1. Pucker free energy surfaces such as the one presented here can be a useful tool in the analysis of enzymatic reactions involving molecular ring pucker in the Michaelis complex.

**Keywords:** free energy; carbohydrate ring pucker; adaptive forces; QM/MM; FEARCF; reaction coordinate

## 1. Introduction

The pathways and manner in which molecular rings deform plays a critical role in key molecular reactions and processes. A notable case is the puckering of a carbohydrate ring when forming a glycosidic bond with another carbohydrate (in the case of oligosaccharides) or a chemical moiety such as a nucleobase (in the case of nucleosides) or degrading the already formed glycosidic linkage through hydrolyses. Glycosyl transferases catalyse the synthesis of glycosidic bonds to form oligosaccharides [1] while glycosidases hydrolyse glycosidic bonds [2, 3]. Both enzyme classes whether promoting the formation of the glycosidic bond or degrading of it have a strong affinity for the reducing carbohydrate to take on high energy conformations [4]. An oxocarbenium ion forms in the Michaelis complex accompanied by the distortion of the saccharide ring. The nature of the distortion is complex and poorly understood. This is mainly because enzyme substrates often contain saccharide rings which exhibit several forms of puckering in the transition state [5, 6]. The puckering conformation of saccharides in the transition state of various glycosidases results in selective reaction pathways and specific reaction products. In the case of pyranose rings competing puckerers have been observed for several members of the glycoside hydrolase family [5, 7]. The puckering conformers observed in X-ray experiments

and proposed from computational analysis include  ${}^1S_5$ ,  $B_{3,0}$ ,  ${}^0S_2$  [7, 8].

The analysis and categorisation of the puckering of molecular rings is marked by the seminal work of Cremer and Pople [9]. Since then several researchers [7, 8, 10] have shown interest in the calculation and designation of puckering for monosaccharides. Here we report an adaptive force based method for calculating the potential of mean force of ring pucker. We reported the calculation of the potential of mean force (PMF) for glycosidic linkage rotation [11] and chemical reactions [12] using this approach which we now term Free Energies from Adaptive Reaction Coordinate Forces (FEARCF). We will detail the FEARCF method here in the context of ring puckering and then calculate the three-dimensional PMF for the  $\beta$ -D-glucose pyranose ring. Although we will present the minima pucker conformations in this paper we will not discuss the stereoelectronic rationale for their existence as this will be reported elsewhere [13].

## 2. Method and theory

The force biasing method presented here is a generalisation of the method used to produce two dimensional conformational [11,14] and reaction [12] free energy surfaces. The concept of adjusting the biasing potential

\*Corresponding author. Email: Kevin.Naidoo@uct.ac.za

iteratively to evolve the free energy was first presented by Mezei as adaptive umbrella sampling [15]. Other methods have subsequently employed this concept and extended it using either forces to directly bias the system or potential biasing. A popular adaptive umbrella potential method developed by Laio and Parrinello uses a combination of Gaussians as a biasing potential from which coarse-grained forces are calculated to direct the system to regions that have not been previously sampled [16]. Of particular relevance here is Darve and Pohorille's adaptive biasing force (ABF) routine for thermodynamic integration that continuously updates the biasing forces at every step during the simulation [17]. The biasing force is estimated locally from sampled conformations of the system gained from previous steps in the simulation. The bias at a specified point in the reaction coordinate is applied as soon as the bin for that point has *enough samples* to estimate the mean force. These forces are then applied along generalised system coordinates. The result is that the instantaneous force acting on the coordinates of interest when subtracted from the equations of motion results in zero acceleration along the reaction coordinate. The choice of sample size is recommended to be small ( $\sim 100$ ) and the sampling is then improved by applying a biasing force of the form  $\lambda \nabla \xi$  where  $\lambda$  is a Lagrange multiplier that can be related to the derivative of the free energy  $dA/d\xi$ .

The *FEARCF* method is based on probability distributions and histograms. The reaction coordinate space is a discretised  $n$ -dimensional grid where the sampling frequency for a bin site is recorded for each simulation. In the case of ring puckering the dimensionality of the grid is  $N - 3$  where  $N$  = number of ring atoms. The population of this grid represents a running tally of the reaction coordinate probability density, derived from the history of simulations to that point. It is used as input for a multidimensional cubic-spline interpolation from which the reaction coordinate biasing forces are calculated. These forces are applied to all atoms used in the reaction coordinate definition to bias the next simulation's reaction coordinate trajectory away from previously sampled areas. The entire reaction coordinate space is equally sampled when the biasing forces are derived from the true PMF. The method requires no intervention from the user other than a physical understanding of the complexity of the computer experiment, which is used to make a judicious choice of reaction coordinate, and simulation length at each update of the biasing force. A combination of relatively short simulation times at the start of *FEARCF* experiment increasing to longer

times (three or more times as long) as convergence is reached is recommended.

### 2.1. Hill–Reilly pucker coordinates

Hill and Reilly proposed an intuitive description of ring puckering [18]. For an  $N$  membered ring they reduced the Cartesian coordinates to  $N - 3$  ring pucker coordinates. Similar to the classic Cremer–Pople definition a puckering conformation is described as the combination of deviations of each coordinate from a mean plane. The Cremer–Pople pucker coordinates (amplitude–phase pairs  $(q_i, \phi_i)$  and a single puckering coordinate  $q_{N/2}$  for even numbered rings) are usually translated to a spherical polar set to describe the pucker conformation in the case of six membered rings. This set consists of a pucker amplitude ( $Q$ ), a phase angle ( $\phi, 0 \leq \phi < 2\pi$ ) which in combination with another angle ( $\theta, 0 \leq \theta \leq \pi$ ), where  $q_2 = Q \sin \theta$  and  $q_3 = Q \cos \theta$ , describe the chair, twist and boat conformations.

The Hill–Reilly definitions are derived from triangular decomposition of an  $N$  membered ring, which is easily visualised in terms of deviations of the atoms that comprise the ring, from mean planes. A monocyclic ring is decomposed into rotatable planes and a reference plane. The pucker coordinates are specified by the angle the rotatable planes make with respect to the reference plane. In this paper we consider only six membered rings and so we illustrate the Hill–Reilly definitions in Figure 1. For  $N=6$  there are three puckering coordinates  $\theta_0, \theta_1, \theta_2 \in [-\pi, \pi]$  where a positive angle indicates rotation of that ring plane above the reference plane. The angle of puckering is calculated from

$$\theta_i = \pi/2 - \cos^{-1}[(q_i \cdot n) \cdot (||q_i|| \cdot ||n||)^{-1}] \quad (1)$$

where  $q_i$  is a vector normal to the rotatable plane  $i$  and  $n$  is the vector normal to the reference plane. We define a three dimensional reaction coordinate ( $\xi$ ) as the combination of these pucker coordinates, i.e.  $\xi = (\theta_0, \theta_1, \theta_2)$ .

### 2.2. Free Energies from Adaptive Reaction Coordinate Forces (FEARCF)

The potential of mean force,  $W(\xi)$  is calculated as a function of the three-dimensional coordinate set  $\xi$  and is related to the probability density,  $P(\xi)$  by

$$W(\xi) = -k_B T \ln P(\xi) \quad (2)$$

where  $k_B$  is the Boltzmann constant and  $T$  is the temperature in kelvin. The PMF is derived in a

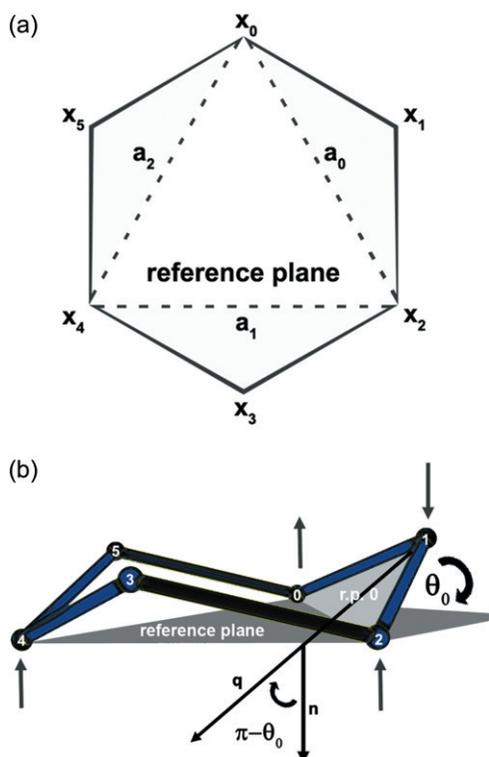


Figure 1. Definition for the Hill-Reilly ring puckering coordinates. (a) Illustrates an overhead view of a 6-membered ring with  $x$ , the ring atoms and  $a$ , the puckering axes. (b) Description of pucker angle relative to the reference plane and rotatable plane 0 (r.p.0). A simple example of a negative force (to bend the flap towards the reference plane) applied to theta and perpendicular to the reference plane will apply forces to atoms 0, 1, 2 and 4 as indicated.

canonical ensemble (constant NVT), which gives the Helmholtz free energy puckering surface. A correct analysis of free energy pathways is only possible if the reaction coordinate is sampled uniformly. To approach uniform sampling we used a variant of Mezei's adaptive umbrella sampling [15] which we applied to a two dimensional dihedral angle PMF earlier [11] and more completely to a four-dimensional intermolecular association hypersurface [19].

To traverse barrier heights larger than  $3k_B T$  a biasing potential,  $U(\xi)$  is applied. The best choice for  $U(\xi)$  would be

$$U(\xi) = k_B T \ln P(\xi). \quad (3)$$

The objective is to bias the new trajectory away from reaction coordinate space that has been sampled in previous simulations towards reaction coordinates that have not been sampled. The concept of iteratively adapting an umbrella potential by using the information collected from previous simulations to uniformly sample a single degree of freedom was proposed by

Mezei [15]. In this procedure the umbrella potential  $U(\xi)$  is adapted from the first guess until it is equal to the inverse of the PMF. At the final stage of the adaptive umbrella sampling when  $U(\xi)$  is applied to the system the entire configurational space (high and low energy regions) will be sampled [15]. The method has been extended to multidimensional conformational space for proteins [20].

We modified the macromolecular program CHARMM [21] to calculate the effect of the perturbing forces generated from the surface of the adapting umbrella potential for three independent reaction coordinates [11]. We have generalised the method to multi-dimensions [19] and described it here for the three-dimensional pucker free energy calculation. At each step of the simulation the biasing force for  $\theta_i$  applied to the atoms involved in the rotatable plane is calculated from the gradient of the umbrella potential for that puckering coordinate. For three-dimensional surfaces these are

$$-\frac{\partial U(\xi)}{\partial \theta_i} = F(\theta_i) \quad (4)$$

where the force  $F$  is calculated as the partial derivative  $\theta_i$  of the applied biasing reaction coordinate umbrella potential. The result is the  $i$ th ring plane is biased away from the areas it has sampled with a force opposite to the accumulated sampling density as shown in Equation (3). When implemented in a molecular dynamics (MD) routine the biasing forces for each puckering coordinate is added to the forces from the equilibrium Boltzmann dynamic forces applied to each atom. The biasing forces on each ring atom can be recovered from the reaction coordinate forces by recasting them in terms of the PMF:

$$\frac{\partial W(\xi)}{\partial \theta_i} = F_i \quad (5)$$

which can be rewritten in terms of the vector  $q_i$  using the chain rule:

$$\frac{\partial W(\xi)}{\partial \theta_i} \frac{\partial \theta_i}{\partial q_i} = \frac{\partial W(\xi)}{\partial q_i}. \quad (6)$$

Similarly the forces can be written in terms of the vector  $n$ ,

$$\frac{\partial W(\xi)}{\partial \theta_i} \frac{\partial \theta_i}{\partial n} = \frac{\partial W(\xi)}{\partial n}. \quad (7)$$

When these are combined (*and including all other contributions*), the Cartesian biasing forces  $(\partial W(\xi)/\partial x_i)$  on the ring atoms  $x_i$  originating from the pucker axis  $a_i$  (see Figure 1(a)) are calculated as

$$\frac{\partial W(\xi)}{\partial a_i} \frac{\partial a_i}{\partial x_{2(i+1)}} = \frac{\partial W(\xi)}{\partial x_{2(i+1)}}; \quad \frac{\partial W(\xi)}{\partial a_i} \frac{\partial a_i}{\partial x_{2i}} = \frac{\partial W(\xi)}{\partial x_{2i}} \quad (8)$$

which are due to the reaction coordinate  $\theta_i$ . The total biasing force is calculated by summing the contributions from each reaction coordinate which is then applied to the ring atoms. This can be calculated for each of the ring atoms  $i$  where  $i$  are the labels (1–6) for the ring atoms (C1,C2,C3,C4,C5,O5).

At the start of the *FEARCF* simulations  $W(\xi)$  and likewise  $P(\xi)$  are not known. To initialise the simulations a zero biasing potential is applied ( $U(\xi) = 0$ ). The resulting probability distribution is used as a first guess for  $W(\xi)$  from which an improved biasing potential  $U(\xi)$  for the next simulation is obtained. This process continues iteratively until the pucker conformational space is adequately sampled.

The *FEARCF* method is similar to the ABF method in that it uses external forces derived from the derivative of the free energy to achieve unbiased sampling of the reaction coordinate. The ABF method however, calculates the adaptive forces on the fly at each bin and uses a ramp function to reduce errors emanating from the incomplete sampling of bins. The ABF method therefore could converge faster in cases where the statistical errors arising from sampling not being optimal for all bins. We have however, not directly compared the two methods to make any definitive statements on this. The *FEARCF* method calculates new biasing forces over all bins from the histogram  $P(\xi)$ , which is constructed after each iteration together with all previous iterations using the WHAM procedure. This is done numerically using a multidimensional cubic-spline routine. The major advantage of the *FEARCF* method is therefore that it is easily extended to multiple dimensions as we show here while in the ABF method it is not straightforward to reconstruct  $A(\xi)$  from its derivatives.

### 3. Simulations details for $\beta$ -D-glucose

We modelled the glucose molecule with the PM3CARB-1 [22] semi-empirical potential and conducted *in vacuo* dynamics in CHARMM 33b2 [21]. Each simulation was of 0.2 ns in length (first 17 iterations and from then on 0.6 ns), at a temperature of 298.15K. Velocity-verlet dynamics were implemented with the Nosé temperature thermostat. The potential of mean force was calculated using an iterative procedure. To further increase the reaction coordinate sampling, batches of eight simulations starting from various pucker conformations (initially all  ${}^4C_1$  and then the last conformer sampled from the previous run) were used with an initial biasing potential of zero. The sampled areas were collated to acquire  $P(\xi)$  from which the first guess for  $W(\xi)$  was

calculated and the associated  $U(\xi)$  applied to the next batch of eight simulations. From this the biasing forces were derived from the reaction coordinate surface as described above to improve the sampling. This iterative procedure was completed 24 times totaling a sampling of 51.2 ns ( $17 \times 0.2 \times 8 + 0.6 \times 5 \times 8$ ).

Summing and weighting the histograms of all the biased simulations appropriately further improved the sampling rate of the reaction coordinate space. This was done using the Weighted Histogram Analysis Method (WHAM) [23]. The WHAM equations were applied iteratively as shown previously [24]; until the maximum difference (tolerance) between the previous and current iteration weighting coefficients ( $|f_j^i - f_j^{i-1}|$ ) was less than 0.001.

### 4. Free energy of $\beta$ -D-glucose ring pucker

A pyranose carbohydrate ring has 38 characteristic puckering conformers that are strainless [25]. These can be reduced to simpler unique classifications such as chairs (C), half chairs (H), boats (B), skews (S), and envelopes (E), which are illustrated in Figure 2. The individual conformers are named with a superscript indicating a ring atom positioned above a plane formed by the remaining atoms while a subscript indicates a position below this plane. It should be noted that this relative plane is not the same as the one we use in the Hill–Reilly description (Figure 1). It is not surprising that the three dimensional free energy surface,  $W(\xi)$ , which we calculated using the *FEARCF* method for  $\beta$ -D-glucose is complex (Figure 3). The reaction coordinate axes ( $\theta_0, \theta_1, \theta_2$ ) are orthogonal to each other and the fourth dimension, free energy (FE) is drawn in colour. In this figure, blue represents low energy conformers that are close in energy to the global minimum and red represents conformations that are at

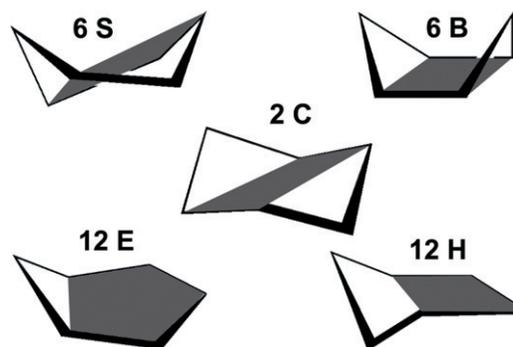


Figure 2. The 38 canonical conformers available to a six-membered ring: Two chairs (C), six boats (B), six twists (S), twelve half-chairs (H) and twelve envelopes (E). The shaded area connects atoms in the same plane.

least  $30 \text{ kcal mol}^{-1}$  greater in energy than the global minimum conformation. To enhance the viewing of the conformers relative to each other we project two contoured slices  $W(\theta_0, \theta_1)$  and  $W(\theta_1, \theta_2)$  across the FE volume. Two perspectives are given, with Figure 3(a) displaying the chair conformations while Figure 3(b) shows the  $B_{1,4}$  conformation on the opposite side of the FE volume.

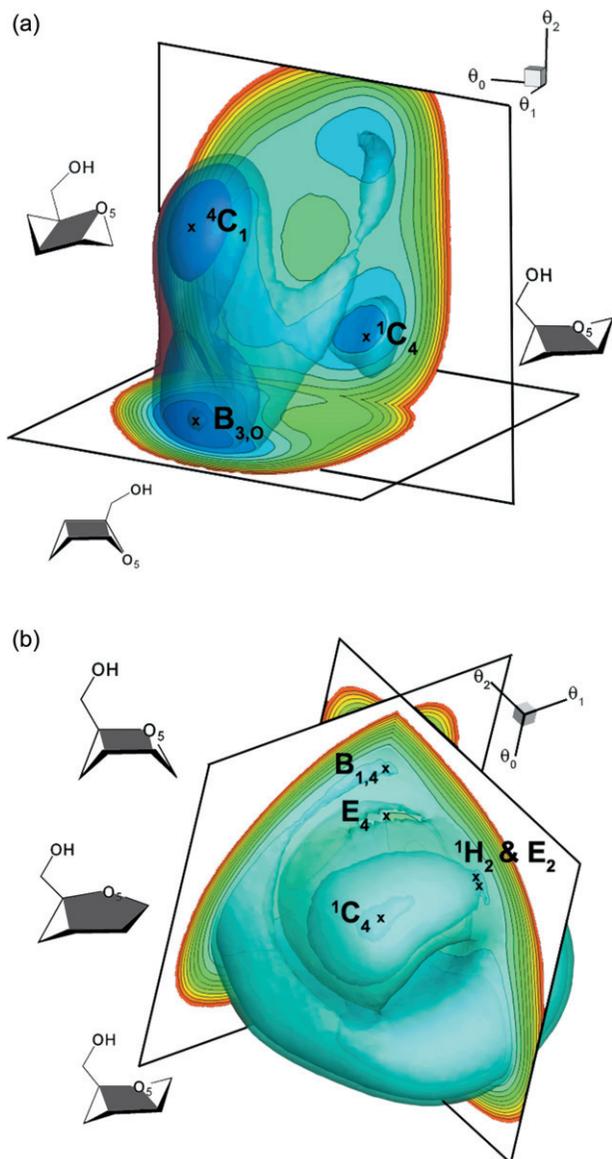


Figure 3. (a) The three-dimensional free energy surface for glucose calculated using the Hill–Reilly reduced coordinate set. The transition path from  ${}^1C_4$  through  $B_{3,0}$  until  ${}^4C_1$  is indicated. (b) Another viewpoint of the three-dimensional free energy surface for glucose calculated using the Hill–Reilly reduced coordinate set. The transition path from  $B_{1,4}$  through  $E_4$  until  ${}^1C_4$  is indicated. Another pathway is also possible; the transition conformers for this ( $E_2$ ,  ${}^1H_2$ ) are indicated.

All the stable low energy conformers have been identified and are shown in Table 1. First we consider the pathways between the two chair conformations. The optimal path goes via a minima boat conformation (Figure 3(a)). A significant advantage of the four-dimensional free energy volume is the opportunity to accurately locate minima free energy pathways that connect the various low energy conformations. We extract the values for conformations linking the two chairs where the  ${}^4C_1$  conformer is the lowest in energy ( $0.00 \text{ kcal mol}^{-1}$ ) followed by the  $B_{3,0}$  ( $2.00 \text{ kcal mol}^{-1}$ ) and  ${}^1C_4$  ( $4.86 \text{ kcal mol}^{-1}$ ) pucker conformations and plot this in a single conflated dimension (Figure 4(a)). In the context of glycosyl transferases [1] and glycosyl hydrolases [3, 26] reaction mechanisms this process of identifying pucker transitions from the stable  ${}^4C_1$  conformer to the thermodynamically less favoured E and H puckers surface can be an important tool in the analysis of catalytic reaction mechanisms. A conformational change in the ring of the  $\beta$ -D-pyranoside substrate is required for nucleophilic assistance of the cleavage according to an intermediate  $S_N1/S_N2$  mechanism. Heterolytic cleavage of the acetal C–O bond requires an antiperiplanar orientation of a doubly occupied, non-bonding orbital. This is referred to as the antiperiplanar lone pair hypothesis (ALPH) and implies that hydrolysis of  $\beta$ -D-pyranoside requires a conformational change of the tetrahydropyran ring from a chair to a twist-boat or boat resulting in a pseudoaxial orientation of the aglycon [26, 27].

Table 1. Conformers sampled and free energies for  $\beta$ -D-glucose.

Conformer	Reaction coordinates			Free energy ( $\text{kcal mol}^{-1}$ )
	$\theta_0$	$\theta_1$	$\theta_2$	
${}^4C_1$	30.0	32.5	35.0	0.00
$E_3$	32.5	40.0	-12.5	4.38
$B_{3,0}$	25.0	40.0	-52.5	2.00
${}^5E$	25.0	-37.5	-30.0	7.93
${}^1H_2$	-45.0	15.0	-7.5	8.32
$E_2$	-47.5	5.0	10.0	8.30
${}^1C_4$	-30.0	-22.5	-22.5	4.86
$E_4$	5.0	-47.5	12.5	9.49
$B_{1,4}$	22.5	-50.0	40.0	7.91
$E_1$	30.0	-47.5	42.5	8.05
$E_O$	-2.5	7.5	-42.5	11.42
Planar	2.5	-2.5	7.5	16.97
${}^0E$	-2.5	-10.0	50.0	10.83
${}^1S_5$	-42.5	50.0	5.0	5.64
${}^0S_2$	-45.0	0.0	57.5	5.48
${}^3,0B$	-27.5	-22.5	65.0	6.05
${}^3E$	-47.5	27.5	52.5	7.39

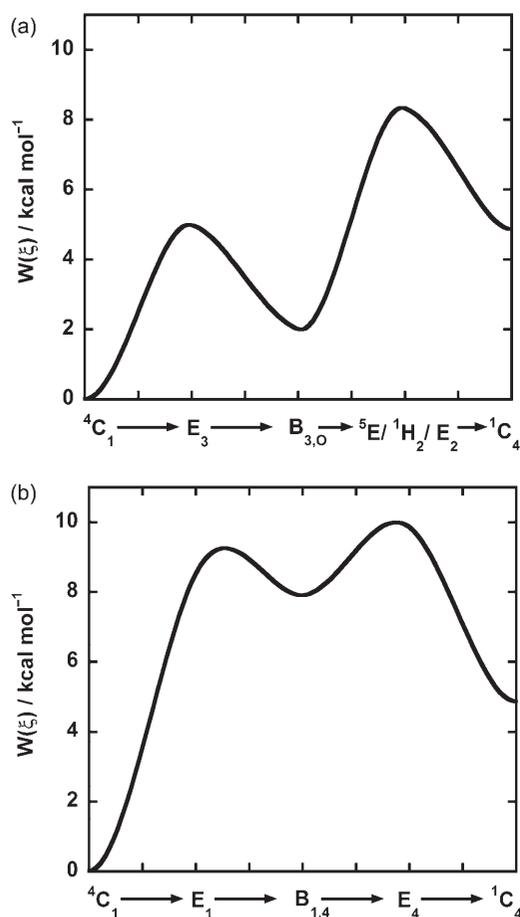


Figure 4. (a) A one dimensional curve generated by selecting points from the three-dimensional energy path shows the free energies required to traverse from the  ${}^1C_4$  to the  ${}^1C_4$  conformer.  $5E/{}^1H_2/E_2$  indicates that there are three traversable routes from the  $B_{3,0}$  to the  ${}^1C_4$  conformer with almost the same energy barrier (although different paths). (b) Another possible pathway between the  ${}^1C_4$  and  ${}^1C_4$  conformers. This path differs from (a) as the traversal between  ${}^1C_4$  and  ${}^1C_4$  is via the  $B_{1,4}$  not  $B_{3,0}$  conformer.

The ALPH is supported by the crystal structures of three *endo*-glycosidases in complex with substrate analogues showing a skew boat or a flattened boat conformation of the tetrahydropyran ring [28].

Atomic force microscopy (AFM) experiments were performed on an amylose strand and the yielding behavior has been attributed to transitions of the  $\alpha$ -linked glucose rings from the  ${}^4C_1$  chair to a twist-boat conformation [29]. Previously using a classical carbohydrate force field potential [30] we calculated AFM profiles for amylose. In contrast to the earlier proposal our classical AFM simulations revealed that the unfolding mechanism of amylose was principally due to the rotation of glycosidic bonds and not the puckering of individual glucose units [31].

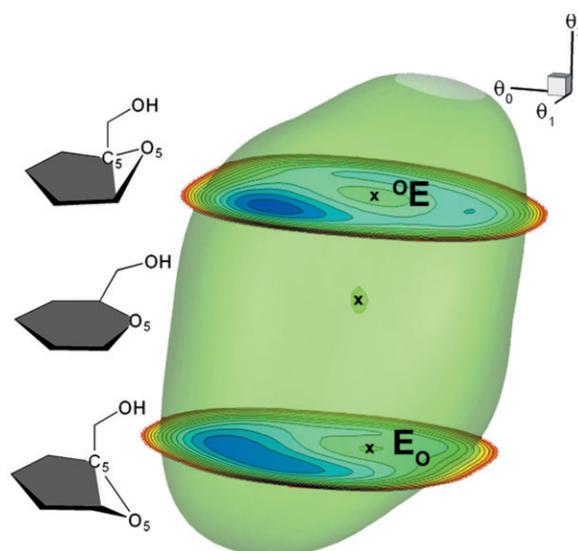


Figure 5. A viewpoint of the three-dimensional free energy surface for glucose indicating the high in energy planar conformation, which is sampled. A transition from  $O_E$  to  $E_O$  can occur via this high energy planar conformer.

The  $B_{1,4}$  conformation which we observe here from the semi-empirical FE pucker volume resides in a shallow equilibrium well that is  $7.91 \text{ kcal mol}^{-1}$  above the  ${}^4C_1$  chair conformation and separated from it by a barrier height of  $9.49 \text{ kcal mol}^{-1}$  (Figure 4(b)). The high barriers separating the boat from the chair conformations supports our previous interpretation of the AFM experiment. That is, that the pyranose transition to boat conformers early on in the amylose pulling experiment is not possible.

The accuracy of the free energy volume is validated when all states in the reaction coordinate space are adequately sampled. Adequate sampling is generally accepted when a ratio of most sampled to least sampled states is at least 1:50 [20, 32]. However, using the adaptive reaction coordinate forces to bias the simulation away from equilibrium favoured sites, sampling ratios of 1:5 have been calculated [11, 14, 24, 30, 31]. When we contour the high energy pucker conformations (Figure 5), it is apparent that these bound the reaction coordinate volume, with a singular high energy conformer in the centre. Upon inspection of the high energy conformers on the periphery, we note that they are not of chemical interest, as they represent conformers where the pyranose ring is 'wrung' out or 'folded' up. The single conformer in the centre of the volume is planar with the closest equilibrium conformers being  $O_E$  above it and  $E_O$  below it.

The extent of conformational space is shown from the first unbiased simulation through to the 24th

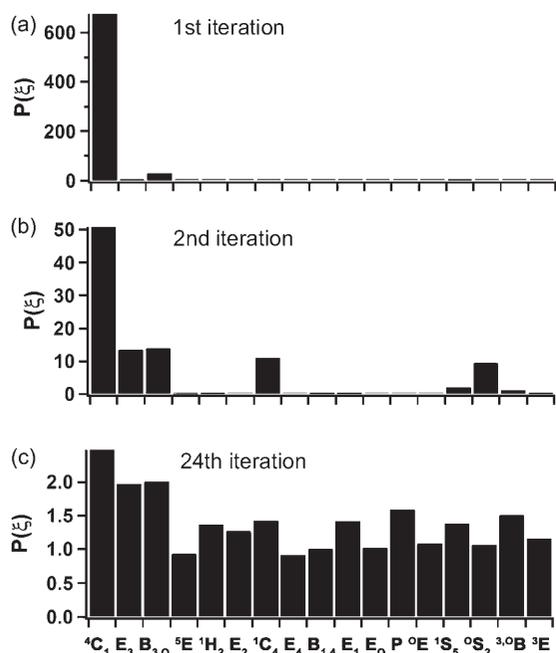


Figure 6. Histograms showing the conformer sampling distributions ( $P(\xi)$  on the scale  $10^{-4}$ ) taken from (a) the first, (b) the second and (c) the 24th PMF iterative calculation. Both low energy and high energy conformers are shown to illustrate the improvement in sampling from the first until to the final iteration.

iteration (Figure 6). The  ${}^4C_1$  conformer is mostly sampled when there are no biasing forces applied with very few transitions to the  $E_3$ ,  $B_{3,O}$ , and  ${}^1S_5$  conformations. After application of the biasing forces, calculated from a first guess at the PMF, the ratio of sampling of  $E_3$ ,  $B_{3,O}$ , and  ${}^1S_5$  improve considerably. More of  $\xi$  space is explored with conformers  ${}^3O_2$ ,  ${}^3OB$  and  ${}^1C_4$  now evident. After 24 iterations we achieve a ratio of 1:1.7 where the least sampled planar conformer is nearly equal to  ${}^4C_1$ , the most stable conformer. This far exceeds the 1:50 ratio used in other studies as a measure of convergence. The high energy planar and  $E^4$  conformers have similar sampling.

Sampling ratios for several conformers relative to the lowest energy conformer ( ${}^4C_1$ ) are even more favourable (Table 2). We compare the free energies for several conformers calculated here with the FEARCF with those reported using a metadynamics routine [7]. The energy differences between the favoured  ${}^4C_1$  conformer are similar to that found by Parrinello *et al.* using their metadynamics routine. However, in the latter case the free energies of the less favoured conformers are slightly higher than the FEARCF free energies. This difference may be due to either the difference in theory used (i.e. density functional theory

Table 2. Conformer sampling ratios and comparison between FEARCF and metadynamics energies for selected conformers.

Conformer	FEARCF		Metadynamics Free energy (kcal mol <sup>-1</sup> )
	Relative sampling	Free energy (kcal mol <sup>-1</sup> )	
${}^4C_1$	1.00	0.00	0
$B_{3,O}$	0.81	2.00	2.6
${}^1C_4$	0.58	4.86	–
$B_{1,4}$	0.41	7.90	7.9
${}^1S_5$	0.56	5.64	5.8
${}^3O_2$	0.61	6.05	7.2

Note: Sampling ratios relative to the lowest energy conformer –  ${}^4C_1$  and averaged over the 24th batch).

in the case of metadynamics simulations compared with the carbohydrate specific PM3 semi-empirical quantum description employed here) or it may be due to the extent of reaction coordinate sampling (sampling ratios were not reported in the metadynamics study).

## 5. Conclusions

Details of the adaptive reaction coordinate force method, which we have previously developed and used to derive specific conformational, configurational and reaction free energy surfaces, have been presented here. Specifically, we generalised the method for free energies from adaptive reaction coordinate forces method (FEARCF) and implemented it in CHARMM showing it possible to converge to the full free energy volume. While any PMF can be calculated using this method its use is particularly significant for the calculation of multidimensional free energy surfaces. To illustrate the accuracy and versatility of the method we apply it to a demanding problem of exploring the hexapyranose ring pucker conformational free energy space using a semi-empirical PM3 quantum Hamiltonian. Excellent sampling of the conformational space is achieved demonstrating how high energy conformers are accessible using this adaptive reaction coordinate force to bias the trajectory about the reaction coordinate multidimensional volume. Although the FEARCF application presented here is *in vacuo* using a PM3-CARB1 semi-empirical level of theory it is easily applied to condensed phase explicit solvent simulations using either classical or higher level quantum theory. Furthermore, while this implementation has been done in CHARMM the method could be incorporated into any molecular dynamics algorithm.

There are several pathways from the global minima  ${}^4C_1$  chair to  ${}^1C_4$  chair conformation. Two possible pathways go via  $B_{3,0}$  or  $B_{1,4}$  boat conformations with half chair and envelope conformational barriers reaching more than  $10 \text{ kcal mol}^{-1}$ . This illustrates that necessary high energy glucopyranose ring conformations observed in glycosylase pockets requires the enzyme to induce and preserve high energy half chair and envelope conformations required for successful hydrolyses and synthesis of the glycosidic bond.

### Acknowledgements

This work is based upon research supported by the South African Research Chairs Initiative (SARChI) of the Department of Science and Technology and the National Research Foundation (NRF) awarded to K.J.N. CB would like to thank fellow group member Johan Strümpfer for helpful discussions and the National Research Foundation (Pretoria) for PhD fellowship support. Kevin J. Naidoo is pleased to present this work in honour of Henry Fritz Schaefer's 65th birthday.

### References

- [1] L.L. Lairson, B. Henrissat, G.J. Davies, and S.G. Withers, *Ann. Rev. Biochem.* **77**, 521 (2008).
- [2] M.L. Sinnott, *Chem. Rev.* **90**, 1171 (1990).
- [3] A. Vasella, G. J. Davies, and M. Böhm, *Curr. Opin. Chem. Biol.* **6**, 619 (2002).
- [4] V.L. Schramm, *Ann. Rev. Biochem.* **67**, 693 (1998).
- [5] S. Fushinobu, B. Mertz, A. D. Hill, M. Hidaka, M. Kitaokac, and P.J. Reilly, *Carbohydr. Res.* **343**, 1023 (2008).
- [6] K.S. Sidhu, J.M. Goodfellow, and J.Z. Turner, *J. Chem. Phys.* **110**, 7943 (1999).
- [7] X. Biarnes, A. Ardevol, A. Planas, A. Rovira, A. Laio, and M. Parrinello, *J. Am. Chem. Soc.* **129**, 10686 (2007).
- [8] A.R. Ionescu, A. Berces, M.Z. Zgierski, D.M. Whitfield, and T. Nukada, *J. Phys. Chem. A* **109**, 8096 (2005).
- [9] D. Cremer and J.A. Pople, *J. Am. Chem. Soc.* **96** (6), 1354 (1975).
- [10] A. Berces, G. Enright, T. Nukada, and D.M. Whitfield, *J. Am. Chem. Soc.* **123**, 5460 (2001).
- [11] K.J. Naidoo and J.W. Brady, *J. Am. Chem. Soc.* **121** (10), 2244 (1999).
- [12] R. Rajamani, K.J. Naidoo, and J. Gao, *J. Comput. Chem.* **24**, 1775 (2003).
- [13] C.B. Barnett and K.J. Naidoo, unpublished (2009).
- [14] M.M. Kuttel and K.J. Naidoo, *J. Phys. Chem. B* **109**, 7468 (2005).
- [15] M. Mezei, *J. Comput. Phys.* **68**, 237 (1987).
- [16] A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci.* **99** (20), 12562 (2002).
- [17] E. Darve and A. Pohorille, *J. Chem. Phys.* **115** (20), 9169 (2001); E. Darve, D. Rodríguez-Gómez, and A. Pohorille, *J. Chem. Phys.* **128** (144120(1–13)) (2008).
- [18] A.D. Hill and P.J. Reilly, *J. Chem. Inf. Model* **47** (3), 1031 (2007).
- [19] J. Strumpfer and K.J. Naidoo, *J. Comput. Chem.* (2008), in review.
- [20] C. Bartels and M. Karplus, *J. Comput. Chem.* **18** (12), 1450 (1997).
- [21] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.* **4** (2), 187 (1983).
- [22] J.P. McNamara, A. Muslim, H. Abdel-Aal, H. Wang, M. Mohr, I.H. Hillier, and R.A. Bryce, *Chem. Phys. Lett.* **394**, 429 (2004).
- [23] S. Kumar, D. Bouzida, R.H. Swendsen, P.A. Kollman and J.M. Rosenberg, *J. Comput. Chem.* **13**, 1011 (1992); S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, *J. Comput. Chem.* **16** (11) 1339 (1995); S. Kumar, P.W. Payne and M. Vásquez, *J. Comput. Chem.* **17** (10) 1269 (1996).
- [24] C.B. Barnett and K.J. Naidoo, *J. Phys. Chem. B* **112**, 15450 (2008).
- [25] R.S. Shallenberger, *Advanced Sugar Chemistry* (Avi Publishing Company Incorporated, Westport, Connecticut, 1982).
- [26] D.L. Zechel and S.G. Withers, *Acc. Chem. Res.* **33** (1), 11 (2000).
- [27] A. White and D.R. Rose, *Curr. Opin. Struct. Biol.* **7** 645 (1997); E. Lorthiois, M. Meyyappan and A. Vasella, *Chem. Commun.* **19** 1829 (2000).
- [28] G.J. Davies, L. Mackenzie, A. Varrot, M. Dauter, A.M. Brzozowski, M. Schülein, and S.G. Withers, *Biochemistry* **37** (34), 11707 (1998).
- [29] H. Li, M. Rief, F. Oesterhelt, H.E. Gaub, X. Zhang and J. Shen, *Chem. Phys. Lett.* **305** 197 (1999); P.E. Marszalek, A.F. Oberhauser, Y.-P. Pang and J.M. Fernandez, *Nature*, **396** 661 (1998); P.E. Marszalek, H. Li, A.F. Oberhauser and J.M. Fernandez, *Proc. Nat. Acad. Sci.* **99** (7) 4278 (2002); B. Heymann and H. Grubmuller, *Chem. Phys. Lett.* **305** 202 (1999); P. O'Donoghue and Z.A. Luthey-Schulten, *J. Phys. Chem. B* **104** (44) 10398 (2000); Z. Lu, W. Nowak, G. Lee, P.E. Marszalek, and W. Yang, *J. Am. Chem. Soc.* **126** (29) 9033 (2004).
- [30] M.M. Kuttel, J.W. Brady, and K.J. Naidoo, *J. Comput. Chem.* **23** (13), 1236 (2002).
- [31] M.M. Kuttel and K.J. Naidoo, *J. Am. Chem. Soc.* **127**, 12 (2005).
- [32] C. Bartels and M. Karplus, *J. Phys. Chem. B* **102**, 865 (1998).