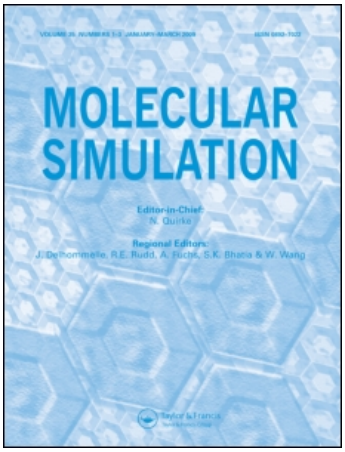

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Conformational flexibility of sulphur linked saccharides a possible key to their glycosidase inhibitor activity

Gerhard A. Venter[a]; Richard P. Matthews[a]; Kevin J. Naidoo[a]
[a] Department of Chemistry, University of Cape Town, Rondebosch, Republic of South Africa

To cite this Article Venter, Gerhard A. , Matthews, Richard P. and Naidoo, Kevin J.(2008) 'Conformational flexibility of sulphur linked saccharides a possible key to their glycosidase inhibitor activity', Molecular Simulation, 34: 4, 391 — 402
To link to this Article: DOI: 10.1080/08927020701765621
URL: http://dx.doi.org/10.1080/08927020701765621

Molecular Simulation
Vol. 34, No. 4, April 2008, 391–401

Taylor & Francis
Taylor & Francis Group

Conformational flexibility of sulphur linked saccharides a possible key to their glycosidase inhibitor activity

Gerhard A. Venter[a], Richard P. Matthews[a] and Kevin J. Naidoo[ab]*

[a]*Department of Chemistry, University of Cape Town, Rondebosch, Republic of South Africa;* [b]*Centre for High Performance Computing, Rosebank, Western Cape, Republic of South Africa*

(*Received 25 July 2007; final version received 22 October 2007*)

We investigate the reasons why sulphur linked saccharides are good inhibitors of retaining β-glycosidases. A comparison of the conformational space and electronic profile of the oxygen and sulphur linked oligosaccharides using HF/6-31G** reveals that they are electronically very similar and have identical conformational preferences. However, the conformational barriers separating the minima are at least 3 kcal mol^{-1} lower for the sulphur linkage implying a greater conformational flexibility. Furthermore, we find using natural bond orbital analysis that the sulphur linkage is significantly less open to acid hydrolysis than the oxygen, found in the natural sugar on which retaining β-glycosidases act. Our nanosecond molecular dynamics studies of *Bacillus agaradhaerens* glycosidase reveal that the thio-cellotriose binds the enzyme 6 kcal mol^{-1}more strongly than does cellotriose. A comparison of the conformational space of the sulphur linkage with that of the oxygen glycosidic linkage provides an explanation for the stronger binding which appears to be due to the greater flexibility of the sulphur glycosidic linkage.

Keywords: glycosidic bond; glycosidase; enzyme reaction

1. Introduction

Glycosidases hydrolyse the glycosidic bonds of saccharides. There are more than 2000 different glycoside hydrolases divided into 76 sequence-distinct families. Structural information on these enzymes is being constantly monitored, updated and made available on http://afmb.cnrs-mrs.fr/ ~ pedro/CAZY/db.html [1]. The functions of these enzymes range from the simple hydrolysis of stored glycosides to viral invasion processes and the control and mediation of cell–cell interactions [2]. Therefore the catalytic mechanisms of these enzymes are of great interest not only because of their value as a pedagogical tool used to described enzyme action in numerous textbooks but also because of the widespread occurrence of genetically inherited disorders of glycoside hydrolysis, and the potential to develop new therapeutic agents from the transition state analogues (TSAs) [3].

The most effective TSAs are able to preferentially bind with the glycosidase and be simultaneously non reactive in the enzyme binding pocket. To efficiently design a TSA it is important to approximate and improve on the electronic and conformational profile of the native saccharide in its transition sate. Here we show that the sulphur linked saccharides are electronically and conformationally very similar to the natural 1–4 linked sugars. We also investigate their effectiveness as inhibitors of glycosidases using a variety of computational methods. In particular we test our newly derived carbohydrate solution force field (CSFF) [4] consistent sulphur parameters, by comparing the performance of thio-cellotriose (glc-o-glc-s-glc) with cellotriose (glc-o-glc-o-glc) in the active site of the *Bacillus agaradhaerens* (BAG) glycosidase [5].

1.1 Transition state structures in glycosidases

The mechanism of retaining β-glycosidase hydrolysis of glycosidic bonds is a controversial subject. However, it is agreed that the retention reaction occurs through two consecutive inversions. The first part results in the attachment of the saccharide to the enzyme and the second part is when the glycosyl is transferred to the acceptor. This double displacement mechanism possibly involves both S_N1 and S_N2 character and two transition state (TS) structures [6]. In this paper we consider only the first in this two part reaction mechanism which is shown in Figure 1. A significant feature of the TS structure is the conformational change in the ring of the non-reducing β-D-pyranose substrate that is required for nucleophile assistance of the glycosidic bond cleavage. 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 β-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 [6–8]. The ALPH is supported by the crystal structures of three *endo*-glycosidases in complexes with substrate analogues

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

ISSN 0892-7022 print/ISSN 1029-0435 online

DOI: 10.1080/08927020701765621
http://www.informaworld.com

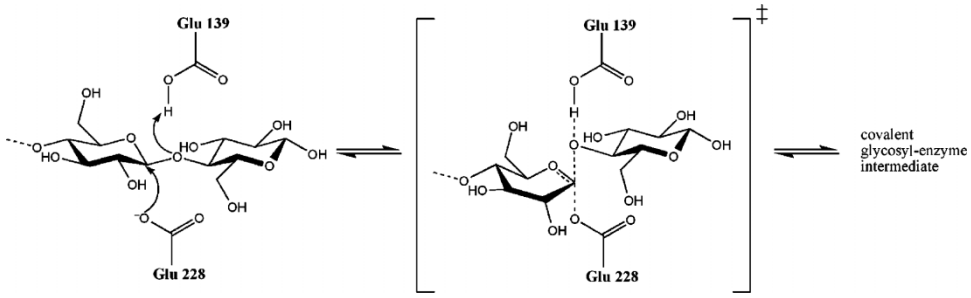

Figure 1. Catalytic reaction mechanism for retaining β-glycoside hydrolase showing the formation of the glycosyl-enzyme transition state (TS) leading to the formation of a covalently bound glycosyl-enzyme intermediate.

showing a skew boat or a flattened boat conformation of the tetrahydropyran ring [5].

Much effort is now expended on developing TSAs for retaining β-glycosidases that are inhibitors. Not surprisingly the success of these attempts are varied [3] despite the targeting of constrained skew boat or boat conformations [8]. An understanding of the shape and charge distribution of the TS is of primary importance in the design of inhibitors. The conformational and electronic information of the TS is presently only accessible through computational means.

The state of the art in the design of inhibitors of the glycoside hydrolysis reaction has generally been attempted by mimicking the assumed half-chair conformation of the TS or its assumed charge [3]. Here the thio-glycosides, containing sulphur, rather than an oxygen heteroatom in the glycosidic linkage have been shown to be stable analogues of their naturally occurring saccharides.

2. Methods

2.1 Conformations from HF calculations

To explore the conformational space of cellobiose and thio-cellobiose we used acetal and thioacetal fragments (Figure 2) as analogues for each of them, respectively. Torsional freedom around the glycosidic linkage is defined in terms of the H1—C1—X—C4′ and C1—X—C4′—H4′ angles as ϕ_H and ψ_H, respectively.

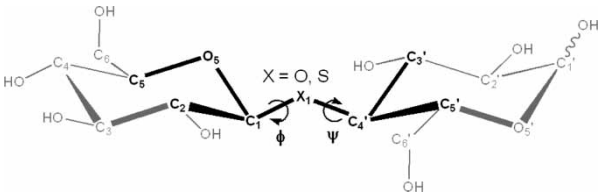

Figure 2. Definition of names for relevant atoms and glycosidic torsion angles ϕ_H = H1—C1—X1—C4′ and ψ_H = C1—X—C4′—H4′ for a representative disaccharide. The acetal/thioacetal fragment used as an analogue for the full disaccharide is shown in bold.

All electronic structure calculations discussed here have been done with GAMESS-UK [9]. The analogues were first optimised using Hartree–Fock (HF) theory, with a 6-31G** basis set. Previously this level of theory has proven reliable compared with methods taking electron correlation into account [10,11]. We confirmed this by optimising the same fragments using density functional theory (DFT), with the three-parameter hybrid exchange functional of Becke [12] in combination with the correlation functional of Lee, Yang and Parr (B3LYP) [13] and a slightly larger basis set, 6-31 + G**. Aside from including the effects of exchange and correlation implicitly this functional also includes 20% exact HF exchange. The two levels of theory gave equivalent results, however, both methods produced a stationary point on the potential energy surface of the β-analogues in which the C5—O5—C1—C2 dihedral did not correspond to a $^4C_1$ glucopyranose-ring geometry, but rather $^1C_4$.

Adiabatic conformational maps in ϕ, ψ space can be derived from rigid and relaxed search procedures. In the rigid approach the total molecular geometry is kept frozen while calculating the conformational energy as a function of ϕ, ψ dihedrals. The relaxed method in contrast includes a geometry optimisation at each ϕ, ψ grid point of the map. Recently da Silva [14] showed for β-lactose that both procedures produced the same stability regions in the same locations. Consequently, to conserve computational time we used the frozen procedure for the *ab initio* calculated conformational maps reported in this paper at values corresponding to stationary points on the potential energy surface of the full disaccharides, calculated at HF/6-31G**. The only degree of freedom allowed was rotation of the methyl groups. An added benefit of the rigid approach is that a change in C5—O5—C1—C2 dihedral angles leading to a ring flip away from $^4C_1$ is not allowed. The maps approximating $^4C_1$ disaccharide ring conformations were calculated at intervals of 20° from − 180° to 180°, for both ϕ and ψ. To improve the accuracy of the low energy conformational energies we performed partial geometry optimisations at HF/6-31G** of ϕ and ψ starting from guessed values in

the low energy regions. The energy of these points were then calculated at B3LYP/6-31 + G**.

2.2 *Electronic structure from natural bond orbitals*

Natural bond orbital (NBO) analysis [15] determines high-occupancy natural orbitals by reversibly transforming the delocalised molecular orbitals into a minimal set of fastest converging one- or two-centre atomic or bond orbitals (in relevant cases, three-centre bonds can also be identified). An important consequence of requiring that the natural orbitals be of high occupancy is that the resulting set shows a high degree of comparison with Lewis structures, consisting of lone pairs and bonds. Delocalisation from the optimal Lewis structure is apparent in occupation of ideally empty anti-bonding or Rydberg-type orbitals. These delocalisation effects can be investigated by considering all possible interactions between filled and empty orbitals and estimating their energetic importance by means of second-order perturbation theory

$$E^{(2)} = -n_i \frac{\langle \sigma_i | F | \sigma_j^* \rangle}{\varepsilon_j^* - \varepsilon_i} \quad (1)$$

where σ_i and σ_j^* are donor and acceptor orbitals, F is the Fock operator in the NBO basis and ε_i and ε_j^* are the orbital energies of σ_i and σ_j^*, respectively. The occupation of σ_i is given by n_i. NBO analyses described here were done on structures optimised at B3LYP/6-31 + G**.

2.3 *Molecular mechanics*

In the present studies of BAG glycosidase, molecular dynamics simulations of the oligosaccharides, protein, and surrounding solvent water molecules were performed using the CHARMM31b programme [16], in which an empirical energy function is used that contains terms for both internal and external interactions, as shown in Equation (2).

$$U(\vec{R}) = \sum_{\text{bonds}} K_b (b - b_0)^2 + \sum_{\text{angle}} K_\theta (\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_\phi \left(1 + \cos(n\phi - \delta)\right) + \sum_{\text{impropers}} K_\omega (\omega - \omega_0)^2 + \sum_{\text{non-bond}} \varepsilon \left[\left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^{12} - \left(\frac{R_{\min_{ij}}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\varepsilon_1 r_{ij}}. \quad (2)$$

The carbohydrate was modelled using an extended CSFF [4], to include oligosaccharides that have sulphur and disulphur glycosidic linkages. In Table 1 we present these new parameters corresponding to Equation (2) describing only the performance of the β(1–4) sulphur-linked saccharides, which is of relevance here. A detailed description of the parameterisation and corresponding NMR experiments is reported elsewhere [17]. The bond and angle reference values were obtained from the HF/6-31G** thioacetal *ab initio* calculations. The C—S bond and C—S—C angle force constants were fitted so that the CHARMM vibrational frequencies obtained for the glycosidic bonds and angles in the thioacetal analogues compared with those produced from HF/6-31G** frequency analysis. All bond and angle force constants, other than those for the S—C bond and C—S—C angle were found to be consistent with previously reported (CHARMM) protein parameters [18]. The protein was modelled using the CHARMM force field [18] along with a TIP3P water model [19] adapted for CHARMM [20].

The enzyme simulations were initiated using the crystal coordinates for BAG obtained from the Brookhaven Protein Data Bank (4A3H). The structure of the enzyme had been resolved at 1.65 Å while the 2,4-dinitrophenyl 2-deoxy-2-fluoro-β-cellobioside complex

Table 1. CSFF-consistent parameters describing the C1-S1-C4′ linkage. Unlisted values (including van der Waals parameters) were kept unchanged from the original CSFF. Also included are parameters for a 1,1 disulphide linkage C-C1-SS-SS-C1-C.

Atom	Charge		
S1	−0.32		
C1	0.12		
H1	0.17		
C4′	0.14		
H4′	0.09		
Bond	K_b	b_0	
S1—C*X*	232.0	1.830	
SS—C1[†]	214.0	1.814	
Angle	K_θ	θ_0	
C1—S1—C4′	37.5	98.55	
H*X*—C*X*—S1	38.0	111.0	
C*X*—C*X*—S1	58.0	112.0	
O5—C1—S1	58.0	112.0	
C—C1—SS[†]	58.0	109.0	
Dihedral	K_ϕ	n	δ
O5—C1—S1—C4′	0.1095	1	0.0
	−0.6860	2	0.0
	0.7820	3	0.0
C2—C1—S1—C4′	−0.7257	1	0.0
	−0.1997	2	0.0
	0.3474	3	0.0
H1—C1—S1—C4′	0.1854	3	0.0
C—C1—SS—SS[a]	0.2800	3	0.0
C1—SS—SS—C1[a]	3.6000	2	0.0

[a] Disulphide linkage parameters, unlisted values kept unchanged from the CHARMM22 forcefield [18].

showed some disorder [5]. Both the thio-cellotriose and the cellotriose were built by modifying the cellobioside complex appropriately.

Two MD simulations were performed. One on the cellotriose and followed by the thio-cellotriose trapped in the protein active site. A water sphere of radius 23.4 Å containing 1941 TIP3P water molecules, centred on the saccharide's centre of mass was used to simulate a solvated enzyme saccharide complex. Those water molecules that overlapped with any of the carbohydrate heavy atoms were removed, and a spherical boundary force consistent with a water density of $1.0\,g\,cm^{-3}$ was applied to the surface of the droplet. The water molecules located in the crystal diffraction study were retained in preference to those from the sphere of pure water, and the same boundary force was applied to all water molecules. A Langevin thermal bath was used to control the temperature of the "surface" atoms in the outer 2 Å thick shell of the reaction sphere. The buffer boundary was set at between 21 Å and the sphere boundary where atoms were propagated under Langevin dynamics [21]. A heat bath of temperature 300 K coupled to the buffer atoms kept the system at thermal equilibrium. The lengths of chemical bonds involving hydrogen atoms were kept fixed using the constraint algorithm SHAKE [22] with a 1 fs integration time step. Long range interactions were made to go to zero between 12 and 13 Å using the CHARMM switching function functions applied on a group-by-group basis [23].

2.4 *Electrostatic free energy of binding*

Electrostatic free energies of binding for the cellotriose followed by the thiocellotriose to the enzyme in solution were calculated by applying the Poisson–Boltzmann equation (PBE) [24]. This describes the solvent as a dielectric continuum with regions inside the molecular surface assigned a low dielectric constant. The solute is introduced as a fixed charge distribution in the form of point charges. Additionally, the effect of a non-zero salt-concentration can also be included [25]. Solutions to the nonlinear PBE were obtained by finite difference techniques using the program Delphi 4.1 [26,27].

Representative structures were taken for both the sulphur- and oxygen-linked trisaccharides from the molecular dynamics trajectories where the interaction energy (the sum of the Coulomb and van der Waals terms) was a minimum. All explicit solvent and crystal water molecules were removed. The PB binding free energy between the enzyme and substrate in its complexed state is calculated as

$$\Delta G_{\text{binding}} = \Delta G_{\text{complex}} - (\Delta G_{\text{enzyme}} + \Delta G_{\text{substrate}}). \quad (3)$$

We use values as suggested by Moreira et al. [28] in our computations and coulombic boundary conditions. The grid size was chosen so that each of systems filled 90% of the grid at a scale of 2.5 points per Ångstrom and the convergence criteria for calculating the potential was set to $10^{-3}\,kT\,e^{-1}$. The solvent was modelled using a dielectric constant of 80 while the interior of the molecule was represented using a value of 4.

3. Results

3.1 *Conformations of the oxygen and sulphur linked saccharides*

The Ramachandran adiabatic conformational maps for the maltose and thio-maltose analogues, $^{\alpha}O$ and $^{\alpha}S$, and the cellobiose and thio-cellobiose analogues, $^{\beta}O$ and $^{\beta}S$ are shown in Figure 3. Values for the four lowest energy conformations for the α-linked (Figure 3(a) and (b)), as well as the β-linked species (Figure 3(c) and (d)), are given in Table 2. There is a high degree of conformational similarity between the α-linked and between the β-linked species.

For both the α-linked species the lowest energy structures are found in the central region of the maps. There is flat region around (−30, 0) encasing two closely spaced stationary points. These two structures for $^{\alpha}O$ at (−30.1, −25.0) and (−26.7, 16.9), are $0.01\,kcal\,mol^{-1}$ apart in energy while the difference between $^{\alpha}S_0$ at (−54.6, −34.8) and $^{\alpha}S_1$ at (−45.7, 35.9) is slightly larger, $0.12\,kcal\,mol^{-1}$. These results are consistent with the HF/6-31G* calculations of French et al. [11] on a glycosidically linked tetrahydropyran dimer aiming to approximate maltose that yielded a minimum at (−47.0, −32.9).

The difference between $^{\alpha}S_2$ and $^{\alpha}S_3$ is smaller, $6\,kcal\,mol^{-1}$, compared to $^{\alpha}O_2$ and $^{\alpha}O_3$, $9\,kcal\,mol^{-1}$. The lowest lying of these two sulphur-linked structures ($^{\alpha}S_2$) is $2.79\,kcal\,mol^{-1}$ higher in energy than the lowest energy structure, this difference is almost doubled in the oxygen-linked ($^{\alpha}O_2$) case, $5.43\,kcal\,mol^{-1}$. In general energy barriers separating the low energy conformations in the oxygen linkage are at least $3\,kcal\,mol^{-1}$ greater than the corresponding thioacetal conformational barriers. Furthermore the central energy plateau is larger for the sulphur than oxygen linkage. The ϕ, ψ potential energy basin that can be sampled via thermal energy (3 kT) at biological temperatures is therefore significantly larger in the sulphur-linked species compared with the oxygen glycosidic linkage. In an explicit comparison of oxygen, sulphur and selenium in the glycosidic linkage of maltose heteroanalogues Weimar et al. have shown, using NOE data from NMR measurement and molecular mechanics calculations, that substitution of oxygen by sulphur to selenium leads to increased flexibility [28].

The reason for this increase in conformational flexibility at the glycosidic linkage is possibly due to the larger size of the linkage atom, which leads to significantly longer C1—S1 (glycosidic) and C4′—S1 (aglycon) bonds and thereby helping to reduce steric clashes between

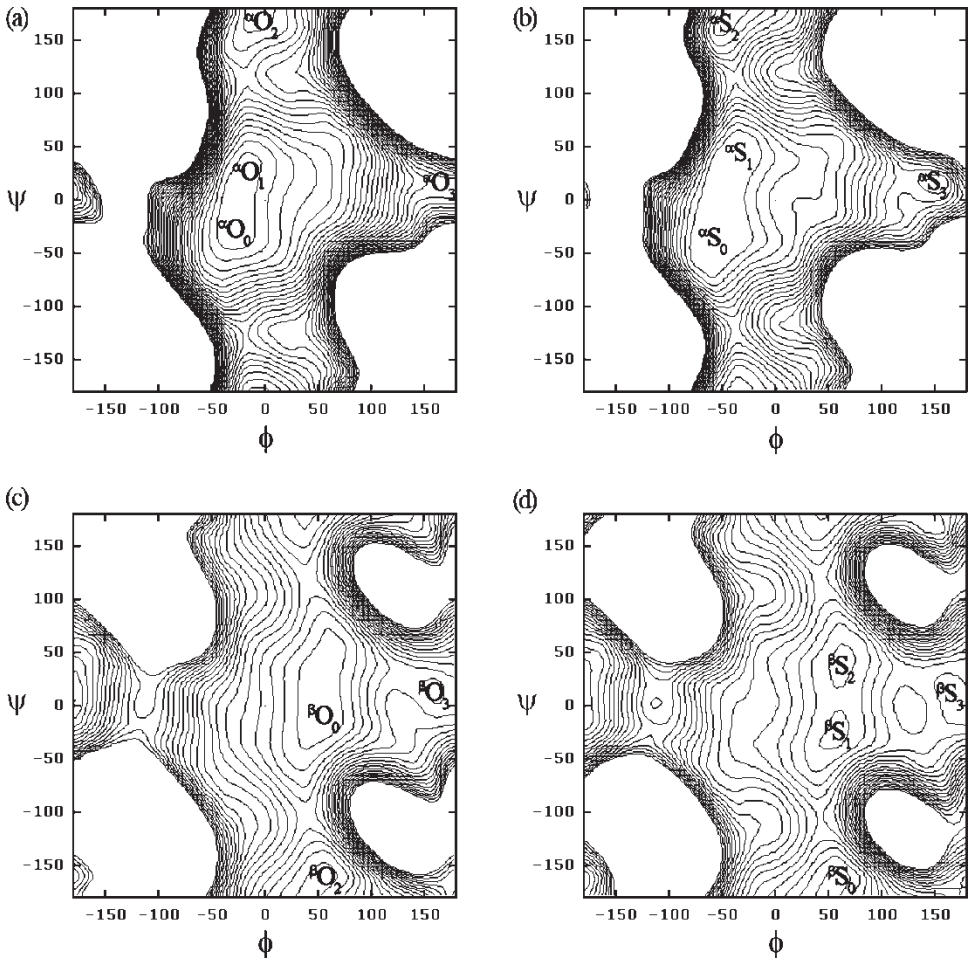

Figure 3. HF/6-31G** energy surfaces for the analogues of (a) maltose, (b) thiomaltose, (c) cellobiose and (d) thiocellobiose. Selected minima are indicated (see Table 1 for values). Contour levels are spaced at 1 kcal mol^{-1}.

Table 2. Energies and glycosidic torsion angles ϕ and ψ for stationary points on the potential energy surface of the maltose ($^{\alpha}O_n$), thiomaltose ($^{\alpha}S_n$), cellobiose ($^{\beta}O_n$) and thiocellobiose ($^{\beta}S_n$) analogues.

	B3LYP/6-31 + G**			extended CSFF		
	ϕ	ψ	$E_{relative}$ (kcal mol^{-1})	ϕ	ψ	$E_{relative}$ (kcal mol^{-1})
$^{\alpha}O_0$	−30.1	−25.0	0.00[a]	−36.2	−30.5	0.00[a]
$^{\alpha}O_1$	−26.7	16.9	0.01	−38.4	23.2	0.52
$^{\alpha}O_2$	−7.9	174.9	5.43	−6.3	−178.9	1.42
$^{\alpha}O_3$	176.9	13.2	14.44	63.2	54.6	6.12
$^{\alpha}S_0$	−54.6	−34.8	0.00[a]	−68.2	−63.1	0.00[a]
$^{\alpha}S_1$	−45.7	35.9	0.12	−58.2	40.2	0.23
$^{\alpha}S_2$	−55.2	159.4	2.79	−43.5	118.7	3.29
$^{\alpha}S_3$	157.4	18.5	8.59	−172.6	6.8	6.51
$^{\beta}O_0$	54.6	−12.9	0.00[a]	52.4	−8.3	0.00[a]
$^{\beta}O_1$[b]						
$^{\beta}O_2$	59.7	−156.5	0.83	31.6	−170.8	1.23
$^{\beta}O_3$	162.2	14.8	5.14	164.5	3.4	3.64
$^{\beta}S_0$	62.0	−162.4	0.00[a]	63.2	−158.9	0.00[a]
$^{\beta}S_1$	57.2	−30.7	0.11	54.5	−23.8	0.15
$^{\beta}S_2$	62.5	39.9	0.15	64.8	36.2	0.01
$^{\beta}S_3$	164.7	18.7	2.93	175.2	30.4	8.91

[a] The energies of stationary points within a specific glycosidic atom and linkage type are relative to the lowest energy structure of each. [b] The second stationary point in the central region of the potential energy map was not obtained.

outlying groups. The full geometry optimisation of thiomaltose at HF/6-31G**, used to determine the fragment geometries, calculated the glycosidic and aglyconic bonds to be 1.835 and 1.831 Å, respectively, as opposed to 1.394 and 1.406 Å for maltose. Although the C1—S1—C4′ angle is smaller in the thio-derivative (101.3°) compared to maltose (119.1°) the larger bond lengths still lead to the C1 and C4′ atoms being 0.4 Å further apart in thiomaltose. A more fundamental explanation of this flexibility difference has been the presence of the exo-anomeric effect in oligosaccharides [30]. Employing a natural orbital analysis the exo-anomeric effect is observed in the delocalisation of lone pair density of the exocyclic glycosidic atom to the endocyclic O—C antibonding natural orbital [31]. Both α and β thio-analogues show considerably less $n_x \rightarrow \sigma^*_{c-o}$ hyperconjugative stabilisation estimated through second-order perturbative analysis [15]. The α thio-analogue is stabilised by 7.92 kcal mol^{-1} compared with 14.64 kcal mol^{-1} in the oxygen linked case and similarly the β thio-analogue is stablised by 8.76 kcal mol^{-1} as opposed to 15.46 kcal mol^{-1} for the oxygen linkage.

In the β-linked analogues of cellobiose and thiocellobiose the location of lowest energy conformations are very similar. There is a minor difference that the lowest stationary point for the oxygen linked case $^{\beta}O_0$ is in the central region (54.6, − 12.9) compared with the bottom–middle of the map (62.0, − 162.4) in the $^{\beta}S_0$ case. It is noteworthy that the range of energy differences between the lowest three structures, $^{\beta}S_0$ to $^{\beta}S_2$, is merely 0.15 kcal mol^{-1}. As before the overall trend is that the energy surface for sulphur is flatter compared with the surface for oxygen. An analysis of the geometry of the stationary points for the disaccharides, with glycosidic bonds of 1.381 and 1.407 Å to the non-reducing and reducing sugars for cellobiose and 1.801 and 1.825 Å for thiocellobiose. In addition the glycosidic angles C1—X—C4′ are 121.6 and 121.2° for X═S and O, respectively. We believe this to be the underlying reason for the greater conformational flexibility in sulphur in both the β- and the α-linkages.

Adiabatic contoured energy surfaces for the acetals and thioacetals were calculated as a function of the glycosidic dihedrals ϕ and ψ and compared with the *ab initio* adiabatic maps. We used a simulated annealing method developed previously [32] for disaccharides to calculate these relaxed adiabatic maps. In addition adiabatic maps for maltose, thiomaltose (see additional material), cellobiose and thiocellobiose (Figure 6) were constructed. To summarise we note that the conformational minima locations for the oxygen and sulphur linkages are almost identical. However, the conformational adiabatic surfaces for the sulphur linkages are significantly shallower indicating that these linkages produce oligosaccharides that are more flexible at biological temperatures than the naturally occurring oxygen linked saccharides.

3.2 Electronic structure of the cellotriose and thiocellotriose transition stage analogues

As we mentioned earlier a potential enzyme inhibitor must conformationally and electronically mimick the native ligand. We have shown above that the conformational preferences of the native oxygen linked saccharides and sulphur linked saccharides are very similar. Here we compare the electronic profile of the oxygen and Sulphur β-linked molecules when they are in the TS in the BAG active site.

It has been established that in the reaction shown in Figure 1 for BAG the glycosidic bond is broken by a Glu 228 nucleophilic attack on the C1 carbon and a Glu 139 acid hydrolysis on the glycosidic oxygen [33]. The exact sequence and nature of the reaction mechanism is not understood and is the subject of our research programme using a multidimensional reaction dynamics method developed by us [34]. However, we will not discuss the reaction mechanism of the hydrolysis of cellotriose here, instead we will only describe the geometries and electronic structures of simplified analogues of the oxygen- and sulphur-linked Michaelis complexes, that includes the amino acid residues Glu 139 and Glu 228. For this purpose the oxygen-linked trisaccharide Michaelis complex from the crystal structure (PDB 4A3H) was used as starting geometry. In constructing the analogues of the two trisaccharides the central sugars were retained while the two end sugars were simplified with propane (R═CH(CH$_3$)$_2$) groups in both cases (See Figure 4). Using the same crystal structure and replacing the glycosidic oxygen with a sulphur atom we generated a sulphur analogue.

We then optimised these geometries using B3LYP/6-31 + G**. The root mean square difference between the distorted skew-boat glucopyranose rings of the oxygen and sulphur structures is 0.033 Å. The natural population analysis (NPA) scheme was used to estimate the atomic partial charges for these two structures, which we show in Table 3. Except for the glycosidic atoms C1—O1—C4′ and C1—S1—C4′ the charges for the two species are almost

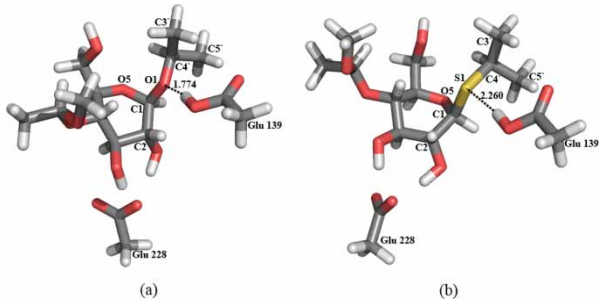

Figure 4. B3LYP/6-31 + G** optimised structures of the TS analogues of (a) cellotriose and (b) thiocellobiose in the catalytic cavity of *B. agaradhaerens*. The acid/base Glu 139 and nucleophile Glu 228 are included.

Table 3. Atomic partial charges as calculated by the Natural Population Analysis (NPA) scheme at B3LYP/6-31 + G** for the cellobiose (left) and thiocellbiose (right) substrate TS analogues.

	Charge (e)	
	Oxygen	Sulphur
O1 / S1	−0.646	0.145
C1	0.376	−0.038
C2	0.023	0.032
O2	−0.821	−0.819
C3	0.033	0.023
O3	−0.806	−0.805
C4	0.040	0.040
O4	−0.584	−0.585
C5	0.040	0.040
O5	−0.608	−0.600
C6	−0.132	−0.133
O6	−0.789	−0.786
C4′	0.060	−0.409
C3′	−0.721	−0.701
C5′	−0.711	−0.702

identical indicating that the difference in glycosidic linkage has not affected the electron distribution in the sugar rings.

3.3 Glycosidic bond reactivity

Even though the charge distribution is comparable across the two molecules the significant difference in charge distribution about atoms involved in the glycosidic bond is a clue to consider the comparable ease with which the glycosidic atom (oxygen or sulphur) can be hydrolysed. We calculated the hydrogen bond strength between the Glu 139 and the linkage atom O1 and S1, and show the optimised structures in Figure 4. This is done by calculating the difference in energy between the substrate–Glu139 complex and that of the isolated amino acid Glu 139 and isolated substrate, both at the geometry they have in the complexed form. It was left uncorrected for the basis set superposition error (BSSE). We found the O1···Glu 139 hydrogen bond to be stronger ($-10.25\,\mathrm{kcal\,mol^{-1}}$) than the S1···Glu 139 hydrogen bond ($-8.80\,\mathrm{kcal\,mol^{-1}}$). Interestingly, even though the energy difference is small the O1···HO_2C—Glu 139 distance (1.774 Å) differs significantly from the geometry of the sulphur linked optimised Michaelis complex where the S1···HO_2C—Glu 139 distance is 2.260 Å. This difference in hydrogen bond strength and overall geometry between the native saccharide and the stability of the thioglycosidic bond or increased difficulty with which these bonds are cleaved have previously been attributed to the lower proton affinity of sulphur over oxygen, resulting in inefficient acid catalysis of the departing aglycon [35,36].

Furthermore the NPA scheme showing a greater negativity on the oxygen ($-0.646\,e$) compared with the sulphur ($0.145\,e$) implies different electron distributions in the C—O and C—S bonds. NBO analysis of the bond composition shows that both glycosidic C—O bonds are polarised in equal amounts towards the oxygen by *ca*. 20% (giving a 70/30 composition), whereas the equivalent C—S bonds remain mainly unpolarised with a 50/50 partitioning. This observation is predicted on the basis of their Pauling electronegativity values alone, oxygen being 3.5 and sulphur 2.5. The sulphur atom has a considerably higher contribution from p-orbitals (±84%) compared with the oxygen atom (±71%). However, in both cases the carbons are sp^3 hybridised. The higher p-character of the sulphur atom results in increased spatial and energetic similarity with the sp^3-hybridised carbons, further explaining the less polarisation observed in the C—S bonds.

A consequence of the greater polarisation of the electron density towards oxygen is the ability to produce a stronger electrostatic interaction with the partially positive hydrogen atom. A hydrogen bond X···HO_2C—Glu 139 between the hydrogen donor on Glu 139 and the acceptor glycosidic X=O, S is represented as the $n_x \rightarrow \sigma_{OH}^*$ charge transfer interaction. We have estimated interactions of 9.54 and 12.21 kcal mol^{-1} for the two lone pairs on the oxygen while for the two lone pairs on the sulphur we get 2.17 and 20.67 kcal mol^{-1}. The second lone pair dominates the interaction when sulphur is involved, whereas with oxygen there is an equivalent contribution from both lone pairs. The interplay between overlap and the energy gap between the natural orbitals show that for oxygen both lone pairs are overlapping to an equal extent and the difference in energy is also comparable. However, with the sulphur donor the first lone pair shows both less overlap and a larger energy gap than the second counterpart. Summing the energetic estimates gives a slightly stronger interaction for $n_S \rightarrow \sigma_{OH}^*$, 22.84 kcal mol^{-1} than for $n_O \rightarrow \sigma_{OH}^*$, 21.74 kcal mol^{-1}. These values, however, do not account for steric interactions and are not a complete account of the total strength of interaction [37].

3.4 Comparative enzyme binding

Finally we evaluate the extent to which BAG binds thiocellotriose in preference to cellotriose. The BAG inhibitor in the crystal structure PDB 4A3H had a 2,4-dinitrophenyl aglycon unit. We replaced this unit with a glucopyranose ring and docked the trisaccharide into the enzyme reactive site. To prevent an overlap of the additional glucopyranose atoms with amino acid residues; Glu 136, Trp 175, Tyr 199, and Ala 231 the conformation of the newly formed glycosidic bond had to be altered resulting in dihedrals $\phi = 60.0°$ and $\psi = -15.0°$. These alterations were made

in the initial setup phase using the dihedral constraint command in CHARMM. The constraints were then removed and subsequent minimisation, heating, equilibration and production phases were conducted.

A key characteristic of an enzyme inhibitor is the extent to which it can preferentially bind to the protein compared with the oxocarbenium ion TS of the oligocellobiose. We have estimated the total interaction energy for the nanosecond long MD trajectories of thio-cellotriose and cellotriose bound to the enzyme. A list of the strongest interactions ($>2\,\text{kcal mol}^{-1}$) between amino acids and the oxygen and sulphur linked substrates is given in Table 4. Average values sampled at intervals of 10 ps as well as the standard deviation are given. Furthermore we performed a PBE binding energy calculation on both systems at maximum interaction points along each trajectory and found the sulphur linked substrate to bind BAG $6.94\,\text{kcal mol}^{-1}$ more strongly than the cellotriose.

Table 4. Average total interaction energy and maximum interaction energy between selected residues of *B. agaradhaerens* Cel5A and the (top) cellotriose and (bottom) thiocellotriose substrates, calculated from MD simulations.

Amino Acid	E_{Total} (kcal mol^{-1})	Std. Deviation (σ)	E_{Max} (kcal mol^{-1})
Oxygen			
His 35	−3.18	2.83	−10.11
Tyr 66	−10.04	1.70	−15.02
Ser 69	−2.81	2.38	−7.99
His 101	−13.72	3.34	−18.67
Glu 139	−2.67	2.31	−10.72
Trp 178	−4.73	2.07	−10.52
Gln 180	−0.12	0.37	−1.32
His 200	−0.03	0.04	−0.17
Phe 201	−0.49	0.31	−1.36
Tyr 202	−5.41	1.56	−12.32
Thr 205	−5.46	4.05	−10.80
Glu 228	−5.40	3.49	−29.07
Ser 232	−4.99	3.23	−11.16
Ala 233	−4.62	2.41	−11.62
Ala 234	−6.56	2.16	−17.40
Trp 262	−6.45	1.61	−10.43
Lys 267	−3.22	8.92	−34.10
Glu 269	−2.72	2.13	−9.79
Sulphur			
His 35	−11.09	2.82	−16.99
Tyr 66	−10.09	1.76	−14.82
Ser 69	−0.52	0.56	−2.26
His 101	−4.28	1.42	−8.70
Glu 139	−7.74	1.58	−16.36
Trp178	−5.91	2.16	−10.14
Gln 180	−2.60	3.00	−10.74
His 200	−3.28	2.02	−8.90
Phe 201	−4.19	1.68	−10.42
Tyr 202	−5.40	1.74	−12.83
Thr 205	−0.50	0.58	−4.02
Glu 228	−29.40	7.14	−50.08
Ser 232	0.07	0.38	−1.32
Ala 233	0.06	0.88	−3.91
Ala 234	−3.11	1.52	−9.59
Trp 262	−7.65	1.03	−10.39
Lys 267	−10.05	11.52	−38.92
Glu 269	−7.44	2.57	−13.23

In a preliminary investigation we identified hydrogen bonds using a geometric definition for the D···H—A hydrogen bond complex. The cutoff criteria for the D—A distance was 3.5 Å and a D—H—A angle of 120°. In this limited evaluation we find a larger number of amino acids that form hydrogen bonds to the oxygen-linked saccharide. We show a schematic of the major interactions in Figure 5 for both cellotriose and thio-cellotriose. Classical MD simulations are unable to simulate bond-breaking mechanisms for example where the nucleophile Glu 228 forms a bond with C1. Instead, we observe that Glu 228 strongly interacts with the hydroxyl on C2 leading to hydrogen bond formation (Figure 5). In the case of thiocellotriose an additional interaction with the hydroxyl on C2 of the reducing sugar occurs intermittently (not shown in Figure 5) contributing significantly to the higher average interaction with the thio-derivative (see Table 4). Furthermore the acid Glu 139 does not interact strongly with the glycosidic oxygen during the 1ns classical force field simulation; instead it regularly forms a hydrogen bond with the hydroxyl on C2.

There is only one atom difference between cellotriose and thio-cellotriose yet they exhibit significant differences in their binding to the BAG glycosidase. In analysing the detailed interactions between the surrounding amino acids and each of the two trisaccharides, a snapshot of which is shown in Figure 5, it is immediately apparent that the differences do not stem from the non-bonded interactions to only the glycosidic oxygen and sulphur. Indeed while there are common amino acids binding to the two trisaccharides there are many different ones that bind in the two cases. The difference in binding interaction of largely similar molecules could most likely only originate from a difference in binding conformation. To this end we analysed the conformational space of cellobiose (Figure 6(a)) and thio-celliobiose (Figure 6(b)). We used a simulated annealing approach to produce these adiabatically relaxed conformational energy maps [31]. The sulphur linked disaccharide has an energetically much lower topography than does cellobiose indicating that this linkage may lead to greater molecular flexibility. The conformational energy barrier separating the central global minima region at around (50, −20) and the minima region at around (180,0) is $4\,\text{kcal mol}^{-1}$ greater for cellobiose than for thio-cellobiose. Interestingly the MD starting conformation ($\phi = 60.0$ and $\psi = -15.0$) of the glycosidic linkage between subsites +1 and −1, upon which the enzyme acts, is at the disaccharide global minimum, for both sugars. In addition we mark (X) on the maps the glycosidic linkage conformation ($\phi = -33$, $\psi = -47$) between the +1 and −1 subsites of a thio-DP5

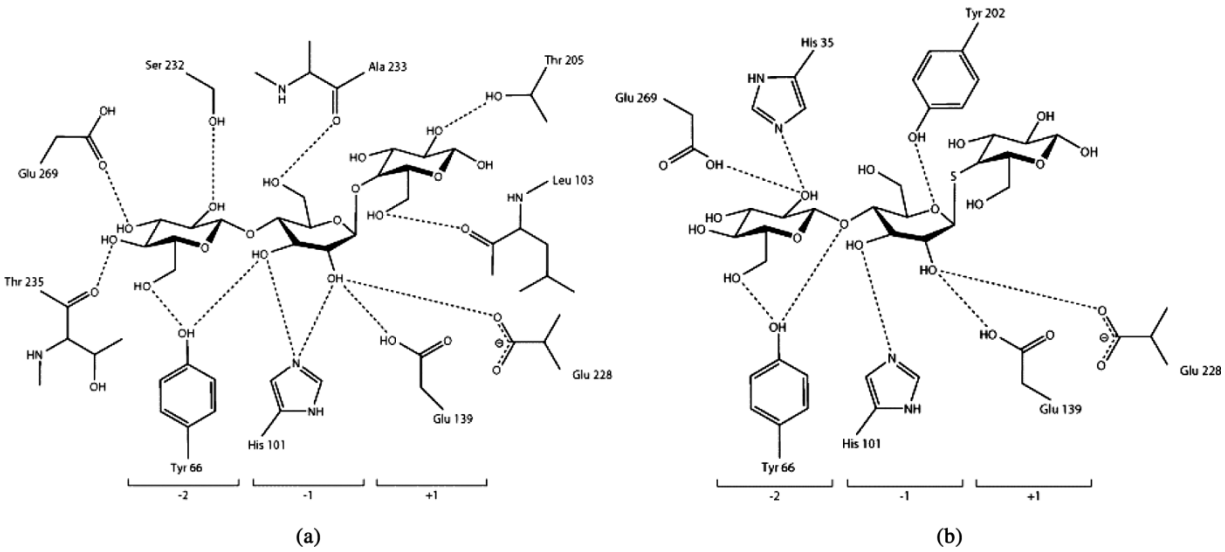

Figure 5. A schematic representation of probable hydrogen bonding of (a) cellotriose and (b) thiocellotriose with *B. agaradhaerens* at the maximum total interaction during the MD trajectory.

inhibitor [methyl-*S*- → -D-glucopyranosyl-(1 → 4)-*S*-4-thio-β-D-glucopyranosyl-(1 → 4)-*S*-4-thio-β-D-glucopyranosyl-(1 → 4)-*S*-4-thio-β-D-glucopyranosyl-(1 → 4)-4-thio-R-D-glucopyranoside)] from the crystal structure of *Fusarium oxysporum* [38]. In this case the inhibitor conformation of subsites +1 and −1 approximates the TS conformation of cellulose.

We have contoured the ϕ, ψ probability surfaces $P(\phi, \psi)$ from the 1ns simulations of the cellotriose (Figure 6(c)) and thio-cellotriose (Figure 6(d)) in the BAG active site in gradients differing by 10% in sampling probability levels from the most populous conformations. The inner contours represent the highest conformational probability for the linkages in ϕ, ψ space but this is at the expense of a greater distribution of conformations. The sulphur linkage in thio-cellotriose samples much larger regions of conformational space than does the corresponding oxygen linkage in cellotriose. This is due to the greater barrier heights separating the minima in cellotriose (14.01 kcal mol^{-1}) compared with thio-cellotriose (9.37 kcal mol^{-1}) shown in Figure 6(a) and (b), respectively. The statistics reveal that the sulphur linkage adjusts to the enzyme environment very soon into the simulation while the oxygen linkage took 250 ps before adjusting into the enzyme pocket from its starting coordinates. The most probable glycosidic conformations for subsites +1 and −1 for the cellotriose ($\phi = -39$, $\psi = -25$) and thio-cellotriose ($\phi = -72$, $\psi = -16$) saccharides are remarkably similar to that found for the thio-DP5 inhibitor ($\phi = -33$, $\psi = -47$). These conformations are in the region of the barrier separating their respective global minimas and their minimas in the region at around (180,0). The enzyme environment therefore appears to drive the glycosidic linkage conformation between the +1 and −1 subsite toward the high energy ϕ, ψ region possibly inducing the ring flip away from $^4C_1$ to the half chair leading to the quasi axial C1 glycosidic bond at the −1 subsite.

While we do not present here exhaustive binding energy studies our present conformational investigations along with the *ab initio* adiabatic maps for the two linkage types indicate that the greater conformational flexibility of the sulphur glycosidic linkage may be a significant factor in its effectiveness as a glycosidase inhibitor.

4. Conclusions

Thiosaccharides have been shown experimentally to be effective inhibitors of retaining β-glycosidases. Here we used a suite of computational methods to understand why this is the case. Our investigations show that the sulphur glycosidic linkage produces oligosaccharides that are conformationally and electronically very similar to the natural saccharides that undergo hydrolysis of their glycosidic linkage. However, the sulphur linkage is less reactive because it is much less susceptible to acid attack as is evidenced by the weaker electronic interaction between Glu 139 and the sulphur compared with oxygen. Furthermore the Sulphur carbon bonds in the glycosidic linkage are stronger due to greater p-orbital character in the bond compared with the oxygen of cellotriose. Finally, the thiocellotriose binds the enzyme more stongly than does the cellotriose apparently because of its greater conformational flexibility.

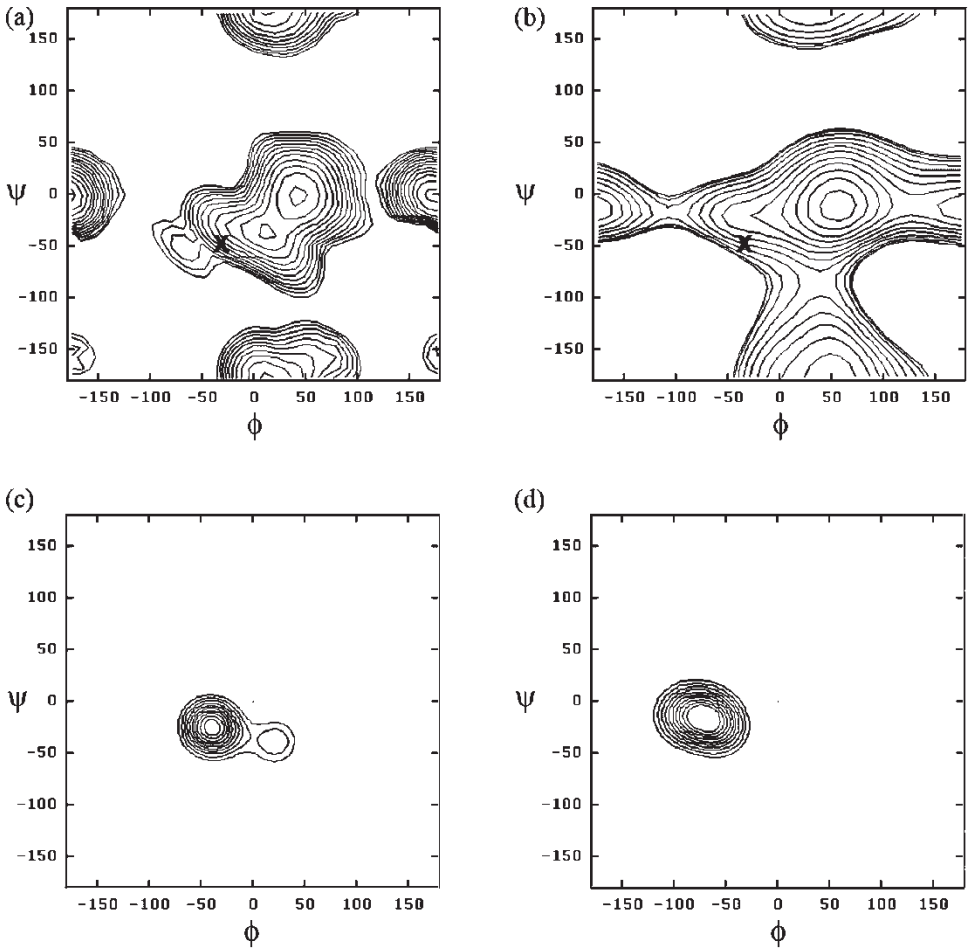

Figure 6. Relaxed adiabatic energy maps for (a) cellobiose and (b) thiocellobiose calculated on a $20° \times 20°$ grid using the extended CSFF force field. Contoured at 1 kcal mol^{-1} intervals for energies. The conformation of the $+1$ and -1 subsites of the thio-DP5 inhibitor [36] of the *Fusarium oxysporum* crystal structure is marked **x** on the maps. A graphic representation of the (ϕ, ψ) statistics for the glycosidic bond linking pyranose units 2 and 3 from the simulations for (c) cellotriose and (d) thiocellotriose. The statistics plots are contoured at intervals of 10% probability.

Acknowledgements

This work is based upon research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation to KJN.

References

[1] B. Henrissat and A. Bairoch, *Updating the sequence-based classification of glycosyl hydrolases*, Biochem. J. 316 (1996), p. 695.
[2] G.J. Davies and B. Henrissat, *Structures and mechanisms of glycosyl hydrolases*, Structure 3 (1995), p. 853.
[3] V.H. Lillelund, H.H. Jensen, X. Liang, and M. Bols, *Recent developments of transition-state analogue glycosidase inhibitors of non-natural product origin*, Chem. Rev. 102 (2002), p. 515.
[4] M.M. Kuttel, J.W. Brady, and K.J. Naidoo, *Carbohydrate solution simulations: producing a force field with experimentally-consistent hydroxyl rotational frequencies*, J. Comput. Chem. 23 (2002), p. 1236.
[5] G.J. Davies, L. Mackenzie, A. Varrot, M. Dauter, A.M. Brzozowski, M. Schülein, and S.G. Withers, *Snapshots along an enzymatic reaction coordinate; analysis of a retaining β-glycoside hydrolase*, Biochemistry 37 (1998), p. 11707.
[6] D.L. Zechel and S.G. Withers, *Glycosidase mechanisms: anatomy of a finely tuned catalyst*, Acc. Chem. Res. 33 (2000), p. 11.
[7] A. White and D.R. Rose, *Mechanism of catalysis by retaining β-glycosyl hydrolases*, Curr. Opin. Struct. Biol. 7 (1997), p. 645.
[8] E. Lorthiois, M. Meyyappan, and A. Vasella, *β-Glycosidase inhibitors mimicking the pyranoside boat conformation*, Chem. Commun. (2000), p. 1829.
[9] M.F. Guest, I.J. Bush, H.J.J. van Dam, P. Sherwood, J.M.H. Thomas, J.H. van Lenthe, R.W.A. Havenith, and J. Kendrick, *The GAMESS-UK electronic structure package: algorithms, developments and applications*, Mol. Phys. 103 (2005), p. 719.
[10] A.D. French and G.P. Johnson, *Advanced conformational energy surfaces for cellobiose*, Cellulose 11 (2004), p. 449.
[11] A.D. French, A. Kelterer, G.P. Johnson, M.K. Dowd, and C.J. Cramer, *HF/6-31G* energy surfaces for disaccharide analogs*, J. Comput. Chem. 22 (2000), p. 65.
[12] A.D. Becke, *Density-funtional thermochemistry. III. The role of exact exchange*, J. Chem. Phys. 98 (1993), p. 5648.
[13] C. Lee, W. Yang, and R.G. Parr, *Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B 37 (1988), p. 785.
[14] C.O. da Silva, *Carbohydrates and Quantum chemistry: how useful is this combination?* Theor. Chem. Accounts 116 (2006), p. 137.

[15] F. Weinhold, in *Encyclopedia of Computational Chemistry*, P.v.R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F., Schaefer III and P.R. Schreiner, eds., Wiley, New York, 1998, p. 1792.

[16] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *CHARMM: A program for macromolecular energy, minimization, and dynamics calculations*, J. Comput. Chem. 4(2) (1983), p. 187.

[17] R.P. Matthews, K. Fehér, L. Szilagyi, K.E. Kövér, and K.J. Naidoo, *Three-bond interglycosidic linkages: NMR and molecular modelling studies of the diglycosyl disulfides*, (2007), *Unpublished Results in preparation*.

[18] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins*, J. Phys. Chem. B 102 (1998), p. 3586.

[19] W.L. Jorgensen and C. Jenson, *Temperature dependence of TIP3P, SPC, and TIP4P water form NPT Monte Carlo simulations: seeking temperature of maximum density*, J. Comput. Chem. 19 (1998), p. 1179.

[20] P.J. Steinbach and B.R. Brooks, *Protein hydration elucidation by molecular dynamics simulations*, Proc. Natl. Acad. Sci. (USA) 90 (1993), p. 9135.

[21] S.E. Feller, Y. Zhang, R.W. Pastor, and B.R. Brooks, *Constant pressure molecular dynamics simulation: the Langevin piston method*, J. Chem. Phys. 103 (1995), p. 4613.

[22] W.F. van Gunsteren and H.J.C. Berendsen, *Algorithms for macromolecular dynamics and constraint dynamics*, Mol. Phys. 34 (1977), p. 1311.

[23] K. Tasaki, S. McDonald, and W. Brady, *Observations concerning the treatment of long-range interactions in molecular dynamics simulations*, J. Comput. Chem. 14 (1993), p. 278.

[24] B. Honig and A. Nicholls, *Classical electrostatics in biology and chemistry*, Science 268 (1995), p. 114.

[25] W. Rocchia, E. Alexov, and B. Honig, *Extending the applicability of the nonlinear Poisson–Boltzmann equation: multiple dilectric constants and multivalent ions*, J. Phys. Chem. B 105 (2001), p. 6507.

[26] A. Nicholls and B. Honig, *A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation*, J. Comput. Chem. 12 (1991), p. 435.

[27] W. Rocchia, S. Sridharan, A. Nicholls, E. Alexov, A. Chiabrera, B. Honig, *Rapid grid-based construction of the molecular surface for both molecules and geometric objects: applications to the finite difference Poisson–Boltzmann methods*, J. Comput. Chem. 23 (2002), p. 128 .

[28] I.S. Moreira, P.A. Femandes, and M.J. Ramos, *Accuracy of the numerical solution of the Poisson–Boltzmann equation*, J. Mol. Struct. (THEOCHEM) 729 (2005), p. 11.

[29] T. Weimar, U.C. Kreis, J.S. Andrews, and B.M. Pinto, *Conformational analysis of maltoside heteroanalogues using high-quality NOE data and molecular mechanics calculations. Flexibility as a function of the interglycosidic chalcogen atom*, Carbohydr. Res. 315 (1999), p. 222.

[30] R.U. Lemieux and S. Koto, *Conformational properties of glycosidic linkages*, Tetrahedron 30 (1974), p. 1933.

[31] M.L. Trapp, J.K. Watts, N. Weinberg, and M. Pinto, *Component analysis of the X—C—Y anomeric effect (X═O, S; Y═F, OMe, NHMe) by DFT molecular orbital calculations and natural bond orbital analysis*, Can. J. Chem. 84 (2006), p. 692.

[32] K.J. Naidoo and J.W. Brady, *The application of simulated annealing to the conformational analysis of disaccarides*, Chem. Phys. 224 (1997), p. 263.

[33] G.J. Davies, M. Dauter, A.M. Brzozowski, M.E. Bjornvad, K.V. Anderson, and M. Schulein, *Structure of the* Bacillus agaradherans *family 5 endoglucanase at 1.6. Å and its cellobiose complex at 2.0. Å resolution*, Biochemistry 37 (1998), p. 1926.

[34] R. Rajamani, K.J. Naidoo, and J. Gao, *Implementation of an adaptive umbrella sampling method for the calculation of multidimensional potential of mean force of chemical reactions in solution*, J. Comput. Chem. 24 (2003), p. 1775.

[35] J.L. Jensen, W.P. Jencks et al., *Hydrolysis of benzaldehyde O,S-acetals'*, J. Am. Chem. Soc. 101 (1979), p. 1476.

[36] T.H. Fife and T.J. Przystas *Intramolecular electrostatic and general acid catalysis in the hydrolysis of O,S-thioacetals*, J. Am. Chem. Soc. 102 (1980), p. 292.

[37] F. Weinhold and C. Landis, *Valency and Bonding. A Natural Bond Orbital Donor–Acceptor Perspective*, Cambridge University Press, Cambridge, 2005.

[38] G. Sulzenbacher, H. Driguez, B. Henrissat, M. Schulein, and G.J. Davies, *Structure of the* Fusarium oxysporum *endoglucanase I with a nonhydrolyzable substrate analogue: substrate distortion gives rise to the preferred axial orientation for the leaving group*, Biochemistry 35 (1996), p. 15280.