Glycosidase inhibitors as conformational transition state analogues†

Ian R. Greig* and Ian H. Williams*

Received (in Cambridge, UK) 13th April 2007, Accepted 13th June 2007
First published as an Advance Article on the web 28th June 2007
DOI: 10.1039/b705607a

A method for estimating the conformational similarity between hexopyranose rings is presented and used to probe the behaviour of various glycosyl hydrolase inhibitors as conformational transition state analogues.

The glycosyl hydrolases (GH) have proved popular targets for the rational design of inhibitors that mimic both conformational and electrostatic features of their reactions' oxacarbenium ion-like transition states. Potent inhibitors of the GHs have been reported whose design rationale has been based on both the electrostatic and conformational similarities between the inhibitors and putative oxacarbenium ion-like transition states.

X-Ray crystallographic studies of GHs in complex with non-hydrolysable substrate analogues, products and intermediate species are assumed to represent snapshots along reaction coordinates. These structures are then used to deduce a reaction's conformational itinerary.[1] Little attention, however, has been paid to the precise conformation of high energy species found along the reaction coordinate or the dynamic sampling of conformations found in both stable and unstable states. At the heart of many discussions of inhibitor behaviour lies the notion of conformational similarity: the extent to which an inhibitor resembles the conformation of both stable and unstable species along a reaction pathway. This communication addresses the question of whether a particular inhibitor is intrinsically a good mimic of a reaction's transition state through the use of hybrid quantum mechanics/molecular mechanics (QM/MM) molecular dynamics (MD) simulations.

Family GH-84[2] glycosyl hydrolases are *N*-acetyl-β-glucosamidases, with human *O*-GlcNAcase being responsible for the removal of *N*-acetylglucosamine moieties from protein serine and threonine residues using a mechanism of substrate-assisted catalysis (Fig. 1(A)).[3] Physiological and proteomic interests surrounding protein *O*-GlcNAcylation have led to the recent development of selective inhibitors of enzymes belonging to the GH-84 family. Both 1,2-dideoxy-2′-methyl-α-D-glucopyranoso-[2,1-*d*]-Δ2′-thiazoline (thiazoline **1**) and *O*-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-*N*-phenylcarbamate (PUGNAc **2**) are nanomolar inhibitors of family GH-84 human *O*-GlcNAcase. Further studies have probed the inhibition of human and bacterial GH-84 hexosaminidases using tetrahydroimidazo[1,2-*a*]pyridine-derivatives (nagstatins **3a** and **3b**). Nagstatin **3a** is a micromolar inhibitor of human *O*-GlcNAcase whilst **3b** is a picomolar inhibitor of a bacterial *O*-GlcNAcase.[4]

Department of Chemistry, University of Bath, Bath, UK BA2 7AY. E-mail: i.greig@bath.ac.uk; chsihw@bath.ac.uk; Fax: 44 1225 386231; Tel: 44 1225 386625

† Electronic supplementary information (ESI) available: Detailed modelling and conformational analysis procedures. See DOI: 10.1039/b705607a

In this communication, intrinsic differences in the conformational space sampled by thiazoline **1**, PUGNAc **2** and model nagstatin **3c** are examined. Knowledge of the intrinsic conformational behaviour of enzyme inhibitors is a useful guide to how successful a design approach has been in mimicking a particular (desired) conformation and, in combination with knowledge of the bound conformation may provide an indication of how important, or unimportant, conformational aspects of inhibitor design are. The conformational behaviour of these three inhibitors is compared to the conformational itinerary determined for a simple model reaction, the rate determining step for the specific acid catalyzed hydrolysis of *N*-acetylglucosamine hemiacetal (Fig. 2).

A free-energy pathway for the specific acid catalyzed oxazolinium ion formation was generated using the potential of mean force approach and the weighted histogram analysis. The QM core, consisting of the protonated *N*-acetylglucosamine-derived hemiacetal, was treated at the AM1 level of theory and solvated in a 40 Å × 40 Å × 40 Å box of TIP3P water molecules. Whilst well-parameterised MM methods provide the most accurate treatments of ground state carbohydrate conformations questions of conformational transition state analogy require the use of (QM) methods able to treat both ground state and transition state ensembles. The AM1 method has been shown to reasonably reproduce energetic trends associated with intramolecular hydrogen bonding and the anomeric effect.[5] The reaction coordinate was defined as the difference in distances ($d_1 - d_2$) between anomeric carbon–leaving group oxygen (d_1) and anomeric carbon–acetamido oxygen (d_2) bond lengths. A broad plateau region is

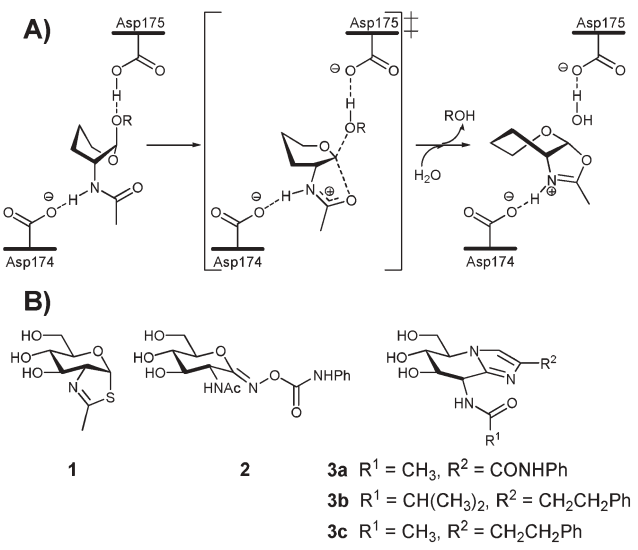

Fig. 1 (A) Human *O*-GlcNAcase utilizes a double-displacement mechanism involving substrate-assisted catalysis to hydrolyse *N*-acetylglucosaminides. (B) Inhibitors of family GH-84 enzymes.

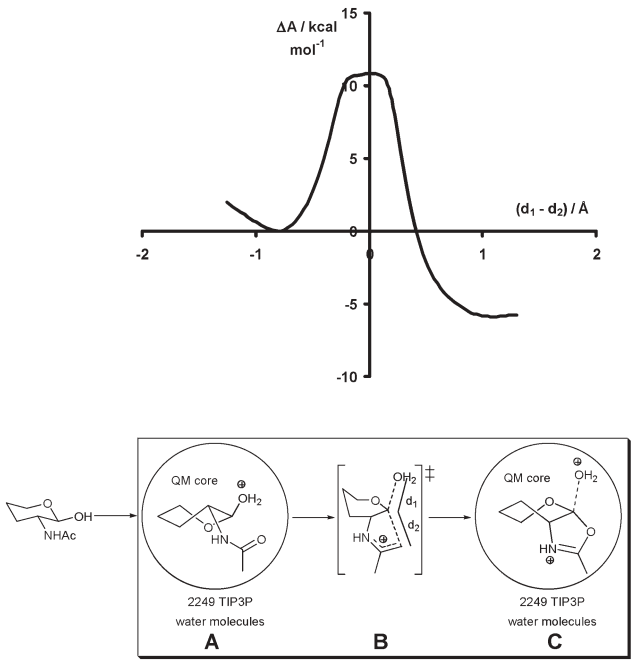

Fig. 2 QM/MM calculated variation in (Helmholtz) free energy with reaction coordinate for the chosen model reaction. The QM/MM system described is highlighted at the bottom of the figure (hydroxyl groups omitted for clarity) and represents the specific acid catalyzed *O*-glycosidic bond cleavage of protonated hemiacetal **A** involving intramolecular *N*-acetamido participation.

found at the peak of the free energy profile, consistent with the borderline associative–dissociative pathway expected for glycoside hydrolyses. Whilst reactant and product state conformations are likely to differ significantly between *O*-GlcNAcase-catalyzed and specific acid catalyzed pathways, the two reactions will share conformationally similar transition states.

Extended (600 ps) MD trajectories of (unconstrained) reactant (protonated hemiacetal **A**) and product species (oxazolinium ion **C**) were produced in addition to (unconstrained) MD trajectories of thiazoline **1**, PUGNAc **2** and nagstatin **3c**. The ensemble of conformations of the carbohydrate in the region of the transition state was taken from the six constrained molecular dynamics trajectories (each of 100 ps) associated with the range of reaction coordinates -0.15 Å $\leqslant (d_1 - d_2) \leqslant 0.1$ Å, for which the free energy of the system varies by only 0.1 kcal mol^{-1}.

Hexopyranose ring conformations were determined for snapshots from trajectories for each of these ensembles and were described following the approach of Bérces *et al.*[6] A conformation was specified in terms of a polar angle, θ ($0^\circ \leqslant \theta \leqslant 180^\circ$), and an equatorial angle, φ ($0^\circ \leqslant \varphi < 360^\circ$). Variation in the polar angle specifies a $^1C_4$ chair ($\theta = 0^\circ$), envelope/half-chair ($\theta = 45^\circ$), boat/skew-boat ($\theta = 90^\circ$), envelope/half-chair ($\theta = 135^\circ$), and $^4C_1$ chair ($\theta = 180^\circ$) transformation; whereas variation in the equatorial angle specifies a boat–skew-boat (for $\theta = 90^\circ$) or envelope–half-chair (for $\theta = 45$ or 135°) pseudorotational itinerary. Angles θ and φ are sufficient to specify which of the 38 canonical conformations a molecule is closest to. A diagrammatic representation of hexopyranose ring conformations may be found in the ESI.†

Arrays describing the fractional occupancies for each of the 38 regions of hexopyranose conformational space were determined

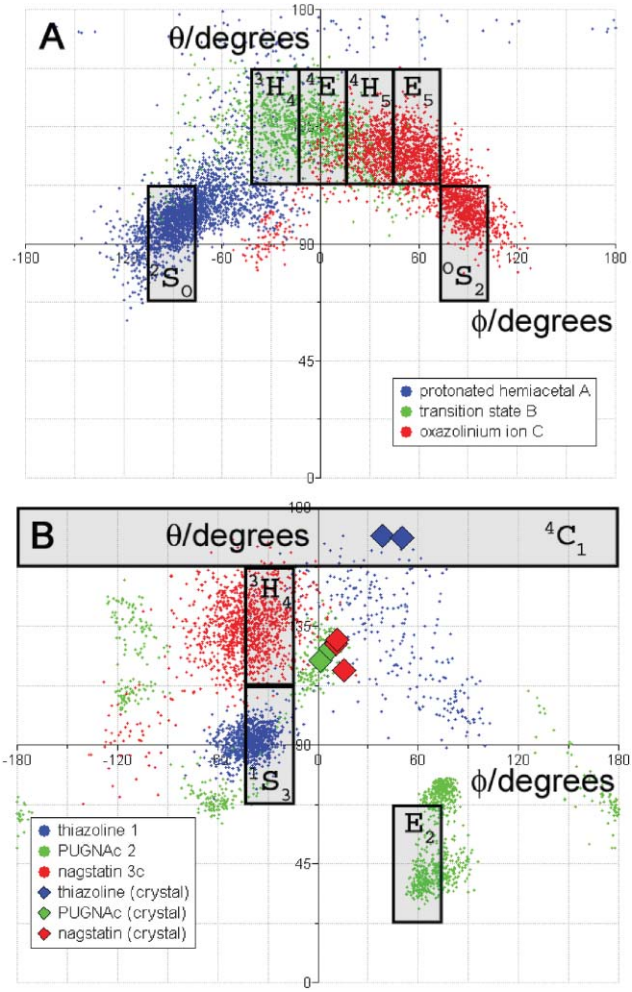

Fig. 3 Regions of hexopyranose conformational space sampled by the MD trajectories for: (A) regions of the reaction coordinate and (B) various inhibitors of GH-84 enzymes.

for each MD trajectory. Furthermore the extent of hexopyranose ring conformational similarity between individual species (protonated hemiacetal **A**, transition state **B**, oxazolinium ion **C**, thiazoline **1**, PUGNAc **2** and nagstatin **3c**) was quantified according to a similarity index (SI).‡

Fig. 3(A) displays the regions of conformational space sampled by protonated hemiacetal **A**, transition state **B** and oxazolinium ion **C**. Protonated hemiacetal **A** is predominantly found in the $^2S_0$ conformation (fractional occupancy 0.50) which, as the reaction proceeds, is distorted to a transition state that may be characterized as primarily falling in the $^3H_4$ and ^4E regions of conformational space (fractional occupancies of 0.29 and 0.36, respectively). The transition state for our simple model reaction therefore samples regions of conformational space consistent with those postulated for the enzymic reaction.[7] Finally, a conformationally mobile oxazolinium ion is formed which samples a broad range of hexopyranose conformations (principally the $^OS_2$, E_5 and $^4H_5$ conformations).

The conformational changes occurring on passage from transition state to product state are related to the intramolecular nature of the reaction under consideration (itself unusual in the context of the GHs). The conformation of the transition state is

influenced not only by the C2–C1–O5–C5 coplanarity expected of an oxacarbenium ion-like species but also by a tendency towards C3–C2–C1–O5 coplanarity predicated by incipient five-membered oxazolinium ion ring formation. This latter influence is at its strongest in the fully formed oxazolinium ion intermediate and is a result of the sp^2 hybridized character shared by the 2-acetamido-derived nitrogen, carbon and oxygen centers. Our model reaction is characterized by a 'late' transition from the perspective of conformational changes occurring along the reaction coordinate. These findings may be placed in a quantitative framework as the value of the SI relating transition state and oxazolinium ion product (0.24) is significantly greater than the SI relating transition state and reactant protonated hemiacetal (0.08).

Fig. 3(B) shows the intrinsic conformational properties of thiazoline **1**, PUGNAc **2** and nagstatin **3c**. Of the three inhibitors studied, nagstatin **3c** is found to be the best conformational analogue of the transition state of our model reaction (SI = 0.60): enforced C2–C1–O5–C5 co-planarity that arises due to the oxacarbenium ion-like transition state for the reaction is mimicked by C2–C1–N5–C5 planarity enforced by the aromatic five-membered ring.

PUGNAc **2** is highly conformationally mobile (the highest fractional occupancy being 0.27 found for the E_2 conformation) and displays modest similarities with transition state and product state conformations. The deficiency of PUGNAc **2**, when compared to nagstatin **3c** in mimicking transition state conformation arises as a planar (sp^2 hybridized) geometry is only enforced at the C1 position rather than at both the C1 and O5 positions found in an oxacarbenium ion-like transition state.

Thiazoline **1** (found principally in the $^1S_3$ conformation with a fractional occupancy of 0.56) is conformationally a very weak transition state analogue (SI = 0.08) and only displays modest similarities with reactant (SI = 0.22) and product states (0.19). Further to these theoretical studies, an analysis of reported enzyme-bound conformations of the inhibitors studied was carried out from published X-ray crystallographically-determined structures.[7] These structures determined are determined at greater than Ångstrom resolution and so there may be significant errors in atomic position (and hence ring conformations). PUGNAc **2** (PDB accession code 2CBJ), nagstatin **3a** (2J47) and nagstatin **3b** (2J62) are reported to bind in the ^4E conformation, consistent with significantly populated regions of conformational space found for those inhibitors in solution.

The reported $^4C_1$ conformation of bound thiazoline **1** (2CHN), however, is unlike reactant, product or transition states of our model reaction and is only sparsely sampled by the inhibitor in solution. Preliminary computational studies of *Bacteriodes thetaiotaomicron O*-GlcNAcase-bound thiazoline **1** in which the inhibitor is treated using the AM1 semi-empirical method and the enzyme using the OPLS forcefield (see ESI†) are consistent with the crystallographically-observed conformation indicating an increased occupancy of the $^4C_1$ conformation as well as a substantial occupancy of the B_{O3} conformation.

A recent report of Vocadlo and co-workers establishes thiazoline **1** as a genuine transition state analogue according to the free-energy relationship-based definition of Bartlett.[8] On the basis of these results, crystallographic studies of bound inhibitors, and calculated molecular electrostatic potential surfaces, Vocadlo and co-workers suggested that, in the case of GH-84 enzymes at least, the most successful transition state analogues are not necessarily those that bear the greatest conformational similarity to the reaction's transition state. Our results expand on those empirical studies and support the notion that thiazoline **1** is a poor conformational mimic of the reaction's transition state in both its bound and unbound states.

We have outlined a simple method for the quantitative description and comparison of ensembles of glycoside conformers derived from MD simulations. These methods have been used to probe the dynamic conformational behaviour of a series of glycosidase inhibitors. This approach to transition state analogy complements existing (static) data available from X-ray crystallographic and KIE-based studies, as well as theoretical studies of electrostatic similarity between inhibitors and transition states.[9] This method may be extended and advanced by considering the dynamic behaviour of bound inhibitors, the determination of detailed conformational itineraries of enzyme-catalyzed reactions, the use of longer time-scales and improved theoretical descriptions of the quantum mechanical system.

We thank the University of Bath for funding and Westgrid for access to resources.

Notes and references

‡ The conformational similarity between two ensembles of structures, A and B, was defined by their similarity index, SI(A, B), according to the equation below.

$$\mathrm{SI}(\mathrm{A},\mathrm{B}) = \frac{2\sum_{c=1}^{38} f_{\mathrm{A}}(c) f_{\mathrm{B}}(c)}{\sum_{c=1}^{38} (f_{\mathrm{A}}(c))^2 + \sum_{c=1}^{38} (f_{\mathrm{B}}(c))^2}$$

The similarity index of A and B is therefore related to their associated fractional occupancy arrays, $f_A(c)$ and $f_B(c)$, which are functions of the 38 canonical hexopyranose conformations. Identical systems will therefore have a similarity index of 1 while systems that sample entirely separate regions of conformational space with have a similarity index of 0.

1 A. Vasella, G. J. Davies and M. Böhm, *Curr. Opin. Chem. Biol.*, 2002, **6**, 619.

2 P. M. Coutinho and B. Henrissat, in *Recent Advances in Carbohydrate Bioengineering*, ed. H. J. Gilbert, G. Davies, B. Henrissat and B. Svensson, The Royal Society of Chemistry, Cambridge, 1999, pp. 3–12.

3 M. S. Macauley, G. E. Whitworth, A. W. Debowski, D. Chin and D. J. Vocadlo, *J. Biol. Chem.*, 2005, **280**, 25313.

4 (*a*) K. A. Stubbs, N. Zhang and D. J. Vocadlo, *Org. Biomol. Chem.*, 2006, **4**, 839; (*b*) B. Shanmugasundaram, A. W. Debowski, R. J. Dennis, G. J. Davies, D. J. Vocadlo and A. Vasella, *Chem. Commun.*, 2006, 4372; (*c*) H. C. Dorfmueller, V. S. Borodkin, M. Schimpl, S. M. Shepherd, N. A. Shpiro and D. M. F. van Aalten, *J. Am. Chem. Soc.*, 2006, **128**, 16484.

5 (*a*) M. Khalil, R. J. Woods, D. F. Weaver and V. H. Smith, Jr., *J. Comput. Chem.*, 1991, **12**, 584; (*b*) C. J. Cramer and D. G. Truhlar, *J. Am. Chem. Soc.*, 1993, **115**, 5745–5753; (*c*) R. J. Woods, W. A. Szarek and V. H. Smith, Jr., *J. Chem. Soc., Chem. Commun.*, 1991, 334.

6 A. Bérces, D. M. Whitfield and T. Nukada, *Tetrahedron*, 2001, **57**, 477.

7 (*a*) R. J. Dennis, E. J. Taylor, M. S. Macauley, K. A. Stubbs, J. P. Turkenburg, S. J. Hart, G. N. Black, D. J. Vocado and G. J. Davies, *Nat. Struct. Mol. Biol.*, 2006, **13**, 365; (*b*) F. V. Rao, H. C. Dorfmueller, F. Villa, M. Allwood, I. M. Eggleston and D. M. F. Van Aalten, *EMBO J.*, 2006, **25**, 1569.

8 G. E. Whitworth, M. S. Macauley, K. A. Stubbs, R. J. Dennis, E. J. Taylor, G. J. Davies, I. R. Greig and D. J. Vocadlo, *J. Am. Chem. Soc.*, 2007, **129**, 635.

9 C. K. Bagdassarian, V. L. Schramm and S. D. Schwartz, *J. Am. Chem. Soc.*, 1996, **118**, 8825.