COMMUNICATION

ChemComm
www.rsc.org/chemcomm

Distortion of a cellobio-derived isofagomine highlights the potential conformational itinerary of inverting β-glucosidases†

Annabelle Varrot,[a] James Macdonald,[b] Robert V. Stick,[b] Gavin Pell,[c] Harry J. Gilbert[c] and Gideon J. Davies*[a]

[a] *Structural Biology Laboratory, Department of Chemistry, The University of York, Heslington, York, UK YO10 5YW. E-mail: davies@ysbl.york.ac.uk; Fax: 44 1904 328266; Tel: 44 1904 328260*

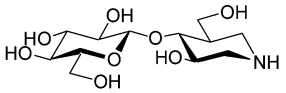
[b] *School of Biomedical and Chemical Sciences, The University of Western Australia, 35 Stirling Hwy, Crawley WA 6009, Australia. E-mail: rvs@chem.uwa.edu.au; Fax: 61 8 9380 3993; Tel: 61 8 9380 3200*

[c] *School of Cell and Molecular Biosciences, University of Newcastle upon Tyne, The Agriculture Building, Newcastle upon Tyne, UK NE1 7RU. E-mail: H.J.Gilbert@ncl.ac.uk; Fax: 44 191 222 8684; Tel: 44 191 222 6962*

Received (in Cambridge, UK) 11th February 2003, Accepted 6th March 2003
First published as an Advance Article on the web 20th March 2003

A cellobio-derived isofagomine glycosidase inhibitor (K_i ~ 400 nM) displays an unusual distorted ^{2,5}B (boat) conformation upon binding to cellobiohydrolase Cel6A from *Humicola insolens*, highlighting the different conformational itineraries used by various glycosidases, with consequences for the design of therapeutic agents.

The enzymatic hydrolysis of glycosides is performed *via* transition states that display substantial oxocarbenium-ion like character. This demands that at, or extremely close to, this transition state, C5, O5, C1 and C2 lie approximately co-planar, an arrangement accommodated only by $^4H_3$ and $^3H_4$ half-chair and ^{2,5}B and $B_{2,5}$ classical boat conformations.[1] Here we demonstrate that the cellobio-derived isofagomine, (3*R*,4*R*,5*R*)-4-(β-D-glucopyranosyl)oxy-3-hydroxy-5-(hydroxymethyl)piperidine, **1**)[2] is distorted towards a ^{2,5}B conformation upon binding to the inverting β-glucosidase, Cel6A, which together with the ^2S_O conformation recently revealed for non-hydrolysable "Michaelis" complexes of this enzyme, supports a novel conformational agenda for these enzymes. The widely held, but incorrect, belief that all glycosidases perform catalysis *via* $^4H_3$-configured transition states must be readdressed in light of recent structural insight.



1

Over 7000 glycosidase sequences are known. These enzymes have been classified into over 90 "GH" families based upon amino acid sequence similarities.[3] A feature of these families is that the configuration of both substrate and subsequent product is conserved within a family. The nature of the conformational pathways harnessed by particular families and which of the four potential conformations is adopted at the oxocarbenium-ion-like transition states remains unclear. Family "GH-6" is a grouping of related endoglucanases and cellobiohydrolases that hydrolyse the β-1,4 linkages of gluco-oligo- and polysaccharides with inversion of anomeric configuration. 3-D structural and kinetic analyses have led to a dissection of the binding subsites and some components of the catalytic apparatus.[1] Trapping of non-hydrolysable thio-oligosaccharides revealed an unusual ^2S_O skew-boat conformation for the "Michaelis" complex in the "catalytic" −1 subsite,[4,5] although displacement of the catalytic acid away from the active centre in these structures confuses interpretation.

† Electronic supplementary information (ESI) available: details of data and structure quality for complex of cel6A with **1**. See http://www.rsc.org/suppdata/cc/b3/b301592k/

Compound **1** is a competitive inhibitor of Cel6A with an apparent K_i of 412 nM.‡ The 3-D structure of the *Humicola insolens* Cel6A (D416A mutant), at 1.3 Å resolution,‡ reveals two molecules in the active-centre tunnel occupying the −2,−1 and +1,+2 subsites. The isofagomine moiety in the catalytic −1 subsite is distorted to a conformation between ^2S_O skew-boat and ^{2,5}B boat (closer to the latter), Fig. 1. This places C6 pseudo-axial and would similarly place O1 pseudo-axial were such an atom present. Distortion appears to be driven, in part, through the steric barrier posed by Tyr174, Fig. 2, "behind" the −1 subsite and through optimal hydrogen-bonding from O3 to the side-chain of invariant Asp405. A feature of the complex with **1**, absent from previous studies, is the correct positioning of the Brønsted acid, Asp226, *via* a 2.7 Å hydrogen-bond to the O4 atom of the +1 subsite sugar.

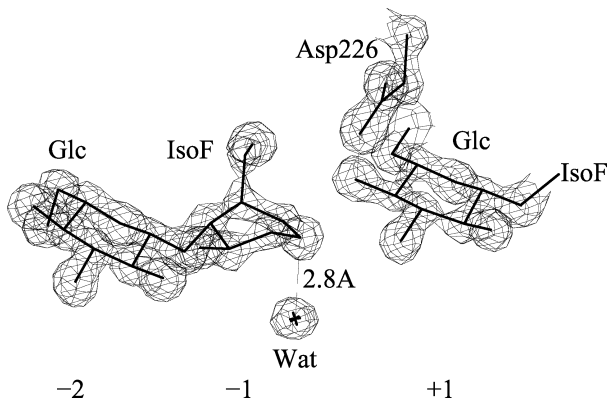


Fig. 1 Electron density for **1** bound to Cel6A. The map is a $2F_{obs}$–F_{calc} synthesis at 1.6 e Å^{-3}. Only the −2,−1 and +1 sites are shown.

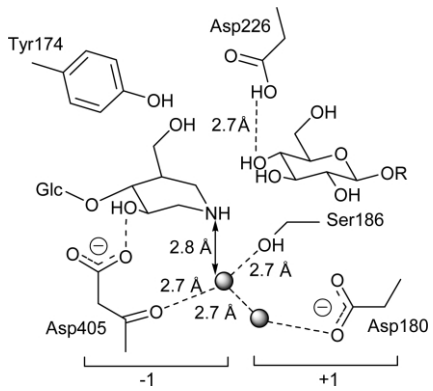


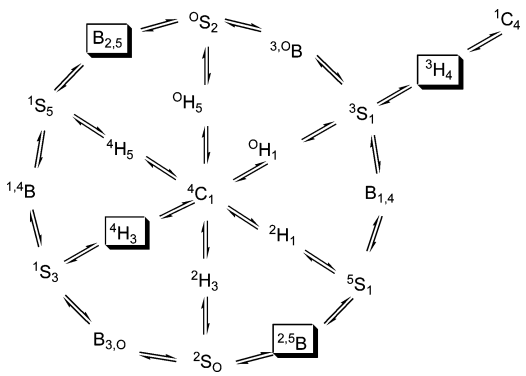
Fig. 2 Interactions between Cel6A and the two molecules of **1**. Only the −1 and +1 subsites are shown for simplicity. Water molecules are shown as shaded spheres.

DOI: 10.1039/b301592k

In addition to posing questions about distortion during catalysis, discussed below, the complex also sheds light onto the nature of the catalytic apparatus in this family where the identity of a residue to provide Brønsted base-catalysed assistance to the hydrolytic water molecule has remained elusive. Several candidates have been discussed[6] with a "Grotthus"-style mechanism *via* a string of solvent molecules also a possibility.[7] The Cel6A complex with **1** reveals a solvent molecule ideally positioned for "inverting" attack at the anomeric centre, Fig. 2, 2.8 Å from the imino nitrogen. There is no candidate base within hydrogen-bonding distance of this water molecule in this complex. Instead, the water molecule interacts with the peptide main-chain carbonyl of Asp405, the side-chain of Ser186 and a second solvent molecule, which is itself hydrogen-bonded to the carboxylate of Asp180. In the absence of conformational change of either substrate or enzyme, deprotonation of the attacking water *via* a solvent chain is most consistent with the 3-D structure of Cel6A with **1**.

It has been estimated that β-glucosidases hydrolyse the glycosidic bond with rates up to 10^{17}-fold above the uncatalysed rate. This is one of the most powerful enzymatic rate enhancements known and implies a 10^{-22} M affinity for the enzyme transition state complex,[8] yet despite this affinity, and the potential to exploit glycosidases as therapeutic targets, the exact conformation of the transition state harnessed by each enzyme has remained elusive. 3-D structural analysis of enzymes trapped in various, necessarily stable, states along the reaction coordinate has supported a $^4H_3$ half-chair transition state for a number of retaining β-glucosidases and related enzymes.[9] Recent trapping of a β-mannanase in $^1S_5$ conformation and its subsequent covalent intermediate in $^OS_2$ suggest a different catalytic strategy[10] *via* a $B_{2,5}$ transition state. A feature of both these conformational agendas is a close interaction between a pseudo-equatorial 2-OH substituent and the carbonyl group of the catalytic nucleophile, which in some cases is believed to contribute in excess of 40 kJ mol^{-1} to catalysis.[11] Some enzymes do not appear to harness such interactions (Antoni Planas, personal communication), indeed inverting β-glucosidases are neither restricted to such a conformational pathway nor are they in a position to harness such an interaction, since the attacking group is a water molecule and not an enzymatic carboxylate.

It seems likely that all the possible transition state conformations, Scheme 1, are harnessed by different enzyme classes.[9] In the case of family GH-6 cellulases, observation of a ^2S_O "Michaelis" complex and a distorted $^{2,5}B$ conformation for **1** is most simply interpreted as reflecting an active-centre environment that favours binding of the $^{2,5}B$ transition state; perhaps *via* a $^2S_O \rightarrow {}^{2,5}B$ itinerary. Use of an alternate conformation would require conformational change of the active centre groups, particularly those that favour a pseudo-axial position for C6–O6. Furthermore, a recent 0.8 Å structure of an unrelated inverting β-glucosidase, the endoglucanase Cel(8)A from *Clostridium thermocellum*, also reveals a $^{2,5}B$ conformation[12] for unhydrolysed substrate, whilst family GH-11 xylanases reveal a similarly conformed covalent intermediate[13] supporting a wider applicability for such a transition state.



Scheme 1 Partial pyranose-ring pseudo-rotational itinerary (adapted from Stoddart[15]). Potential transition state conformations that possess C5, O5, C2 and C1 co-planar are shown boxed.

Piperidine derivatives, whilst often powerful inhibitors, are not considered "transition-state mimics" since their most-favoured ($^4C_1$) conformation does not reflect any potential transition-state shape.[14] The synthesis and analysis of conformationally-locked inhibitors may therefore provide a powerful route to novel, enzyme-specific inhibitors as opposed to those that are generic or merely fortuitous.

Novozymes A/S, in particular the late Martin Schülein, are thanked for provision of Cel6A. The Biotechnology and Biochemical Sciences Research Council and the European Union are thanked for funding. G.J.D. is a Royal Society University Research Fellow.

Notes and references

‡ *Experimental*: K_i determination from a plot of K_M/k_{cat} *vs.* [I] followed the method of I. Matsui, K. Ishikawa E. Matsui, S. Miyairi, S. Fukui and K. Honda *J. Biochem. (Tokyo)*, 1991, **109**, 566. Kinetics were performed with Cel6A at 25 nM, 500 μM cellopentaose in 2 ml of 50 mM sodium phosphate buffer pH 7.0 containing 1 mg ml^{-1} bovine serum albumin. **1** was added at final concentrations of 0, 500 nM, 1 μM and 2 μM, respectively. Parallel reactions, at 37 °C, were sampled at regular time intervals (up to 40 min) in 200 μl aliquots, the enzyme was inactivated by boiling and the reaction products were quantified by high performance anion-exchange chromatography (as in J. Hall, G. W. Black, L. M. Ferreira, S. J. Millward-Sadler, B. R. Ali, G. P. Hazlewood and H. J. Gilbert, *Biochem. J.*, 1995, **309**, 749). The D416A mutant of *H. insolens* Cel6A, chosen for ease of crystallisation,[5] was crystallised from 21% polyethylene glycol monomethyl ether 5K with 200 mM calcium acetate in 100 mM HEPES pH 7.5 following a 1 h pre-incubation with 5 mM **1**. Data were collected at 100 K at the European Synchrotron Radiation Facility beamline ID14-EH2 (λ = 0.934 Å). Data were processed with MOSFLM and other computing used the CCP4 package (Collaborative Computational Project, Number 4, *Acta Crystallogr., Sect. D*, 2000, **50**, 760). The structure was solved by molecular replacement using PDB code 1BVW as the search model (wild-type Cel6A). The structure was refined using REFMAC (G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr., Sect. D*, 1997, **53**, 240). Coordinate accession code 1ocn.

1 G. Davies, M. L. Sinnott and S. G. Withers, in *Comprehensive Biological Catalysis*, ed. M. L. Sinnott, Academic Press, London, 1997, **vol. 1**, p. 119.
2 J. M. Macdonald, R. V. Stick, D. M. G. Tilbrook and S. G. Withers, *Aust. J. Chem.*, 2002, **55**, 747.
3 http://afmb.cnrs-mrs.fr/CAZY/.
4 J.-Y. Zou, G. J. Kleywegt, J. Ståhlberg, H. Driguez, W. Nerinckx, M. Claeyssens, A. Koivula, T. T. Teeri and T. A. Jones, *Structure*, 1999, **7**, 1035.
5 A. Varrot, T. Frandsen, H. Driguez and G. J. Davies, *Acta Crystallogr., Sect. D*, 2002, **58**, 2201.
6 A. Varrot, M. Schülein and G. J. Davies, *Biochemistry*, 1999, **38**, 8884.
7 A. Koivula, L. Ruohonen, G. Wohlfahrt, T. Reinikainen, T. T. Teeri, K. Piens, M. Claeyssens, M. Weber, A. Vasella, D. Becker, M. L. Sinnott, J. Y. Zou, G. J. Kleywegt, M. Szardenings, J. Stahlberg and T. A. Jones, *J. Am. Chem. Soc.*, 2002, **124**, 10015.
8 R. Wolfenden, X. Lu and G. Young, *J. Am. Chem. Soc.*, 1998, **120**, 6814.
9 G. J. Davies, V. M.-A. Ducros, A. Varrot and D. L. Zechel, *Biochem. Soc. Trans.*, 2003, in press.
10 V. Ducros, D. L. Zechel, G. N. Murshudov, H. J. Gilbert, L. Szabo, D. Stoll, S. G. Withers and G. J. Davies, *Angew. Chem., Int. Ed.*, 2002, **41**, 2824.
11 M. N. Namchuk and S. G. Withers, *Biochemistry*, 1995, **34**, 16194.
12 D. M. A. Guerin, M. B. Lascombe, M. Costabel, H. Souchon, V. Lamzin, P. Beguin and P. M. Alzari, *J. Mol. Biol.*, 2002, **316**, 1061.
13 E. Sabini, G. Sulzenbacher, M. Dauter, Z. Dauter, P. L. Jørgensen, M. Schülein, C. Dupont, G. J. Davies and K. S. Wilson, *Chem. Biol.*, 1999, **6**, 483.
14 S. G. Withers, M. Namchuk and R. Mosi, *Potent glycosidase inhibitors: transition state mimics or simply fortuitous binders?*, in *Iminosugars as glycosidase inhibitors: nojirimycin and beyond*, ed. A. E. Stutz, Wiley-VCH, Weinheim, 1999, p. 188.
15 J. F. Stoddart, *Stereochemistry of Carbohydrates*, Wiley Interscience, Toronto, 1971.