Lateral Protonation of a Glycosidase Inhibitor. Structure of the *Bacillus agaradhaerens* Cel5A in Complex with a Cellobiose-Derived Imidazole at 0.97 Å Resolution

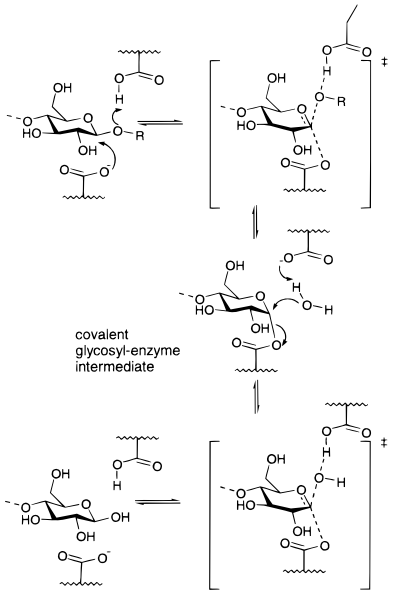
Annabelle Varrot,[§] Martin Schülein,[||] Muriel Pipelier,[†] Andrea Vasella,[†] and Gideon J. Davies*[,§]

Department of Chemistry, University of York Heslington, York YO10 5DD, England
Novo-Nordisk A/S, Novo Allé, DK-2880 Bagsvaerd, Denmark
Laboratorium für Organische Chemie, ETH-Zentrum Universitätstrasse 16, CH-8092 Zürich, Switzerland

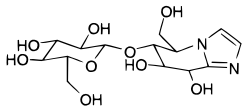
Received December 7, 1998

Inhibitors of glycosidase activity are of interest as both therapeutic agents and mechanistic probes.[1−3] The enzymes cleaving glycosidic bonds, the glycoside hydrolases, are currently classified into over 71 sequence-based families,[4] with 3-dimensional structural representatives available for over 28 of these.[5,6] These enzymes have also been classified into 4 groups, based upon whether they invert or retain the anomeric configuration and whether they cleave axial or equatorial glycosides.[7,8] The paradigmatic mechanism for retaining enzymes was provided by Koshland:[9] a covalent glycosyl-enzyme intermediate is formed and subsequently hydrolyzed, with general acid/base assistance, via oxocarbenium-ion-like transition states, Scheme 1. Mechanistic schemes for the cleavage of β-D-glycosides generally indicate protonation of the glycosidic oxygen from above, and perpendicular to, the sugar plane. The study of inhibitors, however, has now shown that a large number of glycosidases do not protonate from this direction.[10] Many, if not all, utilize side-on or "lateral" protonation in which a proton is transferred in the plane of the glycon. This proton donation may take place from one of two directions with respect to the plane defined by O(1), C(1), and H(1), i.e. it may be either *anti* or *syn* to the pyranoside endocyclic O(5)−C(1) bond. Not surprisingly, there is a correlation between glycosidase families[4−6] defined by the amino acid sequence and hence 3-D structure, and their *syn*- or *anti*-protonation trajectory.[11] On the basis of complexed crystal structures for β-D-glycosidases, families 1, 2, 5, and 10, all belonging to the GH-A clan,[4,12] may be classified as *anti*-protonators as may the family 18 Chitinase[13] and the family 20 chitobiase.[14] *Syn*-protonators are found in families 7, 11, 12, 16, 22, 23, and 45.[15−21]

Scheme 1. Canonical Mechanism for Retaining β-Glycoside Hydrolases such as Cel5A



The success of *gluco*-configured imidazoles as glycosidase inhibitors[23−26] led us to synthesize the cellobiose-derived imidazole (**1**) as a potential glycosidase inhibitor.[22] The disaccharide moiety serves to improve binding and to characterize enzymes possessing multiple subsites. Two key features of these compounds contribute to their binding efficiency. They possess a $^4H_3$ or 4-*sofa* conformation mimicking the one expected for the oxocarbenium-ion-like transition state, while the imidazole ring provides a nonbonding doubly occupied nitrogen orbital for (partial) *anti*-protonation (in the σ-plane) and is sufficiently electronegative to permit a simultaneous interaction with the catalytic nucleophile (in the π-plane).[27] Such compounds are excellent glycosidase inhibitors but, until now, their expected mode of binding has not been established by a 3-D structure. Here we present the X-ray structure of the cellobiose-derived imidazole (**1**) bound to Cel5A from *B. agaradhaerens* at atomic (0.97 Å) resolution.



1

* Corresponding author. E-mail: davies@yorvic.york.ac.uk. Fax: 44-1904-410519. Phone: 44-1904-432596.
[§] University of York.
[||] Novo Allé.
[†] ETH-Zentrum.

(1) Legler, G. *Adv. Carbohyd. Chem. Biochem.* **1990**, *48*, 319−385.
(2) von Itzstein, M.; Colman, P. *Curr. Opin. Struct. Biol* **1996**, *6*, 703−709.
(3) Jacob, G. S. *Curr. Opin. Struct. Biol.* **1995**, *5*, 605−611.
(4) Henrissat, B.; Bairoch, A. *Biochem. J.* **1996**, *316*, 695−696.
(5) Davies, G.; Henrissat, B. *Structure* **1995**, *3*, 853−859.
(6) Henrissat, B.; Davies, G. J. *Curr. Opin. Struct. Biol.* **1997**, *7*, 637−644.
(7) Sinnott, M. L. Glycosyl group transfer. In *Enzyme Mechanisms*; Page, M. I., Williams, A., Eds.; Royal Society of Chemistry: London, 1987; pp 259−297.
(8) Davies, G.; Sinnott, M. L.; Withers, S. G. Glycosyl Transfer. In *Comprehensive Biological Catalysis*; Sinnott, M. L., Ed.; Academic Press: London, 1997; Vol. 1, pp 119−209.
(9) Koshland, D. E. *Biol. Rev.* **1953**, *28*, 416−436.
(10) Heightman, T. D.; Locatelli, M.; Vasella, A. *Helv. Chim. Acta* **1996**, *79*, 2190−2200.
(11) Heightman, T.; Vasella, A. *Angew. Chem.* In press.
(12) Henrissat, B.; Callebaut, I.; Fabrega, S.; Lehn, P.; Mornon, J.-P.; Davies, G. *P.N.A.S. (USA)* **1995**, *92*, 7090−7094.
(13) Terwisscha van Scheltinga, A. C.; Armand, S.; Kalk, K. H.; Isogai, A.; Henrissat, B.; Dijkstra, B. W. *Biochemistry* **1995**, *34*, 15619−15623.
(14) Tews, I.; Perrakis, A.; Oppenheim, A.; Dauter, Z.; Wilson, K. S.; Vorgias, C. E. *Nat. Struct. Biol.* **1996**, *3*, 638−648.
(15) Sulzenbacher, G.; Driguez, H.; Henrissat, B.; Schülein, M.; Davies, G. J. *Biochemistry* **1996**, *35*, 15280−15287.
(16) Törrönen, A.; Harkki, A.; Rouvinen, J. *EMBO J.* **1994**, *13*, 2493−2501.
(17) Sulzenbacher, G.; Shareck, F.; Morosoli, R.; Dupont, C.; Davies, G. J. *Biochemistry* **1997**, *36*, 16032−16039.
(18) Strynadka, N. C. J.; James, M. N. G. *J. Mol. Biol.* **1991**, *220*, 401−424.
(19) Weaver, L. H.; Grütter, M. G.; Matthews, B. W. *J. Mol. Biol.* **1995**, *245*, 54−68.
(20) Kuroki, R.; Weaver, L. H.; Matthews, B. W. *Science* **1993**, *262*, 2030−2033.
(21) Davies, G. J.; Tolley, S. P.; Henrissat, B.; Hjort, C.; Schülein, M. *Biochemistry* **1995**, *34*, 16210−16220.
(22) Piens, K.; Pipelier, M.; Vonhoff, S.; Vasella, A. *Helv. Chim. Acta.* In press.
(23) Granier, T.; Panday, N.; Vasella, A. *Helv. Chim. Acta* **1997**, *80*, 979−987.
(24) Tatsuta, K.; Ikeda, Y.; Miura, S. *J. Antibiot.* **1996**, *49*, 836−838.
(25) Tatsuta, K.; Miura, S.; Ohta, S.; Gunji, H. *Tetrahedron Lett.* **1995**, *36*, 1085−1088.
(26) Tatsuta, K.; Miura, S.; Ohta, S.; Gunji, H. *J. Antibiot.* **1995**, *48*, 286−288.
(27) Heightman, T. D.; Vasella, A.; Tsitsanou, K. E.; Zographos, S. E.; Skamnaki, V. T.; Oikonomakos, N. *Helv. Chim. Acta* **1998**, *81*, 853−864.

10.1021/ja984238n CCC: $18.00
Published on Web 03/09/1999

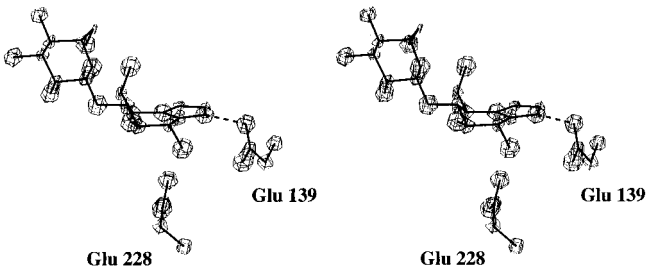


Figure 1. Electron density for the Cel5A cellobiose-derived imidazole complex at 0.97 Å resolution. The catalytic acid/base is Glu 139 and the nucleophile is Glu 228. The map is a maximum-likelihood weighted $2F_{obs} - F_{calc}$, α_{calc} synthesis, contoured at 2.4 electrons/Å^3, and is in divergent stereo.

Cel5A is a typical β-retaining (e−e) glycoside hydrolase belonging to Clan GH-A utilizing an *anti*-protonation trajectory whose native structure has been determined at atomic (0.95 Å) resolution.[28] The reaction pathway has been studied kinetically and by the crystal structure analysis of a series of complexes including analogues of substrate, intermediate, and product.[29] This revealed key features of the glycosidase mechanism such as pyranoside-ring distortion to a $^1S_3$ skew-boat to both facilitate in-line nucleophilic attack and provide beneficial stereoelectronics. The ligands in these complexes, however, all possess a tetrahedral anomeric center. The cellobiose-derived imidazole possessing a trigonal anomeric center is ideally suited to characterize further the catalytic machinery since compounds of this type are (partial) transition-state analogues, as quantified for a *gluco*-derived tetrazole.[30]

The cellobiose-derived imidazole (**1**) inhibits Cel5A with a K_i of 88 μM,[31] in the range observed for imidazo-pyridine-type inhibitors,[23] while the K_i of cellotriose is just 4.5 mM, in keeping with the expected character of the imidazole as a partial transition-state analogue. The 3-D structure of Cel5A, in complex with **1**, has been solved at atomic (0.97 Å) resolution.[32] It binds in the −2 and −1 subsites. The imidazole nitrogen accepts a proton from the enzymatic acid, Glu 139, with a N−O distance of 2.58 Å, Figure 1. The protonation trajectory lies in the σ-plane of the imidazole, as predicted. Residual electron density (not shown) indicates the position of the proton and shows that, in keeping with the approximate pK_a values of the carboxyl group and the imidazolinium-ion, the proton lies closer to the imidazole nitrogen.[31] The catalytic nucleophile is located below the plane of the imidazole with a distance of 3.15 Å between the OE2 of the carboxylate and the "anomeric" carbon of the inhibitor, Figure 2. The conformation of the −1 subsite pyranoside ring is closest to a 4-*sofa* with C(6)−O(6) in a *trans-gauche* configuration. The two saccharide rings are twisted relative to each other, as expressed by the torsion angles φ (O-5, C-1, O-4, C-4) of −76.3° and ψ (C-1, O-4, C-4, C-3) of +139.3°, as compared to φ = −94.4° and ψ = +91.7° obtained for methyl β-cellotriose.[33] This twist leads to an extension of the O-3···O-5 intramolecular H-bond from 2.84 Å in cellotriose to 3.42 Å in the imidazole−enzyme complex. Such a disruption of the O-3···O-5 interaction has been observed in many cellulase oligosaccharide complexes and is considered to facilitate hydrolysis of cellulose by destabilizing the intramolecular association.

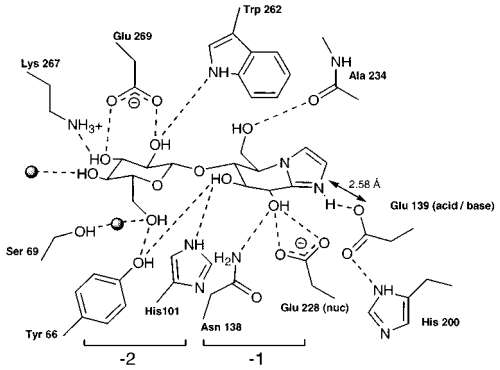


Figure 2. Schematic representation of the interactions between Cel5A and the cellobiose-derived imidazole inhibitor.

The cellobiose−imidazole complex is complementary to the DNP-2F cellobioside structure.[29] The two structures display identical positions for the distal sugar moiety in the −2 subsite, and for the O(4), C(4), C(3), O(3), and C(5) atoms of the −1 subsite ring and C(6)−O(6) move only slightly. The main changes between the 4-sofa and $^1S_3$ skew-boat conformations manifest themselves at O(5) and O(1). The trigonal "anomeric" carbon of the imidazole implies a planar, i.e., close to equatorial orientation of the "glycosidic" nitrogen, while the sp^3 anomeric carbon in the DNP-2F cellobioside structure defines the conformational change leading to a pseudoaxial orientation of the C(1)−O bond. The partial TS character of the imidazole thus results from two different structural aspects: the sp^2 anomeric center and hence half-chair conformation plus the lateral protonation of an equatorial alkoxy group. These two features possibly mimick different points along the reaction coordinate.

Surprisingly, the protonation apparatus remains relatively static in the native, substrate, and imidazole complexes. The glycosidic oxygen moves 1.4 Å above the ring plane in the (distorted) $^1S_3$ substrate complex, yet still accepts a 2.9 Å hydrogen bond from the similarly positioned catalytic acid. This leads us to suggest that one reason for the universal occurrence of carboxylates as the protonation machinery in glycoside hydrolases is their ability to protonate a center whose position changes markedly along the reaction trajectory. The Cel5A cellobiose-imidazole structure confirms the concept of side-on "lateral" protonation for glycoside hydrolases and demonstrates the potential of imidazoles as inhibitors and mechanistic probes. Appropriate glycosidase−inhibitor complexes can thus reveal complementary structural aspects of catalysis and define different positions along the enzymatic reaction coordinate.

Acknowledgment. The authors thank the European Union and the BBSRC for support. Data collection at the EMBL Hamburg Outstation was supported by the TMR/LSF program. G.J.D. is a Royal Society University Research Fellow.

JA984238N

(28) Davies, G. J.; Dauter, M.; Brzozowski, A. M.; Bjornvad, M. E.; Anderson, K. V.; Schülein, M. *Biochemistry* **1998**, *37*, 1926−1932.

(29) Davies, G. J.; Mackenzie, L.; Varrot, A.; Dauter, M.; Brzozowski, A. M.; Schülein, M.; Withers, S. G. *Biochemistry* **1998**, *37*, 11707−11713.

(30) Emmert, P.; Vasella, A.; Weber, M.; Rupitz, K.; Withers, S. G. *Carbohydr. Res.* **1993**, *250*, 113−128.

(31) Cel5A inhibition was determined by steady-state kinetics with reduced cellopentaose as substrate in a linked assay with cytochrome *c* and the *Humicola insolens* cellobiose dehydrogenase. (a) Schou, C.; Christensen, M. H.; Schülein, M. *Biochem. J.* **1998**, *330*, 565−571. (b) Schülein, M. *J. Biotechol.* **1997**, *57*, 71−81. The K_i for cellotriose was determined by using PNP−cellobiose as substrate. The pK_a value of the glucose-derived imidazole is 6.1 (Panday, N.; Vasella, A., in preparation).

(32) Crystals of the orthorhombic form of Cel5A were bathed overnight in a solution containing the cellobiose-imidazole. Powdered inhibitor was added to the edge of the droplet and allowed to diffuse slowly into the crystal under constant observation. Data to 0.97 Å resolution were collected at 100 K on beamline BW7B of the EMBL Hamburg outstation. Data were scaled and reduced with use of the HKL suite: (a) Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology*; Carter, C. W., Jr., Sweet, R. M., Eds.; Academic Press: London and New York, 1997; Vol. 276, pp 307−326. The structure was refined with REFMAC: (b) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr.* **1997**, *D 53*, 240−255. The final data are 93% complete with R_{merge} of 0.032, a mean $I/\sigma I$ of 45, and a multiplicity of 5.4 observations/reflection. The final model structure has 2472 protein atoms, 499 solvent waters, 7 glycerol molecules, single sulfate and acetate ions and 25 atoms for the cellobiose-derived inhibitor. The R_{cryst} is 0.10 with R_{free} = 0.12, and coordinates have been deposited with the Brookhaven Protein Databank under accession code 8A3H.

(33) Raymond, S.; Henrissat, B.; Qui, D. T.; Kvick, A.; Chanzy, H. *Carbohydr. Res.* **1995**, *277*, 209−229.