Probing the β-1,3:1,4 glucanase, *Ct*Lic26A, with a thio-oligosaccharide and enzyme variants†

Victoria A. Money,[a] Alan Cartmell,[b] Catarina I. P. D. Guerreiro,[c] Valerie M.-A. Ducros,[a] Carlos M. G. A. Fontes,[c] Harry J. Gilbert[b] and Gideon J. Davies*[a]

Received 13th December 2007, Accepted 10th January 2008
First published as an Advance Article on the web 4th February 2008
DOI: 10.1039/b719288f

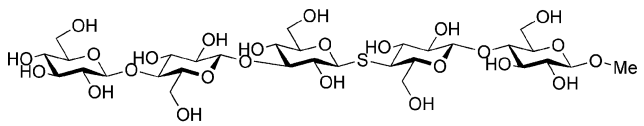
The substrate binding regions of a β-1,3:1,4 glucanase are revealed through structural analysis with a thio-oligosaccharide and kinetics of enzyme variants.

The rekindling of interest in enzymatic biomass conversion, to address the pressing energy needs of the 21st century, yet again highlights the need to dissect and understand the complex enzymology of this process. One class of substrates, and their degradation, stands out for examination: plant polysaccharides. We know, from Wolfenden's seminal calculations,[1] that plant polysaccharide hydrolysing enzymes are amongst the most efficient enzymatic catalysts, as judged by the ratio of the catalysed and uncatalysed rate constants k_{cat}/k_{uncat}. Yet, for many classes of enzyme, it is still unclear how binding energy, from diverse substrates, is reflected in catalysis. Despite a plethora of 3-D structures for polysaccharidases, insightful complexes remain the exception rather than the rule. The plant polysaccharide β-1,3:1,4 glucan is a candidate for degradation, *via* smaller oligosaccharides and glucose, to biofuels. Mixed linked β-1,3:1,4 glucans occur predominantly in cereal grains, but are also abundant in mosses and lichens and show a higher degree of solubility than many similar plant polysaccharides such as cellulose, a β-1,4 glucose polymer. The enzymatic hydrolysis of β-1,3:1,4 glucan is performed by a battery of glycoside hydrolases (hereafter GHs), notably the *endo*-acting β-1,3:1,4 glucanases, which are found in a variety of sequence-based families (www.cazy.org).

Attention in our laboratory has focussed on a family "GH26" β-1,3:1,4 glucanase (hereafter *Ct*Lic26A[2]) from the organism *Clostridium thermocellum*, whose genome sequence has recently been determined (http://genome.jgi-psf.org/) in response to the potential utility of the enzymatic consortia of this bacterium for biomass conversion to biofuels.[3–5] Indeed the inclusion of *Ct*Lic26A in the toolbox of enzymes used in biofuel production from biomass will make a significant contribution in accessing mixed-linked glucans, which are abundant in many cereals.

*Ct*Lic26A hydrolyses both polymeric β-1,3:1,4 glucans and mixed linkage β-1,3:1,4 aryl oligosaccharides such as methylumbelliferyl di and trisaccharides.[2] Catalysis is performed with net retention of anomeric configuration *via* a covalent glycosyl-enzyme intermediate flanked by oxocarbenium-ion-like transition states (for a mechanistic review see ref. 6). The 3-D structure has been solved in native form, with a number of imino/aza sugar inhibitors based upon extended isogfagomine and noeuromycin templates,[2,7] and in a series of complexes shot along the reaction coordinate including the unhydrolysed Michaelis complex and the trapped covalent glycosyl enzyme intermediate.[8] Yet, despite this array of information we have, until now, been unable to dissect the interactions of distal subsites, partially reflecting our failure to map "leaving group" positive subsite interactions through previous approaches (subsite nomenclature described in ref. 9).

One approach that has proved insightful to study glycosidase–ligand interactions is the use of thio-oligosaccharide substrate mimics in which one, or more, of the labile *O*-glycosidic bonds is replaced by a non-hydrolysable *S*-glycosidic linkage. Such a strategy[10] has revealed many different binding modes, most notably distorted ligands spanning the active centre (examples include ref. 11,12) but also undistorted ligands evading the catalytic apparatus on enzymes from different classes.[13–15] A series of *S*-linked β-1,3:1,4 gluco-oligosaccharides, for the study of β-1,3:1,4 glucanases,[16] has been reported. Unlike the majority of previous thio-oligosaccharides, K_i values for these mixed linkage species were high, reflecting poor binding. Here we report the 3-D structure, at 1.5 Å resolution, of *Ct*Lic26A in complex with one such thiopentasaccharide **1**, and harness the structural data to construct seven active-site variants, which have been used to probe substrate binding and catalysis in this important class of polysaccharide-degrading enzymes.



1

Crystallisation conditions were screened, by co-crystallisation, for a variety of different β-1,3:1,4 gluco-oligosaccharides (synthesis described in ref. 16). Crystals were obtained with only one compound (**1**) but this form allowed excellent high resolution 3-D structural analysis,[17] Table 1, Fig. 1. Crystals are in space group $P2_12_12_1$ with cell dimensions of $a = 49.3$ Å, $b = 63.0$ Å and $c = 78.3$ Å and with a single molecule of *Ct*Lic26A in the asymmetric unit. Electron density revealed that although **1** did not span the active-centre, it did reveal the protein–ligand interactions of the leaving group subsites +1 and +2. Electron density, Fig. 1(b), is most appropriately modelled as two molecules of **1**. The first has its three non-reducing end sugars disordered in solvent with

[a]*Structural Biology Laboratory, Department of Chemistry, The University of York, Heslington, York, U.K. YO10 5YW. E-mail: davies@ysbl.york.ac.uk; Fax: 44-1904-32826; Tel: 44-1904-328260*
[b]*Institute for Cell and Molecular Biosciences, Newcastle University, 2nd floor Catherine Cookson Building, The Medical School Framlington Place, Newcastle upon Tyne, U.K. NE2 4HH*
[c]*CIISA-Faculdade de Medicina Veterinária, Universidade Técnica de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal*

† Coordinates of the *Ct*Lic26A structure in complex with compound **1** have been deposited with the PDB with accession code 2vi0.

Table 1 X-Ray data and structure quality statistics

	*Ct*Lic26A with **1**
Resolution of data (Å)	34.88–1.51
R_{merge}	0.09 (0.496)
Mean $I/\sigma I$	11.8 (2.0)
Completeness %	95.8 (72.0)[a]
Multiplicity	5.1 (3.0)
R_{cryst}	0.16
R_{free}	0.19
r.m.s.d 1–2 bonds (Å)	0.010
r.m.s.d 1–3 bonds (°)	1.3

[a] High resolution incompleteness reflects integration into the corners of a square detector (the 1.7 to 1.6 Å shell is 99.9% complete).

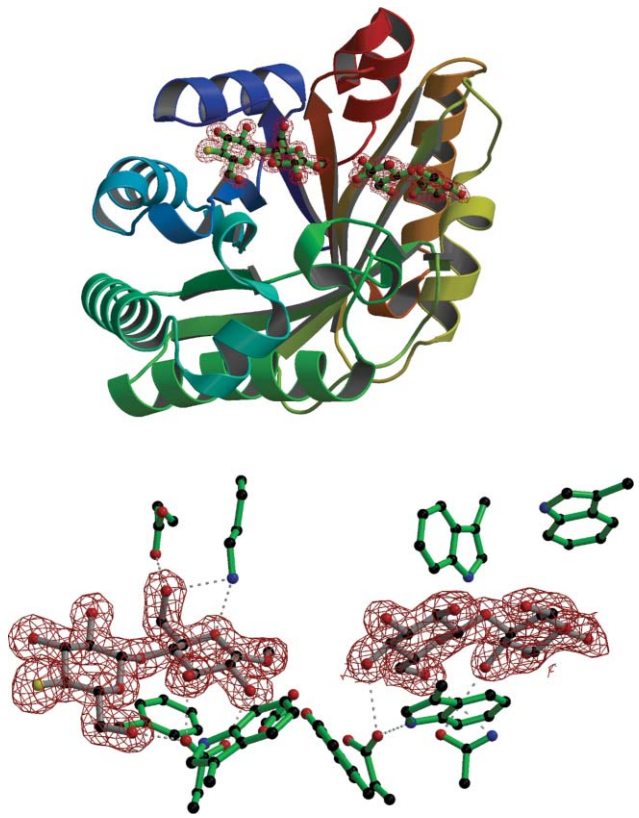


Fig. 1 3-D Structure of *Ct*Lic26A with 1: (A) overall protein cartoon (B) electron density for **1**, in the −3/−2 and +1/+2 subsites.

the two reducing-end (β-1,4 linked) glucosides occupying subsites −3 and −2. In contrast, the second molecule binds with its non-reducing end disaccharide in subsites +1 and +2, but then becomes disordered in solvent.

Unlike the majority of thio-based glucan mixed-linked inhibitors, molecule **1** displayed an affinity for *Ct*Lic26A, (K_i of 75 μM) that was similar (assuming $K_m \sim K_s$ when $k_2 \ll k_{-1}$) to the corresponding oligosaccharide substrate, Glc-β1,4-Glc-β1,3-Glc-methylumbelliferyl, which has a K_m of 112 μM (Fig. 2). Kinetic[18] analyses were performed both on high viscosity β-glucan and on Glc-β1,4-Glc-β1,3-Glc-methylumbelliferyl, Table 2. Both substrates reveal the importance of aromatic residues both in providing hydrophobic platforms for sugar binding (for example, the variants of F41 and W114 show decreases in the apparent

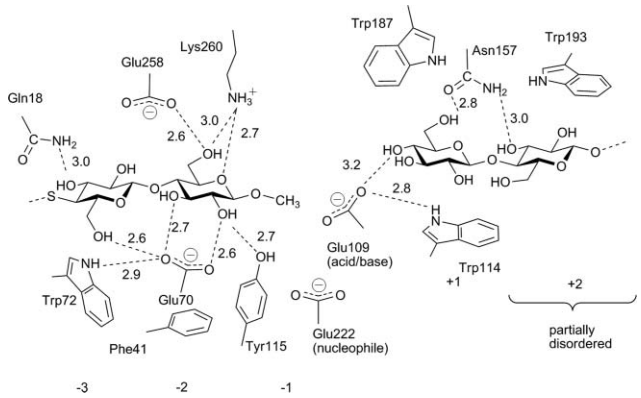


Fig. 2 Schematic diagram of the interactions of *Ct*Lic26A with **1**.

second order rate constant k_{cat}/K_m of 200-fold to essentially inactive on the methylumbelliferyl trisaccharide) and through hydrogen bonding exemplified by Y115 whose mutation to alanine reduces k_{cat}/K_m, relative to wild-type, by 15-fold on β-glucan and up to 1200-fold on the aryl oligosaccharide. The influence of the residues that comprise the hydrophobic platform on catalysis and substrate binding, however, are quite different. Thus, the F41A mutation does not alter K_m but has a dramatic effect on k_{cat} for both the oligosaccharide and polysaccharide substrates. These data indicate that the hydrophobic interaction of the substrate at the −2 subsite plays a critical role in stabilizing the transition state, possibly because it assists in distorting the sugar at the −1 subsite into its $^4H_3$ transition state[8] configuration by tethering the O3 bond. The K260A mutation reduces binding at the −2 subsite through the loss of polar interactions with O6 and the endocyclic oxygen causing a ~10-fold increase in K_m and a similar reduction in k_{cat}. The W14A and W72A amino acid substitutions, which influence binding at the +1 and −3 subsites, respectively, result in a substantial decrease in oligosaccharide substrate affinity ($K_m \gg 1$ mM), but do not have a significant effect on the hydrolysis of the polysaccharide. These hydrophobic platforms do not directly contribute to maintaining the conformation of the transition state, but likely stabilize the Michaelis complex of the enzyme. The large increase in K_m and decrease in k_{cat} for the oligosaccharide substrate in the −3 mutant Q18A also suggests that this distal subsite contributes to both substrate binding in the ground state and transition state stabilization.

As with the majority of *endo*-acting glycoside hydrolases (for example ref. 19,20), disruption of subsite interactions has a far greater effect on oligosaccharide substrates than highly polymeric glucans. Mutations in the negative subsites are, in the vast majority of cases, more deleterious than those in the positive subsites. Furthermore, the W114 mutation, which disrupts hydrophobic stacking in +1, is considerably more damaging to catalysis than mutation of N157; the only residue which hydrogen-bonds to the +1/+2 sugars. Indeed, in retaining glycoside hydrolases, mutations in the distal positive binding sites can sometimes lead to an elevation in k_{cat} by increasing the efficiency of product release after glycosylation,[20] while the removal of amino acids at the +1 subsite generally cause a significant reduction in catalytic activity against all substrates.[19–21]

Although the thio-oligosaccharides screened did not deliver an active-centre spanning complex, they did unveil the interactions of

Table 2 Enzyme kinetics of mutant and wild-type (WT) proteins on fluorescent oligosaccharide and polysaccharide substrates

	MeUMB			β-Glucan		
Enzyme	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m (mM^{-1} min^{-1})	K_m (mg mL^{-1})	k_{cat} (min^{-1})	k_{cat}/K_m (mL mg^{-1} min^{-1})
WT	110 ± 18	23 ± 1.5	205	0.34 ± 0.06	$74 \times 10^3 \pm 6 \times 10^3$	219×10^3
Q18A	1763 ± 890	2.1 ± 0.44	1.2	0.67 ± 0.1	$23 \times 10^3 \pm 2 \times 10^3$	35×10^3
F41A	139 ± 23	0.001 ± 0.0001	0.001	1.3 ± 0.36	$1.6 \times 10^3 \pm 0.1 \times 10^3$	1×10^3
E70A	155 ± 18	0.005 ± 0.0003	0.32	1.3 ± 0.13	$8.1 \times 10^3 \pm 2 \times 10^3$	6×10^3
W72A	—	—	0.66 ± 0.08[a]	0.46 ± 0.08	$135 \times 10^3 \pm 5 \times 10^3$	294×10^3
W114A	—	—	0.16 ± 0.002[a]	1.4 ± 0.14	$85 \times 10^3 \pm 4 \times 10^3$	62×10^3
Y115A	260 ± 56	0.05 ± 0.004	0.17	1.0 ± 0.18	$15 \times 10^3 \pm 0.4 \times 10^3$	15×10^3
N157A	11 ± 2	0.73 ± 0.04	64	0.49 ± 0.09	$78 \times 10^3 \pm 9 \times 10^3$	158×10^3
E258A	188 ± 24	1.6 ± 0.06	8.3	0.55 ± 0.07	$26 \times 10^3 \pm 1 \times 10^3$	46×10^3
K260A	1162 ± 232	3.2 ± 0.25	2.8	0.65 ± 0.12	$5 \times 10^3 \pm 0.1 \times 10^3$	7×10^3

[a] Estimated from linear regression.

the +1 and +2 subsites which had previously proved recalcitrant to observation despite our previous analysis of a large panel of ligands. Thus, although somewhat serendipitous, the *Ct*Lic26A complex reported here highlights the benefits both of thio-oligosaccharides and of the deployment of a toolbox of approaches to study protein–carbohydrate interactions. If enzymatic biomass conversion is to become both facile and economically viable, our understanding of protein–carbohydrate interactions will have the continuing development of such probes, to dissect molecular interactions at its core.

Acknowledgements

The authors would like to thank the BBSRC and Fundação para a Ciência e a Tecnologia (POCI/BIA-PRO/59118/2004 and SFRH/BD/16731/2004 for CG) for funding. GJD is a Royal Society-Wolfson Research Merit Award recipient. We thank Dr Hugues Driguez for provision of compound **1** and related inhibitors.

Notes and references

1 R. Wolfenden, X. D. Lu and G. Young, *J. Am. Chem. Soc.*, 1998, **120**, 6814–6815.
2 E. J. Taylor, A. Goyal, C. I. P. D. Guerreiro, J. A. M. Prates, V. Money, N. Ferry, C. Morland, A. Planas, J. A. Macdonald, R. V. Stick, H. J. Gilbert, C. M. G. A. Fontes and G. J. Davies, *J. Biol. Chem.*, 2005, **280**, 32761–32767.
3 E. A. Bayer, R. Lamed and M. E. Himmel, *Curr. Opin. Biotechnol.*, 2007, **18**, 237–245.
4 A. M. Boudet, S. Kajita, J. Grima-Pettenati and D. Goffner, *Trends Plant Sci.*, 2003, **8**, 576–581.
5 A. J. Ragauskas, C. K. Williams, B. H. Davison, G. Britovsek, J. Cairney, C. A. Eckert, W. J. Frederick, Jr., J. P. Hallett, D. J. Leak, C. L. Liotta, J. R. Mielenz, R. Murphy, R. Templer and T. Tschaplinski, *Science*, 2006, **311**, 484–489.
6 G. J. Davies, M. L. Sinnott and S. G. Withers, Glycosyl Transfer in *Comprehensive Biological Catalysis*, ed. M. L. Sinnott, Academic Press, London, 1997, vol. 1, pp. 119–209.
7 P. J. Meloncelli, T. M. Gloster, V. A. Money, C. A. Tarling, G. J. Davies, S. G. Withers and R. V. Stick, *Aust. J. Chem.*, 2007, **60**, 549–565.
8 V. A. Money, N. L. Smith, A. Scaffidi, R. V. Stick, H. J. Gilbert and G. J. Davies, *Angew. Chem., Int. Ed.*, 2006, **45**, 5136–5140.
9 G. J. Davies, K. S. Wilson and B. Henrissat, *Biochem. J.*, 1997, **321**, 557–559.
10 H. Driguez, *ChemBioChem*, 2001, **2**, 311–318.
11 A. Varrot, S. Leydier, G. Pell, J. M. Macdonald, R. V. Stick, H. J. Gilbert and G. J. Davies, *J. Biol. Chem.*, 2005, **280**, 20181–20184.
12 G. Sulzenbacher, H. Driguez, B. Henrissat, M. Schulein and G. J. Davies, *Biochemistry*, 1996, **35**, 15280–15287.
13 M. Hrmova, R. De Gori, B. J. Smith, J. K. Fairweather, H. Driguez, J. N. Varghese and G. B. Fincher, *Plant Cell*, 2002, **14**, 1033–1052.
14 K. A. Watson, C. McCleverty, S. Geramia, S. Cottaz, H. Driguez and L. N. Johnson, *EMBO J.*, 1999, **18**, 4619–4632.
15 G. Parsiegla, M. Juy, C. Reverbel-Leroy, C. Tardif, J. P. Belaich, H. Driguez and R. Haser, *EMBO J.*, 1998, **17**, 5551–5562.
16 V. Moreau, J. L. Viladot, E. Samain, A. Planas and H. Driguez, *Bioorg. Med. Chem.*, 1996, **4**, 1849–1855.
17 Crystals of *Ct*Lic26A were grown as previously from 0.15 M ammonium sulfate, 30% PEG 5 K MME 0.1 M MES pH 6.5 and ~10 mM compound **1**. Data, to 1.5 Å resolution, were collected on beamline ID14–3 of the European Synchrotron Radiation Source. Data reduction and all subsequent processing involved the CCP4 suite (Collaborative Computational Project Number 4, *Acta Crystallogr.*, 1994, **D50**, 760–763), with structure solution performed with AMoRE (J. Navaza and P. Saludijan, *Methods Enzymol.*, 1997, **276**, 581–594) using the native Lic26A protein-only coordinates as the search model. The structure was refined with REFMAC (G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr.*, 1997, **D53**, 240–255) with manual correction using COOT (P. Emsley and K. Cowtan, *Acta Crystallogr.*, 2004, **D60**, 2126–2132.) Structure and density figures were drawn with BOBSCRIPT (R. M. Esnouf, *J. Mol. Graphics Modell.*, 1997, **15**, 132–134).
18 Site directed variants were made as described previously.[2] Enzyme kinetics were measured both against high viscosity β-glucan (Megazyme, Eire) through detection of reducing sugars using the dinitrosalicylic acid method(G. L. Miller, *Anal. Chem.*, 1959, **31**, 426–428) at 60 °C in 50 mM sodium phosphate buffer, pH 7.0, containing 1 mg ml^{-1} BSA, β-glucan that ranged in concentration from 0.1 mg ml^{-1} to 6 mg ml^{-1} and an appropriate concentration of enzyme (5 nM for wild-type and 600 nM for the least active mutant) and against the fluorescent substrate Glc-β1,4-Glc-β1,3-Glc-methylumbelliferyl with excitation at 365 nM and emission at 440 nM, relating fluorescence to the concentration of methylumbelliferone deploying a standing curve ranging from 1–6 μM. Where possible, the kinetic parameters were determined using seven substrate concentrations that straddled the K_m. The K_i for **1** was determined using standard procedures (T. M. Gloster, P. Meloncelli, R. V. Stick, D. Zechel, A. Vasella and G. J. Davies, *J. Am. Chem. Soc.*, 2007, **129**, 2345–2354).
19 D. Hogg, E. J. Woo, D. N. Bolam, V. A. McKie, H. J. Gilbert and R. W. Pickersgill, *J. Biol. Chem.*, 2001, **276**, 31186–31192.
20 S. Armand, S. R. Andrews, S. J. Charnock and H. J. Gilbert, *Biochemistry*, 2001, **40**, 7404–7409.
21 G. Pell, E. J. Taylor, T. M. Gloster, J. P. Turkenburg, C. M. Fontes, L. M. Ferreira, T. Nagy, S. J. Clark, G. J. Davies and H. J. Gilbert, *J. Biol. Chem.*, 2004, **279**, 9597–9605.