Structure of the *Humicola insolens* cellobiohydrolase Cel6A D416A mutant in complex with a non-hydrolysable substrate analogue, methyl cellobiosyl-4-thio-β-cellobioside, at 1.9 Å

Annabelle Varrot,[a] Torben P. Frandsen,[b]† Hugues Driguez[c] and Gideon J. Davies[a]*

[a]Structural Biology Laboratory, Department of Chemistry, The University of York, Heslington, York Y010 5YW, England, [b]Novozymes A/S, Smoermosevej 25, DK 2880 Bagsvaerd, Denmark, and [c]Centre de Recherches sur les Macromolécules Végétales, CNRS, BP 53, 38041 Grenoble CEDEX 9, France

† Current address: PANTHECO A/S, Danish Science Park, Bøge Alle 3, DK 2970 Hørsholm, Denmark.

Correspondence e-mail: davies@ysbl.york.ac.uk

The enzymatic degradation of cellulose continues to be one of the most important enzyme-catalysed reactions. Glycoside hydrolases from family GH-6 hydrolyse cellulose with inversion of the configuration of the anomeric carbon. Whilst the catalytic proton donor has been clearly identified (Asp226 in *Humicola insolens* Cel6A), the identification and even the existence of a potential Brønsted base remains unclear. Equally controversial is the role of surface-loop flexibility. Here, the structure of the D416A mutant of the *H. insolens* cellobiohydrolase Cel6A in complex with a non-hydrolysable thiooligosaccharide methyl cellobiosyl-4-thio-β-cellobioside at 1.9 Å resolution is presented. Substrate distortion in the -1 subsite, to a $^2S_0$ skew-boat conformation, is observed, similar to that seen in the analogous *Trichoderma reesei* Cel6A structure [Zou *et al.* (1999), *Structure*, **7**, 1035–1045], but the active-centre N-terminal loop of the *H. insolens* enzyme is found in a more open conformation than described for previous structures.

1. Introduction

The enzymatic hydrolysis of cellulose remains one of the most important enzyme-catalysed reactions. It is integral to biosphere maintenance through the recycling of plant biomass; it also finds numerous and increasing industrial applications, whilst glucosidases in general are often considered as model systems for enzymatic action (Zechel & Withers, 2000). Over 87 families of glycoside hydrolases (GH) are known, with cellulases described in at least 12 distinct families (see http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html; Coutinho & Henrissat, 1999). Family GH-6 contains both endoglucanases and cellobiohydrolases, which are believed to act on so-called 'amorphous' and crystalline cellulose, respectively. These classes differ mainly through the topography of their substrate-binding channels: open grooves in endoglucanases and enclosed tunnels conferred by two extended surface loops in the cellobiohydrolases (Davies & Henrissat, 1995; Rouvinen *et al.*, 1990). The flexibility of these loops may account for many of the unusual properties of these enzymes, which are discussed more fully in Varrot *et al.* (2000).

GH-6 enzymes perform the hydrolysis of cellulose with inversion of anomeric configuration (Rouvinen *et al.*, 1990; Schou *et al.*, 1993) in a single-displacement mechanism. This canonical mechanism is generally believed to demand the presence of at least two carboxylate groups: a Brønsted acid to protonate the glycosidic bond and assist leaving-group departure, and a catalytic base to activate a hydrolytic water molecule for nucleophilic attack at the anomeric centre (Koshland, 1953). For family GH-6 enzymes, the acid has been identified both through three-dimensional structure determination of enzyme complexes and elegant mutagenesis experiments (Damude *et al.*, 1995, 1996; Varrot, Hastrup *et al.*, 1999; Varrot, Schülein *et al.*, 1999; Zou *et al.*, 1999). The catalytic proton donor of Cel6A is Asp226 (Varrot, Schülein *et al.*, 1999). Although pH profiles reveal both acid and base components to catalysis, consistent with the paradigmatic inversion mechanism, the identity and even the existence of a classical Brønsted base in the family GH-6 enzyme mechanism remains unclear (Damude *et al.*, 1995, 1996; Varrot, Hastrup *et al.*, 1999; Varrot, Schülein *et al.*, 1999; Wolfgang & Wilson, 1999; Zou *et al.*, 1999). The residue equivalent to Asp405 of Cel6A has been considered to be a candidate in many systems (Damude *et al.*, 1996; Spezio *et al.*, 1993; Varrot, Schülein *et al.*, 1999), but other data argue against this proposal (Rouvinen *et al.*, 1990; Wolfgang & Wilson, 1999; Zou *et al.*, 1999) and consequently none of the potential candidates are convincing.

The identification of base catalysts in inverting glycoside hydrolases is often ambig-

uous (Nurizzo, Nagy *et al.*, 2002; Nurizzo, Turkenburg *et al.*, 2002). The inconsistencies between reported mutagenesis data for potential base mutants of family GH-6 enzymes (described above) led us to investigate the family 6 enzymes from *H. insolens* in more detail. Mutation of the *H. insolens* endoglucanase Cel6B putative base Asp (corresponding to Asp265 of the *T. fusca* E2, Asp405 of *H. insolens* Cel6A and Asp401 of the *T. reesei* Cel6A) led to an essentially inactive enzyme, consistent with results on the *C. fimi* endoglucanase (Damude *et al.*, 1995, 1996). Conversely, mutation of D405 of the *H. insolens* cellobiohydrolase Cel6A to either Asn or Ala, however, resulted in an enzyme with activity on reduced cellohexaose reduced by just 100–300-fold (lecture by the late Martin Schülein to the 2001 Carbohydrate Bioengineering Meeting, Stockholm). These perplexing results, mirroring inconsistencies across the whole family, led us to consider whether additional basic residues could 'rescue' phenotypic differences in the cellobiohydrolase members of the family. Asp416 of Cel6A lies in a position where it could potentially activate a water for 'inverting' attack at C1 should the true base be deleted. It is conserved in all the family GH-6 cellobiohydrolases, but is on a loop absent in the endoglucanase members of this family.

Here, we present the three-dimensional structure of the D416A mutant of *H. insolens* Cel6A in complex with a non-hydrolysable thiooligosaccharide, methyl cellobiosyl-4-thio-β-cellobioside (hereafter called Glc_2-S-Glc_2), at 1.9 Å resolution. The ligand spans the active centre and displays distortion of the −1 subsite sugar to an unusual $^2S_0$ skew-boat conformation similar to that observed for the corresponding enzyme from *T. reesei* (Zou *et al.*, 1999), but displaying a much more open substrate tunnel than in previous Cel6A structures.

2. Experimental

The D416A mutant of Cel6A from *H. insolens* was a kind gift from the late Martin Schülein at Novozymes A/S. Cloning was performed so that only the signal peptide and the catalytic module were expressed after removal of the DNA corresponding to the N-terminal cellulose-binding domain and linker region. The mutant protein was expressed in *Aspergillus oryzae*, purified from the extracellular fluid and treated with endoglucanase F as described previously for the wild-type enzyme (Varrot, Hastrup *et al.*, 1999). Anion-exchange chromatography was then performed using a Q-Sepharose

column in Tris–HCl buffer pH 8.0 and a sodium chloride gradient (0–500 mM) for elution. The protein was finally concentrated to 7 mg ml^{-1} in water. Cel6A was first incubated with 1 mM Glc_2-S-Glc_2 for 1 h prior to crystallization in hanging drops from a solution containing 20% PEG 5K MME, 100 mM NaOAc pH 4.6 and 200 mM MgOAc. 20% glycerol was added to the crystallization condition as a cryoprotectant prior to crystal mounting in rayon-fibre loops and freezing in an N_2 stream at 100 K.

X-ray diffraction data were collected from a single crystal at 100 K at ESRF (Grenoble), beamline ID14-2 (λ = 0.933 Å), using an ADSC Quantum 4 CCD as detector. Data were processed and reduced using the *HKL* suite (Otwinowski & Minor, 1997). All further computing used the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994), unless otherwise stated. The Cel6A D416A mutant in complex with Glc_2-S-Glc_2 was solved by molecular replacement using the coordinates of the native wild type (PDB code 1bvw) as a search model (Varrot, Hastrup *et al.*, 1999) with the program *AMoRe* (Navaza & Saludjian, 1997). 5% of the observations were immediately set aside for cross-validation analysis (Brünger, 1992) and were used to monitor various refinement strategies such as the weighting of geometrical and temperature-factor restraints and the insertion of solvent water during maximum-likelihood refinement using *REFMAC* (Murshudov *et al.*, 1997). Manual corrections of the model using the X-FIT routines of the program *QUANTA* (Accelrys, San Diego, CA, USA) were interspersed with cycles of maximum-likelihood refinement. 'Riding' H atoms were only included when their position were definable. Water molecules were added in an automated manner using *REFMAC/ARP* (Lamzin & Wilson, 1993; Murshudov *et al.*, 1997) and verified manually prior to deposition of coordinates. The incorporation of the ligand was performed after inspection of the $mF_o − DF_c$ weighted map.

3. Results and discussion

Cel6A mutant D416A possesses approximately 10% residual activity (unpublished data) and in order to obtain a complex, a non-hydrolysable thiooligosaccharide (methyl cellobiosyl-4-thio-β-cellobioside, Glc_2-S-Glc_2 whose synthesis is described in Reverbel-Leroy *et al.*, 1998) was harnessed. D416A Cel6A–Glc_2-S-Glc_2 complex data were collected to 1.9 Å and had an R_{merge} of 0.062 (0.28), a mean $I/\sigma(I)$ of 21.2 (4.6), a completeness of 99.1% (98.0) and a multi-

plicity of observation of 3.8 (3.7) in the resolution range 30.0–1.90 Å (data in parentheses are for the outer resolution shell, 1.93–1.90 Å). The final model structure has an R_{cryst} value of 0.166, with an R_{free} of 0.211 for data in the resolution range 30.0–1.9 Å. This model contains 2827 protein atoms, two sodium ions, one *N*-acetylglucosamine, one Glc_2-S-Glc_2 and 270 water molecules. It displays deviations from stereochemical target values of 0.014 Å for bonds and 1.57° for angles, with mean B values of 18.2 Å^2 (r.m.s.d. B for bonded atoms = 0.83 Å^2) and 20.7 Å^2 (r.m.s.d. B for bonded atoms = 2.2 Å^2) for the main chain and side chains, respectively. The mean solvent B is 25 Å^2. There are no residues in forbidden regions of the Ramachandran plot; 90% of the residues are in the most favoured regions and 10% are in the additional allowed regions (calculated using *PROCHECK*; Laskowski *et al.*, 1993).

Inspection of the initial electron-density map for the Cel6A D416A mutant confirmed the mutation and revealed clear density for all four glucose moieties of the thiooligosaccharide. The tetrasaccharide spans the active centre and binds in subsites −2 to +2 (nomenclature according to Davies *et al.*, 1997) (Fig. 1). Glucosides in subsites −2, +1 and +2 are in the $^4C_1$ (chair) conformation. That in the catalytic −1 subsite is distorted to an unusual $^2S_0$ (skew-boat) conformation (described below), as has previously been reported for similar complexes of the *T. reesei* Cel6A wild type (PDB code 1qjw) and the Y169F mutant (1qk2) from *T. reesei* (Zou *et al.*, 1999). No significant structural changes result from the mutation; indeed, the three equivalent complexes superpose well with an r.m.s.d. of 0.6 Å for 357 equivalent C$^\alpha$ atoms (calculated with *LSQMAN*; Kleywegt & Jones, 1994). The only significant differences between the *T. reesei* and *H. insolens* complexes are observed at the level of the two active-site loops which enclose the cellobiohydroase active-centre tunnel.

Family 6 cellobiohydrolases are characterized by the two extended loops that together form an enclosed tunnel, which is believed to maintain a single glucan chain through numerous catalytic events whilst preventing it re-adhering to crystalline cellulose (Davies & Henrissat, 1995; Rouvinen *et al.*, 1990). These loops, although often assumed to be static, are observed in a range of conformations in different family GH-6 structures. The N-terminal loop is found 'closed' in the *T. reesei* wild-type structure (Rouvinen *et al.*, 1990) but more open in the Y169F structure (PDB code
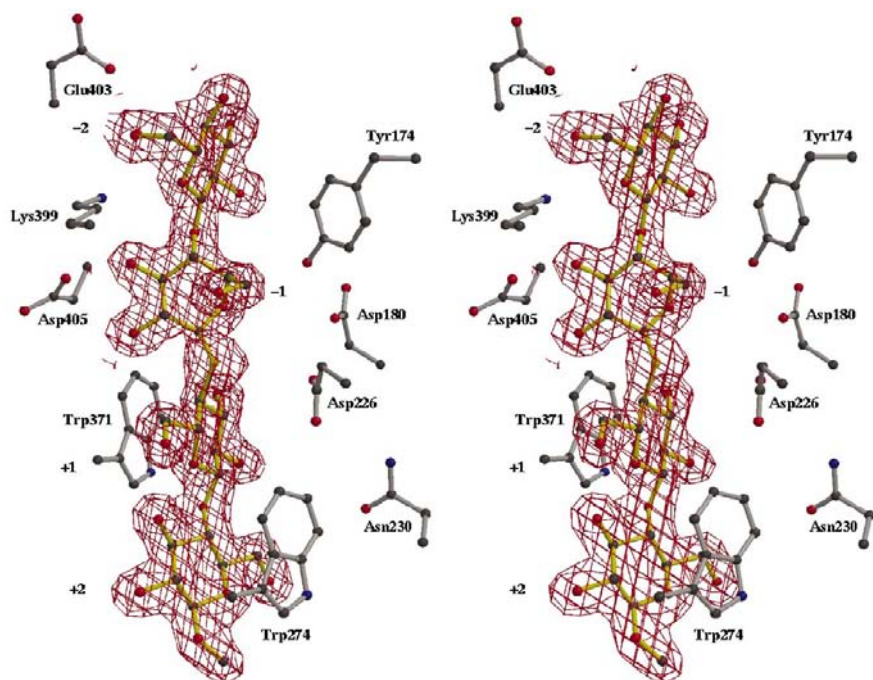
Figure 1
Divergent stereographic representation of the maximum-likelihood-weighted $2mF_o - DF_c$ electron-density map contoured at 0.42 e Å^{-3} around the methyl cellobiosyl-4-thio-β-cellobioside bound to Cel6A. The subsites and residues involved in the interactions with the substrates are labelled.
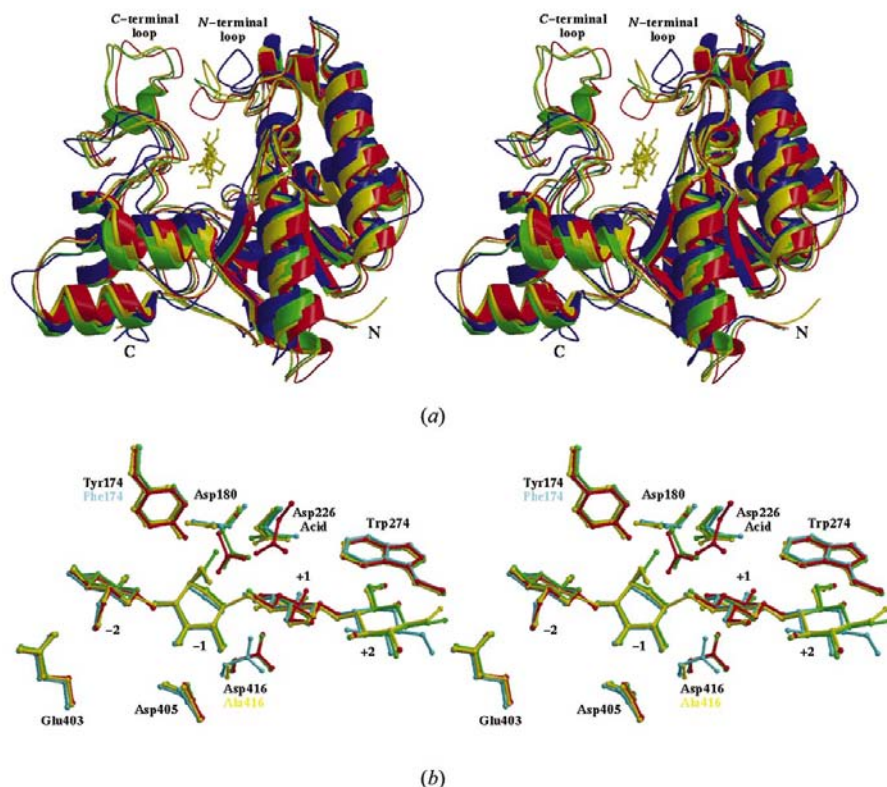

(a)

(b)

Figure 2
(*a*) Overlap of the D416A Cel6A–Glc$_2$-*S*-Glc$_2$ complex structure (yellow) with the native wild-type (green) and the glucose–cellotetraose complex (red) structures of Cel6A from *H. insolens* and with the native structure of *H. insolens* Cel6B (blue), in divergent stereo. (*b*) Overlap of the D416A Cel6A–Glc$_2$-*S*-Glc$_2$ complex structure (yellow) with the wild-type glucose-cellotetraose complex (red) structures of Cel6A from *H. insolens* and with the Glc$_2$-*S*-Glc$_2$ complex structure of wild-type (green) and the Y169F mutant (cyan) Cel6A from *T. reesei*. These figures were drawn with *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merritt & Bacon, 1997).

1cb2; Koivula *et al.*, 1996). Similarly, partially open and closed structures have been reported for the wild-type *H. insolens* Cel6A structures in the apo form (PDB code 1bvw; Varrot, Hastrup *et al.*, 1999) and in a glucose-cellotetraose complex (PDB code 2bvw; Varrot *et al.*, 2000), respectively. In the corresponding endoglucanase Cel6B structure (PDB code 1dys; Davies *et al.*, 2000) this loop is fully open (Fig. 2). The *H. insolens* D416A Cel6A structure reveals an opening of the 'N-terminal' loop of approximately 30°, resulting from 15–25° changes in the main-chain conformational torsion angles between residues 181–183 and 186–187. We believe that this is the most open conformation for this loop witnessed in a cellobiohydrolase thus far. Furthermore, the C-terminal loop is also slightly shifted by a rigid-body movement of approximately 1 Å in order to avoid a steric clash of the main chain of Ala413 with the side chain of Ala185.

Within the active centre and substrate tunnel few significant changes are noted, with the exception of the −1 subsite, as described below. The glucosyl moieties in the −2, +1 and +2 subsites are exactly as observed in the glucose-cellotetraose complex (PDB code 2bvw; Davies *et al.*, 2000). In the −1 subsite, the glucose moiety is distorted from a $^4C_1$ chair to a $^2S_0$ skew-boat conformation, coupled with an approximate 90° rotation of the pyranoside ring relative to its neighbours. In this conformation, the only direct hydrogen bond to protein is from the O3 hydroxyl to Asp405; all other interactions are mediated by water molecules. It is unclear whether the $^2S_0$ conformation reflects a genuine conformational point along the reaction trajectory or if it is merely an 'artefact' necessary to accommodate the hydroxymethyl group lest it clash with the side chain of Tyr174. Distorted Michaelis complexes of retaining glycoside hydrolases (for an example, see Davies *et al.*, 1998, and references therein) are widely believed to reflect the adjacent transition-state conformation for glycoside hydrolysis. The nearest potential transition-state conformation (*i.e.* that with C5, O5, C1 and C2 coplanar) to the $^2S_0$ conformation is the $^{2,5}B$ (boat). On superficial inspection this would seem catalytically unlikely for glucoside hydrolysis, since it displays unfavourable 'bowsprit' interactions between the C5 hydroxymethyl substituent and H2 (Sabini *et al.*, 1999). However, recent evidence for unusual $B_{2,5}$ (Ducros *et al.*, 2002) and $^3H_4$ (Lobsanov *et al.*, 2002) transition states for β- and α-mannosidases, respectively, coupled with catalysis via a $^{2,5}B$ state in

short communications

family 11 xylanases (Sabini *et al.*, 1999, 2001; Sidhu *et al.*, 1999) suggests that no potential transition-state conformation should be ruled out. The question of whether the $^2S_0$ conformation observed is catalytically relevant is further complicated by the displacement of the catalytic acid, Asp226, away from the glycosidic S atom in this complex and that reported for the *T. reesei* Cel6A (PDB code 1qjw; Zou *et al.*, 1999). In contrast to the orientation observed in the *H. insolens* Cel6A–glucose-cellotetraose complex (Varrot, Schülein *et al.*, 1999), Asp226 lies 4.5 Å away from the glycosidic S atom, with inappropriate geometry for proton transfer (Fig. 2*b*). The displaced position of Asp226 reflects its close interaction with Asp180; this latter residue, implicated in pK_a elevation (Damude *et al.*, 1996), interacts with the substrate in the +1 subsite in the glucose-cellotetraose complex. Furthermore, in the D416A Cel6A complex with Glc$_2$-*S*-Glc$_2$ we see no water molecule poised for inverting nucleophilic attack at C1.

Catalysis by family GH-6 members remains difficult to dissect on both macroscopic and microscopic levels. The D416A Cel6A complex reported here clearly demonstrates the propensity for loop conformational change, reinforcing earlier work (Varrot, Schülein *et al. et al.*, 1999; Zou *et al.*, 1999). Despite these structural observations on enzymes from different sources, loop movements are still dismissed by many authors. The role of a catalytic base is certainly confusing. The residue equivalent to Asp405 of Cel6A had seemed appropriately placed from a geometrical view in the different structures (Rouvinen *et al.*, 1990; Spezio *et al.*, 1993; Varrot, Schülein *et al.*, 1999) and was consistent with site-directed mutagenesis on the endoglucanase A from *C. fimi* (Damude *et al.*, 1995). However, mutagenesis on other systems appears to contradict this (Wolfgang & Wilson, 1999), as does the position of Asp405 in the plane of the ring in the Glc$_2$-*S*-Glc$_2$ complexes presented both here and elsewhere (Zou *et al.*, 1999) (Fig. 2*b*). The complex hydrogen bonding within the active centre of Cel6 enzymes, together with the potential for catalytic rescue by other bases within the binding milieu and the absence of a complex which has all components of the catalytic machinery in the correct location, makes interpretation of site-directed mutagenesis data of partially active enzymes fraught with danger. A great deal of work remains necessary before catalysis by family GH-6 enzymes and their role in cellulose degradation is fully understood.

Note added in proof: Since acceptance of this paper, Koivula and colleagues have published even more open Cel6A complexes, for the enzyme from *Trichoderma reesei*. Modelling studies support a $^{2,5}B$ transition-state conformation. These authors also propose that attack of water is aided by a solvent-mediated deprotonation in a Grotthus-type mechanism (Koivula *et al.*, 2002). In an unrelated study, Geurin and colleagues observed a $^{2,5}B$ conformation for a glucosyl moiety in the active centre of an inverting family 8 endoglucanase also lending credence to a $^{2,5}B$ transition-state conformation (Guerin *et al.*, 2002).

References

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Coutinho, P. M. & Henrissat, B. (1999). *Carbohydrate-Active Enzymes: An Integrated Approach in Recent Advances in Carbohydrate Engineering*, edited by H. J. Gilbert, G. J. Davies, B. Svensson & B. Henrissat, pp. 3–12. Cambridge: Royal Society of Chemistry.

Damude, H. G., Ferro, V., Withers, S. G. & Warren, R. A. J. (1996). *Biochem. J.* **315**, 467–472.

Damude, H. G., Withers, S. G., Kilburn, D. G., Miller, R. C. Jr & Warren, R. A. J. (1995). *Biochemistry*, **34**, 2220–2224.

Davies, G. J., Brzozowski, A. M., Dauter, M., Varrot, A. & Schülein, M. (2000). *Biochem. J.* **348**, 201–207.

Davies, G. & Henrissat, B. (1995). *Structure*, **3**, 853–859.

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

Davies, G. J., Wilson, K. S. & Henrissat, B. (1997). *Biochem. J.* **321**, 557–559.

Ducros, V., Zechel, D. L., Murshudov, G. N., Gilbert, H. J., Szabo, L., Stoll, D., Withers, S. G. & Davies, G. J. (2002). *Angew. Chem. Int. Ed. Engl.* **41**, 2824–2827.

Guerin, D. M. A., Lascombe, M. B., Costabel, M., Souchon, H., Lamzin, V., Beguin, P. & Alzari, P. M. (2002). *J. Mol. Biol.* **316**, 1061–1069.

Kleywegt, G. J. & Jones, T. A. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 9–14.

Koivula, A., Reinikainen, T., Ruohonen, L., Valkeajärvi, A., Claeyssens, M., Teleman, O., Kleywegt, G. J., Szardenings, M., Ruovinen, J., Jones, T. A. & Teeri, T. T. (1996). *Protein Eng.* **9**, 691–699.

Koivula, A., Ruohonen, L., Wohlfahrt, G., Reinikainen, T., Teeri, T. T., Piens, K., Claeyssens, M., Weber, M., Vasella, A., Becker, D., Sinnott, M. L., Zou, J. Y., Kleywegt, G. J., Szardenings, M., Stahlberg, J. & Jones, T. A. (2002). *J. Am. Chem. Soc.* **124**, 10015–10024.

Koshland, D. E. (1953). *Biol. Rev.* **28**, 416–436.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.

Laskowski, R. A., McArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lobsanov, Y. D., Vallee, F., Imberty, A., Yoshida, T., Yip, P., Herscovics, A. & Howell, P. L. (2002). *J. Biol. Chem.* **277**, 5620–5630.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–594.

Nurizzo, D., Nagy, T., Gilbert, H. J. & Davies, G. J. (2002). *Structure*, **10**, 547–556.

Nurizzo, D., Turkenburg, J. P., Charnock, S. J., Roberts, S. M., Dodson, E. J., McKie, V. A., Taylor, E. J., Gilbert, H. J. & Davies, G. J. (2002). *Nature Struct. Biol.* **9**, 665–668.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Reverbel-Leroy, C., Parsiela, G, Moreau, V., Juy, M., Tardif, C., Driguez, H., Belaich, J. P. & Haser, R. (1998). *Acta Cryst.* D**54**, 114–118.

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. K. C. & Jones, T. A. (1990). *Science*, **249**, 380–386.

Sabini, E., Sulzenbacher, G., Dauter, M., Dauter, Z., Jørgensen, P. L., Schülein, M., Dupont, C., Davies, G. J. & Wilson, K. S. (1999). *Chem. Biol.* **6**, 483–492.

Sabini, E., Wilson, K. S., Danielsen, S., Schülein, M. & Davies, G. J. (2001). *Acta Cryst.* D**57**, 1344–1347.

Schou, C., Rasmussen, G., Kaltoft, M.-B., Henrissat, B. & Schülein, M. (1993). *Eur. J. Biochem.* **217**, 947–953.

Sidhu, G., Withers, S. G., Nguyen, N. T., McIntosh, L. P., Ziser, L. & Brayer, G. D. (1999). *Biochemistry*, **38**, 5346–5354.

Spezio, M., Wilson, D. B. & Karplus, P. A. (1993). *Biochemistry*, **32**, 9906–9916.

Varrot, A., Hastrup, S., Schülein, M. & Davies, G. J. (1999). *Biochem. J.* **337**, 297–304.

Varrot, A., Schülein, M. & Davies, G. J. (1999). *Biochemistry*, **38**, 8884–8891.

Varrot, A., Schülein, M. & Davies, G. J. (2000). *J. Mol. Biol.* **297**, 819–828.

Wolfgang, D. E. & Wilson, D. B. (1999). *Biochemistry*, **38**, 9746–51.

Zechel, D. L. & Withers, S. G. (2000). *Acc. Chem. Res.* **33**, 11–18.

Zou, J.-Y., Kleywegt, G. J., Ståhlberg, J., Driguez, H., Nerinckx, W., Claeyssens, M., Koivula, A., Teeri, T. T. & Jones, T. A. (1999). *Structure*, **7**, 1035–1045.