The 1.62 Å structure of *Thermoascus aurantiacus* endoglucanase: completing the structural picture of subfamilies in glycoside hydrolase family 5

Leila Lo Leggio*, Sine Larsen

Centre for Crystallographic Studies, Department of Chemistry, Chemical Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark

Received 25 March 2002; revised 23 May 2002; accepted 6 June 2002

First published online 25 June 2002

Edited by Irmgard Sinnig

Abstract The crystal structure of *Thermoascus aurantiacus* endoglucanase (Cel5A), a family 5 glycoside hydrolase, has been determined to 1.62 Å resolution by multiple isomorphous replacement with anomalous scattering. It is the first report of a structure in the subfamily to which Cel5A belongs. Cel5A consists solely of a catalytic module with compact eight-fold β/α barrel architecture. The length of the tryptophan-rich substrate binding groove suggests the presence of substrate binding subsites −4 to +3. Structural comparison shows that two glycines are completely conserved in the family, in addition to the two catalytic glutamates and six other conserved residues previously identified. Gly 44 in particular is part of a type IV C-terminal helix capping motif, whose disruption is likely to affect the position of an essential conserved arginine. One aromatic residue (Trp 170 in Cel5A), not conserved in term of sequence, is nonetheless spatially conserved in the substrate binding groove. Its role might be to force the bend that occurs in the polysaccharide chain on binding, thus favoring substrate distortion at subsite −1.

Key words: Glycoside hydrolase; Family 5; Subfamily; Tryptophan; Clan GH-A; 4/7 Superfamily

1. Introduction

Thermoascus aurantiacus, a moderately thermophilic fungus [1], produces several plant cell wall degrading enzymes including endoglucanases [2–4]. Endoglucanases hydrolyze internal β-1,4-glycosidic bonds in mixed glucans, carboxymethylcellulose (CMC) and in some cases cellulose. Here we report structure determination of a β-1,4-endoglucanase from this fungus, which according to its abundance and characteristics [5] is probably identical or similar to endocellulases identified by others [2–4,6]. These enzymes all showed an apparent molecular weight of approximately 32–34 kDa, high temperature stability and optimum (around 70°) and low optimum pH (3–4.5).

N-terminal sequencing [7] demonstrated that *T. aurantiacus* endoglucanase belongs to glycoside hydrolase family 5 (GH5) [8] and thus it will be referred as Cel5A in the text. The retaining mechanism in family 5 has been extensively studied by chemical modification, mutagenesis and X-ray crystallography (three excellent examples are in [9–11]). GH5 belongs to glycoside hydrolase clan A (clan GH-A), also referred to as 4/7 superfamily [12–13], a superfamily of glycoside hydrolases with conserved overall structure and mechanism but divergent substrate specificity.

GH5 in itself is a highly divergent family in terms of sequence and function and thus has been divided in subfamilies, which were originally defined as groups of sequences sharing less than 25% sequence identity [14]. Five subfamilies were originally identified. On the basis of the N-terminal sequence we previously suggested [7] that Cel5A formed a new subfamily 6 of GH5 sequences, as no sequences in the original subfamilies description [14] were matched in a BLAST search. This was due to an unfortunate oversight, as Wang et al. [9] had subsequent to the first subfamily description identified other members of subfamily 5 which did match our search. Since others have since adopted our numbering (two mannanase subfamilies, 5-7 and 5-8, have been later classified [15]) we shall refer to this subfamily as subfamily 5-5/6 to avoid further confusion. The subfamily including the *Candida albicans* exoglucanase [16] is referred here as subfamily 5-9.

Structures have been published for all GH5 subfamilies known except 5-5/6 [15–23]. The structure of *T. aurantiacus* Cel5A reported here fills in this gap in the structural knowledge of GH5.

2. Materials and methods

2.1. Crystallization and data collection

Crystallization, data collection and processing has been described in detail elsewhere [7]. The crystals belonged to space group $P2_12_12_1$ with $a = 76.4$ Å, $b = 85.7$ Å, $c = 89.5$ Å and 2 molecules in the asymmetric unit. Native data collected at room temperature extend to 1.62 Å resolution (Table 1).

2.2. Structure determination

Four heavy atom derivatives were used in structure determination by multiple isomorphous replacement with anomalous scattering (MIRAS). MIRAS phases were calculated in SHARP (version 1.3.4, [24]). Experimental conditions, data collection and phasing statistics are summarized in Table 1.

*Corresponding author. Fax: (45)-35320299.
E-mail address: leila@ccs.ki.ku.dk (L. Lo Leggio).

Abbreviations: Cel5A, *Thermoascus aurantiacus* endoglucanase; CMC, carboxymethylcellulose; GHX, glycoside hydrolase family number X; GH-A, glycoside hydrolase clan A; MIRAS, multiple isomorphous replacement with anomalous scattering; NCS, non-crystallographic symmetry; *R*-factor, crystallographic *R*-factor for the working set; *R*-free, crystallographic *R*-factor for the test set; rmsd, root men square deviation

0014-5793/02/$22.00
PII: S0014-5793(02)02954-X

Table 1
Summary of data collection and phasing statistics

	Native[a] (CuKα)	Native[a] (Photon Factory)	$SmCl_3$	$HgAc_2$	UO_2Ac_2	K_2PtCl_4
Maximum resolution (Å)	2.35	1.62	2.7	3.3	2.7	3.2
Completeness (%)	91.3	81.6	71.7	91.5	88.3	82.7
Multiplicity	4.5	5.3	4.3	3.1	2.6	2.9
R_{sym}	0.077	0.056	0.096	0.097	0.126	0.075
Soak conditions	–	–	10 mM, 2.5 h	75 mM, 20 h	10 mM, 3 days	10 mM, 7 days
No. sites	–	–	2	4	2	4
R_{cullis} (acentric)	–	–	0.83	0.72	0.88	0.86
R_{cullis} (centric)	–	–	0.82	0.71	0.83	0.83
R_{cullis} (anom.)	–	–	0.86	0.97	0.98	0.98
Phasing power (acentric)	–	–	1.43	2.15	1.29	1.46
Phasing power (centric)	–	–	1.25	1.78	1.08	1.21

[a]The native data set used during structure refinement consisted of the synchrotron data set where missing low resolution data had been 'refilled' with reflections from the in-house data, as described in [7]. The final data set had 88.5% overall completeness (99.2% overall completeness between 35.3 and 5.1 Å).

An initial model was constructed in O [25] based on maps calculated after cycles of density modification including two-fold averaging in DM [26], and was greatly improved using ARP [27]. However progress was initially hampered by lack of complete sequence information (sequence was only available for 35 N-terminal residues). Recently sequence information has became available [28] and allowed full refinement of the model in CNS [29]. All reflections were used for refinement except for 10% of total, which were reserved for calculation of *R*-free (crystallographic *R*-factor for the test set, where *R*-factor is the crystallographic *R*-factor for the working set).

Non-crystallographic symmetry (NCS) restraints were used in all but the last few cycles of refinement and were progressively relaxed as refinement progressed. The two molecules in the asymmetric unit are very similar. NCS related water molecules were identified with the program WATNCS [30] and were also NCS restrained initially. The final model contains 304 residues per molecule in the asymmetric unit, with one N-terminal residue not visible in the electron density map. Fourteen side chains in molecule A and 18 side chains in molecule B were modeled in double conformations. 389 water molecules were present in the final model (including 129×2 NCS related waters). The model has very good geometry with an overall PROCHECK [31] *G*-factor of 0.28, bond length root men square deviation (rmsd) of 0.005 Å and bond angle rmsd of 1.25°. 90.4% of all residues were in the most favored regions of the Ramachandran plot. Asn 60 lies in the disallowed region of the Ramachandran plot, but its conformation is confirmed by good quality electron density. The final *R*-factor was 15.9% with an *R*-free of 17.7%. Coordinates and structure factors have been deposited in the PDB with code 1GZJ.

3. Results and discussion

3.1. Overall structure

Cel5A is one of the most compact family 5 enzymes for which structure is available. First of all, unlike many GH5 enzymes, Cel5A consists only of the catalytic domain. Secondly, unlike many GH5 enzymes, which have long excursions in the β/α-loops, the eight-fold β/α-barrel in Cel5A is very regular and compact with short loops. The structure has only few extra elements of secondary structure compared to the canonical eight-fold β/α-barrel, a short two stranded antiparallel β-sheet in β/α-loop 3 and three one-turn helices (Figs. 1a and 2).

A search against representative structures in FSSP [32] using DALI [33] showed that GH5 subfamilies 4, 3, 1, 2, 7 and 8 (in this order) are the closest structural neighbors of Cel5A with 15–20% sequence identity over 240–280 residues structurally aligned. The rmsd for Cα positions calculated by DALI between Cel5A and subfamily 4 *Clostridium cellulolyticum* CelCCA was 2.7 Å over 280 aligned residues. The exo-glucanase structure in subfamily 9 [16] is a more distant relative, with only 11% sequence identity. The rmsds for Cα positions with members of GH5 were all between 2.7 and 3.0 Å. The following, lower score hits include members of the 4/7 superfamily (clan GH-A) [12–13] and β-amylases [34], with rmsds of 2.9–3.7 Å. β-amylases are not usually considered to be part of the conventional 4/7 superfamily because of their inverting mechanism, but are actually more structurally similar to Cel5A than some of the members of the 4/7 superfamily, as has been previously observed in comparisons with members of GH10, GH5-1, and GH2 [12,17,35].

Cel5A has three *cis*-peptides per monomer, between Leu21 and Pro 22, Trp 273 and Ala 274, Pro 288 and Asp 289. The *cis*-peptide involving Trp 273 is conserved in family 5 and some other members of the 4/7 superfamily [17,36] and is thought to be important in positioning the Trp side chain, which is important in substrate binding. A disulfide bridge in Cel5A between Cys 212 and Cys 249 is not conserved in the rest of the family.

3.2. The active site and substrate binding groove

The catalytic residues of Cel5A are easily identified by homology to the other GH5 structures as Glu 133 (acid-base) and Glu 240 (nucleophile) and are shown in Fig. 1a. Although the pH of crystallization is far from the acidic pH optimum of Cel5A, and despite the absence of substrate, the active site appears to be in an active conformation, at least with respect to the position of the two catalytic glutamates, unlike for example in the subfamily 3 *Clostridium thermocellum* CelC [36].

The substrate binding groove of Cel5A is shallow and long. Using as reference the complex of *Acidothermus cellulolyticus* complex with cellotetraose [17] the existence of substrate binding subsites −4 to +3 can be postulated. This is in good agreement with the reported endo mode of action on CMC, and the preference for central linkages in reduced cellulodextrins up to DP 6 [2,5]. A large number of aromatic residues line the groove (Fig. 1b), including tryptophans 278, 279, 273, 170 and 174 going from the non-reducing to the reducing end of the substrate.

3.3. Comparison between the eight subfamilies in family 5

As well as the catalytic residues, the other six established conserved residues in GH5 [9,17] are also present in Cel5A: Asn 132, Arg 49, His 93, His 198, Tyr 200 and Trp 273, as shown in Fig. 1c and in the structure based sequence alignment in Fig. 2.

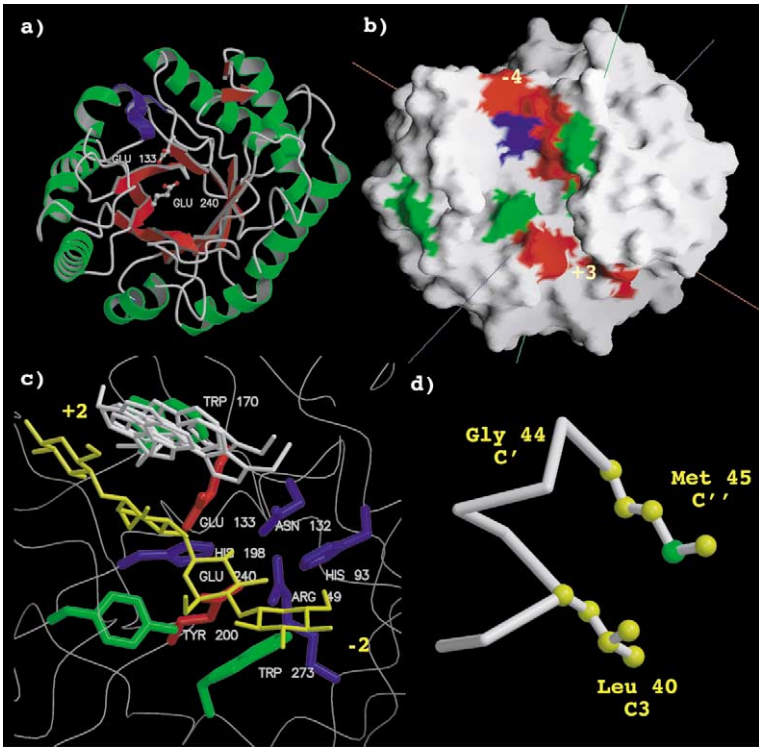

Fig. 1. Views of molecule A in the Cel5A structure. a: Overall view of the βα-barrel with the two catalytic glutamates in ball and stick representation. 3_{10} helices are shown in blue. b: View of the molecular surface with aromatic residues in color (Trp in red, Tyr in green and Phe in blue). Putative substrate binding subsites are marked. c: Conserved active site residues in GH5 shown in red (catalytic glutamates), green (aromatics) and blue (polar). The spatially conserved aromatic residues corresponding to Trp 170 are also shown for representatives of the other GH5 subfamilies: *A. cellulolyticus* endocellulase E1 in subfamily 1 [17]; *Bacillus agaradherens* endoglucanase in subfamily 2 [18]; *C. thermocellum* CelC in subfamily 3 [21]; *C. cellulolyticum* CelCCA in subfamily 4 [22]; *T. fusca* mannanase in subfamily 7 [15]; *T. reesei* mannanase in subfamily 8 [23]; *C. albicans* exoglucanase in subfamily 9 [16]. Cellotetraose bound to endocellulase E1 is shown for reference in yellow [17]. d: The C-terminal type IV capping motif involving Gly 44, conserved in GH5. This figure was made using the programs Molscript [51], Raster3D [52] and GRASP [53].

The structure based alignment shows that the subfamily 8 *Trichoderma reesei* β-mannanase [23] lacks the histidine at the end of beta strand 3 (His 93 in Cel5A) which is conserved in all the other structures, and the conserved acidic/polar residue two residues before (Asp 91 in Cel5A) which is present in the other structures. His 93 is therefore not essential for activity, consistent with its role in substrate binding to the sugar moiety O3 at subsite −1 (for example in [11,16,17,36]) and not in catalysis. This is further supported by the finding that a Phe substitution of the equivalent histidine in the *Erwinia chrisanthemi* cellulase [10] results in an enzyme with significant residual catalytic activity. Loss of a hydrogen bond with the substrate at the −1 subsite might explain the binding of mannobiose and mannotriose at the +1/+2 subsites in the *T. reesei* β-mannanase [23] as opposed to the minus end of the substrate binding groove, as preferentially seen for the *Thermomonospora fusca* β-mannanase structure [15].

The alignment identifies two completely conserved glycines in GH5, Gly 8 and Gly 44. Gly 8 is found in the middle of barrel strand 1, and the reason for its conservation is not clear. Gly 44 is part of a type IV or Schellman C-terminal capping motif [37] (Fig. 1d) which is conserved in the GH5 structures. This motif involves a hydrophobic interaction between the third last residue of the alpha helix (C3) and the second residue (C″) after the the Ccap residue, with a glycine in position C′. Gly 44 is also highly conserved in GH5 enzymes for which only the sequence is available. Structurally, Gly 44 is at a tight αβ-loop preceding strand 2 in the barrel. Strand 2 bears one of the absolutely conserved residues in GH5 (Arg 49 in Cel5A), which has a proposed role in stabilizing the negative charge on the nucleophile. This arginine is also conserved in the 4/7 superfamily, except in family 10 and family 26 [17,38] where residues from other loops (a histidine from strand 6 and an arginine from strand 4) take over its function. Substitution of Gly 44 in GH5 is likely to disrupt the capping motif and with it the positioning of the conserved arginine. Thus it seems to have been disfavored by evolution.

Inspection of the members of the 4/7 superfamily represented in FSSP [32] supports this role for the conserved capping motif containing Gly 44, as the presence of the motif is correlated with conservation of Arg 49. The type IV C-terminal capping motif is conserved in those representative members which have the conserved Arg: family 1 *Thermosphaera aggregans* β-glycosidase [39] (a similar motif, but with an Asn in place of a Gly is found in myrosinase from the same family [40]), family 2 human β-glucuronidase [41], and family 17 barley glucanases [42]. However, the capping motif is not conserved in members lacking the conserved Arg, family 10 *T. aurantiacus* xylanase [43] or family 26 *Pseudomonas cellulosa* mannanase [38] (which however has a Gly at the end of the corresponding helix as part of a a type VIa capping motif).

Several other glycines throughout the GH5 sequences are semi-conserved in the structure based sequence alignment. In

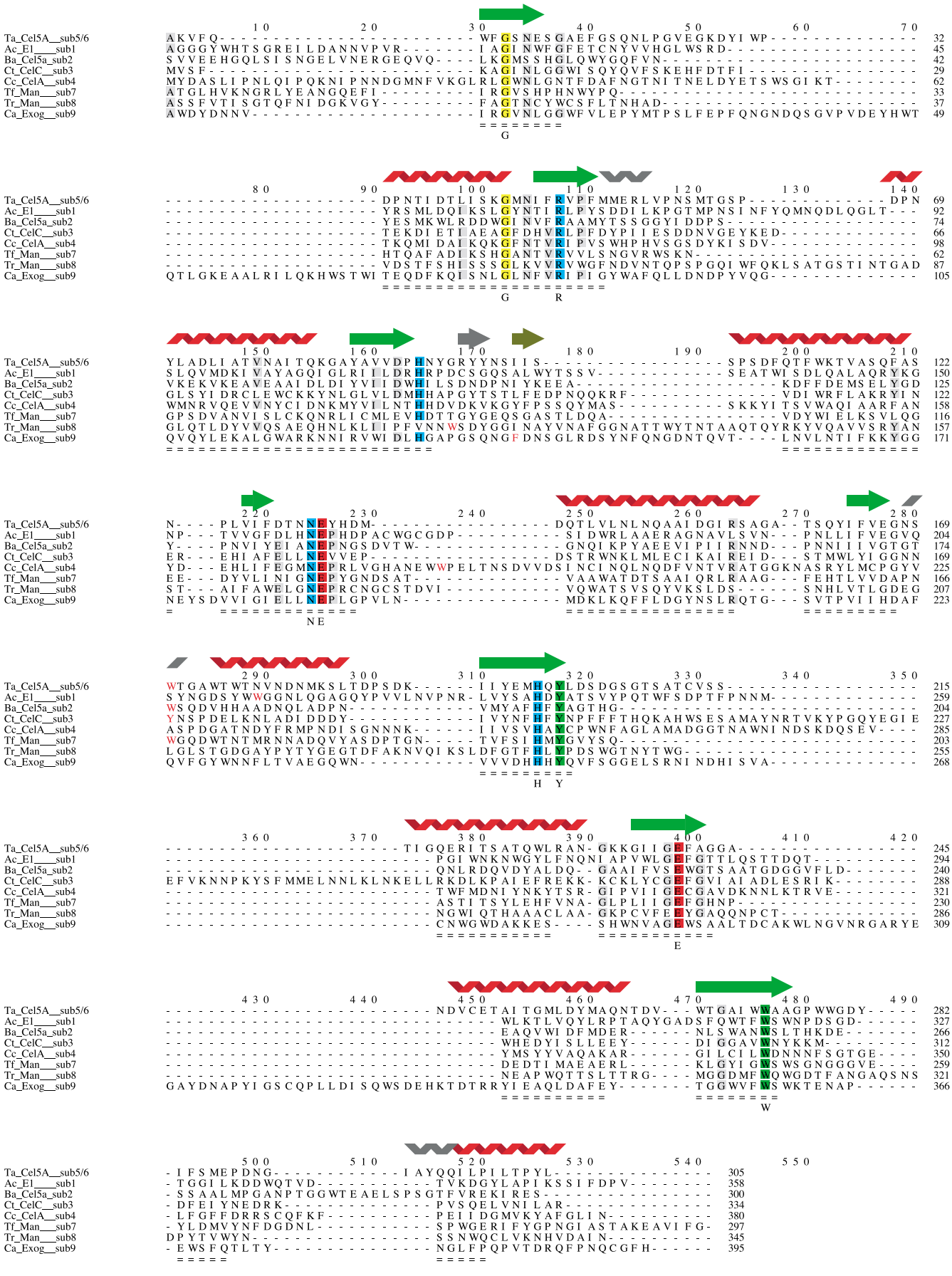
Ta_Cel5A__sub5/6
Ac_E1____sub1
Ba_Cel5a_sub2
Ct_CelC__sub3
Cc_CelA__sub4
Tf_Man___sub7
Tr_Man___sub8
Ca_Exog__sub9
G
G R
NE
H Y
E
W

Fig. 2. Structure based alignment of the GH5 subfamilies (Cel5A and the representatives in Fig. 1c). The structure based alignment was output by STRUPRO [54] and displayed in Indonesia (Madsen, D., http://xray.bmc.uu.se/dennis/manual/). Elements of secondary structures outside the βα-barrel are shown in gray. The conserved active site residues in GH5 are color coded as in Fig. 1c, while the conserved glycine are shaded in yellow. The spatially conserved aromatic residues at subsites +1/+2 are typed in red. Other semi-conserved residues are shaded in gray.
←

one other case a semi-conserved glycine (Gly 233) is in the αβ-loop preceding strand 7, which bears the loop containing the catalytic nucleophile. This Gly is also present in family 1 *T. aggregans* β-glycosidase and family 26 *P. cellulosa* mannanase. A negatively charged residue (or Asn in the *T. fusca* mannanase) at the end of strand 4 (Asp 129 in Cel5A) is also conserved in GH5 and serves to orient the conserved Arg 49.

3.4. A spatially conserved aromatic residue on the plus side of the substrate binding groove

Although all the enzymes in GH5 have many aromatic residues in their substrate binding grooves, only Tyr 200 and Trp 273 at the −2/−1 subsites are strictly conserved on the basis of both sequence and structure. A structure based comparison shows however that one other aromatic residue at subsites +1/+2 (using the *A. cellulolyticus* endoglucanase as reference) is spatially conserved in GH5 (Fig. 1c) although it maybe contributed by different loops in the different subfamilies (Fig. 2). In Cel5A this residue is Trp 170 and is contributed by the 5th βα-loop. Chemical modification of the corresponding Trp residue in *T. reesei* endoglucanase III, an enzyme in the same subfamily as Cel5A, decreased the k_{cat}/K_m to approximately half its original value [44].

The plane of the ring lies approximately in the same orientation in all subfamilies, and parallel to the monosaccharide ring at subsite +1/+2 where experimental complexes are available [17,23,45–46]. In a view of catalysis where the extended substrate binds in a native-like conformation, but then gets bent with distortion of the sugar moiety at subsite −1 so that the glycosidic bond to be cleaved is in an axial conformation [46–48], the role of this residue could be precisely to impose the bend in the polysaccharide chain.

Both mannan and cellulose have ribbon-like structures in which the monosaccharide rings are oriented at 180° with respect to each other [49,50], therefore it is reasonable that in both β-1,4-endoglucanases and β-mannanases the plane of this conserved aromatic residue is in a similar orientation. In the *P. cellulosa* family 26 mannanase Trp 217 is also in a similar orientation. Substitution of this residue with Ala caused a four-fold decrease in k_{cat} with carob galactomannan as a substrate [38], suggesting that the catalytic process is affected, and not just substrate binding. This residue is not spatially conserved in the other representative structures in the 4/7 superfamily examined [39–43].

The aromatic residue is however conserved in the subfamily GH5-9 exo-β-1,3-glucanase, expected to bind to a substrate with quite different conformation. Cutfield and coworkers [16] have pointed out that in this enzyme the spatially conserved aromatic (Phe 144), which in this enzyme is shifted more towards the +1 subsite, is involved in a 'clamp' formed by two antiparallel phenylalanines and suggest again a role in forcing substrate distortion from an equatorial to an axial scissile bond. Thus a similar spatial orientation of this aromatic side chain can be exploited in enzymes with different specificities.

Acknowledgements: Drs. M.K. Bhat and N. Parry are thanked for provision of protein for crystallization. The research was funded by the Danish National Research Foundation and the EU under contract CT93-1272. We acknowledge the Photon Factory, Tzukuba, Japan, for synchrotron time, and our colleagues, Drs. J. Jenkins, G. Harris and O. Mayans for help with data collection. In particular we wish to thank Dr. R.W. Pickersgill for fruitful discussions throughout the project and Drs. D. Madsen and E. Johansson for their help with making the structure based sequence alignment.

References

[1] Deploey, J.J. (1995) Mycologia 87, 362–365.
[2] Shepherd, M.G., Tong, C.C. and Cole, A.L. (1981) Biochem. J. 193, 67–74.
[3] Feldman, K.A. and Lovett, J.S. (1988) Enzyme Microb. Technol. 10, 262–272.
[4] Khandke, K.M., Vithayathil, P.J. and Murthy, S.K. (1989) Arch. Biochem. Biophys. 274, 491–500.
[5] Parry, N.J. (1996) PhD thesis, Reading University, England.
[6] Shepherd, M.G., Cole, A.L. and Tong, C.C. (1988) Methods Enzymol. 160, 300–307.
[7] Lo Leggio, L., Parry, N.J., Van Beeumen, J., Claeyssens, M., Bhat, M.K. and Pickersgill, R.W. (1997) Acta Crystallogr. D53, 599–604.
[8] Coutinho, P.M. and Henrissat, B. (1999) in: Recent Advances in Carbohydrate Bioengineering, pp. 3–12, The Royal Society of Chemistry.
[9] Wang, Q., Tull, D., Meinke, A., Gilkes, N.R., Warren, R.A.J., Aebersold, R. and Withers, S.G. (1993) J. Biol. Chem. 268, 14096–14102.
[10] Bortoli-German, I., Haiech, J., Chippaux, M. and Barras, F. (1995) J. Mol. Biol. 246, 82–94.
[11] Davies, G.J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A.M., Schülein, M. and Withers, S.G. (1998) Biochemistry 37, 11707–11713.
[12] Jenkins, J., Lo Leggio, L., Harris, G.W. and Pickersgill, R.W. (1995) FEBS Lett. 362, 281–285.
[13] Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.-P. and Davies, G. (1995) Proc. Natl. Acad. Sci. USA 92, 7090–7094.
[14] Beguin, P. (1990) Annu. Rev. Microbiol. 44, 219–248.
[15] Hilge, M., Gloor, S.M., Rypniewski, W., Sauer, O., Heightman, T.D., Zimmermann, W., Winterhalter, K. and Piontek, K. (1998) Structure 6, 1433–1444.
[16] Cutfield, S.M., Davies, G.J., Murshudov, G., Anderson, B.F., Moody, P.C.E., Sullivan, P.A. and Cutfield, J.F. (1999) J. Mol. Biol. 294, 771–783.
[17] Sakon, J., Adney, W.S., Himmel, M.E., Thomas, S.R. and Karplus, P.A. (1996) Biochemistry 35, 10648–10660.
[18] Davies, G.J., Dauter, M., Brzozowski, A.M., Bjørnvad, M.E., Andersen, K.V. and Schülein, M. (1998) Biochemistry 37, 1926–1932.
[19] Shirai, T., Ishida, H., Noda, J., Yamane, T., Ozaki, K., Hakamada, Y. and Ito, S. (2001) J. Mol. Biol. 310, 1079–1087.
[20] Chapon, V., Czjzek, M., El Hassouni, M., Py, B., Juy, M. and Barras, F. (2001) J. Mol. Biol. 310, 1055–1066.
[21] Dominguez, R., Souchon, H., Spinelli, S., Dauter, Z., Wilson, K.S., Chauvaux, S., Béguin, P. and Alzari, P.M. (1995) Nat. Struct. Biol. 2, 569–576.
[22] Ducros, V., Czjzek, M., Belaich, A., Gaudin, C., Fierobe, H.-P., Belaich, J.-P., Davies, G.J. and Haser, R. (1995) Structure 3, 939–949.
[23] Sabini, E., Schubert, H., Murshudov, G., Wilson, K.S., Siika-Aho, M. and Penttilä, M. (2000) Acta Crystallogr. D 56, 3–13.
[24] de La Fortelle, E. and Bricogne, G. (1997) Methods Enzymol. 276, 472–494.

[25] Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991) Acta Crystallogr. A47, 110–119.
[26] Cowtan, K. (1994) Joint CCP4 and ESF-EACBM Newslett. Protein Crystallogr. 31, 34–38.
[27] Lamzin, V.S. and Wilson, K.S. (1993) Acta Crystallogr. D49, 129–147.
[28] Hong, J., Tamaki, H. and Kumagai, H. (2001) AC:AAL16412.1 (GI:16356671) direct database submission.
[29] Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Acta Crystallogr. D54, 905–921.
[30] Lu, G. (2001) CCP4 Program Manuals, version 4.1.
[31] Laskowski, R.A., Moss, D.S. and Thornton, J.M. (1993) J. Appl. Crystallogr. 26, 283–291.
[32] Holm, L. and Sander, C. (1996) Science 273, 595–602.
[33] Holm, L. and Sander, C. (1993) J. Mol. Biol. 233, 123–138.
[34] Mikami, B., Degano, M., Hehre, E.J. and Sacchettini, J.C. (1994) Biochemistry 33, 7779–7787.
[35] Juers, D.H., Huber, R.E. and Matthews, B.W. (1999) Protein Sci. 8, 122–136.
[36] Dominguez, R., Souchon, H., Lascombe, M.-B. and Alzari, P.M. (1996) J. Mol. Biol. 257, 1042–1051.
[37] Aurora, R. and Rose, G.D. (1998) Protein Sci. 7, 21–38.
[38] Hogg, D., Woo, E.J., Bolam, D.N., McKie, V.A., Gilbert, H.J. and Pickersgill, R.W. (2001) J. Biol. Chem. 276, 31186–31192.
[39] Chi, Y.-I., Martinez-Cruz, L.A., Swanson, R.V., Robertson, D.E. and Kim, S.-H. (1999) FEBS Lett. 445, 375–383.
[40] Burmeister, W.P., Cottaz, S., Rollin, P., Vasella, A. and Henrissat, B. (2000) J. Biol. Chem. 275, 39385–39393.
[41] Jain, S., Drendel, W.B., Chen, Z.W., Mathews, F.S., Sly, W.S. and Grubb, J.H. (1996) Nat. Struct. Biol. 3, 375–381.
[42] Varghese, J.N., Garrett, T.P.J., Colman, P.M., Chen, L., Høj, P.B. and Fincher, G.B. (1994) Proc. Natl. Acad. Sci. USA 91, 2785–2789.
[43] Lo Leggio, L., Kalogiannis, S., Bhat, M.K. and Pickersgill, R.W. (1999) Proteins 36, 295–306.
[44] Macarrón, R., Henrissat, B. and Claeyssens, M. (1995) Biochim. Biophys. Acta 1245, 187–190.
[45] Fort, S., Varrot, A., Schülein, M., Cottaz, S., Driguez, H. and Davies, G.J. (2001) ChemBioChem 2, 319–325.
[46] Varrot, A., Schülein, M., Fruchard, S., Driguez, H. and Davies, G.J. (2001) Acta Crystallogr. D57, 1739–1742.
[47] Divne, C., Ståhlberg, J., Teeri, T.T. and Jones, T.A. (1998) J. Mol. Biol. 275, 309–325.
[48] Lo Leggio, L., Jenkins, J., Harris, G.W. and Pickersgill, R.W. (2000) Proteins 41, 362–373.
[49] Kroon-Batenburg, L.M.J. and Kroon, J. (1995) Carbohydr. Eur. 12, 15–19.
[50] Dorset, D.L. and McCourt, M.P. (1993) J. Struct. Biol. 111, 118–124.
[51] Kraulis, P.J. (1991) J. Appl. Crystallogr. 24, 946–950.
[52] Merritt, E.A. and Bacon, D.J. (1997) Methods Enzymol. 277, 505–524.
[53] Nicholls, A., Sharp, K.A. and Honig, B. (1991) Proteins 11, 281–296.
[54] Kleywegt, G.J., Zou, J.Y., Kjeldgaard, M. and Jones, T.A. (2001) International Tables for Crystallography (Rossmann, M.G. and Arnold, E., Eds.), Vol. F., pp. 353–356, 366–367.