High-resolution native and complex structures of thermostable β-mannanase from *Thermomonospora fusca* – substrate specificity in glycosyl hydrolase family 5

Mark Hilge[1†*], Sergio M Gloor[1], Wojciech Rypniewski[2], Oliver Sauer[3], Tom D Heightman[4], Wolfgang Zimmermann[5], Kaspar Winterhalter[1] and Klaus Piontek[1*]

Background: β-Mannanases hydrolyse the O-glycosidic bonds in mannan, a hemicellulose constituent of plants. These enzymes have potential use in pulp and paper production and are of significant biotechnological interest. Thermostable β-mannanases would be particularly useful due to their high temperature optimum and broad pH tolerance. The thermophilic actinomycete *Thermomonospora fusca* secretes at least one β-mannanase (molecular mass 38 kDa) with a temperature optimum of 80°C. No three-dimensional structure of a mannan-degrading enzyme has been reported until now.

Results: The crystal structure of the thermostable β-mannanase from *T. fusca* has been determined by the multiple isomorphous replacement method and refined to 1.5 Å resolution. In addition to the native enzyme, the structures of the mannotriose- and mannohexaose-bound forms of the enzyme have been determined to resolutions of 1.9 Å and 1.6 Å, respectively.

Conclusions: Analysis of the −1 subsite of *T. fusca* mannanase reveals neither a favourable interaction towards the axial HO–C(2) nor a discrimination against the equatorial hydroxyl group of gluco-configurated substrates. We propose that selectivity arises from two possible mechanisms: a hydrophobic interaction of the substrate with Val263, conserved in family 5 bacterial mannanases, which discriminates between the different conformations of the hydroxymethyl group in native mannan and cellulose; and/or a specific interaction between Asp259 and the axial hydroxyl group at the C(2) of the substrate in the −2 subsite. Compared with the catalytic clefts of family 5 cellulases, the groove of *T. fusca* mannanase has a strongly reduced number of aromatic residues providing platforms for stacking with the substrate. This deletion of every second platform is in good agreement with the orientation of the axial hydroxyl groups in mannan.

Addresses: [1]Laboratorium für Biochemie, ETH Zentrum, Universitätstrasse 16, CH-8092 Zürich, Switzerland, [2]European Molecular Biology Laboratory, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany, [3]Abteilung für Strukturbiologie, Institut für Physikalische Chemie, Karl-Franzens-Universität Graz, A-8010 Graz, Austria, [4]Laboratorium für Organische Chemie, ETH Zentrum, Universitätstrasse 16, CH-8092 Zürich, Switzerland and [5]Biotechnology Laboratory, Aalborg University, Sohngaardsholmsvej 57, DK-9000 Aalborg, Denmark.

[†]Present address: Biophysical Structural Chemistry, Gorlaeus Laboratories, Leiden University, PO Box 9502, 2300 RA Leiden, The Netherlands.

*Corresponding authors.
E-mail: hilge@chem.leidenuniv.nl
klaus.plontek@bc.biol.ethz.ch

Key words: crystal structure, family 5, glycosyl hydrolase, mannanase, *Thermomonospora fusca*

Received: **12 August 1998**
Revisions requested: **9 September 1998**
Revisions received: **17 September 1998**
Accepted: **18 September 1998**

Structure 15 November 1998, **6**:1433–1444
http://biomednet.com/elecref/0969212600601433

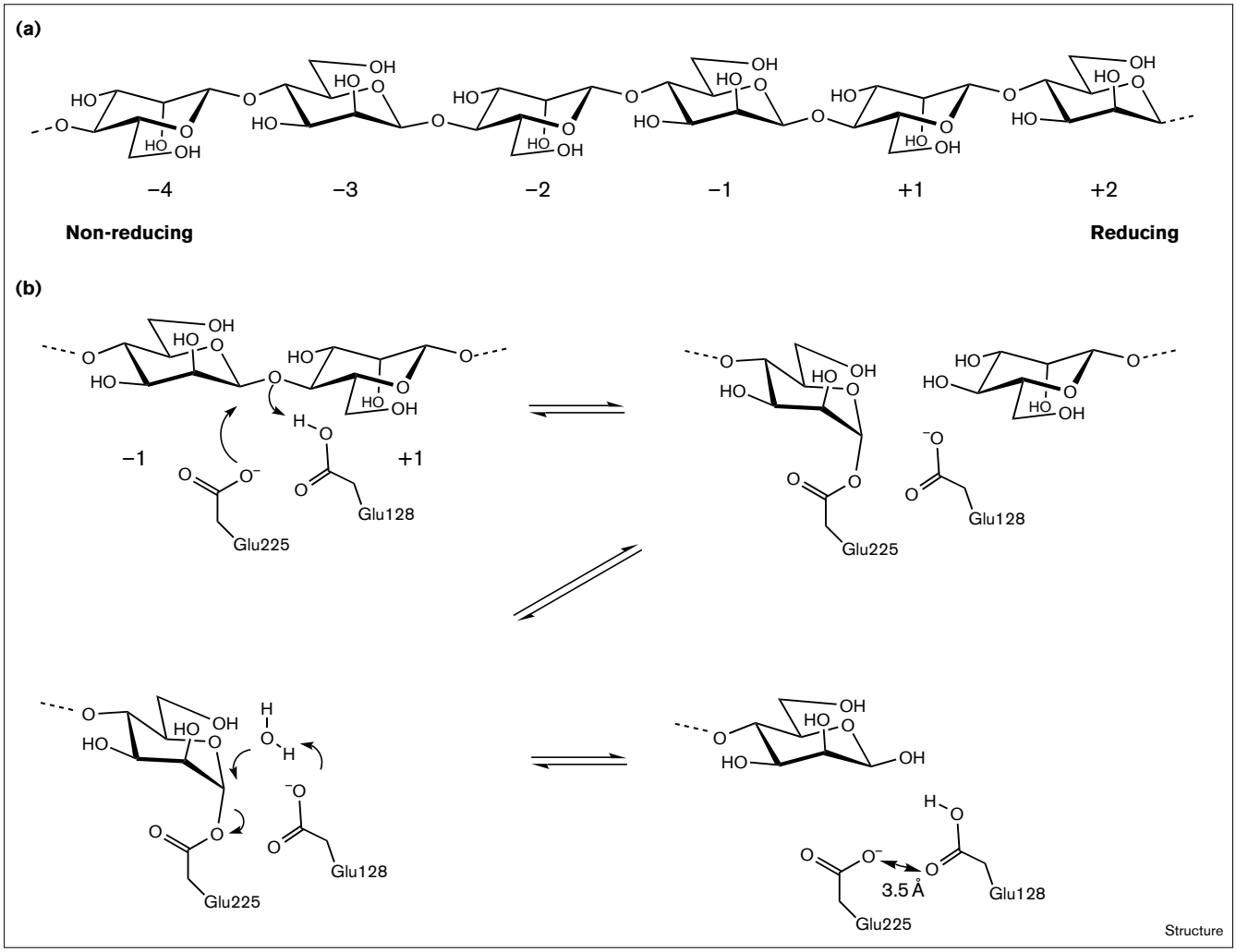
Introduction

Hemicelluloses are linear or branched polysaccharides that are mostly found as heteroglycans in higher terrestrial plants. Depending on their sugar backbone composition they are classified as either xylans, mannans, arabinogalactans or arabinans. The two most important representative hemicelluloses are the hetero-1,4-β-D-xylans and the hetero-1,4-β-D-mannans. In hetero-1,4-β-D-mannans, the primary hydroxyl group of the backbone residues is substituted by α-linked galactose; the degree of substitution is dependent on the source. Together with cellulose and lignin, hemicelluloses form the most abundant structural components of plant cell walls [1].

Microorganisms that degrade hemicelluloses frequently appear in soil, compost, and in the rumen of animals. To completely hydrolyse (hetero)mannans, fungi and bacteria have to produce at least one mannanase (EC 3.2.1.78), one β-mannosidase (EC 3.2.1.25) and an α-galactosidase (EC 3.2.1.22) [2]. In the process of degradation mannanases hydrolyse the 1,4-β-mannopyranoside bond in galacto-, gluco- and galactoglucomannans to mannooligomers, mannobiose and mannose, respectively. The degree of hydrolysis of galactomannan decreases with increasing substitution by galactose, suggesting that galactose residues obstruct the enzymic cleavage of the mannan backbone [2].

On the basis of sequence comparisons by homology and hydrophobic cluster analysis [3,4], the catalytic domains of glycosyl hydrolases have been classified into 64 families (http://expasy.hcuge.ch/cgi-bin/lists?glycosid.txt). Despite acting on three different substrates, endo-1,3-β-glucanases, endo-1,4-β-glucanases and endo-1,4-β-mannanases have so far all been assigned to glycosyl hydrolase family 5. As the

Figure 1



The reaction catalysed by β-mannanase. **(a)** The nomenclature for sugar-binding subsites in glycosyl hydrolases [15]. **(b)** The retaining mechanism for *T. fusca* mannanase, in which the glycosidic oxygen is protonated by Glu128 (proton donor) and the anomeric carbon atom is attacked by Glu225 (nucleophile). The resulting mannosyl–mannanase intermediate is then hydrolysed by a water molecule, generating a product with the same anomeric configuration as the substrate [17].

sequence identity among the members of this family is rarely higher than 20%, an additional classification into six subfamilies, A1–A6, has been introduced [5,6]. Enzymes within subfamilies show at least 25% sequence identity and may display similar substrate specificities. For instance, many enzymes of subfamily A3 additionally reveal lichenase activity, whereas some members of subfamily A4 show considerable xylanase activity [5].

So far the crystal structures of four family 5 cellulases, three in combination with a substrate, have been solved: *Clostridium thermocellum* cellulase (CelC) which belongs to subfamily A3 [7,8]; *Clostridium cellulolyticum* endoglucanase A (CelCCA), subfamily A4 [9]; *Acidothermus cellulolyticus* endocellulase (E1cd), subfamily A1 [10]; and *Bacillus agaradherans* endoglucanase (Cel5A), subfamily A2 [11]. With the enzymes of the glycosyl hydrolase families 1, 2, 5, 10, 17, 26, 30, 35, 39 and 42 they form the GH-A clan [12]. Enzymes of this clan share an $(\alpha/\beta)_8$-barrel fold and with the exception of family 26 three conserved active-site residues (one asparagine and two glutamate residues), and follow a retaining cleavage mechanism (Figure 1).

The few mannanase sequences reported permit assignment of these enzymes to either family 5 or 26. Recent work using hydrophobic cluster analysis has identified four conserved residues in family 26, and suggests that these mannanases also belong to the GH-A clan [13]. Mannanases within family 5 share eight strictly conserved amino acid

residues with more than 60 cellulases [10,14]. As deduced from the available cellulase crystal structures, it is obvious that all these conserved residues are located at or close to the active site. They either participate in binding of the sugar in the –1 subsite [15] or stabilise the position and protonation state of the catalytic glutamate residues [7–10].

Despite the lack of the major part of the primary structure, the 42 N-terminal amino acids of *Thermomonospora fusca* β-mannanase Q1.1 allowed the assignment of this enzyme to glycosyl hydrolase family 5 [16]. Here we report the high-resolution structures of Q1.1, in native and complexed form, and provide a rationale for the substrate specificity of this, as yet, structurally uncharacterised class of enzymes. In addition, we propose two new subfamilies for family 5.

Results and discussion

Overall structure description

The crystal structure of *T. fusca* β-mannanase Q1.1 was determined by the multiple isomorphous replacement (MIR) method at 2.2 Å resolution and experimental phases were extended to 1.5 Å. The final model, consisting of 302 amino acids, has a crystallographic R factor of 11.9% and an R free value of 17.6%. The overall structure of β-mannanase Q1.1 is presented in Figure 2a. The protein shows the architecture of the classical $(\alpha/\beta)_8$-barrel motif. Among the 28 families of glycosyl hydrolases for which a three-dimensional structure has been determined, families 1, 2, 5, 6, 10, 13, 14, 17, 18 and 20 display this $(\alpha/\beta)_8$-barrel motif [17]. The molecule has approximate dimensions of $45 \times 45 \times 40$ Å and its most obvious feature is a prominent cleft which is about 15 Å deep, 15 Å wide and 30 Å long (Figure 2b). The surface electrostatic potential is dominated by negative charges, which is consistent with the rather acidic pI (3.9) of the protein. This may relate to the stability of the enzyme in alkaline solutions.

Aside from the structural elements of the barrel, two short β strands at the N terminus form the bottom of the barrel (blue arrows; Figure 2a). These two β strands are also found in endocellulase E1cd from *A. cellulolyticus* [10]. The connections between the αβ repeats (at the bottom of the barrel) are generally very short (4–6 residues). The only disulphide bond of Q1.1, between Cys74 and Cys81, is located at the end of helix 2 and the beginning of strand 3, and might permit the tight turn from α2 to β3. This disulphide bond does not exist in the structures of CelC, CelCCA and E1cd; although Cys74 is conserved in CelCCA and CelC, there is no counterpart for Cys81.

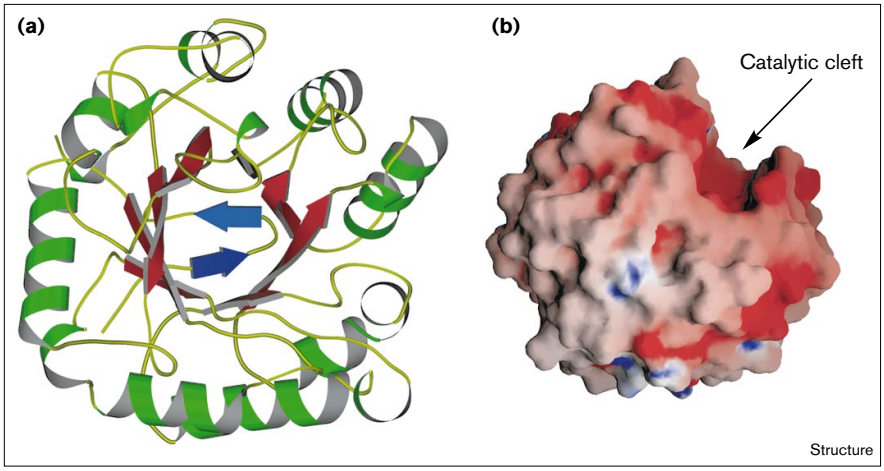
Two new family 5 subfamilies

An alignment of all the available bacterial and eukaryotic mannanase amino acid sequences in family 5 is shown in Figure 3. The alignment reveals the existence of two distinct groups of mannanases, which we propose as subfamilies A7 (eukaryotic mannanases) and A8 (bacterial mannanases). Among subfamily A8 members the sequence identity is above 43% (i.e. substantially higher than the homology within other subfamilies of glycosyl hydrolase family 5. Between members of subfamilies A7 and A8 the sequence identity is below 20%. In a pairwise alignment of *T. fusca* mannanase Q1.1 with all family 5 sequences deposited in the SWISS-PROT data bank, the highest identities were shown with the endo-1,4-β-glucanases from *Cellulomonas fimi* (23.2%; SWISS-PROT code P50400), *Ruminococcus albus* (22.7%; SWISS-PROT code Q07940) and *Streptomyces lividans* (21.6%; SWISS-PROT code P27035).

Comparison with other glycosyl hydrolases

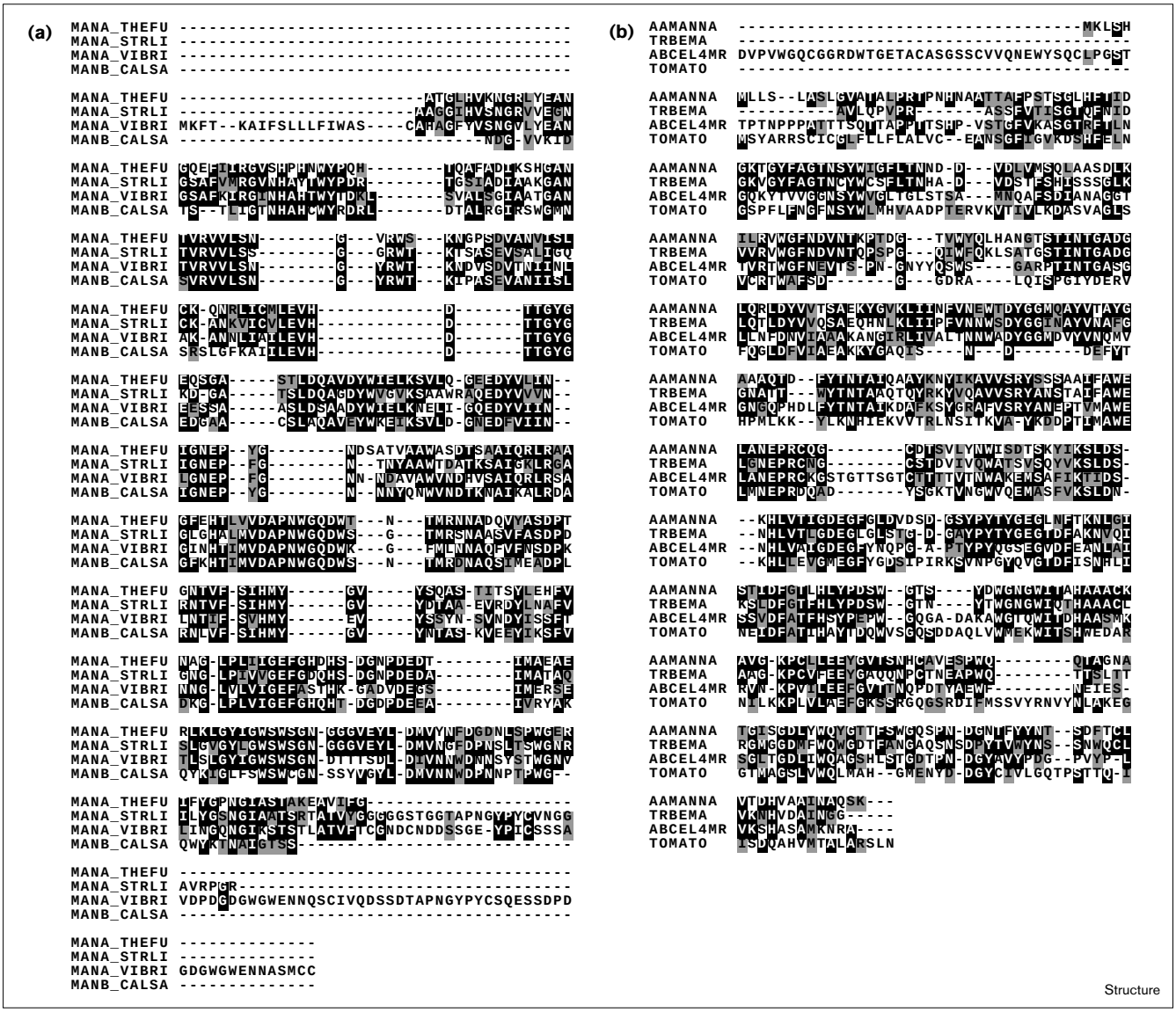
The structure of *T. fusca* β-mannanase was compared with structures deposited in the Protein Data Bank (PDB) using the program DEJAVU [18]. Thereafter, potentially

Figure 2



The overall fold of *T. fusca* β-mannanase. **(a)** Schematic representation of the structure of *T. fusca* β-mannanase. β Strands β1 and β0, which form the bottom of the barrel, are shown as blue arrows; all other β strands are shown as red arrows. Helices are depicted as green spirals. The figure was generated using the program MOLSCRIPT [43]. **(b)** Surface electrostatic potential distribution for *T. fusca* β-mannanase. Positive potential is shown in blue and negative potential in red. The figure shows the catalytic cleft at the top right of the molecule. The figure was prepared using the program GRASP [44].

Figure 3



Amino acid sequence alignment of family 5 β-mannanases. Identical amino acids are indicated by black shading and similar amino acids by grey shading. **(a)** The sequences of bacterial β-mannanases: MANA_THEFU from *Thermomonospora fusca* (EMBL code, AJ006227); MANA_STRLI from *Streptomyces lividans* (SWISS-PROT code, P51529); MANA_VIBRI from *Vibrio sp.* (EMBL code, D86329); and MANB_CALSA from *Caldocellum saccharolyticum* (SWISS-PROT code, P22533). **(b)** The sequences of eukaryotic β-mannanases: AAMANNA from *Aspergillus aculeatus* (EMBL code, L35487); TRBEMA from *Trichoderma reesei* (EMBL code, L25310); ABCEL4MR from *Agaricus bisporus* (EMBL code, Z50095); and TOMATO from *Lycopersicon esculentum* (EMBL code, AF 017144). To allow comparison between (a) and (b) gaps were introduced into the bacterial sequences. The alignment was generated using the CLUSTALW algorithm, and shading was carried out with BOXSHADE.

related structures were more precisely analysed with LSQMAN [19] (Table 1), which eliminated false positive hits. The overall fold of β-mannanase Q1.1 was, as expected, most similar to those of endocellulase E1cd [10], endoglucanase CelCCA [9] and endocellulase CelC [7,8]. Lower similarity could be detected to the catalytic domains of β-1,4-glycanase CEX from *C. fimi* [20] (PDB code 1EXP), 1,3-β-D-glucanase and 1,3;1,4-β-D-glucanase from plants [21] (PDB codes 1GHR and 1GHS), xylanase A from *S. lividans* [22] (PDB code 1XAS) and xylanase from *C. thermocellum* [8] (PDB code 1XYZ), which are all members of the GH-A clan. Besides the three key catalytic residues (Asn127, Glu128 and Glu 225), 1EXP, 1GHR and 1GHS also show the conserved Arg50 residue, which is replaced in family 10 xylanases by a valine or threonine. The sugar stacking in the –1 subsite of family

Table 1

Structural alignment of *Thermomonospora fusca* β-mannanase and related enzymes.

Protein	Family	N_r	1MAN			1EDG			1ECE			1EXP			1GHR		
			N_m	rmsd	N_c	N_m	rmsd	N_c	N_m	rmsd	N_c	N_m	rmsd	N_c	N_m	rmsd	N_c
1MAN	5	302															
1EDG	5	380	92	1.085	29												
1ECE	5	358	102	0.996	32	115	1.104	34									
1CEN	5	340	112	1.011	25	162	0.923	42	143	1.020	35						
1EXP	10	312	68	1.232	11												
1GHR	17	306	64	0.947	13							53	1.061	6			
1XAS	10	299	63	1.181	11							264	0.810	144	59	1.096	6

Pairwise root mean square deviations (rmsds; in Å) on Cα atoms for *T. fusca* β-mannanase Q1.1 (PDB code, 1MAN); *C. cellulolyticum* endoglucanase CelCCA [9] (PDB code, 1EDG); *A. cellulolyticus* endocellulase E1cd [10] (PDB code, 1ECE); *C. thermocellum* endocellulase CelC [7] (PDB code, 1CEN); *C. fimi* β-1,4-glycanase CEX [20] (PDB code, 1EXP); *Hurdeum vulgare* 1,3;1,4-β-D-glucanase [21] (PDB code, 1GHR); and *S. lividans* xylanase A [22] (PDB code, 1XAS). The alignments were carried out with the program LSQMAN [19]. A 2.0 Å cut-off was used to define structurally equivalent residues. N_r, number of residues in the protein; N_m, number of matched residues; N_c, number of matched residues of identical type in both proteins.

10 enzymes is still maintained by a tryptophan residue, whereas in 1EXP, 1GHR, 1GHS this function is achieved by a phenylalanine.

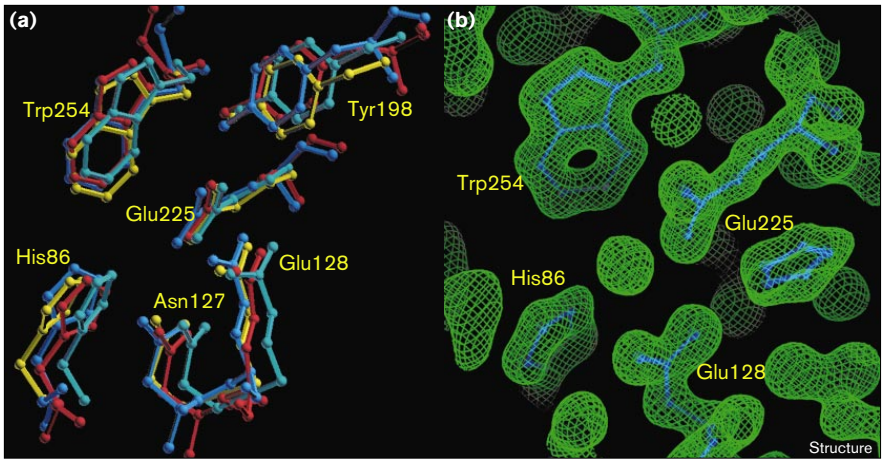
As in the other structures of enzymes from families 1, 2, 5, 17 and 18 (but surprisingly not 10), a rare nonproline *cis* peptide bond [23] occurs following the aromatic residue interacting with the –1 sugar. This *cis* peptide bond appears to be essential for the enzyme function because it constrains the position of the preceding tryptophan or phenylalanine residue [7]. In the *T. fusca* mannanase this *cis* peptide bond is formed by a serine, as it is for E1cd, the representative for subfamily 1.

The active site

In the –1 subsite, the arrangement of residues involved in the attack of the glycosidic bond is well maintained among Q1.1 and the three other family 5 enzyme structures (Figure 4a). In particular, Arg50, His86, Asn127, Glu128, His196, Tyr198, Glu225 and Trp254 are strictly conserved in family 5 enzymes.

By analogy with the substrate complexes of E1cd and CelC, Glu128 and Glu225 perform the roles of catalytic proton donor and nucleophile, respectively. His86 is responsible for the hydrogen bond mediated recognition of the HO–C(3), and Trp254 provides a hydrophobic surface for the α-face of the pyranoside ring (Figure 4b). The other conserved residues, Arg50, Asn127, His196 and Tyr198, stabilise the active-site environment and are likely to influence the protonation state of the two catalytic glutamate residues. Arg50 forms a salt bridge with Glu225 and a hydrogen bond with Asn127. Interestingly, Asn127, which may interact with the equatorial HO–C(2) in cellulases [7,10], is also present at an identical position in Q1.1, which raises the question of the role of this

Figure 4



The active site of β-mannanase. **(a)** Structural superposition of the active-site residues of family 5 glycanases: β-mannanase Q1.1 is depicted in yellow, endoglucanase CelCCA in red, endocellulase E1cd in blue, and endocellulase CelC in green. The superposition was made with the program LSQMAN [19] and displayed in the graphics program O [45]. Residues Arg50 and His196 are omitted for clarity. **(b)** $3F_o–2F_c$ electron-density map contoured at 1.5σ, displaying the β-mannanase active site. The labels correspond to the numbering in *T. fusca* β-mannanase.

Table 2

Data collection and MIR phasing statistics.

	Native 1	Mannotriose	Mannohexaose	TMLA	HgI_4	Baker	$IrCl_6$	Native 2*	Xenon*
λ (Å)	0.99	1.00	0.83	1.00	0.90	1.00	0.90	0.9	1.54
Soak concentration (mM)		100	100	100	1	2	5		
Soak pressure (bar)	–	–	–	–	–	–	–	–	50
Soak time		1.5 h	30 h	40 h	60 h	72 h	84 h		3 min
No. of sites (major/minor)				4/10	4/–	1/4	–/5		2/2
Resolution (Å)	1.80	1.9	1.6	2.40	2.35	3.0	2.50	1.5	2.2
Redundancy	2.90	3.34	3.65	3.32	3.22	3.12	2.91	3.22	2.98
Completeness (%)†	93.4 (84.8)	99.3 (97.0)	83.9 (84.5)	98.8 (97.2)	98.8 (99.2)	98.3 (98.8)	97.8 (95.9)	97.2 (92.2)	97.4 (95.8)
I/σ†	12.4 (3.5)	10.6 (2.9)	18.0 (2.6)	13.9 (6.9)	16.74 (9.1)	12.8 (6.9)	11.6 (4.9)	17.2 (2.9)	16.9 (10.5)
R_{merge} (I)†	7.3 (30.9)	10.9 (39.9)	6.2 (47.2)	8.4 (17.6)	6.1 (12.2)	9.0 (17.2)	9.5 (23.5)	6.2 (33.1)	6.2 (10.8)
Phasing power (iso/ano)				2.18/0.80	1.18/1.17	1.37/1.01	0.70/0.43		1.19/0.72

*Data collected under cryogenic conditions; all other data were collected at room temperature. †Values in parentheses correspond to those for the highest resolution shell.

residue in substrate binding and/or transition-state stabilisation (see below). His196 and Tyr198 are hydrogen bonded to Glu128 and Glu225, respectively. In contrast to the three cellulases, the highly similar arrangement of the strictly conserved sidechains in Q1.1 is achieved by an energetically unfavourable course of the backbone. This is emphasised by the Ramachandran plot where Glu128 and Tyr198 lie in the generously allowed regions.

Substrate binding

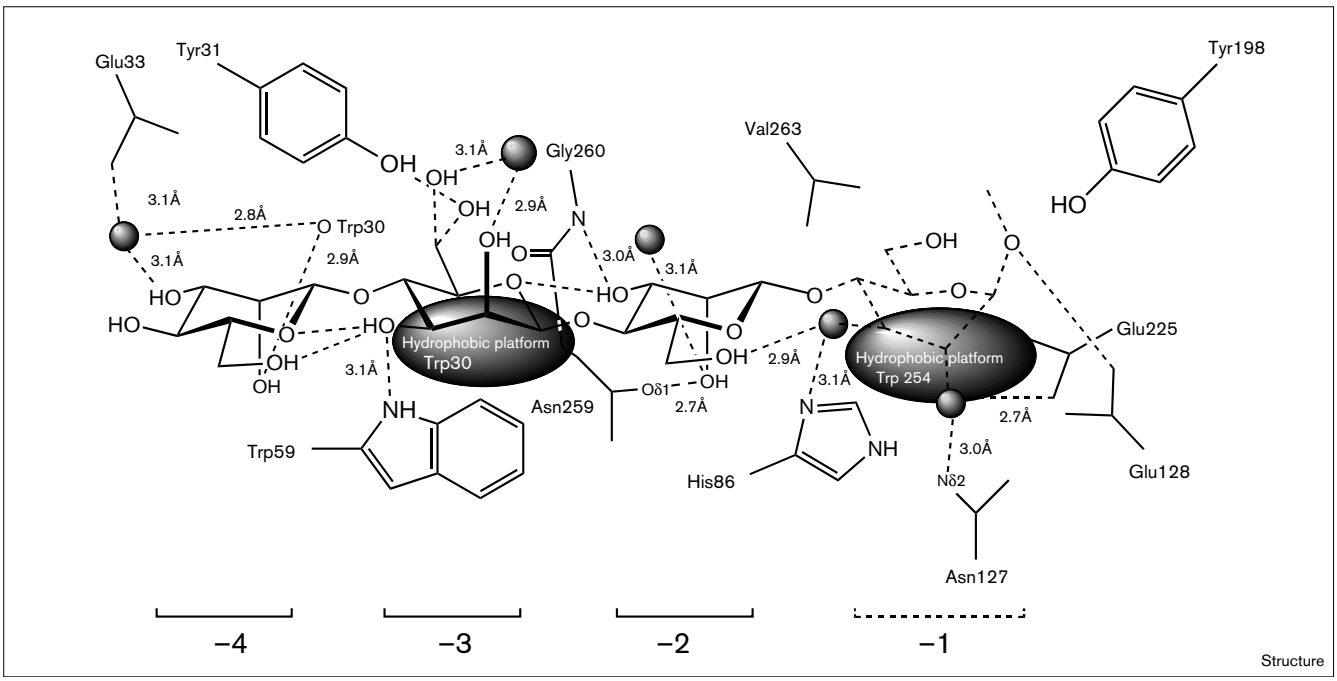
High-resolution data were collected from crystals soaked in mannotriose and mannohexaose (Table 2). The interactions between the substrate residues and *T. fusca* mannanase are shown in Figure 5. Both substrate complexes show well defined electron density, revealing an identical occupation of the –2 and –3 subsites and a potential –4 subsite (Figure 6). Similar binding to the –2 and –3 subsites was found for cellobiose in cellulase Cel5A from *Bacillus agaradherans*, unlike the cellobiose complexes of E1cd and CelC where the –1 subsite is occupied. This vacancy is consistent with the view that the –1 subsite should favour the transition state rather than substrate binding. As E1cd and CelC do not have a –3 subsite, cellobiose binding equivalent to that seen in Q1.1 and Cel5A would be unfavourable.

Whereas the –1 subsites of the three cellulases and Q1.1 have very similar structures, the –2 subsites differ considerably. In Q1.1, the groove narrows at the –2 subsite and lacks an aromatic platform compared to the cellulase structures in family 5. The only direct interactions between the –2 mannosyl residue and the protein are a 2.7 Å hydrogen bond formed between Asn259 and the axial HO–C(2) (Figure 6c) and a 3.0 Å hydrogen bond between HO–C(3) and the backbone nitrogen atom of Gly260. In E1cd, Phe29 forms a hydrophobic platform at the bottom of the relatively wide –2 subsite, so that the –2 and –1 glucosyl rings are approximately coplanar. The phenyl ring plane is nearly parallel to the –2 sugar ring, at a distance of ~5 Å. In CelC the substrate-binding groove also narrows at the –2 subsite forcing the glucose residue to twist such that its ring plane is approximately 90° to that of the –1 residue. The platform Trp212 in CelC is appropriately positioned to the side of the cleft forming van der Waals contacts with the α-face of the –2 residue.

Similarly, only a few interactions are present at the –3 subsite. The mannosyl residue is primarily bound through hydrophobic stacking with Trp30 and is laterally restricted through Trp59. In CelCCA, Trp57 shows a similar orientation to Trp30 in Q1.1. The hydrogen bond between Nε1 of Trp59 and HO–C(3) seems to be the only interaction with the protein in the 1.9 Å mannotriose complex. However, in the mannohexaose complex the hydroxymethyl group has a double conformation: in the *gauche–trans* conformation it forms a hydrogen bond with Tyr31, whereas in the *gauche–gauche* conformation it interacts via a water-mediated hydrogen bond with the axial hydroxyl group of the –3 mannosyl residue.

A novel feature in the mannotriose and mannohexaose complexes is the existence of a –4 subsite. Compared to the –3 and –2 subsites the electron density is weaker, presumably due to even fewer interactions. The 2.9 Å hydrogen bond between the axial HO–C(2) and the backbone carbonyl oxygen of Trp30 and the water-mediated

Figure 5



Schematic representation of the *T. fusca* mannanase–mannotriose complex interactions. The positions of the four subsites are indicated below. As the −1 mannosyl residue is not clearly visible in the electron-density map, it is depicted with dashed lines at the expected position. The five water molecules are represented by spheres. The two water molecules in the −1 subsite have low B factors and may represent the positions of HO–C(3) and HO–C(2). In the −3 subsite the double conformation of the primary hydroxyl group is shown.

interaction of HO–C(3) with Glu33 present the only contacts. Depending on the crystal form, additional contacts are established through interactions with a symmetry-related molecule. As a consequence the electron density beyond C(3) is improved.

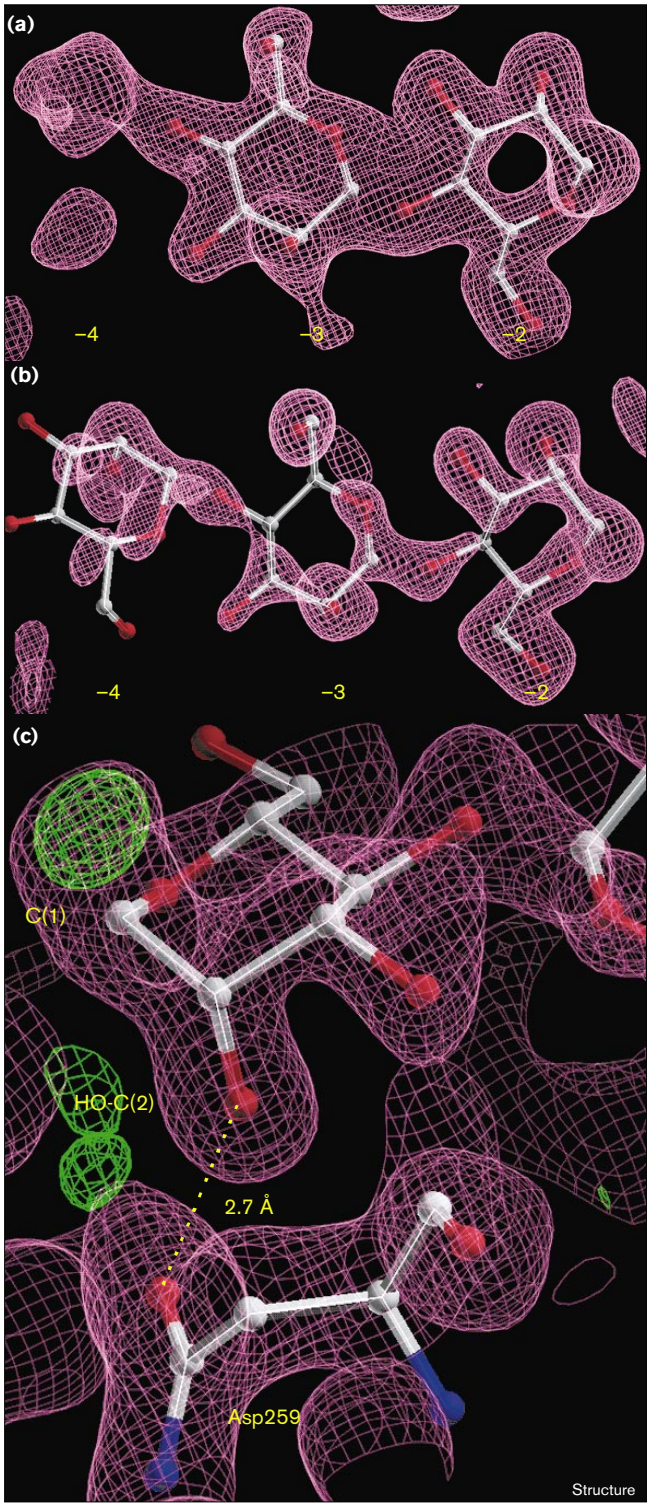
The +1 and +2 subsites of all three cellulases and Q1.1 are similar, suggesting that in Q1.1 the fissile glycosidic linkage is twisted in an equivalent way to the one in E1cd. In Q1.1 Trp167 fulfils the function of Trp213 of E1cd and Tyr176 in CelC, but there is no counterpart to Tyr245 in E1cd and Phe203 in CelC. Trp171 in Q1.1 could form the +3 subsite, allowing the mannosyl residues to adopt a conformation similar to unbound mannan. Trp57 and Trp181 are comparably positioned in CelCCA and in CelC, Phe244 provides a potential platform for the +3 subsite.

The −1 subsite

The mannohexaose complex does not reveal a clear view of the −1 subsite, although strong difference density peaks indicate positions expected for C(2), C(3) and their corresponding hydroxyl groups. The absence of density for the glycosidic bond between the −2 and −1 mannosyl residues and the lack of reliable density for the rest of the pyranoside ring prevented satisfactory construction of the −1 sugar. Superimposing the native structure with the mannotriose and mannohexaose structures revealed a well-matched pattern of electron density that may represent the water structure in the −1 subsite. Indeed, the most prominent water molecules, which probably substitute the positions of HO–C(2) and HO–C(3), are within hydrogen-bonding distance of Asn127 (3.0 Å) and His86 (3.1 Å), respectively.

The asparagine preceding the proton donor is strictly conserved in all glycanases of the GH-A clan and mutations of this residue lead to complete loss of activity [24]. As previously suggested [7,10], this residue might play a more important role in transition-state binding than in substrate recognition, as the interatomic distance between the equatorial HO–C(2) of the substrate and the δN and δO atoms of asparagine is relatively long (~3.3 Å [10]). On forming the glycosyl ester intermediate this distance is reduced, as observed in the 2″,4″-dinitrophenyl-2-deoxy-2-fluoro-β-cellobioside–xylanase/glucanase complex from *C. fimi* (2.8 Å) [25]. Consistent with this assumption is the conservation of Asn127 in Q1.1, although the distance to the axial HO–C(2) is even longer (~4 Å). In the transition state, via which the glycosyl ester is formed, the −1 sugar residue must approach the catalytic nucleophile, and the concomitant ring distortion would place the HO–C(2) in a pseudo-equatorial position. Both the approach of the −1

Figure 6



F_o–F_c electron-density maps contoured at 3σ before the incorporation of the carbohydrate compound. The electron density for **(a)** the mannotriose and **(b)** the mannohexaose complex. **(c)** The positive difference density (in green) for the α-anomer above C(1) in the mannohexaose complex. In addition, the 2.7 Å hydrogen bond between Asp259 and the axial HO–C(2) is shown.

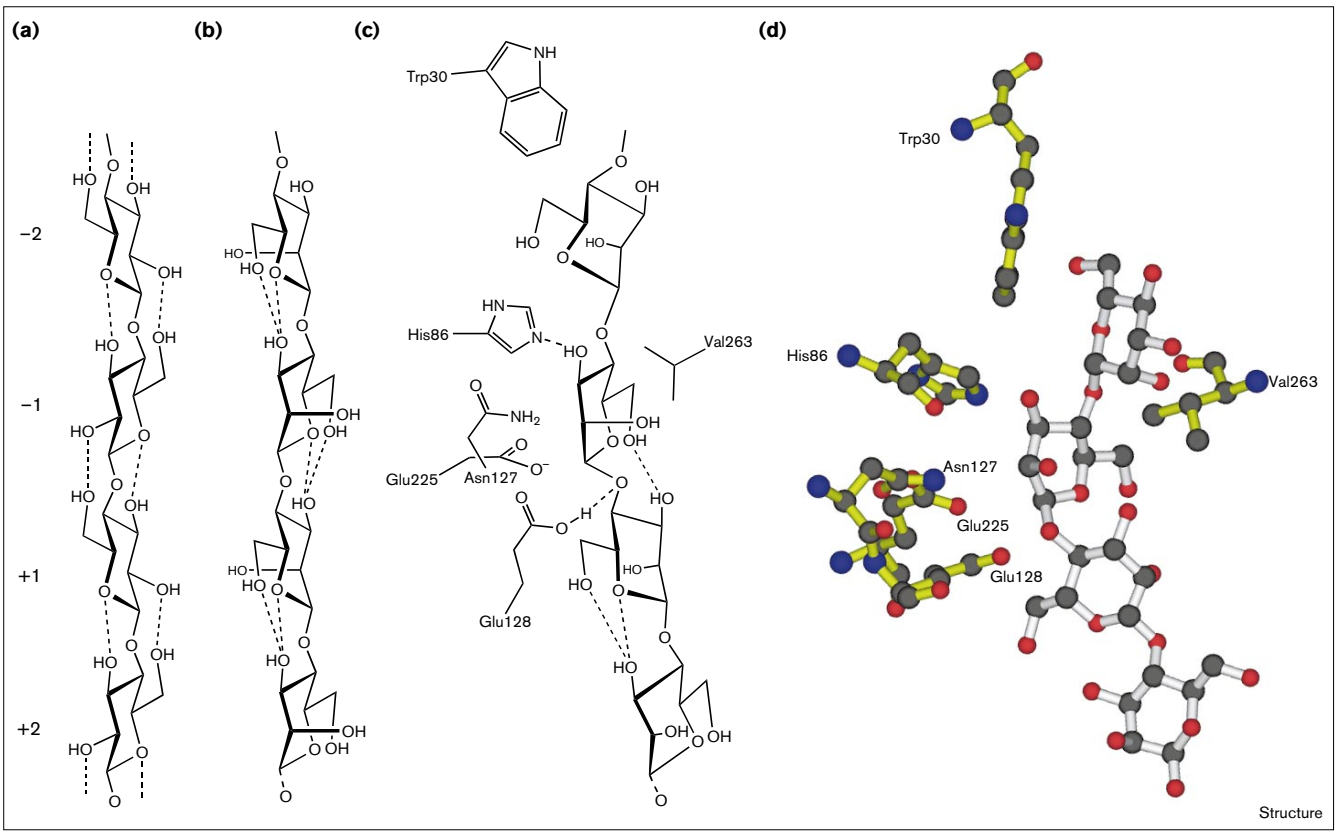
sugar and its ring distortion reduce the distance between HO–C(2) and Asn127. In this way an additional hydrogen bond is formed which contributes to the catalytic effectiveness of the enzyme.

Substrate specificity

Despite the rather limited number of interactions with the protein, the mannosyl residues bind specifically into the catalytic cleft, as indicated by the equivalent sugar atom positions in the mannotriose and mannohexaose complexes (root mean square deviation [rmsd] 0.1 Å). As expected, for mannose the reducing end of the substrate is present in the preferred α-configuration (Figure 6c). In addition, there is some evidence for an α-anomer in the –4 subsite, although not as prominent as for the –2 subsite. The α-anomer seems to be shifted by 0.2 Å from its expected position suggesting that the substrate is not yet properly positioned in the –4 subsite.

Among the various substrates hydrolysed by members of the GH-A clan, mannan is the only compound having the HO–C(2) group in the axial position. As mannan appears at first glance to be otherwise identical to cellulose, the preference of Q1.1 for mannan might be expected to result from either a stabilising interaction towards the axial hydroxyl group at C(2) of the pyranose ring, or a discrimination against the equatorial hydroxyl group of gluco-configurated substrates. At sugar subsite –1, proposed as the catalytic subsite, neither of these criteria are fulfiled in the Q1.1 structure (results based on a cellotetraose model, described below), suggesting that Q1.1 should also hydrolyse gluco-configurated substrates. Indeed, Q1.1 shows activity towards xylan (below 10% compared with mannan), and to an even lesser extent towards carboxymethylcellulose. This is in contrast to the β-mannanase from *S. lividans*, where no xylanase, endocellulase or α-galactosidase activity could be detected [26]. Mannan is still the preferred substrate, however, suggesting that the substrate selectivity may arise from conformational, rather than configurational, differences between mannan and its gluco-configurated analogues. Electron diffraction studies of mannan [27,28] show a conformation similar to that of crystalline cellulose, in which alternating rings are oriented ~180° to each other. In cellulose this conformation is stabilised by hydrogen bonds between HO–C(3) of one residue and O5 of the neighbouring residue, and between the hydroxymethyl group of the first residue and HO–C(2) of the second (Figure 7a). In mannan this second hydrogen bond is not possible because HO–C(2) is axial. The conformations of the glycosidic linkage and the hydroxymethyl group change to accommodate an alternative hydrogen-bonding scheme in which HO–C(3) is hydrogen bonded to both O5 and the hydroxymethyl group of the neighbouring residue (Figure 7b). The main distinguishing features between mannan and cellulose are thus the slightly

Figure 7



The substrate specificity of mannanase Q1.1. A comparison between **(a)** cellulose and **(b)** mannan illustrates the different hydrogen-bond networks and the differing conformations of the hydroxymethyl group. Illustrations in **(c)** and **(d)** show a model of mannotetraose (bonds in white) docked into the active site of Q1.1 (bonds in yellow) based on the mannotriose and mannohexaose complexes. Atoms are shown in standard colours.

shifted torsion angles (in cellulose I $\phi = -98°$, $\psi = -143°$; in mannan $\phi = -90°$, $\psi = -149°$ [29]) of the glycosidic linkage, and the conformation of the hydroxymethyl group (*trans–gauche* in cellulose I, $\chi = -81°$; *gauche–trans* in mannan, $\chi = 175°$).

On the basis of the structure of bound mannotriose, cellotetraose of E1cd was docked into the active site of Q1.1, superimposing the –2 glucosyl residue exactly on the equivalent residue of mannotriose (Figures 7c and 7d). In this position the hydrogen bonds between the catalytic proton donor Glu128 and the glycosidic oxygen and between His86 and HO–C(3) are preserved, as is the hydrophobic interaction with Trp254. In E1cd the hydroxymethyl group of the –1 glucose residue is bound in the *gauche–gauche* conformation and is hydrogen bonded to Asp259. In Q1.1 this position is occupied by Val263, forcing the hydroxymethyl group to bind in the *gauche–trans* conformation which is preferred in the unbound mannan structure (Figures 7c and 7d). Thus, a cellulose chain would have to undergo considerably more conformational distortion and disruption of the hydrogen-bond network in order to be bound to the active site of Q1.1. Xylan, which does not have the hydroxymethyl group, would undergo less hydrogen-bond disruption on binding to Q1.1, as reflected in its higher rate of hydrolysis.

Interestingly, three consecutive glycine residues immediately preceding Val263 are found in a double conformation, forming a flexible loop. Furthermore, the B factors of Val263 are 4 Å^2 lower in the room temperature complex structures than compared to the cryo structure without substrate, suggesting a more fixed position in the complex.

A mannanase-specific feature is the relative paucity, compared to cellulases, of hydrophobic platforms undergoing van der Waals interactions with the substrate. In fact, in *T. fusca* mannanase there are only stacking partners for the –3, –1 and perhaps +1 and +3 subsites, and these are exclusively maintained by tryptophan residues. This is in good agreement with the axial hydroxyl group pointing in the opposite direction of the aromatic residues. In all the other subsites the axial hydroxyl group would be directed towards

the aromatic ring. Thus, in cellulases, the axial hydroxyl group of mannan could cause a negative selection and prevent proper binding to the enzyme. On the other hand, a lower binding affinity for mannanase has to be expected for substrates lacking the axial hydroxyl group, due to the reduced number of hydrophobic platforms.

Finally, the mannanase–substrate complexes indicate that the distinction between mannan and other substrates arises not only in the –1 subsite, but may also result from interactions with the axial HO–C(2) in subsites lacking a hydrophobic platform (Figure 5). Among the three axial hydroxyl groups visible in the electron density, HO–C(2) in the –2 subsite is of special interest, as it forms a strong hydrogen bond with Asp259 (Figure 6c). The hydrogen bond between HO–C(2) and the Nε1 of the strictly conserved Trp254, as found for family 5 cellulases, is not possible for Q1.1. These considerations make Val263 and Asp259 prime candidates for mutations.

Thermostability

Enzymes from thermophiles have considerable potential in biotechnological applications because of their resistance to heat denaturation and consequently lower replacement costs when integrated into high temperature processes. The structural basis of thermophilic properties is still poorly understood, however. As deduced from structures of hyperthermophilic organisms with a growth temperature optimum above 80°C, oligomerisation of the protein, ionic interactions at the protein surface and multiple interactions among different segments appear to be the most important determinants for thermostability [30]. In a survey of 38 high-resolution protein structures it was found that typical mesophilic proteins undergo on average 0.04 ion-pair interactions per residue. This value is up to three times higher for thermostable proteins [31] and is 0.07 for Q1.1. There are also indications that the utilisation of arginine in such ion pairs is strongly preferred. Arginine residues participating in ion pairs show multiple hydrogen bonds between the guanidinium group of the arginine and the carboxylate group of the acidic partner in the ion pair. In ion pairs involving lysine, no equivalent bidentate interactions are observed [32].

Like the endocellulase E1cd from *A. cellolyticus* [10], Q1.1 β-mannanase has a temperature optimum of 80°C. Given their low sequence identity, it is rather difficult to compare these two structures in terms of interactions responsible for thermostability. The *S. lividans* β-mannanase has 54% sequence identity to *T. fusca* β-mannanase, but shows a considerably lower temperature optimum of 58°C. Assuming that mainly salt bridges are responsible for thermostability, the aligned sequences were investigated for differences in their charge: 17 charged residues not present in *S. lividans* mannanase and nine charged residues not present in *T. fusca* mannanase were found. Expecting high structural similarity these charge differences were checked in the atomic model, which revealed that Q1.1 might form ten salt bridges more than *S. lividans* mannanase. In addition, Q1.1 displays four ion triplets, two of them including an arginine forming hydrogen bonds. Site-directed mutagenesis experiments could verify the role of these residues in thermostability.

Biological implications

Thermostable and broad pH-tolerant enzymes would be particularly suited for pulp and paper processing performed at elevated temperature and alkaline pH conditions. Glycosyl hydrolases, such as xylanases, enhance the bleaching, thereby reducing the consumption of chlorine, peroxide and ozone-based bleaching chemicals. The addition of mannanases could improve the bleaching results. Thus *Thermomonospora fusca* β-mannanase Q1.1 is of much interest due to its tolerance to a broad, alkaline pH range and high temperature optimum of 80°C. Knowledge of the tertiary structure of this mannanase, together with the availability of the gene sequence, is essential for the development of engineered enzymes for use in biotechnological processes.

The three-dimensional structures of Q1.1 and its substrate complexes provide insight into the mechanism of this, as yet, uncharacterised class of enzymes and contribute to an understanding of substrate specificities in family 5 glycosyl hydrolases. The arrangement of amino acid sidechains in subsite –1 of the active site is almost identical to those of family 5 cellulases and no specific interactions with the axial HO–C(2) group could be found, suggesting that the enzyme is not specific. Indeed, the Q1.1 mannanase shows considerable activity towards xylan and to a much lesser extent towards cellulose. Mannan, however, is the preferred substrate. The mannotriose complex and a model for the –1 subsite suggest that Val263, conserved in bacterial family 5 mannanases, and a specific hydrogen bond towards the axial hydroxyl group at the C(2) in the –2 subsite discriminate against cellulose. Enzyme-bound mannan is in a conformation close to that of unbound mannan, allowing the conservation of its hydrogen-bonding system, whereas the binding of cellulose would require an unfavourable disruption of its hydrogen-bond network. Site-directed mutagenesis will now be required to test these assumptions experimentally and to further clarify the understanding of substrate specificity.

Materials and methods

Protein purification and crystallisation

β-Mannanases from *T. fusca* were purified by anion-exchange and affinity chromatography. Crystals were grown by the hanging-drop vapour diffusion method at room temperature as described previously [16], using a reservoir solution of 1.85 M ammonium sulphate, 0.1 M HEPES (pH 7.0). The protein crystallised in the orthorhombic space group $P2_12_12_1$ with cell dimensions of either a = 43.70 Å, b = 46.06 Å and

c = 132.51 Å (native) or a = 46.8 Å, b = 71.6 Å and c = 98.1 Å (mannotriose and mannohexaose complex). Both crystal forms have one molecule in the asymmetric unit.

Data collection

For phase determination the method of MIR was used. By soaking native protein crystals in microbridges, heavy-atom derivatives were obtained with trimethyl lead acetate (TMLA), HgI_4, Baker's dimercurial and $IrCl_6$. In addition, a xenon derivative was prepared by applying 50 bar xenon pressure for 3 min [33]. Data sets for native 2 and the xenon derivatives were collected at 100K; data sets for native 1, all the other derivatives and the substrate complexes were collected at room temperature. With the exception of the xenon derivative (for which a rotating anode was used), all data were collected at the EMBL outstation in Hamburg using synchrotron radiation. Data processing and reduction was performed with the HKL package [34]. A summary of data collection statistics is shown in Table 2.

Structure determination

Patterson maps for all derivatives were calculated using the programs in the CCP4 package [35]. Only sites in the xenon and TMLA Patterson maps were unambiguously interpretable. Minor sites in the TMLA derivative and the sites in the other room temperature derivatives were located by means of residual maps implemented in SHARP [36]. The anomalous differences allowed the decision for the correct hand and contributed a considerable amount of phasing power. The statistics for the heavy-atom refinement are given in Table 2. With a SOLOMON [37] solvent-flattening procedure implemented in SHARP, phases were directly extended from 2.44 to 1.5 Å, revealing a map of high quality. Afterwards an electron-density map on a fine grid (0.2 Å) was calculated for input in wARP [38] and a total of 480 cycles were run. A major obstacle in solving the β-mannanase structure was the absence of the complete amino acid sequence. However, the 1.5 Å wARP map allowed for the automatic building of the backbone for 292 of the final 302 residues visible in the electron density; the building was performed using an unpublished program by A Perrakis and V Lamzin. The density was not continuous for only two loop regions (residues 229–233 and residues 260–264) and the C terminus. About 260 amino acids could be assigned to the electron density with high probability. An unambiguous stretch of eight amino acids at the C terminus permitted the design of degenerate downstream primers. Together with a specific upstream primer, approximately 90% of the mannanase gene sequence could be obtained by use of the polymerase chain reaction (PCR) and DNA sequencing. The polypeptide sequence previously assigned from the electron density was then checked and, where necessary, corrected.

Structure refinement

To obtain an unbiased indicator of the refinement progress, 5% of the reflections were set aside for free R factor calculations. Prior to every rebuilding step, the Ramachandran plot was inspected and real-space fit values were calculated with OOPS [39] to highlight trouble spots. The model, consisting of 292 residues, was refined in a combined procedure using REFMAC [40] and ARP [41]. After several rounds of model building and REFMAC/ARP refinement the two loops, which turned out to form double conformations, could be built at least for the higher occupied species, resulting in a model containing 302 amino acids. Extensive trials to build the C terminus failed.

Solutions for the second crystal form, used for the mannotriose and mannohexaose soaks, were easily obtained by molecular replacement with AMORE [42]. The refinement of the substrate complexes revealed the same troublesome regions as found for the native structure. The loop containing the three glycine residues in double conformation could be built, whereas the region of residues 229–233 was too disordered to allow reliable modelling. Finally, the temperature factors for atoms of the 1.5 Å native and 1.6 Å mannohexaose structure were refined anisotropically (REFMAC version 4.0.1 [40]), reducing the free R factor by 2.5 and 3.1%, respectively. For all the structures, no residues were detected in the disallowed regions of the Ramachandran plot; two residues (Glu128 and Tyr198, both involved in catalysis) lie in the generously allowed region. Superposition of the native structure with the two complex structures revealed rms distance values of 0.21 Å for both complexes, the largest differences were found in the two flexible loop regions. A summary of refinement and model statistics is given in Table 3.

Table 3

Refinement and model statistics.

	Native 2	Mannotriose	Manno-hexaose
Resolution range (Å)	32–1.5	30–1.9	18–1.6
Number of reflections work/test	40,351/2146	25,233/1344	36,145/1159
Final R factor (all data) (%)	11.9	17.8	12.8
Last recorded value of R free (all data) (%)	17.6	21.7	16.6
No. of protein atoms	2328	2324	2324
No. of water molecules	426	197	328
Average B factors			
protein atoms (Å^2)	12.4	18.9	18.5
water molecules (Å^2)	31.0	35.4	42.8
−2 mannosyl residue (Å^2)	–	~28	~29
−3 mannosyl residue (Å^2)	–	~28	~28
−4 mannosyl residue (Å^2)	–	–	~26
Rms deviation			
bond length (Å)	0.007	0.010	0.009
bond angles (Å)	0.022	0.027	0.027
Ramachandran plot outliers	none	none	none
Residues with unusual peptide orientations	Ser255	Ser255	Ser255

Restraints used in the isotropic and anisotropic refinement were 0.02 Å for the bond lengths and 0.04 Å for the angles. Anisotropic temperature factors were restrained according to the standard criteria in REFMAC and the weights were optimised to achieve the maximum drop in free R factor.

Accession numbers

The atomic coordinates of native β-mannanase Q1.1 and the two complexes have been deposited with the Protein Data Bank (accession codes 1MAN, 2MAN and 3MAN, respectively). The DNA sequence has been deposited with the EMBL data bank (accession number AJ006227).

Acknowledgements

We are very grateful to Eric de La Fortelle, Anastassis Perrakis and Victor Lamzin for their extensive help and enthusiasm with the programs SHARP, wARP and ARP, respectively. We would also like to thank Christoph Kratky for generously assigned beam time on his rotating anode and Gerlind Sulzenbacher for fruitful discussions on substrate complexes. We particularly thank the people at the EMBL outstation at DESY in Hamburg and at the Swiss Norwegian Beamline at ESRF in Grenoble for their support. Alexander Kraev, Lei Li and Fredi Stutz are thanked for competent help in DNA sequencing. This work was funded by an ETH grant 0-20-158-96 and the company Dr W Kolb AG, CH-8903 Hedingen.

References

1. Zimmermann, W. (1991). Bacterial degradation of hemicelluloses. In *Degradation of Environmental Pollutants by Microorganisms and their Metalloenzymes.* (Sigel, H. & Sigel, A eds), pp. 357-398, Marcel Dekker, New York.
2. McCleary, B.V. (1983). Enzymic interactions in the hydrolysis of galactomannan in germinating guar: the role of exo-β-mannanase. *Phytochemistry* **22**, 649-658.

3. Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309-316.
4. Henrissat, B. & Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid similarities. *Biochem. J.* **293**, 781-788.
5. Béguin, P. (1990). Molecular biology of cellulose degradation. In *Annu. Rev. Microbiol.* **44**, (Ornston, N., ed), pp. 219-248, Annual Reviews Inc., Palo Alto.
6. Lo Leggio, L., Parry, N.J., van Beeumen, J., Claeyssens, M., Bhat, M.K. & Pickersgill, R.W. (1997). Crystallization and preliminary X-ray analysis of the major endoglucanase from *Thermoascus aurantiacus*. *Acta Cryst. D* **53**, 599-604.
7. Dominguez, R., Souchon, H., Lascombe, M.-B. & Alzari, P.M. (1996). The crystal structure of a family 5 endoglucanase mutant in complexed and uncomplexed forms reveals an induced fit activation mechanism. *J. Mol. Biol.* **257**, 1042-1051.
8. Dominguez, R., *et al.*, & Alzari, P.M. (1995). A common protein fold and similar active site in two distinct families of β-glycanases. *Nat. Struct. Biol.* **2**, 569-576.
9. Ducros, V., *et al.*, & Haser, R. (1995). Crystal structure of the catalytic domain of a bacterial cellulase belonging to family 5. *Structure* **3**, 939-949.
10. Sakon, J., Adney, W.S., Himmel, M.E., Thomas, S.R. & Karplus, P.A. (1996). Crystal structure of thermostable family 5 endocellulase E1 from *Acidothermus cellulolyticus* in complex with cellotetraose. *Biochemistry* **35**, 10648-10660.
11. Davies, G.J., Dauter, M., Brzozowski, A.M., Bjørnvad, M.E., Andersen, K.V. & Schülein, M. (1998). Structure of the *Bacillus agardherans* family 5 endoglucanase at 1.6 Å and its cellobiose complex at 2.0 Å resolution. *Biochemistry* **37**, 1926-1932.
12. Henrissat, B. & Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7**, 637-644.
13. Bolam, D.N., *et al.*, & Gilbert, H.J. (1996). Mannanase A from *Pseudomonas fluorescens* ssp. *cellulosa* is a retaining glycosyl hydrolase in which E212 and E320 are the putative catalytic residues. *Biochemistry* **35**, 16195-16204.
14. Wang, Q., *et al.*, & Withers, S.G. (1993). Glu 280 is the nucleophile in the active site of *Clostridium thermocellum* CelC, a family A endo-β-1,4-glucanase. *J. Biol. Chem.* **268**, 14096-14102.
15. Davies, G.J., Wilson, K.S. & Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557-559.
16. Hilge, M., Gloor, S., Winterhalter, K., Zimmermann, W. & Piontek, K. (1996). Crystallization and preliminary crystallographic analysis of two β-mannanase isoforms from *Thermomonospora fusca* KW3. *Acta Cryst. D* **52**, 1224-1225.
17. Davies, G. & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859.
18. Kleywegt, G.J. & Jones, T.A. (1997). Detecting folding motifs and similarities in protein structures. In *Methods in Enzymology Vol. 277*. (Carter, C.W.J. & Sweet, R.M., eds), pp. 525-545, Academic Press, New York.
19. Kleywegt, G.J. & Jones, T.A. (1994). A super position. *ESF/CCP4 Newsletter* **31**, 9-14.
20. White, A., Withers, S.G., Gilkes, N.R. & Rose, D.R. (1994). Crystal structure of the catalytic domain of the β-1,4-glycanase Cex from *Cellulomonas fimi*. *Biochemistry* **33**, 12546-12552.
21. Varghese, J.N., Garrett, T.P.J., Colman, P.M., Chen, L., Høj, P.B. & Fincher, G.B. (1994). Three-dimensional structures of two plant β-glucan endohydrolases with distinct substrate specificities. *Proc. Natl Acad. Sci. USA* **91**, 2785-2789.
22. Derewenda, U., *et al.*, & Derewenda, Z.S. (1994). Crystal structure, at 2.6 Å resolution, of the *Streptomyces lividans* xylanase A, a member of the F family of β-1,4-D-glycanases. *J. Biol. Chem.* **269**, 20811-20814.
23. Stewart, D.E., Sarkar, A. & Wampler, J.E. (1990). Occurrence and role of *cis*-peptide bonds in protein structures. *J. Mol. Biol.* **214**, 253-260.
24. Bortoli-German, I., Haiech, J., Chippaux, M. & Barras, F. (1995). Informational suppression to investigate structural functional and evolutionary aspects of the *Erwinia chrysanthemi* cellulase EGZ. *J. Mol. Biol.* **246**, 82-94.
25. White, A., Tull, D., Johns, K., Withers, S.G. & Rose, D.R. (1996). Crystallographic observation of a covalent catalytic intermediate in a β-glycosidase. *Nat. Struct. Biol.* **3**, 149-154.
26. Arcand, N., Kluepfel, D., Paradis, F.W., Morosoli, R. & Shareck, F. (1993). β-Mannanase of *Streptomyces lividans* 66: cloning and DNA sequence of the *manA* gene and characterization of the enzyme. *Biochem. J.* **290**, 857-863.
27. Chanzy, H., Pérez, S., Miller, D.P., Paradossi, G. & Winter, W.T. (1987). An electron diffraction study of mannan I. Crystal and molecular structure. *Macromolecules* **20**, 2407-2413.
28. Dorset, D.L. & McCourt, M.P. (1993). Electron crystallographic analysis of a polysaccharide structure – direct phase determination and model refinement for mannan I. *J. Struct. Biol.* **111**, 118-124.
29. Chandrasekaran, R. (1997). Polysaccharide helices in oriented fibres. *Adv. Carbohydr. Chem. Biochem.* **52**, 326-333.
30. Ermler, U., Merckel, M.C., Thauer, R.K. & Shima, S. (1997). Formylmethanofuran: tetrahydromethanopterin formyltransferase from *Methanopyrus kandleri* – new insights into salt-dependence and thermostability. *Structure* **5**, 635-646.
31. Barlow, D.J. & Thornton, J.M. (1983). Ion-pairs in proteins. *J. Mol. Biol.* **168**, 857-885.
32. Yip, K.S.P., *et al.*, & Rice, D.W. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure* **3**, 1147-1158.
33. Sauer, O., Schmidt, A. & Kratky, C. (1997). Freeze-trapping isomorphous xenon derivatives of protein crystals. *J. Appl. Cryst.* **30**, 476-486.
34. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology Vol. 276*. (Carter, C.W.J. & Sweet, R.M., eds), pp. 307-326, Academic Press, New York.
35. Collaborative Computational Project, No. 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760-763.
36. de La Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. In *Methods in Enzymology Vol. 276*. (Carter, C.W. & Sweet, R.M., eds), pp. 472-494, Academic Press, New York.
37. Abrahams, J.P. & Leslie, A.G.W. (1996). Methods used in the structure determination of bovine mitochondrial F_1 ATPase. *Acta Cryst. D* **52**, 30-42.
38. Perrakis, A., Sixma, T.K., Wilson, K.S. & Lamzin, V.S. (1997). wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Cryst. D* **53**, 448-455.
39. Kleywegt, G.J. & Jones, T.A. (1996). Efficient rebuilding of protein structures. *Acta Cryst D* **52**, 829-832.
40. Murshudov, G., Vagin, A. & Dodson, E. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst D* **53**, 240-255.
41. Lamzin, V.S. & Wilson, K.S. (1993). Automated refinement of protein models. *Acta Cryst. D* **49**, 129-147.
42. Navaza, J. (1994). AMoRe: an Automated Package for Molecular Replacement. *Acta Cryst. A* **50**, 157-163.
43. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950.
44. Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. *Biophys. J.* **64**, 166.
45. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110-119.