REVIEW

Amylolytic families of glycoside hydrolases: focus on the family GH-57

Štefan JANEČEK[1,2]

[1] *Institute of Molecular Biology, Slovak Academy of Sciences, Dúbravská cesta 21, SK-84551 Bratislava, Slovakia; phone: ++ 421 2 5930 7420, fax: ++ 421 2 5930 7416, e-mail: Stefan.Janecek@savba.sk*

[2] *Department of Biotechnologies, Faculty of Natural Sciences, University of St. Cyril and Methodius in Trnava, Nám. J. Herdu 2, SK-91701 Trnava, Slovakia*

Abstract: The amylolytic and related enzymes have been classified into the families of glycoside hydrolases (GHs). At present there are almost one hundred GH families. The main α-amylase family is the family GH-13 that forms the clan GH-H together with the families GH-70 and GH-77. β-Amylases and glucoamylases have their own families GH-14 and GH-15, respectively. Some amylolytic enzymes are grouped also in the family GH-31. The main topic of this review, the family GH-57, was established as the novel α-amylase family in 1996, based on the two sequences of amylases that were obviously dissimilar to the α-amylases of the family GH-13. The family GH-57 contains also the amylopullulanase, α-galactosidase and 4-α-glucanotransferase specificities in addition to the specificity of α-amylase. The family is remarkable by several facts: (i) it consists of only 10% of real enzymes (90% is formed by the putative proteins); (ii) the protein members originate exclusively from prokaryotic sources of extremophilic nature (many of them are archaeal hyperthermophiles); (iii) the amino acid sequences vary from less than 400 to more than 1,300 residues and are extremely diversified. There are five conserved sequence regions characteristic of the family GH-57. The catalytic domain of a GH-57 member is the $(\beta/\alpha)_7$-barrel with the Glu (nucleophile) and the Asp (proton donor) residues at the β-strands β4 and β7 of the barrel, respectively, employing the retaining mechanism. From the evolutionary point of view, the family GH-57 could be divided into several subfamilies. Importantly, the individual enzyme specificities form their own independent clusters.

Key words: glycoside hydrolases; family GH-57; alpha-amylase; 4-alpha-glucanotransferase; evolutionary relationships.

Introduction

The enormous diversity of saccharides existing on Earth is in a very close relationship with the existence of the enzymes that are able to metabolize them. Glycoside hydrolases (GH) constitute one large group of these enzymes. Due to the evolutionary rules they are mutually interconnected with other enzyme classes, especially with transferases, lyases and isomerases, to form the families widely represented by various enzymes, e.g. the α-amylase family (MACGREGOR et al., 2001). This situation has been most usefully reflected in establishing the sequence-based classification system of glycoside hydrolases (HENRISSAT, 1991), nowadays available under the CAZy web-system (COUTINHO & HENRISSAT, 1999), where the individual enzymes form their own families according to their sequence-structural similarities and catalytic machineries. Currently the number of different glycoside hydrolase families reaches almost one hundred (COUTINHO & HENRISSAT, 1999).

The presented review is focused on the family GH-57 of glycoside hydrolases. This remarkable family was established in 1996 based on the existence of the sequences of two amylases dissimilar to those of α-amylases ordinary at that time (HENRISSAT & BAIROCH, 1996). In order to sketch a more comprehensive picture, however, a brief overview of other GH families containing the amylolytic enzymes is given in the first part of this article.

Amylolytic non-GH-57 families of glycoside hydrolases

Amylolytic enzymes form a substantial portion of all glycoside hydrolases. Generally these enzymes act on starch and related oligo- and polysaccharides.

The clan GH-H – the families GH-13, GH-70 and GH-77

Most of them belong to the so-called α-amylase family (MACGREGOR et al., 2001) that was established as the GH-13 (HENRISSAT, 1991) and now, in the CAZy system (COUTINHO & HENRISSAT, 1999), it constitutes the clan GH-H. This clan covers three GH fami-

lies: GH-13, GH-70 and GH-77 (MacGregor, 2005). All GH-H clan members share several characteristics: (i) their catalytic domain is formed by the $(\beta/\alpha)_8$-barrel fold (i.e. TIM-barrel) with a longer loop connecting the strand β3 to helix α3 known as a distinct domain B (Fig. 1a); (ii) they use a common catalytic machinery consisting of the β4-strand aspartate as a base (nucleophile) and β5-strand glutamate as a proton donor (acid/base catalyst) with the help of the third residue, the β7-strand aspartate, essential for substrate binding (transition state stabiliser); and (iii) they employ the retaining mechanism for the cleavage of the α-glycosidic bonds (Matsuura et al., 1984; Buisson et al., 1987; Machius et al., 1995; Aghajari et al., 1998; Matsuura, 2002). The GH-13 ranks among the largest GH families with almost 30 enzyme specificities and more than 2,000 sequences (Coutinho & Henrissat, 1999; Janecek, 2000; Pujadas & Palau, 2001; Svensson et al., 2002). It is the main family of the entire clan GH-H. In addition to α-amylase (EC 3.2.1.1) it contains cyclodextrin glucanotransferase, α-glucosidase, amylopullulanase, neopullulanase, amylosucrase, branching enzyme, etc. (MacGregor et al., 2001). It seems reasonable to group the very closely related GH-13 members into subfamilies, e.g., the α-glucosidase-like and neopullulanase-like members (Oslancova & Janecek, 2002; Oh, 2003). The GH-70 consists of glucansucrases which synthesize glucans from sucrose. They are remarkable homologues to GH-13 members in that they possess the catalytic $(\beta/\alpha)_8$-barrel domain in a form that is circularly permuted to the α-amylase type $(\beta/\alpha)_8$-barrel (MacGregor et al., 1996). Although 40 primary structures have been described, no three-dimensional structure is, however, available (Kralj et al., 2004a,b). The family GH-77 contains amylomaltases (also known as disproportionating enzymes or 4-α-glucanotransferases) that differ from the main family GH-13 by lacking domain C (Fig. 1b) succeeding in the GH-13 the catalytic $(\beta/\alpha)_8$-barrel (Przylas et al., 2000; Strater et al., 2002; Kaper et al., 2004). One of the most important and characteristic features common to the all GH-H clan members is the presence of between four and seven conserved regions in their amino acid sequences (Janecek, 2002). Of the α-amylase family enzymes, neopullulanase (EC 3.2.1.135) occupies the exceptional position since it was proved (Hondoh et al., 2003) that it could perform in its single active site all the four reactions typical for these enzymes: hydrolysis and transglycosylation of both α-1,4- and α-1,6-glucosidic bonds (Kuriki et al., 2005). It should be pointed out, however, that some α-amylase family members are catalytically active also towards the α-1,1-, α-1,2- and α-1,3-glucosidic bonds (MacGregor et al., 2001; Janecek, 2002). All the subtle sequence and/or structural differences reflected in the flexible active site and resulting in the unusually wide spectrum of reactions catalyzed make the α-amylase family and, in a wider sense the entire clan GH-H, an attractive subject for the basic research studies and also for the industrial exploitation (MacGregor et al., 2001; Park et al., 2000; Nielsen et al., 2000; Pujadas & Palau, 2002; van der Maarel et al., 2002). It is now possible that the catalytic triad may be the only residues invariantly conserved throughout all enzymes of the entire clan GH-H (Machovic & Janecek, 2003).

The families GH-14 and GH-15

There are two other amylolytic GH families in CAZy (Coutinho & Henrissat, 1999), GH-14 and GH-15, covering β-amylases and glucoamylases, respectively. They both employ an inverting mechanism for cleaving the α-glucosidic bonds, i.e., the products of their reactions are β-anomers (Rye & Withers, 2000). Concerning the structures, β-amylase (Fig. 1c) belongs to $(\beta v\alpha)_8$-barrel proteins (Mikami et al., 1993; 1999; Cheong et al., 1995; Oyama et al., 1999) and glucoamylase (Fig. 1d) adopts an $(\alpha/\alpha)_6$-barrel fold (Aleshin et al., 1992; 2003; Sevcik et al., 1998). From an evolutionary point of view, β-amylases seem to be a "solitary" family GH-14 since - at least according to our present knowledge - they do not exhibit any pronounced sequence-structural similarity to a related glycoside hydrolase (Pujadas et al., 1996; Coutinho & Henrissat, 1999). On the other hand, glucoamylases from GH-15 were recognised to form a clan GH-L with GH-65 (Egloff et al., 2001).

The family GH-31

Other amylolytic enzymes have been classified into family GH-31 containing mostly α-glucosidases, α-xylosidases, and glucan lyases (Yu et al., 1999; Lee et al., 2003; 2005). They employ the retaining mechanism like those of the clan GH-H (Chiba, 1997; Nakai et al., 2005). The GH-31 was predicted to join the clan GH-H (Rigden, 2002). This eventuality can be supported by the revealed remote sequence similarities between the GH-13 and the GH-31, thus connecting the two families at a hierarchy higher (wider) than a clan (Janecek et al., unpublished results). Such eventuality has recently been supported by the three-dimensional structure and domain organization of the GH-31 α-xylosidase (Lovering et al., 2005) that both strongly resemble those of a GH-13 enzyme (Fig. 1e).

The family GH-57

For a long time family GH-57 belonged to the most remarkable families of glycoside hydrolases attracting much scientific interest. The family was established in 1996 (Henrissat & Bairoch, 1996) as a family with only two members: a heat-stable eubacterial α-amylase from *Dictyoglomus thermophilum* known from 1988 (Fukusumi et al., 1988) and an extremely thermostable archaeal α-amylase from *Pyrococcus furiosus* determined in 1993 (Laderman et al., 1993a). These

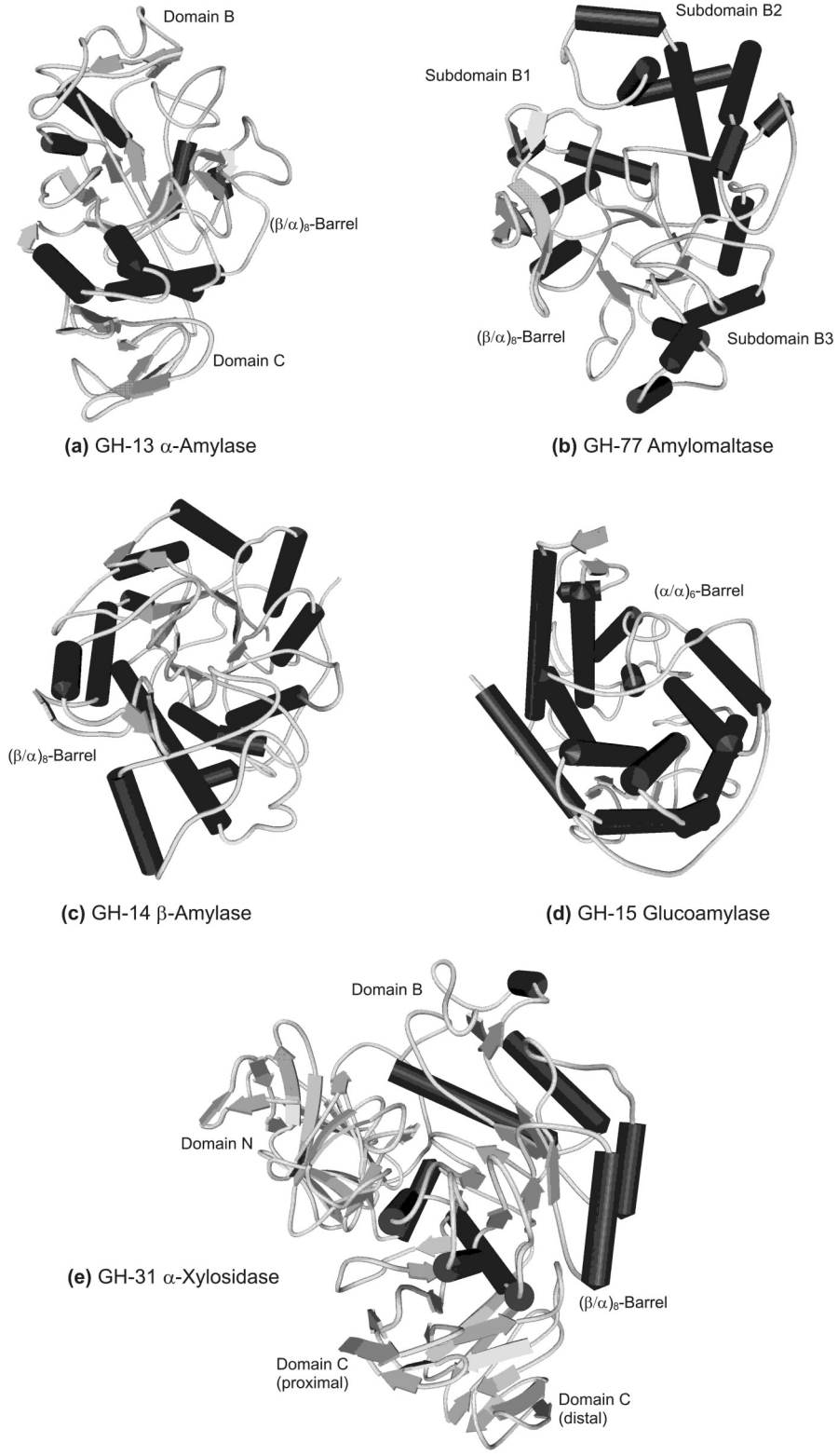


Fig. 1. Three-dimensional structures of enzymes from the glycoside hydrolase families. (***a***) GH-13 α-amylase from *Aspergillus oryzae* (PDB code: 2TAA; Matsuura et al., 1984), (***b***) GH-77 amylomaltase from *Thermus aquaticus* (1CWY; Przylas et al., 2000), (***c***) GH-14 β-amylase from soybean (1BYA; Mikami et al., 1993), (***d***) GH-15 glucoamylase from *Aspergillus awamori* (1AGM; Aleshin et al., 1992), and (***e***) GH-31 α-xylosidase from *Escherichia coli* (1XSI; Lovering et al., 2005).

two mutually similar sequences, although each apparently coding for an amylolytic enzyme (Fukusumi et al., 1988; Laderman et al., 1993a), did not exhibit any of the several conserved sequence regions characteristic of the α-amylase family, i.e. the family GH-13 (Janecek, 2002).

Very diversified sequences

During the time, especially in the last few years, when the complete genomes of many microorganisms were sequenced, the family GH-57 expanded. It has been a purely prokaryotic family with the majority of thermophilic (extremophilic) sources, archaeons being dom-

inant (Zona et al., 2004). In 2004 the GH-57 consisted of almost fifty members according to CAZy (Coutinho & Henrissat, 1999) and even more than sixty if also the sequences from the PFAM database (Bateman et al., 2002) were taken into account. At present the GH-57 contains the sequences that exhibit the enzyme specificities of α-amylase (EC 3.2.1.1), α-galactosidase (EC 3.2.1.22), amylopullulanase (EC 3.2.1.1/41) and 4-α-glucanotransferase (EC 2.4.1.25). It is worth mentioning, however, that the real enzymes form only about 10% of the family members. The vast majority of the GH-57 are hypothetical proteins obtained from whole genome sequences (Zona et al., 2004).

One of the main reasons, why the family has attracted so much scientific interest, is the fact that the sequences of GH-57 members have been extremely diversified. Certain sequences are shorter than 400 residues whereas others may be longer than 1,500 residues (Zona et al., 2004). This fact causes that the sequences cannot be aligned using the routine alignment programs. On the other hand, because these sequences form a sequence-based family, they have to contain at least a few conserved sequence segments or amino acid residues common for the entire family. Based on an exhaustive bioinformatics study focused on all available sequences, five conserved sequence regions in the family GH-57 (Fig. 2a) were proposed (Zona et al., 2004). Importantly these regions contain the GH-57 catalytic residues confirmed by site-directed mutagenesis, the Glu291 (Region 3) and Asp394 (Region 4) in the amylopullulanase from *Thermococcus hydrothermalis* as the nucleophile and the proton donor, respectively (Zona et al., 2004). These results were possible to achieve in conjunction with the identification of the catalytic nucleophile (Glu123) in the 4-α-glucanotransferase of *Thermococcus litoralis* (Imamura et al., 2001) and the determination at that time of the three-dimensional structure of this enzyme (Imamura et al., 2003). The equivalent catalytic residue (Glu117) was identified also in the α-galactosidase from *Pyrococcus furiosus*, although the probable misalignment did not allow the authors to find the catalytic proton donor (van Lieshout et al., 2003). Very recently both catalytic residues, the same as in the case of *Thermococcus hydrothermalis* amylopullulanase (Zona et al., 2004), were successfully determined in the related amylopullulanase from *Pyrococcus furiosus* (Kang et al., 2005).

There are some GH-57 members with no equivalents at the positions corresponding with the catalytic residues (Zona et al., 2004), e.g. the hypothetical protein coded by the ORF TP0147 in the *Treponema palidum* genome (Fig. 2a). Of the two, i.e. Glu and Asp, the latter was found as more variable (Zona et al., 2004). It is worth mentioning, however, that all of the GH-57 members lacking at least one of the catalytic residues or even both of them are only putative proteins without any biochemical characterization.

With regard to the additional residues that should be important for the function and activity of the GH-57 members, these may vary due to various substrates accommodated and reactions performed. This situation can be similar to that in the main α-amylase family (clan GH-H), where the essential catalytic residues are invariantly conserved, whereas the additional functionally important residues may vary from one enzyme to another (Janecek, 2002). Nevertheless, based on the five conserved sequence regions (Fig. 2a), the residues His13, Glu79, Glu216, Asp354 (in *Thermococcus hydrothermalis* 4-α-glucanotransferase) were postulated as important for the GH-57 members. Also the three tryptophanes, Trp120, Trp221 and Trp357 (Fig. 2a,b), as well as the residues in the corresponding positions could be of interest. All these residues and eventually some further ones await their experimental confirmation.

Little structural information

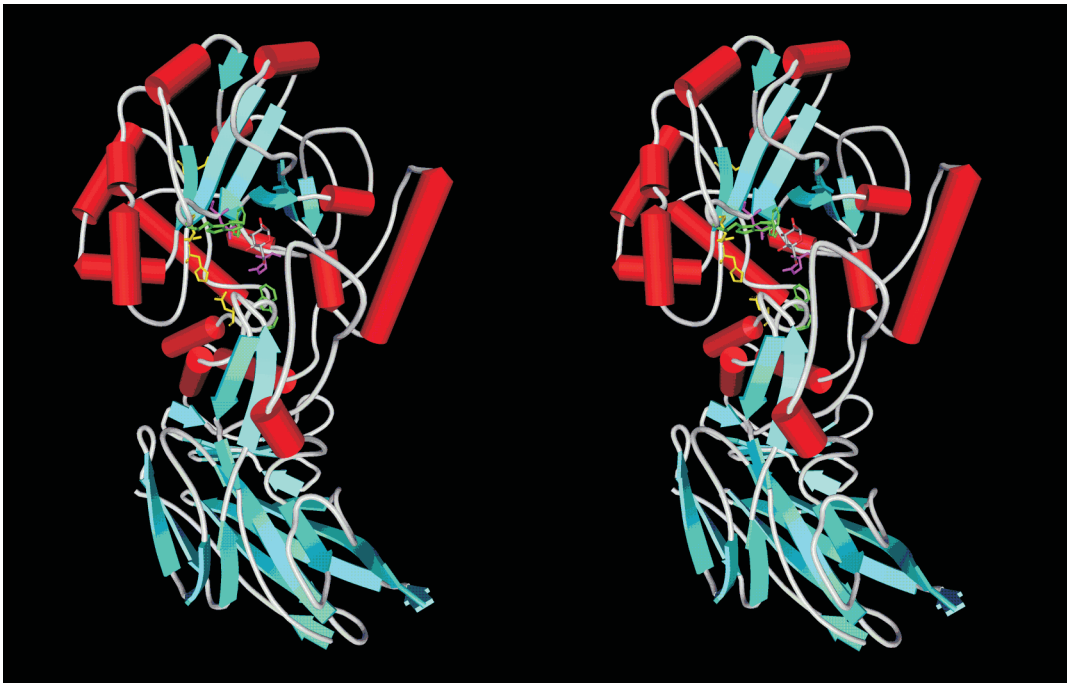
According to the requirements of the CAZy classification system (Coutinho & Henrissat, 1999) all extremely diversified sequences of the family GH-57 should have common catalytic machinery, reaction mechanism and fold. Although the first sequences of α-amylases that were obviously different from those of the α-amylase family GH-13, appeared in the 1988 (Fukusumi et al., 1988) and 1993 (Laderman et al., 1993a), the structure remained unrevealed until 2003 (Imamura et al., 2003). There was a crystallization report in 1995 on an amylopullulanase from *Pyrococcus woesei* (Knapp et al., 1995) of which the sequence possibly belonged to the GH-57. However, the detailed crystallographic analysis of this protein has not been published as yet.

The first three-dimensional structure of a GH-57 member became known only recently (Fig. 2b) when Imamura et al. (2003) published their results on 4-α-glucanotransferase from *Thermococcus litoralis*. The most important feature is the N-terminal catalytic domain that adopts the irregular $(\beta/\alpha)_8$-barrel, i.e. $(\beta/\alpha)_7$-barrel, called also a pseudo TIM-barrel (Fig. 2b) with the catalytic residues. The Glu and Asp are located near the C-terminal ends of the strands β4 and β7 of the barrel, respectively (Imamura et al., 2003). This fact supports the eventuality that the family GH-57 and the main α-amylase family GH-13 (today in a clan GH-H with GH-70 and GH-77) cannot form a clan in terms of the CAZy classification system (Henrissat, 1991; Coutinho & Henrissat, 1999). The efforts aimed at joining the two families into a clan were relevant before the first three-dimensional structure of a GH-57 member became available, i.e. to look for the remote homology at the sequence level only (Dong et al., 1997; Janecek, 1998). The independent position of the GH-57 with regard to GH-13 is not only due to the differences in the catalytic barrel domain, but more importantly, due to a difference in the catalytic

(a)

Enzyme/ORF	Region 1	Region 2	Region 3	Region 4	Region 5	C-ter
O32462_THELI	13_HQP	76_GQLEIV	120_WLTPRV	212_HDDGEKFGVW	350_AQCNDAYWH	659
Q9HHB5_PYRFU	7_HGN	69_GLIEIL	114_WLPPLA	246_GTDIEFLGYR	331_AENSDARGW	364
Q59006_METJA	10_HQP	96_GNVELI	142_RNTPLI	235_YMDYETFGEH	349_LQTSDNLYY	467
Q9Y8I8_THEHY	15_HQP	246_GNVEVT	288_WAAPSA	392_TLDGENPVEN	539_AEASDWFWW	310
Q9V038_PYRAB	12_YQP	97_EIVETL	150_WLPPNV	251_SSDLESLVAN	375_FGVIDMLRT	597
Q8YZ60_ANASP	44_YQP	138_GHGNAI	183_WLAPTA	325_GTDGETFGHH	432_QLFRDPWQA	907
Q97GF3_CLOAB	13_HMP	139_GCVEII	186_WLPPCA	348_PYDTELYGHW	460_AEASDWSFI	527
O83182_TREPA	29_HDP	79_KRLELL	124_FLEASA	246_THTGEIYKEP	382_RELMNMYVQ	619
Q8EC76_SHEON	17_LAF	78_GMCELI	123_YINPQA	223_GSDAEVFDFR	374_EAQQDIQIE	638

(b)



(c)

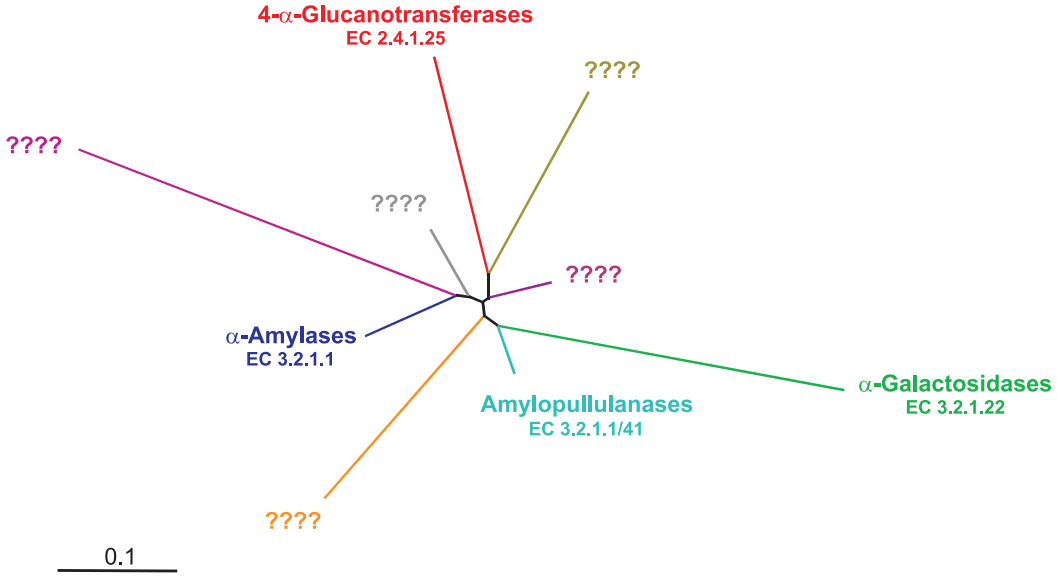


Fig. 2. (***a***) Conserved sequence regions of selected representatives from the family GH-57 (Zona et al., 2004). The catalytic residues are highlighted in pink. The potential functionally important residues are highlighted in yellow (His, 2xGlu, and Asp) and in green (three tryptophanes). The residues conserved at a level of at least 50% are highlighted in grey. The enzymes and/or ORFs are abbreviated according to the UniProt database code (Apweiler et al., 2004) and the colors correspond with the individual enzyme specificities and/or GH-57 subfamilies shown in the evolutionary tree. (***b***) Three-dimensional structural features of GH-57 4-α-glucanotransferase from *Thermococcus litoralis* (PDB code: 1K1W; Imamura et al., 2003). The highlighted residues are colored as in the conserved sequence regions. The protein structure was displayed using the program WebLabViewerLite (Molecular Simulations, Inc.). (***c***) Simplified evolutionary tree of the family GH-57. Only the main branches leading to the clusters of related sequences are indicated. At present the four known enzyme specificities (α-amylase, amylopullulanase, α-galactosidase, and 4-α-glucanotransferase) form their own clusters. The branches leading to the question marks are for the clusters (subfamilies) containing only hypothetical proteins. The tree was calculated using the CLUSTAL program (Thompson et al., 1994) and displayed with the program TreeView (Page, 1996).

machinery: there are not identical catalytic residues positioned at corresponding secondary structure elements. The fact that both families employ the same retaining reaction mechanism (Coutinho & Henrissat, 1999) keeps the possibility to classify them together, however, at a hierarchy higher than a clan.

The 4-α-glucanotransferase from *Thermococcus litoralis* also contains a C-terminal non-catalytic domain (Fig. 2b) that consists of β-strands only adopting a twisted β-sandwich fold (Imamura et al., 2003). It is clear that this domain is missing in some members with shorter amino acid sequences, e.g., in α-galactosidases with less than 370 residues (van Lieshout et al., 2003). On the other hand, some members, especially the extra-long amylopullulanases with more than 1,300 residues (Erra-Pujada et al., 1999) have to contain some additional domains of which the structure remains at present unknown. One kind of the domains could be a longer version of a typical SLH motif (surface layer homology), so-called SLH motif-bearing domain found, e.g., in the amylopullulanase from *Thermococcus hydrothermalis* (Erra-Pujada et al., 1999). This domain has recently been shown to be present also in the GH-15 glucodextranase from *Arthrobacter globiformis* (Mizuno et al., 2004; 2005). Interestingly, within the family GH-57 the presence of this SLH motif-bearing domain is known only in the amylopullulanases, i.e. in the same specificity as it has been revealed for the SLH motifs in the frame of the family GH-13 (Zona & Janecek, 2005).

It should also be pointed out that based on the similarities seen in the three-dimensional structures of catalytic domains of the GH-57 4-α-glucanotransferase from *Thermococcus litoralis* (Imamura et al., 2003) and the GH-38 α-mannosidase from *Drosophila melanogaster* (van den Elsen et al., 2001), the evolution of these two families from a common ancestor has been proposed. However, the catalytic domain of the α-mannosidase from the family GH-38 is not formed by any version of the $(\beta/\alpha)_7$-barrel present in the family GH-57 (Imamura et al., 2003).

Recent evolutionary implications

The most detailed and comprehensive evolutionary relationships within the family GH-57 were first described by Zona et al. (2004). According to that study the entire family GH-57 can be divided into several subfamilies that may reflect the different enzyme specificities. It is worth mentioning that at present the four enzyme specificities (α-amylase, amylopullulanase, α-galactosidase, and 4-α-glucanotransferase) form their own four clusters and the remaining clusters contain only hypothetical proteins (Fig. 2c). This does not necessarily mean that in the future, when the putative GH-57 members from these remaining clusters will be biochemically characterized, some of the present-day specificities will appear in those clusters.

Concerning the enzyme specificities, it should be pointed out that only α-galactosidase, amylopullulanase and 4-α-glucanotransferase were unambiguously demonstrated for the real enzymes. These are the cases of α-galactosidase from *Pyrococcus furiosus* (van Lieshout et al., 2003), amylopullulanases from *Thermococcus hydrothermalis* (Chang-Pi-Hin et al., 2002) and *Pyrococcus furiosus* (Dong et al., 1997), and the 4-α-glucanotransferases from *Thermococcus litoralis* (Jeon et al., 1997) and *Pyrococcus kodakaraensis* (Tachibana et al., 1997). With regard to the specificity of α-amylase, i.e. the enzymes from *Dictyoglomus thermophilum* (Fukusumi et al., 1988) and *Pyrococcus furiosus* (Laderman et al., 1993a), the situation is not so clear. The "α-amylase" from *Pyrococcus furiosus* was shown already in 1993 to display also the 4-α-glucanotransferase activity (Laderman et al., 1993b), whereas for the enzyme from *Dictyoglomus thermophilum* the transglycosylation activity was confirmed only recently (Nakajima et al., 2004). Thus the family GH-57 seemed to miss the pure α-amylase specificity.

The only real α-amylase in the family GH-57 could be the enzyme from the archaeon *Methanococcus jannaschii* (Kim et al., 2001). The MJ1611 locus originally annotated as coding for an α-amylase (Bult et al., 1996) could really encode for an α-amylase without any transglycosylation activity (Li & Peeples, 2004; T. Peeples, The University of Iowa, IA, USA; personal communication).

Conclusions

It can be concluded that the family GH-57 really belongs to a group of remarkable families of glycoside hydrolases. The sequences ranging in lengths from about 400 to about 1500 residues exhibit the common features characteristic for the entire family. The real enzymes form only 10% of the GH-57 members; the rest being formed by hypothetical proteins. The individual enzyme specifictes may establish their own subfamilies. Although the family GH-57 was originally defined as the new α-amylase family (additional to the main α-amylase family GH-13), it is paradoxically just the α-amylase specificity that has to be confirmed to be unambiguously present in the family. Thus although much has became clear in the last years, the family GH-57 still retains many of its secrets until now.

Acknowledgements

This work was financially supported in part by the VEGA grant number 2/5067/25 from the Slovak Grant Agency for Science.

References

Aghajari, N., Feller, G., Gerday, C. & Haser, R. 1998. Structure **6:** 1503–1516.

Aleshin, A.E., Feng, P.H., Honzatko, R.B. & Reilly, P.J. 2003. J. Mol. Biol. **327:** 61–73.

Aleshin, A., Golubev, A., Firsov, L.M. & Honzatko, R.B. 1992. J. Biol. Chem. **267:** 19291–19298.

Apweiler, R., Bairoch, A., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C., Redaschi, N. & Yeh, L.S. 2004. Nucleic Acids Res. **32:** D115–D119.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., Marshall, M. & Sonnhammer, E.L. 2002. Nucleic Acids Res. **30:** 276–280.

Buisson, G., Duee, E., Haser, R. & Payan, F. 1987. EMBO J. **6:** 3909–3916.

Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb, J.F., Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodek, A., Scott, J.L., Geoghagen, N.S. & Venter, J.C. 1996. Science **273:** 1058–1073.

Chang-Pi-hin, F., Erra-Pujada, M., Dauchez, M., Debeire, P., Duchiron, F. & O'Donohue, M.J. 2002. Biologia, Bratislava **57 (Suppl. 11):** 155–162.

Cheong, C.G., Eom, S.H., Chang, C., Shin, D.H., Song, H.K., Min, K., Moon, J.H., Kim, K.K., Hwang, K.Y. & Suh, S.W. 1995. Proteins **21:** 105–117.

Chiba, S. 1997. Biosci. Biotechnol. Biochem. **61:** 1233–1239.

Coutinho, P.M. & Henrissat, B. 1999. Carbohydrate-Active Enzymes server; http://afmb.cnrs-mrs.fr/∼cazy/.

Dong, G., Vieille, C. & Zeikus, J.G. 1997. Appl. Environ. Microbiol. **63:** 3577–3584.

Egloff, M.P., Uppenberg, J., Haalck, L. & van Tilbeurgh, H. 2001. Structure **9:** 689–697.

Erra-Pujada, M., Debeire, P., Duchiron, F. & O'Donohue, M.J. 1999. J. Bacteriol. **181:** 3284–3287.

Fukusumi, S., Kamizono, A., Horinouchi, S. & Beppu, T. 1988. Eur. J. Biochem. **174:** 15–21.

Henrissat, B. 1991. Biochem. J. **280:** 309–316.

Henrissat, B. & Bairoch, A. 1996. Biochem. J. **316:** 695–696.

Hondoh, H., Kuriki, T. & Matsuura, Y. 2003. J. Mol. Biol. **326:** 177–188.

Imamura, H., Fushinobu, S., Jeon, B.S., Wakagi, T. & Matsuzawa, H. 2001. Biochemistry **40:** 12400–12406.

Imamura, H., Fushinobu, S., Yamamoto, M, Kumasaka, T., Jeon, B.S., Wakagi, T. & Matsuzawa, H. 2003. J. Biol. Chem. **278:** 19378–19386.

Janecek, S. 1998. Folia Microbiol. **43:** 123–128.

Janecek, S. 2000. In: Ohnishi, M., Hayashi, T., Ishijima, S. & Kuriki, T. (eds) Glycoenzymes, Japan Academic Societies Center, Tokyo, pp. 19–54.

Janecek, S. 2002. Biologia, Bratislava **57 (Suppl. 11):** 29–41.

Jeon, B.S., Taguchi, H., Sakai, H., Ohshima, T., Wakagi, T. & Matsuzawa, H. 1997. Eur. J. Biochem. **248:** 171–178.

Kang, S., Vieille, C. & Zeikus, J.G. 2005. Appl. Microbiol. Biotechnol. **66:** 408–413.

Kaper, T., van der Maarel, M.J., Euverink, G.J. & Dijkhuizen, L. 2004. Biochem. Soc. Trans. **32:** 279–282.

Kim, J.W., Flowers, L.O., Whiteley, M. & Peeples, T.L. 2001. Folia Microbiol. **46:** 467–473.

Knapp, S., Rudiger, A., Antranikian, G., Jorgensen, P.L. & Ladenstein, R. 1995. Proteins **23:** 595–597.

Kralj, S., van Geel-Schutten, G.H., Dondorff, M.M., Kirsanovs, S., van der Maarel, M.J. & Dijkhuizen, L. 2004a. Microbiology **150:** 3681–3690.

Kralj, S., van Geel-Schutten, G.H., van der Maarel, M.J. & Dijkhuizen, L. 2004b. Microbiology **150:** 2099–2112.

Kuriki, T., Hondoh, H. & Matsuura, Y. 2005. Biologia, Bratislava **60 (Suppl. 16):** 13–16.

Laderman, K.A., Asada, K., Uemori, T., Mukai, H., Taguchi, Y., Kato, I. & Anfinsen, C.B. 1993a. J. Biol. Chem. **268:** 24402–24407.

Laderman, K.A., Davis, B.R., Krutzsch, H.C., Lewis, M.S., Griko. Y.V., Privalov, P.L. & Anfinsen, C.B. 1993b. J. Biol. Chem. **268:** 24394–24401.

Lee, S.S., Yu, S. & Withers, S.G. 2003. Biochemistry **42:** 13081–13090.

Lee, S.S., Yu, S. & Withers, S.G. 2005. Biologia, Bratislava **60 (Suppl. 16):** 137–148.

Li, M. & Peeples, T.L. 2004. J. Chromatogr. **807B:** 69–74.

Lovering, A.L., Lee, S.S., Kim, Y.W., Withers, S.G. & Strynadka, N.C.J. 2005. J. Biol. Chem. **280:** 2105–2115.

MacGregor, E.A. 2005. Biologia, Bratislava **60 (Suppl. 16):** 5–12.

MacGregor, E.A., Janecek, S. & Svensson, B. 2001. Biochim. Biophys. Acta **1546:** 1–20.

MacGregor, E.A., Jespersen, H.M. & Svensson, B. 1996. FEBS Lett. **378:** 263–266.

Machius, M., Wiegand, G. & Huber, R. 1995. J. Mol. Biol. **246:** 545–559.

Machovic, M. & Janecek, S. 2003. Biologia, Bratislava **58:** 1127–1132.

Matsuura, Y. 2002. Biologia, Bratislava **57 (Suppl. 11):** 21–27.

Matsuura, Y., Kusunoki, M., Harada, W. & Kakudo, M. 1984. J. Biochem. **95:** 697–702.

Mikami, B., Hehre, E.J., Sato, M., Katsube, Y., Hirose, M., Morita, Y. & Sacchettini, J.S. 1993. Biochemistry **32:** 6836–6845.

Mikami, B., Yoon, H.J. & Yoshigi, N. 1999. J. Mol. Biol. **285:** 1235–1243.

Mizuno, M., Tonozuka, T., Suzuki, S., Uotsu-Tomita, R., Kamitori, S., Nishikawa, A. & Sakano, Y. 2004. J. Biol. Chem. **279:** 10575–10583.

Mizuno, M., Tonozuka, T., Ichikawa, K., Kamitori, S., Nishikawa, A. & Sakano, Y. 2005. Biologia, Bratislava **60 (Suppl. 16):** 171–177.

Nakai, H., Okuyama, M., Kim, Y.M., Saburi, W., Wongchawalit, J., Mori, H., Chiba, S. & Kimura, A. 2005. Biologia, Bratislava **60 (Suppl. 16):** 131–135.

Nakajima, M., Imamura, H., Shoun, H., Horinouchi, S. & Wakagi, T. 2004. Biosci. Biotechnol. Biochem. **68:** 2369–2373.

Nielsen, J.E. & Borchert, T.V. 2000. Biochim. Biophys. Acta **1543:** 253–274.

Oh, B.H. 2003. Biologia, Bratislava **58:** 299–305.

Oslancova, A. & Janecek, S. 2002. Cell. Mol. Life Sci. **59:** 1945–1959.

Oyama, T., Kusunoki, M., Kishimoto, Y., Takasaki, Y. & Nitta, Y. 1999. J. Biochem. **125:** 1120–1130.

Page, R.D. 1996. Comput. Applic. Biosci. **12:** 357–358.

Park, K.H., Kim, T.J., Cheong, T.K., Kim, J.W., Oh, B.H. & Svensson, B. 2000. Biochim. Biophys. Acta **1478:** 165–185.

Przylas, I., Tomoo, K., Terada, Y., Takaha, T., Fujii, K., Saenger, W. & Sträter, N. 2000. J. Mol. Biol. **296:** 873–886.

Pujadas, G. & Palau, J. 2001. Mol. Biol. Evol. **18:** 38–54.

Pujadas, G. & Palau, J. 2002. Biologia, Bratislava **57 (Suppl. 11):** 43–57.

Pujadas, G., Ramírez, F.M., Valero, R. & Palau, J. 1996. Proteins **25:** 456–472.

Rigden, D.J. 2002. FEBS Lett. **523:** 17–22.

Rye, C.S. & Withers, S.G. 2000. Curr. Opin. Chem. Biol. **4:** 573–580

Sevcik, J., Solovicova, A., Hostinova, E., Gasperik, J., Wilson, K..S. & Dauter, Z. 1998. Acta Crystallogr. **D54:** 854–866.

Svensson, B., Jensen, M.T., Mori, H., Bak-Jensen, K.S., Bonsager, B., Nielsen, P.K., Kramhoft, B., Praetorius-Ibba, M., Nohr, J., Juge, N., Greffe, L., Williamson, G.

& Driguez, H. 2002. Biologia, Bratislava **57 (Suppl. 11):** 5–19.

Sträter, N., Przylas, I., Saenger, W., Terada, Y., Fuji, K. & Takaha, T. 2002. Biologia, Bratislava **57 (Suppl. 11):** 93–99.

Tachibana, Y., Fujiwara, S., Takagi, M. & Imanaka, T. 1997. J. Ferment. Bioeng. **83:** 540–548.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. 1994. Nucleic Acids Res. **22:** 4673–4680.

van den Elsen, J.M.H., Kuntz, D.A. & Rose, D.R. 2001. EMBO J. **20:** 3008–3017.

van der Maarel, M.J.E.C., van der Veen, B., Uitdehaag, J.C.M., Leemhuis, H. & Dijkhuizen, L. 2002. J. Biotechnol. **94:** 137–155.

van Lieshout, J.F.T., Verhees, C.H., Ettema, T.J.G., van der Saar, S., Imamura, H., Matsuzawa, H., van der Oost, J. & de Vos, W.M. 2003. Biocatal. Biotransform. **21:** 243–252.

Yu, S., Bojsen, K., Svensson, B. & Marcussen, J. 1999. Biochim. Biophys. Acta **1433:** 1–15.

Zona, R., Chang-Pi-Hin, F., O'Donohue, M.J. & Janecek, S. 2004. Eur. J. Biochem. **271:** 2863–2872.

Zona, R. & Janecek, S. 2005. Biologia, Bratislava **60 (Suppl. 16):** 115–121.

Received March 29, 2005
Accepted May 06, 2005