J Mol Evol (1997) 45:322–331

JOURNAL OF MOLECULAR EVOLUTION

Domain Evolution in the α-Amylase Family

Štefan Janeček,[1] Birte Svensson,[2] Bernard Henrissat[3]

[1] Institute of Microbiology, Slovak Academy of Sciences, Štefánikova 3, SK-81434 Bratislava, Slovakia
[2] Department of Chemistry, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark
[3] Centre de Recherches sur les Macromolécules Végétales,* C.N.R.S., BP 53, F-38041 Grenoble Cédex, France

Received: 4 December 1996 / Accepted: 13 March 1997

Abstract. The available amino acid sequences of the α-amylase family (glycosyl hydrolase family 13) were searched to identify their domain B, a distinct domain that protrudes from the regular catalytic $(\beta/\alpha)_8$-barrel between the strand β3 and the helix α3. The isolated domain B sequences were inspected visually and also analyzed by Hydrophobic Cluster Analysis (HCA) to find common features. Sequence analyses and inspection of the few available three-dimensional structures suggest that the secondary structure of domain B varies with the enzyme specificity. Domain B in these different forms, however, may still have evolved from a common ancestor. The largest number of different specificities was found in the group with structural similarity to domain B from *Bacillus cereus* oligo-1,6-glucosidase that contains an α-helix succeeded by a three-stranded antiparallel β-sheet. These enzymes are α-glucosidase, cyclomaltodextrinase, dextran glucosidase, trehalose-6-phosphate hydrolase, neopullulanase, and a few α-amylases. Domain B of this type was observed also in some mammalian proteins involved in the transport of amino acids. These proteins show remarkable similarity with $(\beta/\alpha)_8$-barrel elements throughout the entire sequence of enzymes from the oligo-1,6-glucosidase group. The transport proteins, in turn, resemble the animal 4F2 heavy-chain cell surface antigens, for which the sequences either lack domain B or contain only parts thereof. The similarities are compiled to indicate a possible route of domain evolution in the α-amylase family.

Key words: α-Amylase family — Glycosyl hydrolase family 13 — Hydrophobic Cluster Analysis — $(\beta/\alpha)_8$-Barrel — Domain B — Amino acid transport-related proteins — 4F2 Heavy-chain cell surface antigens — Evolutionary relatedness

Introduction

The α-amylase family consists of a large group of starch hydrolases and related enzymes (for reviews, see MacGregor 1993; Janeček 1994a; Svensson 1994) comprising about 20 different enzyme specificities, and is currently known as glycosyl hydrolase family 13 (Henrissat 1991; Henrissat and Bairoch 1993). Three-dimensional structures of α-amylases (Matsuura et al. 1984; Brady et al. 1991; Qian et al. 1993; Kadziola et al. 1994; Brayer et al. 1995; Machius et al. 1995; Ramasubbu et al. 1996), cyclodextrin glucanotransferases (CGTases) (Klein and Schulz 1991; Kubota et al. 1991; Lawson et al. 1994; Knegtel et al. 1996), oligo-1,6-glucosidase (Kizaki et al. 1993), and maltotetraohydrolase (Matsuura 1995) are available, as well as structure predictions based on sequence comparisons (Jespersen et al. 1991, 1993) or homology modeling (Kuriki et al. 1996; Lamminmäki and Vihinen 1996). The members of this family possess related catalytic $(\beta/\alpha)_8$-barrels with a small domain (domain B) protruding between the third β-strand and the third α-helix. Some types of domain B contain several β-strands and one or two α-helices (Klein and Schulz

Abbreviations: CGTase, cyclodextrin glucanotransferase; HCA, hydrophobic cluster analysis; TAA, Taka-amylase A (α-amylase from *Aspergillus oryzae*)

*Affiliated with the Université Joseph Fourier, Grenoble, France

Correspondence to: Š. Janeček

1991; Qian et al. 1993; Kizaki et al. 1993), while others such as in barley α-amylase have no well-defined secondary structure elements (Kadziola et al. 1994).

The α-amylase family thus contains a $(\beta/\alpha)_8$-fold, discovered in chicken muscle triosephosphate isomerase more than 20 years ago (Reardon and Farber 1995). Other $(\beta/\alpha)_8$-barrel folds are present in a large variety of enzymes. The evolution of the entire $(\beta/\alpha)_8$-barrel folding family, currently comprising more than 40 different enzyme specificities and the three proteins, remains unclear (for a recent review, see Janeček and Bateman 1996). Arguments have been given in support of divergent as well as convergent evolution (Lesk et al. 1989; Farber and Petsko 1990; Brändén 1991; Raine et al. 1994; Scrutton 1994; Henrissat et al. 1995; Janeček and Baláž 1995; Jenkins et al. 1995; Reardon and Farber 1995; Janeček 1996).

The α-amylase family itself fulfills several criteria of divergent evolution: (1) a few amino acid sequence similarities that serve as fingerprints are located at or near strands β2, β3, β4, β5, and β7, and at a short sequence near the C-terminus of domain B (Jespersen et al. 1993; Janeček 1995a); (2) a conserved domain organization including the characteristic insertion of domain B between strand β3 and helix α3 of the catalytic $(\beta/\alpha)_8$-barrel (MacGregor 1993; Janeček 1994a; Svensson 1994); (3) invariant catalytic residues (Asp206, Glu230, and Asp297; Taka-amylase A (TAA) numbering) (Qian et al. 1994; Strokopytov et al. 1995); and (4) related functions, that is, α-glucosyl hydrolase/transferase activity (Svensson, 1994; Henrissat 1991; Henrissat and Bairoch 1993; Jespersen et al. 1993).

Evolutionary trees of the α-amylase family clearly respect both taxonomy and variation in enzyme specificity for individual members (Jespersen et al. 1993; Janeček 1994a). These studies addressed a short conserved sequence near the C-terminus of domain B (Janeček 1995a) and sequence similarities at β-strands and their C-terminal extensions in the $(\beta/\alpha)_8$-barrel domain (Jespersen et al. 1993, Janeček 1994a,b, 1995b). Domain B was demonstrated to play an important functional role. It thus controls several of the isozyme specific properties in barley α-amylases, including substrate affinity and sensitivity to the barley α-amylase/subtilisin inhibitor specific to isozyme 2 (Rodenburg et al. 1994; Juge et al. 1995). The purpose of the present work is to investigate the evolution of domain B by addressing the following questions (1) Is the evolution of domain B comparable to that of the $(\beta/\alpha)_8$-barrel structural elements? (2) How many types of domain B are represented in the α-amylase family? (3) Is domain B present in other $(\beta/\alpha)_8$-barrel proteins or in different fold families?

Materials and Methods

Amino acid sequences representing the enzyme specificities from the α-amylase family were extracted from the SwissProt protein and GenBank DNA sequence data bases. Domain B from each sequence was defined by identifying the easily recognized flanking β3 and α3β4 elements of the catalytic $(\beta/\alpha)_8$-barrel (Jespersen et al. 1993). This takes into account that domain B ends by a short conserved sequence comprising the calcium-ligand (Asp175 in TAA) preceding helix α3 (Matsuura et al. 1984; Janeček 1992, 1995a).

First the sequences of domain B of selected family members were visually inspected including use of Hydrophobic Cluster Analysis (HCA; Gaboriaud et al. 1987) and compared with the three-dimensional structures of α-amylases (Matsuura et al. 1984; Qian et al. 1993; Kadziola et al. 1994; Machius et al. 1995), CGTases (Klein and Schulz 1991; Lawson et al. 1994), oligo-1,6-glucosidase (Kizaki et al. 1993), and maltotetraohydrolase (Matsuura 1995). This resulted in a subdivision of the α-amylase family members according to sequence similarities. Size, shape and orientation of clusters of hydrophobic residues are considered to reflect elements of secondary structure in proteins (Woodcock et al. 1992). The HCA motifs thus define groups likely to share secondary and supersecondary structure in domain B.

Since the group represented by the *Bacillus cereus* oligo-1,6-glucosidase has the largest variation of enzyme specificity, it was further analyzed; the sequences, moreover, were used as queries in a BLAST search (Altschul et al. 1990) for occurrence of domain B in other proteins. The HCA similarity scores were calculated for representatives of different enzyme specificities using the program SUNHCA (Lemesle-Varloot et al. 1993). The scores were averaged to obtain a matrix showing the similarity between each pair of enzymes from this group. The program CLUSTAL V (Higgins et al. 1992) was used for sequence alignment.

Results and Discussion

Domain B Similarities and Differences

Visual inspection of the sequences of domain B from representatives of the α-amylase family confirmed that domain B varies greatly in both length and sequence (cf. Jespersen et al. 1993). It therefore makes no sense to produce a sequence alignment that includes all the investigated enzyme specificities. Similarly, HCA plots of domain B in available crystal structures (Fig. 1) agree that there is no common arrangement of secondary structure elements for α-amylase family members. A short stretch (173_LPDLD in TAA) near the C-terminus of domain B comprises a conserved calcium-binding aspartate (Asp175 in TAA) (Janeček 1992, 1995a) and appears to be the best conserved motif in domain B, although this area was not identified for a few specificities, for example, the glycogen branching and the debranching enzymes. It should be pointed out, however, that the glycogen debranching enzymes have an exceptionally long domain B (250 amino acid residues; Jespersen et al. 1993).

Each HCA plot of domain B in Figure 1 represents closely related enzymes with a sequence identity higher than 50% and thus probably the same supersecondary structure in domain B. α-Amylase from *B. licheniformis* covers liquefying bacterial α-amylases, the intracellular α-amylase from *Streptococcus bovis*, and maltohexaohy-

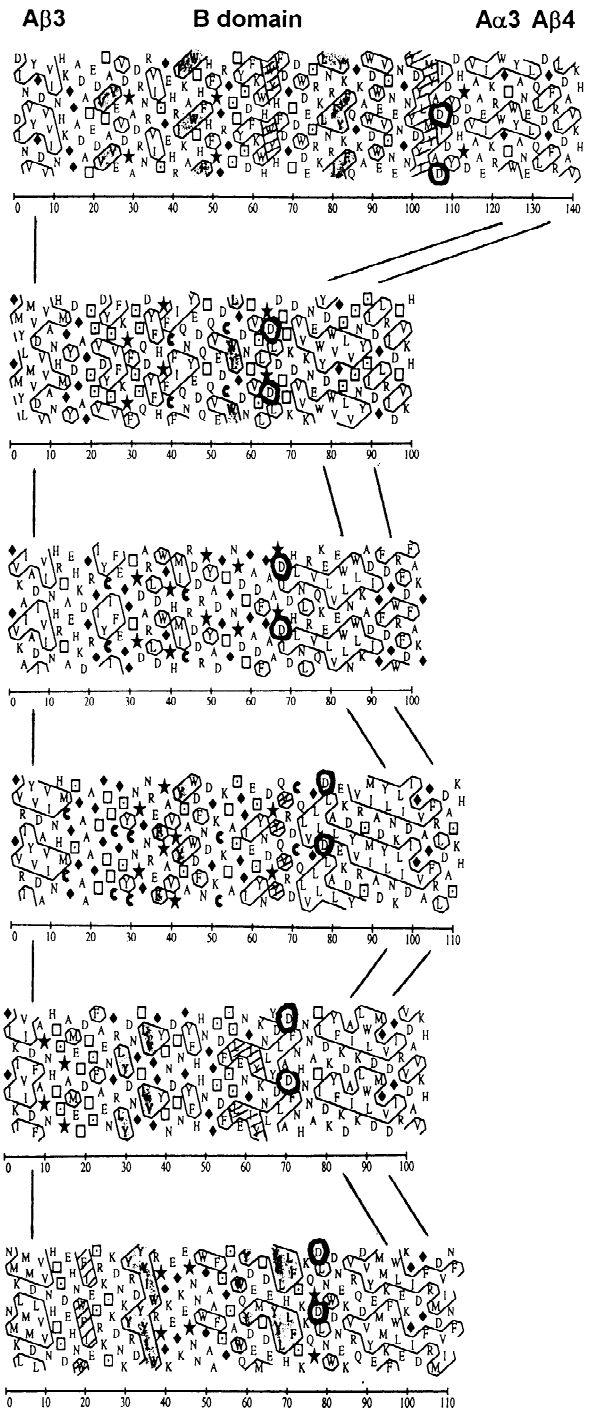

Fig. 1. HCA plots of the domain B region of enzymes in the α-amylase family with known three-dimensional structure. BLI, α-amylase from *Bacillus licheniformis* (Yuuki et al. 1985; Machius et al. 1995); TAA, Taka-amylase A (Toda et al. 1982; Matsuura et al. 1984); BAR, barley α-amylase (Rogers 1985; Kadziola et al. 1994); PPA, α-amylase from pig pancreas (Pasero et al. 1986; Qian et al. 1993); CGT, CGTase from *Bacillus circulans* (Nitschke et al. 1990; Klein and Schulz 1991); OGL, oligo-1,6-glucosidase from *Bacillus cereus* (Watanabe et al. 1990; Kizaki et al. 1993). The length and the position of domain B are given in parentheses. The HCA clusters in domain B corresponding to α-helices (in BLI 3_{10}-helices) are filled with conventional lines and the ones corresponding to β-strands are shaded. The invariant aspartates equivalent to the calcium-binding Asp175 of TAA are encircled. The amino acid residues are presented by the single-letter code except for glycine (♦), proline (★), cysteine (**C**), serine (⊡), and threonine (□).

drolases. TAA represents α-amylases from fungi and yeasts, while barley α-amylase represents the plant enzymes and probably the maltotetraohydrolases. Pig pancreatic α-amylase covers the enzymes from mammals, insects, streptomycetes, *Alteromonas haloplanctis*, and maltopentaohydrolases. CGTase from *B. circulans* comprises different CGTases, including reclassified sequences (Janeček et al. 1995), together with the maltogenic amylase from *B. stearothermophilus.* Finally, the oligo-1,6-glucosidase from *B. cereus* represents oligo-1,6-glucosidases, α-glucosidases, dextran glucosidases, trehalose-6-phosphate hydrolases, cyclomaltodextrinases, neopullulanases and a few bacterial α-amylases. These groups are in accordance with previously reported evolutionary relationships among taxonomically different α-amylases and CGTases (e.g., Hickey et al. 1987; Tsukamoto et al. 1988; MacGregor and Svensson 1989; Janse et al. 1993; Jespersen et al. 1993; Oguma et al. 1993; Feller et al. 1994; Janeček 1994b, 1995b; Whitehead and Cotta 1995; Takii et al. 1996).

In maltotetraohydrolase from *Pseudomonas stutzeri*, whose structure has recently been determined (Matsuura

1995), domain B is stabilized by a disulfide bridge between Cys140 and Cys150, a feature comparable to those of TAA and pig pancreatic α-amylase. But domain B of maltotetraohydrolase has no distinct secondary structure, and resembles barley α-amylase in which a disulfide bridge is not formed between the two cysteines in domain B (Kadziola et al. 1994). In distance trees based mainly on $(\beta/\alpha)_8$-barrel elements (Jespersen et al. 1993; Janeček 1995b), this maltotetraohydrolase was earlier found on a branch adjacent to the cluster of plant α-amylases.

The HCA plots of members of known three-dimensional structure are in agreement with their distinctly different secondary structure in domain B. For instance, *B. licheniformis* α-amylase lacks the helix which is common to domain B in most α-amylases and CGTases (Machius et al. 1995). Remarkably, this liquefying α-amylase seems to be very different from the saccharifying α-amylase from *B. subtilis*, which can be evolutionarily related to α-amylase from *Butyrivibrio fibrisolvens* (based on the similarity of the entire sequences; Janeček 1994b), and also to potato amylomaltase (based on the similarity in the short conserved sequence near the C-terminus of domain B; Janeček 1995a).

The Enzymes with Domain B of the Oligo-1,6-Glucosidase Type

The group defined by *B. cereus* oligo-1,6-glucosidase includes many different amylolytic enzymes, and, surprisingly also three amino acid transport-related proteins, as found using BLAST (Altschul et al. 1990) with relevant domain B sequences as queries (Fig. 2). Because size, shape, and orientation of clusters in an HCA plot correlate with secondary structure elements in the folded protein (Woodcock et al. 1992), the similarity in HCA plots reflects similarity of secondary and supersecondary structure in relevant segments of domain B. In other words, this HCA motif is defined by the secondary structure in domain B identified in the crystal structure of *B. cereus* oligo-1,6-glucosidase, that is, one α-helix and a three-stranded antiparallel β-sheet (Kizaki et al. 1993; Watanabe et al. 1994). The absence or presence of the second β-strand Bβ2 (153_WQYD in the sequence of *B. cereus* oligo-1,6-glucosidase), however, defines two tentative subgroups: cyclomaltodextrinases, neopullulanases, and *D. thermophilum* AmyB α-amylase (Horinouchi et al. 1988) lacking Bβ2, and the remaining proteins including α-amylases from *B. megaterium* (Metz et al. 1988), *D. thermophilum* AmyC (Horinouchi et al. 1988), and *Xanthomonas campestris* K-11151 (Abe et al. 1996), that contain Bβ2 (cf. Fig. 2). These subclasses are supported by the two recent independent alignments of Oguma et al. (1993) and Takii et al. (1996). The short conserved sequence (167_QPDLN; *B. cereus* oligo-1,6-glucosidase numbering) may serve as selection marker (Janeček 1995a), as enzymes lacking Bβ2 have lysine corresponding to Asp175 in TAA (Asp169 in *B. cereus* oligo-1,6-glucosidase).

To analyze the relationship between individual members depicted in Figure 2, HCA similarity scores were calculated for the secondary structure elements Bα1, Bβ1, Bβ3, and presented as a matrix (Table 1). Averaged scores higher than 0.750 reflect a similar folding of the polypeptide chains with root mean square deviations lower than 2Å (Gaboriaud et al. 1987). Domain B from *B. cereus* oligo-1,6-glucosidase and *Streptococcus mutans* dextran glucosidase thus resemble each other very closely.

Amino Acid Transport-Related Proteins and the 4F2 Heavy-Chain Cell Surface Antigens

The 85% identical amino acid transport-related proteins from human, rat, and rabbit kidney contain Bβ2 (264_WHFD; Fig. 2) and were previously reported to resemble α-glucosidases (Wells and Hediger 1992). They clearly possess characteristic features of the α-amylase family, including β-strands β2, β3, β4, β8 of the $(\beta/\alpha)_8$-fold, and the conserved stretch near the C-terminus of domain B (Fig. 3). In addition, the 4F2 heavy-chain cell surface antigens are structurally related to the α-amylase family (Quackenbush et al. 1987; Wells and Hediger 1992), as supported by sequence similarity especially at the above β-strands, but the 4F2 antigen lacks domain B. The β3 → α3 segment 213_GENSWFFTQV in 4F2 antigen, however, has close sequence resemblance to Bβ2 (Fig. 3) preceded by 211_YR. In contrast, some amylolytic enzymes (see above) that contain a proper domain B related to that of oligo-1,6-glucosidase, miss a segment equivalent to Bβ2 (Fig. 2).

Of the three catalytic residues in family 13 glycosyl hydrolases (i.e., Asp206, Glu230, and Asp297 of TAA), only Asp206 of TAA in the β4-strand region can be traced unambiguously in the amino acid transport-related proteins and the 4F2 heavy-chain antigens (Fig. 3). Interestingly, the former proteins in β3-strand region have a sequence Phe-X-Pro, which in family 13 of glycosyl hydrolases is specific to CGTases (MacGregor and Svensson 1989; Jespersen et al. 1993; Janeček 1994a,b; Janeček et al. 1995). This provides support to the view that the so-called *intermediary* sequence features (i.e., features characteristic of a given enzyme specificity exhibited with another enzyme specificity) can be observed among α-amylase family members (Janeček 1994b, 1995a,b; Janeček et al. 1995).

With regard to the evolution of the 4F2 heavy-chain cell surface antigens, two explanations are possible. Either they evolved from a $(\beta/\alpha)_8$-barrel common ancestor to the α-amylase family but diverged before the α-amylase family, with the inserted domain B, became evident,

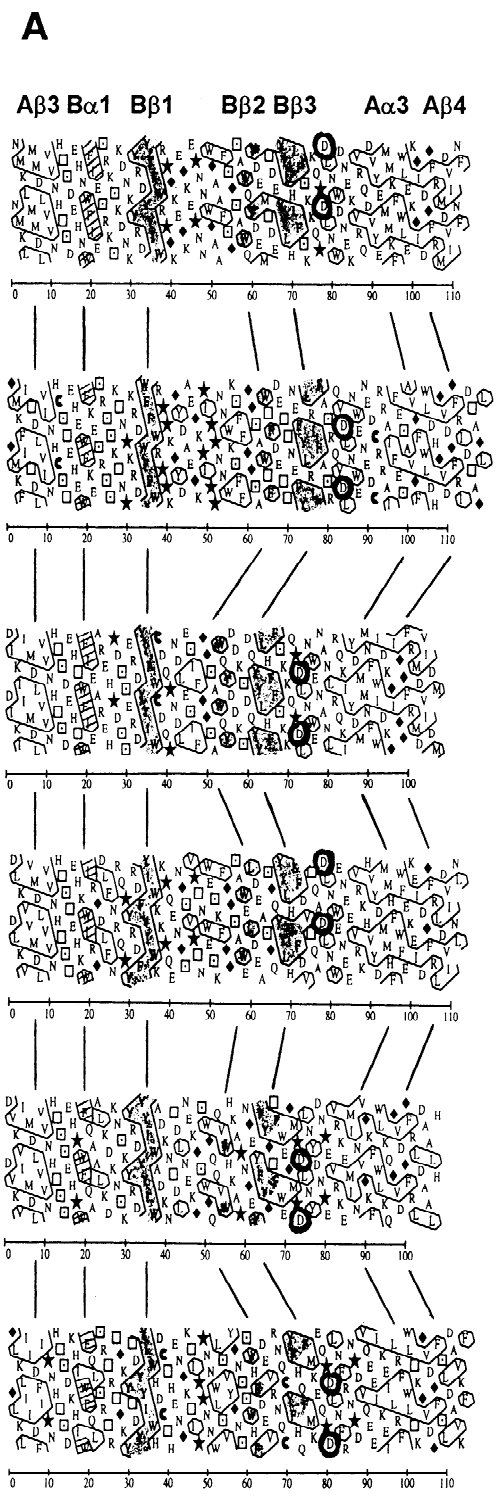

Fig. 2. HCA plots of the representatives that share **A** and resemble **B,** respectively, the structure of domain B from *Bacillus cereus* oligo-1,6-gluosidase. OGL, *Bacillus cereus* oligo-1,6-glucosidase (Watanabe et al. 1990); AGL, *Saccharomyces carlsbergensis* α-glucosidase (Hong and Marmur 1986); DGL, *Streptococcus mutans* dextran glucosidase (Russell and Ferretti 1990); T6P, *Bacillus subtilis* trehalose-6-phosphate hydrolase (Helfert et al. 1995); BME, *Bacillus megaterium* α-amylase (Metz et al. 1988) representing also the α-amylases from *Dictyoglomus thermophilum* AmyC (Horinouchi et al. 1988) and *Xanthomonas campestris* K-11151 (Abe et al. 1996); AAT, human kidney amino acid transport-related protein (Bertran et al. 1993); CMD, *Bacillus sphaericus* cyclomaltodextrinase (Oguma et al. 1993); NPU, *Thermoactinomyces vulgaris* neopullulanase (Tonozuka et al. 1993); DTB, *Dictyoglomus thermophilum* α-amylase AmyB (Horinouchi et al. 1988). The length and the position of domain B are given in parentheses. The HCA clusters in domain B corresponding to α-helix of OGL are filled with conventional lines, and the ones corresponding to β-strands of OGL are shaded. The putative calcium-binding aspartates or corresponding lysines are encircled. The amino acid residues are presented by the single-letter code except for glycine (♦), proline (★), cysteine (**C**), serine (⊡), and threonine (□).

or they lost a major part of the β3 → α3 segment, in this case the domain B type having a three-stranded antiparallel β-sheet. The evolutionary events taking all this into account are outlined in Figure 4. It leaves space for the independent evolution, after recruitment from the oligo-1,6-glucosidase group, of the amino acid transport-related proteins and subsequently the 4F2 heavy-chain cell surface antigens.

Conclusion

The present study focuses on domain B, a distinct 40–250 amino acid long domain inserted between the strand β3 and helix α3 of the $(\beta/\alpha)_8$-barrel in the α-amylase family (glycosyl hydrolases family 13). Previously, the evolutionary relationship was primarily described on the basis of the elements of the $(\beta/\alpha)_8$-motif (Jespersen et al.

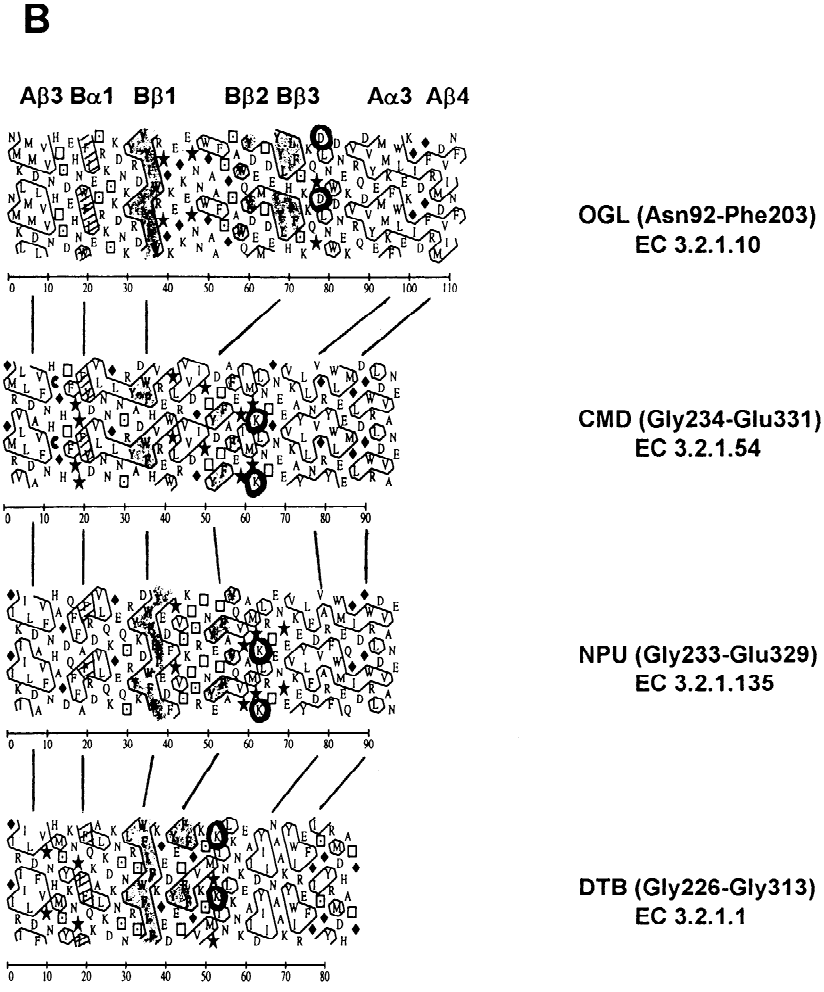

Fig. 2. Continued.

1993; Janeček 1994a). The present sequence comparison analysis revealed that although domain B probably evolved from a single ancestor, as supported by a short conserved sequence near the C-terminus of domain B (cf. Janeček 1992, 1995a), different types of domain B define groups of α-amylase family members, such as bacterial liquefying α-amylases, plant α-amylases, animal α-amylases, and, in particular, the group including the oligo-1,6-glucosidase.

Another argument for a common origin of present-day domain B is that this domain is inserted with conserved topology in the $(\beta/\alpha)_8$-fold in all members of the α-amylase family. Furthermore, the case of the 4F2 heavy-chain cell surface antigens that contain part of a typical domain B sequence (equivalent to one of the β-strands) in the predicted $\beta 3 \rightarrow \alpha 3$ segment, supports the divergent evolution. The insertion of a domain B into an α-amylase-type $(\beta/\alpha)_8$-barrel seems, therefore, to be an early event, preceding the specialization of the α-amylase-type $(\beta/\alpha)_8$-barrel prototype to enzymes with differ-

Table 1. Average scores of the three clusters corresponding to Bα1, Bβ1, and Bβ3 in the HCA plots (Fig. 2) of domain B of the *Bacillus cereus* oligo-1,6-glucosidase group

	OGL[a]	AGL	CMD	DGL	T6P	NPU	BME	DTB
AGL	0.933							
CMD	0.643	0.675						
DGL	1.000	0.933	0.643					
T6P	0.859	0.903	0.624	0.859				
NPU	0.577	0.592	0.821	0.577	0.629			
BME	0.822	0.867	0.643	0.822	0.909	0.577		
DTB	0.704	0.748	0.667	0.704	0.889	0.595	0.815	
AAT	0.915	0.852	0.726	0.915	0.790	0.667	0.746	0.690

[a] The protein sources are abbreviated as follows: OGL, oligo-1,6-glucosidase from *Bacillus cereus;* AGL, α-glucosidase from *Saccharomyces carlsbergensis;* CMD, cyclomaltodextrinase from *Bacillus sphaericus;* DGL, dextran glucosidase from *Streptococcus mutans;* T6P, trehalose-6-phosphate hydrolase from *Bacillus subtilis;* NPU, neopullulanase from *Thermoactinomyces vulgaris;* BME, α-amylase from *Bacillus megaterium;* DTB, α-amylase from *Dictyoglomus thermophilum* (AmyB); AAT, amino acid transport-related protein from human kidney

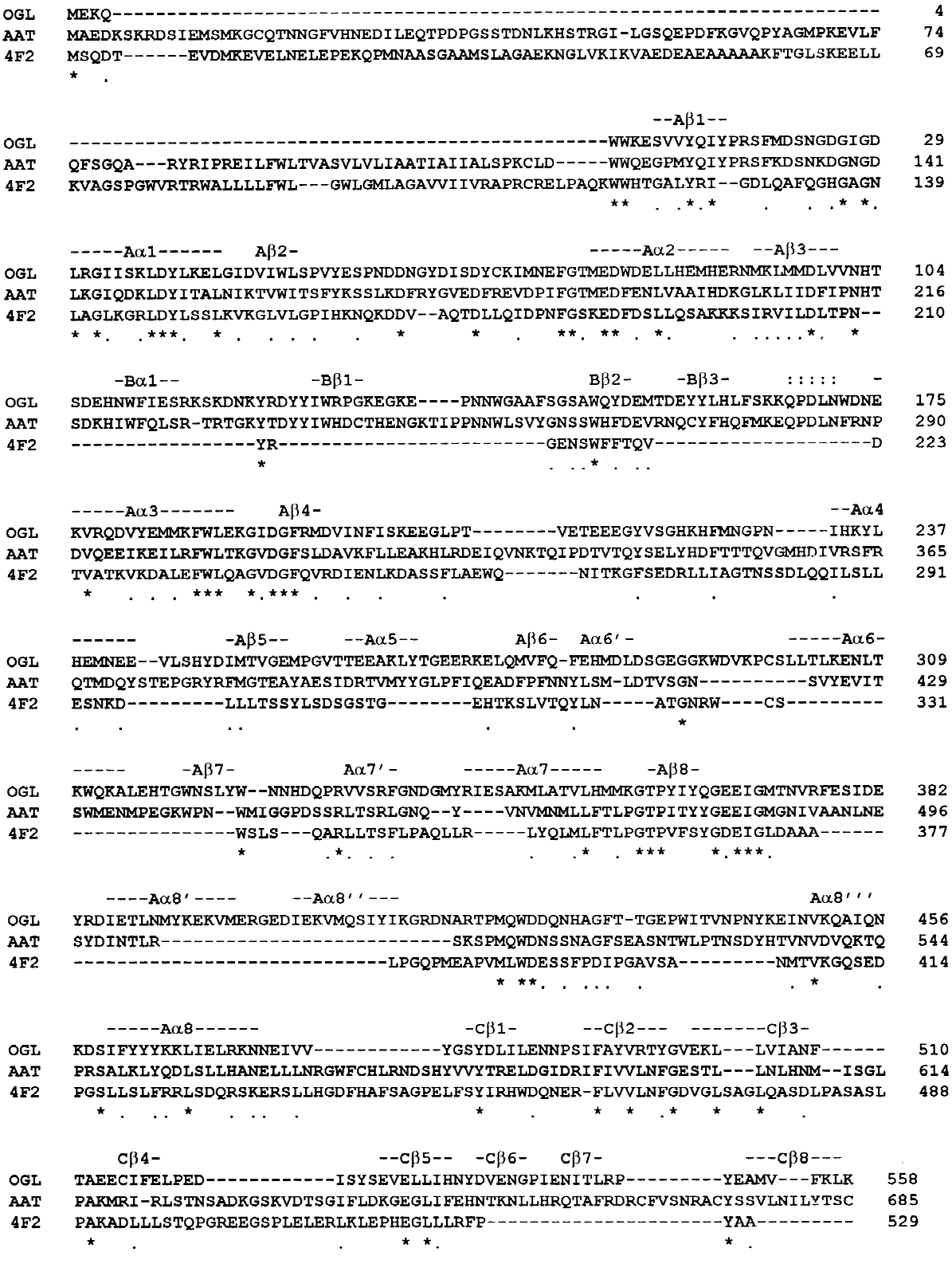

Fig. 3. Amino acid sequence alignment of oligo-1,6-glucosidase from *Bacillus cereus* (OGL; Watanabe et al. 1990), human amino acid transport-related protein (AAT; Bertran et al. 1993), and human 4F2 heavy-chain cell surface antigen (4F2; Gottesdiener et al. 1988). The invariant amino acid residues and conservative substitutions are marked by asterisks and dots, respectively. The secondary structure elements of *B. cereus* oligo-1,6-glucosidase (Watanabe et al. 1994) are indicated above the alignment. The short conserved sequence near the C-terminus of domain B is indicated by colons.

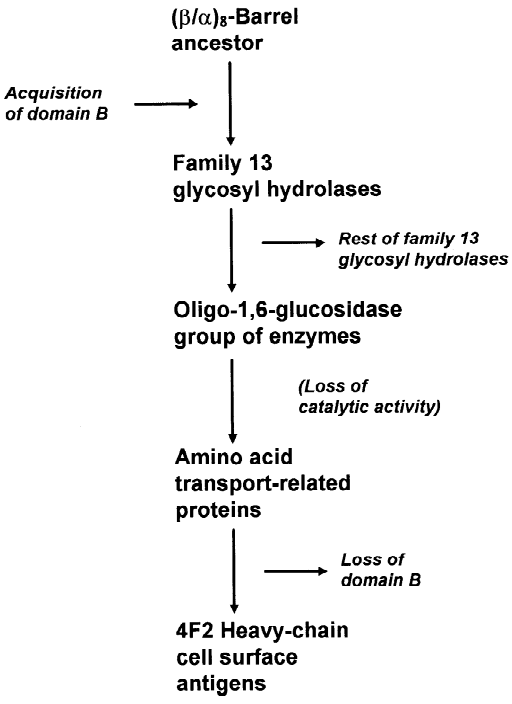

Fig. 4. Possible evolutionary events in the α-amylase family.

ent function. Thus one original type of domain B has adopted various secondary and supersecondary structures in different members of the α-amylase family. Moreover, as demonstrated in Figure 1, the evolution of domain B matches that of the $(\beta/\alpha)_8$-fold, since the groups of enzymes exhibiting high-sequence similarities in domain B and in structural elements of the $(\beta/\alpha)_8$-barrel are essentially the same.

With regard to the presence of domain B either in other $(\beta/\alpha)_8$-barrel proteins or in proteins from different fold families, domain B is found in a mammalian amino acid transport-related protein and as a rudiment in 4F2 heavy-chain cell surface antigen. Both proteins sequences, however, show a distant structural relationship to the family 13 glycosyl hydrolases (Fig. 3). Analogous excursion from the regular $(\beta/\alpha)_8$-barrel domain at the place of $\beta \rightarrow \alpha$ loop 3 has been reported for old yellow enzyme (Fox and Karplus 1994) and tRNA-guanine transglycosylase (Romier et al. 1996). The $\beta 3 \rightarrow \alpha 3$ segment of tRNA-guanine transglycosylase is particularly interesting by having a three-stranded antiparallel β-sheet flanked by α-helices (Romier et al. 1996), and thus a supersecondary structure motif highly similar to that of *B. cereus* oligo-1,6-glucosidase. The old yellow enzyme and tRNA-guanine transglycosylase thus seem excellent candidates for evolutionary studies, although they are distantly related with the α-amylase family. Finally, at present, no protein from a different protein fold family contains a domain that resembles domain B of the family 13 glycosyl hydrolases.

Acknowledgments. ŠJ thanks the Federation of European Biochemical Societies for a Short-Term Fellowship. This work was supported in part by the grant 2/3013/96 from the Slovak Grant Agency for Science (ŠJ) and the EU biotechnology program BIO2-CT94-3008 (BS).

References

Abe J, Shibata Y, Fujisue M, Hizukuri S (1996) Expression of periplasmic α-amylase of *Xanthomonas campestris* K-11151 in *Escherichia coli* and its action on maltose. Microbiology 142:1505–1512

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Bertran J, Werner A, Chillaron J, Nunes V, Biber J, Testar X, Zorzano A, Estivill X, Murer H, Palacin M (1993) Expression cloning of a human renal cDNA that induces high affinity transport of L-cystine shared with dibasic amino acids in *Xenopus* oocytes. J Biol Chem 268:14842–14849

Brady RL, Brzozowski AM, Derewenda ZS, Dodson EJ, Dodson GG (1991) Solution of the structure of *Aspergillus niger* acid α-amylase by combined molecular replacement and multiple isomorphous replacement methods. Acta Crystallogr B47:527–535

Brayer GD, Luo Y, Withers SG (1995) The structure of human pancreatic α-amylase at 1.8 Å resolution and comparisons with related enzymes. Protein Sci 4:1730–1742

Brändén C-I (1991) The TIM barrel—the most frequently occurring folding motif in proteins. Curr Opin Struct Biol 1:978–983

Farber GK, Petsko GA (1990) The evolution of α/β barrel enzymes. Trends Biochem Sci 15:228–234

Feller G, Payan F, Theys F, Qian M, Haser R, Gerday C (1994) Stability and structural analysis of α-amylase from the antarctic psychrophile *Alteromonas haloplanctis* A23. Eur J Biochem 222:441–447

Fox KM, Karplus PA (1994) Old yellow enzyme at 2 Å resolution: overall structure, ligand binding, and comparison with related flavoproteins. Structure 2:1089–1105

Gaboriaud C, Bissery V, Benchetrit T, Mornon JP (1987) Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett 224:149–155

Gottesdiener KM, Karpinski BA, Lindsten T, Strominger JL, Jones NH, Thompson CB, Leiden JM (1988) Isolation and structural characterization of the human 4F2 heavy-chain gene, an inducible gene involved in T-lymphocyte activation. Mol Cell Biol 8:3809–3819

Helfert C, Gotsche S, Dahl MK (1995) Cleavage of trehalose-phosphate in *Bacillus subtilis* is catalysed by a phospho-α-(1-1)-glucosidase encoded by the *treA* gene. Mol Microbiol 16:111–120

Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 280:309–316

Henrissat B, Bairoch A (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 293:781–788

Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP, Davies G (1995) Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. Proc Natl Acad Sci USA 92:7090–7094

Hickey DA, Benkel BF, Boer PH, Genest Y, Abukashawa S, Ben-David G (1987) Enzyme-coding genes as molecular clocks: the molecular evolution of animal α-amylases. J Mol Evol 26:252–256

Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSTAL V: Improved software for multiple sequence alignment. Comput Applic Biosci 8:189–191

Hong SH, Marmur J (1986) Primary structure of the maltase gene of the *MAL6* locus of *Saccharomyces carlsbergensis.* Gene 41:75–84

Horinouchi S, Fukusumi S, Ohshima T, Beppu T (1988) Cloning and expression in *Escherichia coli* of two additional amylase genes of a strictly anaerobic thermophile, *Dictyoglomus thermophilum,* and

their nucleotide sequences with extremely low guanine-plus-cytosine contents. Eur J Biochem 176:243–253

Janeček Š (1992) New conserved amino acid region of α-amylases in the third loop of their $(\beta/\alpha)_8$-barrel domains. Biochem J 288:1069–1070

Janeček Š (1994a) Parallel β/α-barrels of α-amylase, ćyclodextrin glycosyltransferase and oligo-1,6-glucosidase versus the barrel of β-amylase: evolutionary distance is a reflection of unrelated sequences. FEBS Lett 353:119–123

Janeček Š (1994b) Sequence similarities and evolutionary relationships of microbial, plant and animal α-amylases. Eur J Biochem 224: 519–524

Janeček Š (1995a) Close evolutionary relatedness among functionally distantly related members of the $(\alpha/\beta)_8$-barrel glycosyl hydrolases suggested by the similarity of their fifth conserved sequence region. FEBS Lett 377:6–8

Janeček Š (1995b) Tracing the evolutionary lineages among α-amylases and cyclodextrin glycosyltransferases: the question of so-called ''intermediary'' enzymes. Biologia 50:515–522

Janeček Š (1996) Invariant glycines and prolines flanking in loops the strand β2 of various $(\alpha/\beta)_8$-barrel enzymes: a hidden homology? Protein Sci 5:1136–1143

Janeček Š, Baláž Š (1995) Functionally essential, invariant glutamate near the C-terminus of strand β5 in various $(\alpha/\beta)_8$-barrel enzymes as a possible indicator of their evolutionary relatedness. Protein Eng 8:809–813

Janeček Š, Bateman A (1996) The parallel $(\alpha/\beta)_8$-barrel: perhaps the most universal and the most puzzling protein folding motif. Biologia 51:613–628

Janeček Š, MacGregor EA, Svensson B (1995) Characteristic differences in the primary structure allow discrimination of cyclodextrin glucanotransferases from α-amylases. Biochem J 305:685–686

Janse BJH, Steyn, AJC, Pretorius IS (1993) Regional sequence homologies in starch-degrading enzymes. Curr Genet 24:400–407

Jenkins J, Lo Leggio L, Harris G, Pickersgill R (1995) β-Glucosidase, β-galactosidase, family A cellulases, family F xylanases and two barley glycanases form a family of enzymes with 8-fold β/α architecture and with two conserved glutamates near the carboxy-terminal ends of β-strands four and seven. FEBS Lett 362:281–285

Jespersen HM, MacGregor EA, Sierks MR, Svensson B (1991) Comparison of the domain-level organization of starch hydrolases and related enzymes. Biochem J 280:51–55

Jespersen HM, MacGregor EA, Henrissat B, Sierks MR, Svensson B (1993) Starch- and glycogen-debranching and branching enzymes: prediction of structural features of the catalytic $(\beta/\alpha)_8$-barrel domain and evolutionary relationship to other amylolytic enzymes. J Protein Chem 12:791–805

Juge N, Rodenburg KW, Guo X-J, Chaix J-C, Svensson B (1995) Isozyme hybrids within the protruding third loop domain of the barley α-amylase $(\beta/\alpha)_8$-barrel. Implication for BASI sensitivity and substrate specificity. FEBS Lett 363:299–303

Kadziola A, Abe J, Svensson B, Haser R (1994) Crystal and molecular structure of barley α-amylase. J Mol Biol 239:104–121

Kizaki H, Hata Y, Watanabe K, Katsube Y, Suzuki Y (1993) Polypeptide folding of *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase revealed by 3.0 Å resolution X-ray analysis. J Biochem 113:646–649

Klein C, Schulz GE (1991) Structure of cyclodextrin glycosyltransferase refined at 2.0 Å resolution. J Mol Biol 217:737–750

Knegtel RMA, Wind RD, Rozeboom HJ, Kalk KH, Buitelaar RM, Dijkhuizen L, Dijkstra BW (1996) Crystal structure at 2.3 Å resolution and revised nucleotide sequence of the thermostable cyclodextrin glycosyltransferase from *Themoanaerobacterium thermosulfurigenes* EM1. J Mol Biol 256:611–622

Kubota M, Matsuura Y, Sakai S, Katsube Y (1991) Molecular structure of *B. stearothermophilus* cyclodextrin glucanotransferase and analysis of substrate binding site. Denpun Kagaku 38:141–146

Kuriki T, Kaneko H, Yanase M, Takata H, Shimada J, Handa S, Takada T, Umeyama H, Okada S (1996) Controlling substrate preference and transglycosylation activity of neopullulanase by manipulating steric constraint and hydrophobicity in active center. J Biol Chem 271:17321–17329

Lamminmäki U, Vihinen M (1996) Structural consequences of neopullulanase mutations. Biochim Biophys Acta 1295:195–200

Lawson CL, Van Montfort R, Strokopytov B, Rozeboom HJ, Kalk KH, De Vries GE, Penninga D, Dijkhuizen L, Dijkstra BW (1994) Nucleotide sequence and X-ray structure of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 in a maltose-dependent crystal form. J Mol Biol 236:590–600

Lemesle-Varloot L, Gaboriaud C, Pantel G, Morgat A, Mornon JP, Lavaitte S, Lestang F, Henrissat B (1993) MANSEK and SUN-HCA: two interactive programs for the hydrophobic cluster analysis of protein sequences. Comput Applic Biosci 9:37–44

Lesk AM, Brändén C-I, Chothia C (1989) Structural principles of α/β barrel proteins: the packing of the interior of the sheet. Proteins Struct Funct Genet 5:139–148

MacGregor EA (1993) Relationships between structure and activity in the α-amylase family of starch-metabolising enzymes. Starch 7: 232–237

MacGregor EA, Svensson B (1989) A supersecondary structure predicted to be common to several α-1,4-D-glucan-cleaving enzymes. Biochem J 259:145–152

Machius M, Wiegand G, Huber R (1995) Crystal structure of calcium-depleted *Bacillus licheniformis* α-amylase at 2.2 A resolution. J Mol Biol 246:545–559

Matsuura Y (1995) Crystal structure of α-amylases and related enzymes. In: Yamamoto T (ed) Enzyme chemistry and molecular biology of amylases and related enzymes. CRC Press, Ann Arbor, pp 137–145

Matsuura Y, Kusunoki M, Harada W, Kakudo M (1984) Structure and possible catalytic residues of Taka-amylase A. J Biochem 95:697–702

Metz RJ, Allen LN, Cao TM, Zeman NW (1988) Nucleotide sequence of an amylase gene from *Bacillus megaterium.* Nucl Acids Res 16:5203

Nitschke L, Heeger K, Bender H, Schulz GE (1990) Molecular cloning, nucleotide sequence and expression in *Escherichia coli* of the β-cyclodextrin glycosyltransferase gene from *Bacillus circulans* strain no. 8. Appl Microbiol Biotechnol 33:542–546

Oguma T, Matsuyama A, Kikuchi M, Nakano E (1993) Cloning and sequence analysis of the cyclomaltodextrinase gene from *Bacillus sphaericus* and expression in *Escherichia coli* cells. Appl Microbiol Biotechnol 39:197–203

Pasero L, Mazzéi-Pierron Y, Abadie B, Chicheportiche Y, Marchis-Mouren G (1986) Complete amino acid sequence and location of the five disulfide bridges in porcine pancreatic α-amylase. Biochim Biophys Acta 869:147–157

Qian M, Haser R, Payan F (1993) Structure and molecular model refinement of pig pancreatic α-amylase at 2.1 Å resolution. J Mol Biol 231:785–799

Qian M, Haser R, Buisson G, Dueé E, Payan F (1994) The active center of a mammalian α-amylase. Structure of the complex of a pancreatic α-amylase with a carbohydrate inhibitor refined to 2.2-Å resolution. Biochemistry 33:6284–6294

Quackenbush E, Clabby M, Gottesdiener KM, Barbosa J, Jones NH, Strominger JL, Speck S, Leiden JL (1987) Molecular cloning of complementary DNAs encoding the heavy chain of the human 4F2 cell-surface antigen: A type II membrane glycoprotein involved in normal and neoplastic cell growth. Proc Natl Acad Sci USA 84: 6526–6530

Raine ARC, Scrutton NS, Mathews FS (1994) On the evolution of alternate core packing in eightfold β/α-barrels. Protein Sci 3:1889–1892

Ramasubbu N, Paloth V, Luo Y, Brayer GD, Levine MJ (1996) Structure of human salivary α-amylase at 1.6 Å resolution: implications for its role in the oral cavity. Acta Crystallogr D52:435–446

Reardon D, Farber GK (1995) The structure and evolution of α/β barrel proteins. FASEB J 9:497–503

Rodenburg KW, Juge N, Guo X-J, Søgaard M, Chaix J-C, Svensson B (1994) Domain B protruding at the third β strand of the α/β barrel in barley α-amylase confers distinct isozyme-specific properties. Eur J Biochem 221:277–284

Rogers JC (1985) Two barley α-amylase gene families are regulated differently in aleurone cells. J Biol Chem 260:3731–3738

Romier C, Reuter K, Suck D, Ficner R (1996) Crystal structure of tRNA-guanine transglycosylase: RNA modification by base exchange. EMBO J 15:2850–2857

Russell RRB, Ferretti JJ (1990) Nucleotide sequence of the dextran glucosidase (*dexB*) gene of *Streptococcus mutans*. J Gen Microbiol 136:803–810

Scrutton NS (1994) α/β Barrel evolution and the modular assembly of enzymes: emerging trends in the flavin oxidase/dehydrogenase family. BioEssays 16:115–122

Strokopytov B, Penninga D, Rozeboom HJ, Kalk KH, Dijkuizen L, Dijkstra BW (1995) X-ray structure of cyclodextrin glycosyltransferase complexed with acarbose. Implications for the catalytic mechanism of glycosidases. Biochemistry 34:2234–2240

Svensson B (1994) Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity, and stability. Plant Mol Biol 25:141–157

Takii Y, Takahashi K, Yamamoto K, Sogabe Y, Suzuki Y (1996) *Bacillus stearothermophilus* ATCC12016 α-glucosidase specific for α-1,4 bonds of maltosaccharides and α-glucans shows high amino acid sequence similarities to seven α-D-glucohydrolases with different substrate specificity. Appl Microbiol Biotechnol 44:629–634

Toda H, Kondo K, Narita K (1982) The complete amino acid sequence of Taka-amylase A. Proc Japan Acad B58:208–212

Tonozuka T, Ohtsuka M, Mogi S-I, Sakai H, Ohta T, Sakano Y (1993) A Neopullulanase-type α-amylase from *Thermoactinomyces vulgaris* R-47. Biosci Biotech Biochem 57:395–401

Tsukamoto A, Kimura K, Ishi Y, Takano T, Yamane K (1988) Nucleotide sequence of the maltohexaose-producing amylase gene from an alkalophilic *Bacillus* sp. #707 and structural similarity to liquefying type α-amylases. Biochem Biophys Res Commun 151:25–31

Watanabe K, Kitamura K, Iha H, Suzuki Y (1990) Primary structure of the oligo-1,6-glucosidase of *Bacillus cereus* ATCC7064 deduced from the nucleotide sequence of the gene. Eur J Biochem 192:609–620

Watanabe K, Masuda T, Ohashi H, Mihara H, Suzuki Y (1994) Multiple proline substitutions cumulatively thermostabilize *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase. Irrefragable proof supporting the proline rule. Eur J Biochem 226:277–283

Wells RG, Hediger MA (1992) Cloning of a rat kidney cDNA that stimulates dibasic and neutral amino acid transport and has sequence similarity to glucosidases. Proc Natl Acad Sci USA 89: 5596–5600

Whitehead TR, Cotta MA (1995) Identification of intracellular amylase activity in *Streptococcus bovis* and *Streptococcus salivarius*. Curr Microbiol 30:143–148

Woodcock S, Mornon J-P, Henrissat B (1992) Detection of secondary structure elements in proteins by hydrophobic cluster analysis. Protein Eng 5:629–635

Yuuki T, Nomura T, Tezuka H, Tsuboi A, Yamagata H, Tsukagoshi N, Udaka S (1985) Complete nucleotide sequence of a gene coding for heat- and pH-stable α-amylase of *Bacillus licheniformis:* comparison of the amino acid sequences of three bacterial liquefying α-amylases deduced from the DNA sequences. J Biochem 98:1147–1156