Biologia, Bratislava, **57/Suppl. 11:** 29—41, 2002 **Review**

How many conserved sequence regions are there in the α-amylase family?

Štefan Janeček*

Institute of Molecular Biology, Slovak Academy of Sciences, Dúbravská cesta 21, SK-84251 Bratislava, Slovakia; tel.: ++ 421 2 5930 7420, fax: ++ 421 2 5930 7416, e-mail: stefan.janecek@savba.sk

Janeček, Š., How many conserved sequence regions are there in the α-amylase family? Biologia, Bratislava, **57/Suppl. 11:** 29—41, 2002; ISSN 0006-3088.

The α-amylase family, i.e. the clan GH-H of glycoside hydrolases, is the largest family of glycoside hydrolases, transferases and isomerases comprising nearly 30 different enzyme specificities. One of the most interesting features of this family is that its members contain a few very well-conserved sequence regions despite the overall low sequence similarity. It seems that only 4 amino acid residues may be totally invariantly conserved throughout the family (Arg204 plus the three catalytic residues: Asp206, Glu230 and Asp297; Taka-amylase A numbering). Four conserved sequence regions covering the strands β3, β4, β5, and β7 of the catalytic $(\beta/\alpha)_8$-barrel domain were identified and used for defining the α-amylase family. The present review is therefore focused on the three additional conserved sequence regions proposed after the basic characteristics of the family have been established. Two of these three regions cover roughly the strands β2 and β8 of the catalytic $(\beta/\alpha)_8$-barrel and one is located near the C-terminus of domain B (in the β3$\rightarrow$ α3 connection of the catalytic $(\beta/\alpha)_8$-barrel). While the four original conserved sequence regions contain the catalytic and substrate-binding residues of the individual members of the family, the three conserved sequence regions later identified are shown to contain amino acid residues connected to a given enzyme specificity. The problems that may arise with correctly identifying the β5-strand catalytic glutamate, positioned in the conserved sequence region III, are discussed and a way that may lead to a correct solution is indicated. In conclusion, it is proposed that the α-amylase enzyme family should be characterised by as many conserved sequence regions as possible.

Key words: alpha-amylase family, conserved sequence regions, sequence similarities, specificity features, evolutionary relationships.

Abbreviations: CGTase, cyclodextrin glucanotransferase.

Introduction

The α-amylase enzyme family was originally recognised as a group of starch hydrolases and related enzymes that exhibit clear sequence similarities and a predicted common supersecondary fold, a parallel $(\beta/\alpha)_8$-barrel (MacGregor, 1988; Svensson, 1988; MacGregor & Svensson, 1989; Jespersen et al., 1991; 1993). From the beginning it covered several different enzyme speci-

* Corresponding author

ficities, including cyclodextrin glucanotransferase (CGTase), isoamylase, branching enzyme, neopullulanase and oligo-1,6-glucosidase in addition to α-amylase (Binder et al., 1986; Amemura et al., 1988; Romeo et al., 1988; Kuriki & Imanaka, 1989; Watanabe et al., 1990). The α-amylase family has been known as the family 13 of glycoside hydrolases from 1991 when the sequence-based classification of all glycoside hydrolases, transferases, and isomerases was established (Henrissat, 1991).

The concept of this group of enzymes as the α-amylase family was proposed a year later (Takata et al., 1992). According to that definition the members of the α-amylase family are enzymes that satisfy the four requirements: (i) they act on α-glucosidic linkages; (ii) they hydrolyse or form by transglycosylation α-glucosidic linkages; (iii) their amino-acid sequence contains four conserved regions; and (iv) they contain Asp, Glu and Asp residues corresponding to the Asp206, Glu230 and Asp297 of Taka-amylase A.

The current situation in the α-amylase family is as follows: (i) the family contains almost 30 different enzyme specificities covering hydrolases, transferases, and isomerases (Svensson et al., 2002); (ii) the family constitutes a clan GH-H of glycoside hydrolases comprising the families 13, 70, and 77 (Coutinho & Henrissat, 1999); (iii) the enzymes from the family can operate on α-1,1-, α-1,2-, α-1,3- and α-1,5-glucosidic linkages in addition to the α-1,4- and α-1,6-bonds originally considered (MacGregor et al., 2001); (iv) there may be at most only 4 invariant amino acid residues, i.e. 3 catalytic ones plus the Arg in position $i-2$ with respect to the catalytic nucleophile, the aspartate in the β4-strand (Janeček, 2000a); (v) the relationship of the main α-amylase family (clan GH-H) to the "extremophilic" α-amylase family, family 57 of glycoside hydrolases, has still been neither confirmed nor disproved (Janeček, 1998; Imamura et al., 2001a); and (vi) from the points of view of sequence and evolution, some mammalian transport proteins and antigens are perhaps related to the enzymatic members of the family (Janeček, 2000b).

Many specialised reports are available focusing on protein engineering of the individual enzymes and certain groups from the α-amylase family, as well as on the possibilities of tailoring their properties to meet specific needs (e.g., Kuriki et al., 1996; Matsui & Svensson, 1997; Binderup & Preiss, 1998; Ibuka et al., 1998; Wind et al., 1998; Saab-Rincón et al., 1999; Beier et al., 2000; Yamamoto et al., 2000; Gottschalk et al., 2001; Mori et al., 2001; Kim et al., 2001; Leemhuis et al., 2002). The findings have been thoroughly discussed in various reviews (e.g., MacGregor, 1993; Svensson, 1994; Janeček, 1997a; Nielsen & Borchert, 2000; Park et al., 2000; van der Veen et al., 2000; Uitdehaag et al., 2002; van der Maarel et al., 2002). In general, however, the residues playing the most important roles are located in the short sequence stretches conserved throughout the family and known as the conserved sequence regions. This review therefore aims to present and summarise the short sequence regions that are best conserved either throughout the α-amylase family or in several closely-related enzyme groups from the family, and to show the relationships between these conserved sequence regions and the individual enzyme specificities.

From family to clan and from clan to a ... suprafamily ?

When the number of sequences recognised as belonging to the α-amylase family was low, i.e. several tens at the beginning of the 1990s (MacGregor & Svensson, 1989; Henrissat, 1991; Takata et al., 1992), it was clear that, despite the existence of more enzyme specificities, all the sequences constitute a classical enzyme family with members showing divergent evolution (Jespersen et al., 1993; Janeček, 1994a). As one of the main criteria for deciding whether or not a new sequence belongs to the α-amylase family, the presence of the four conserved sequence regions (e.g., for Taka-amylase A – I: 117_DVVANH, II: 202_GLRIDTVKH, III: 230_EVLD, IV: 292_FVENHD) was used as a good criterion (Nakajima et al., 1986).

This basic trend has, in fact, been maintained up to the present day (MacGregor et al., 2001) although the number of sequences has reached almost one thousand (Coutinho & Henrissat, 1999). According to the Henrissat classification, however, the original α-amylase family, family 13 of glycoside hydrolases (Henrissat, 1991), has been changed to a clan, glycoside hydrolases clan GH-H, covering the families 13, 70, and 77 (Bourne & Henrissat, 2001). While the original family 13 and the family 77 seem to be very closely related to each other (Przylas et al., 2000), i.e. they roughly fulfil the attributes of a single family of similar reaction mechanism, fold, and catalytic machinery seen as identical catalytic residues on identical elements of secondary structure, family 70 con-

tains enzymes with a version of the $(\beta/\alpha)_8$-barrel structure that is a circularly permuted form of the classical α-amylase-type $(\beta/\alpha)_8$-barrel (MacGregor et al., 1996). Although an alternative prediction without circular permutation was presented for the family 70 glucosyltransferases (Devulapalle et al., 1997), the predictions taking into account the possibility of circular permutations seem more reasonable and have been supported by mutagenesis (MacGregor et al., 1996; Monchois et al., 1999; 2000; Tsai et al., 2000).

The existence of a single α-amylase family, the family 13 of glycoside hydrolases, has appeared more "complicated" since 1988 and 1993 when two "unusual" amino acid sequences of α-amylases were reported: the AmyA from *Dictyoglomus thermophilum* (Fukusumi et al., 1988) and one from *Pyrococcus furiosus* (Laderman et al., 1993). These and also some other mostly extremophilic amylolytic enzymes (Dong et al., 1997b; Jeon et al., 1997; Erra-Pujada et al., 1999) were found to lack the characteristic conserved sequence regions; this resulted in a new family, family 57 of glycoside hydrolases being established (Henrissat & Bairoch, 1996). Later, however, a distant evolutionary relationship between the major family 13 and family 57 was indicated by using HCA (hydrophobic cluster analysis) on the amino-acid sequence of archaeal *Methanococcus jannaschii* α-amylase (Janeček, 1998). At present family 57 contains members mostly from extremophiles (Coutinho & Henrissat, 1999) having α-amylase, 4-α-glucanotransferase, amylopullulanase and even α-galactosidase (van Lieshout et al., 2001) specificities in addition to numerous hypothetical proteins coming from various complete-genome sequencing projects. Recent determination of the three-dimensional structure of the 4-α-glucanotransferase from *Thermococcus litoralis,* with site-directed mutated glutamate residues, indicated (Imamura et al., 2001a,b) that the HCA-based predictions (Janeček, 1998) may not match the catalytic amino acids and thus may invalidate the postulated close relationship of glycoside hydrolase family 57 to clan GH-H. Despite this recent finding, the common retaining mechanism and a common $(\beta/\alpha)_8$-barrel fold used by both the clan GH-H and the family 57 preserve the possibility of classifying them into an even larger group than the clan, e.g. a suprafamily representing relatedness at a higher hierarchal level than a clan.

Conserved sequence regions

Study of the α-amylase family is practically impossible without taking into account the question of the conserved sequence regions. For a considerable time, the sequence similarity has been known to be extremely low (about 10%) even for the α-amylases alone (i.e. for EC 3.2.1.1) as described from different microorganisms, plants, and animals (Nakajima et al., 1986). Later when the family grew, i.e. when many sequences from various sources and with different enzyme specificities became available, the number of identical residues among the α-amylase family enzymes had decreased to 8-10 amino acids in 1994 (Janeček, 1994a; Svensson, 1994). At present there might be at most only 4 residues that are totally invariantly conserved throughout the α-amylase family, i.e. the 3 catalytic residues corresponding to Asp206, Glu230 and Asp297 in Taka-amylase A plus the β4 arginine equivalent to the Arg204 in Taka-amylase A. The β3 aspartate (Asp117 in Taka-amylase A) is also extremely well conserved (Fig. 1), however, it has been found substituted at least one time (Š. Janeček, unpublished results): by glutamate in trehalose synthase from *Thermus aquaticus* (Tsusaki et al., 1997). Despite this fact, amino-acid sequences of enzymes from the α-amylase family contain several short conserved stretches, the so-called conserved sequence regions (Fig. 1). Of the four above-mentioned invariantly-conserved residues Asp206 (the catalytic nucleophile) and Glu230 (the proton donor) together with the second conserved aspartate, Asp 297, and two histidine residues, His122 and His296, were postulated as being crucial for Taka-amylase A (Matsuura et al., 1984). They form the basis of the four well-known conserved sequence regions in the α-amylase family enzymes (the regions I, II, III and IV in Fig. 1) and are positioned near the C-termini of strands β3, β4, β5, and β7 of the catalytic $(\beta/\alpha)_8$-barrel domain (Nakajima et al., 1986; MacGregor et al., 2001). It is clear, however, that the two histidines are not conserved throughout the family (Fig. 1).

Although Nakajima et al. (1986) established the four conserved sequence regions of α-amylases, the authors paid less attention to the remaining part of their alignment of 11 different α-amylase sequences. This fact was mentioned already by MacGregor (1988), who indicated that there could be additional sequence similarities in α-amylases that were found later in other enzymes belonging to the α-amylase family by Svensson (1988), MacGregor & Svensson (1989) and

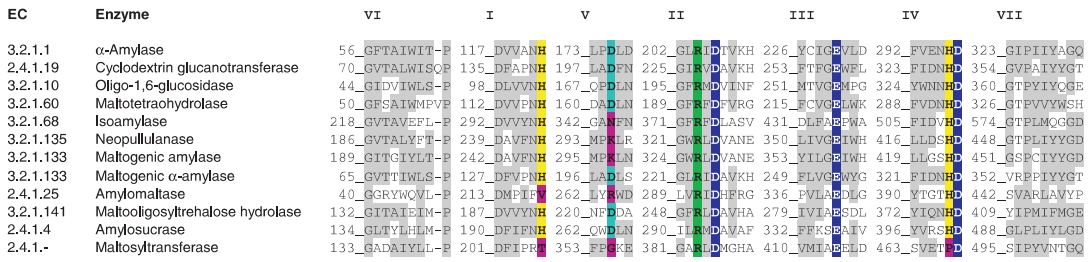

Fig. 1. Conserved sequence regions in the α-amylase family. Sources of the enzymes: α-amylase (*Aspergillus oryzae*; Toda et al., 1982; Matsuura et al., 1984), cyclodextrin glucanotransferase (*Bacillus circulans* strain 8; Nitschke et al., 1990; Klein & Schulz, 1991), oligo-1,6-glucosidase (*Bacillus cereus*; Watanabe et al., 1990; Kizaki et al., 1993), maltotetraohydrolase (*Pseudomonas stutzeri*; Fujita et al., 1989; Morishita et al., 1997), isoamylase (*Pseudomonas amyloderamosa*; Amemura et al., 1988; Katsuya et al., 1998), neopullulanase (*Thermoactinomyces vulgaris* R-47; Tonozuka et al., 1993; Kamitori et al., 1999), maltogenic amylase (*Thermus* sp.IM6501; Kim et al., 1999b; Kim et al., 1999a), maltogenic α-amylase (*Bacillus stearothermophilus*; Diderichsen & Christiansen, 1988; Dauter et al., 1999), amylomaltase (*Thermus aquaticus*; Terada et al., 1999; Przylas et al., 2000), maltooligosyltrehalose hydrolase (*Sulfolobus solfataricus*; Kobayashi et al., 1996; Feese et al., 2000), amylosucrase (*Neisseria polysaccharea*; Potocki De Montalk et al., 1999; Skov et al., 2001), maltosyltransferase (*Thermotoga maritima*; Meissner & Liebl, 1998; Roujeinikova et al., 2001). Colour code: catalytic aspartates and glutamate – blue; functional histidines – yellow; invariant arginine – green; Ca^{2+}-binding aspartate – turquoise; conserved residues – grey; non-conserved residues – pink. The first two columns denote the year of three-dimensional structure determination and the EC number of a given member, respectively. Only the members whose structures have already been determined are shown.

Jespersen et al. (1991, 1993). Based on the exhaustive analyses of a large number of amino-acid sequences of α-amylases and related enzymes, Janeček (1992, 1994a,b, 1995a) proposed three additional conserved sequence regions. The fifth conserved sequence region (Janeček, 1992, 1995a) is located near the C-terminus of domain B around the calcium-binding aspartate, Asp175 in Taka-Amylase A (the region V in Fig. 1). Two additional conserved sequence regions (Janeček, 1994a,b) can be found at strands β2 and β8 of the catalytic $(\beta/\alpha)_8$-barrel domain (the regions VI and VII in Fig. 1). It has been proposed that the conservations may be related to maintenance of structure (MacGregor et al., 2001), but variations within these segments may also be characteristic of certain enzyme specificities (Janeček, 1997a).

Thus, it has become more and more clear that the α-amylase family should be characterised by as many conserved sequence regions as possible because it contains a large number of different members (Svensson et al., 2002). For many sequences, particular amino-acid residues characteristic of the enzyme's specificity are present in all the regions (MacGregor et al., 2001). Even if an enzyme of a given specificity does not contain a given conserved sequence region, this absence may be used as a feature characteristic of that enzyme specificity (Janeček, 2000a).

From the chronological point of view the first three conserved sequence regions covering the strands β3, β4, and β7 (Fig. 1) were originally pointed out by Toda et al. (1982) for Taka-amylase A and pig pancreatic α-amylase. A year later Friedberg (1983) added the regions of *Bacillus amyloliquefaciens* α-amylase, and finally Rogers (1985) completed the picture by describing the regions in barley α-amylase. Thus three of the four conserved sequence regions that were subsequently used for the definition of the α-amylase enzyme family (Takata et al., 1992) were known in 1985 for the α-amylases from bacteria, fungi, plants, and animals (Toda et al., 1982; Friedberg, 1983; Rogers, 1985). In 1986 Nakajima et al. (1986) made a comparative study of eleven amino-acid sequences of α-amylases from different origins (microorganisms, plants, and animals) and identified the fourth conserved sequence region covering the strand β5 (the one with the β5-strand catalytic glutamic acid; Fig. 1). They also established the length of all the four conserved sequence regions. These four regions were progressively identified by sequence comparison and/or prediction for other starch hydrolases and related enzymes (e.g., MacGregor & Svensson, 1989; Jespersen et al., 1991, 1993; MacGregor et al., 1996) and became one of the distinguishing marks of the α-amylase family (Vihinen & Mäntsälä, 1989; Takata et al., 1992; MacGre-

(a)

```
                                  ↓
MSSSCRFRPLSFLLAAALGASAAFAPSVHAADAAPATPRSEGSGVWYEIFVRAWYDTDGD
GIGDLNGVTAKLDYLQSLGVSGIWLMPINPSPSYHGYDITDYEGINPQYGTMADFEKLVS
EAHKRGIEVILDLVINHTSDQHPWFKAALDPKDAHRSWYTWAGPGTNLKAVSAVGGPAWH
ANGKQHYLGDFTGAMPDLNYDEPAVRREMIAVGKFWLDKGADGFRLDAARHIYDDLESDN
GQPAVIARNAQWWNEFRQGLRQVRPDVYLVGEVSAKQPGELAPYLPALGSVFDFPLAEQL
IASAGQEKAGKLPALLTETYAAFRAAAGDDYADAPFLSNHDQERVLSQLGGDLRHMRTAA
AMLLTLPGRPYLYYGEEIGMLGRKPDENLREPMRWQRQAGAPGDSRWKPYSVKQGGEVSV
QAEQDRPDSLLSFYRTLIHWRVEVPALRDGALRVVDTGAPALVAYERVTADSRALVVHNL
SGKPASMKLGADRAKAFAAIRLHSAPGATLTEGRLTVPAYATVVLQ
```

(b)

```
          Region I                                                                    Region V
BceOGL    DLVVNHTSDEHNWFIESRKSKDN-KYRDYYIWRPGKEGKEPNNWGAAFSGSAWQYDEMT---DEYYLHLFSKKQPDLNWDNEKVR 178
XcaAMY    DLVINHTSDQHPWFKAALDPKD--AHRSWYTWA--GPGTNLKAV-SAVGGPAWHANGK-----QHYLGDFTGAMPDLNYDEPAVR 176
BmeAMY    DLVVNHTSSEHPWFQAALKDKNS-KYRDYYIWA--DKNTDLNEK-GSWGQQVWHKAPNG----EYFYGTFWEGMPDLNYDNPEVR 185
DthAMY    DLVVNHTSSRHPWFVSSASSYNS-PYRDYYIWS--TEKPEKNS-------NLWYKKPT-----GYYYALFWSEMPDLNFDNPKVR 192
TmaAMY    DLVINHTSDEHPWFKDAVENTTSSPYWDYYIMS--LEDHSGQD--------HWHWKINSKGQKVWYFGLFGYNMPDLNHDSQKVR 197
          *** ****  * **              *                  *                 *    **** *   **

                              Region II                                                   Region III
BceOGL    QDVYEMMKFWLEKGIDGFRMDVINFISKEEGLPTVETEEEGYVSGHKHFMNGPNIHKYLHEMNEEVLSHYDIMTVGEMPGVTTEE 263
XcaAMY    REMIAVGKFWLDKGADGFRLDAARHIYDDLESDNGQPAVI-----ARNAQWWNEFRQGLRQVRP------DVYLVGEVSAKQPGE 250
BmeAMY    KEMINVGKFWLKQGVDGFRLDAALHIFKGQTPEGA----------KKNILWWNEFRDAMKKENP------NVYLTGEVWD-QPEV 253
DthAMY    EEVKKIAKFWIEKGVDGFRLDAAKHIYDDDS--------------KNIQWWKEFYSYLKSIKP------DVYLVGEVWD-NEYK 255
TmaAMY    EEVKKIVDFWISKGVDGFRIDAAKHIYGWSWDDGI----------QESAEYFEWFRDYVLSKKP------DAILVGEVFSGNTYD 266
                    **    * **** *    *                                                  **

                                                                                Region IV
BceOGL    AKLYTGEERKELQMVFQFEHMDLDSGEGGKWDVKPCSLLTLKENLTKWQKALEHTGWNSLYWNNHD 329
XcaAMY    LAPYLPA----LGSVFDFPLAEQLIASAGQEKAGKLPALLTETYAAFRAAAG-DDYADAPFLSNHD 311
BmeAMY    VAPYYQS----LDSLFNFDLAGKIVSSVKAGNDQGIATAAAATDELFKSYN--PNKIDGIFLTNHD 314
DthAMY    IAEYYKG----LPSNFNFPLSDKIMNSSSKSK-RLRNYRISRLKRLFGENN--TDFADAIFLRNHD 314
TmaAMY    LSLYP-------IPVFNFALMYSIRNYPEGQD-------------GMIENN--WVEESFLFLENHD 310
             *           * *                                    ***
```

Fig. 2. Identifying the β5-strand catalytic glutamate (Region III) where there might be an ambiguity. (**a**) The amino-acid sequence of the periplasmic α-amylase from *Xanthomonas campestris* (Abe et al., 1996) with highlighted conserved sequence regions in yellow except for the region III containing the catalytic Glu. The segment 225_EFRQ (highlighted in pink) was proposed (Abe et al., 1996) as the conserved sequence region III with the β5-strand catalytic glutamate. The start of the mature enzyme is indicated by an arrow. (**b**) Based on the alignment with closely related sequences and, more importantly, taking into account also the characteristic sequences in the other conserved regions, especially in this particular case in the fifth conserved sequence region (165_MPDLN, Region V), the catalytic glutamate is proposed rather to be the Glu242 (238_YLVGEVSA, Region III). The glutamates corresponding to Glu225 of *X. campestris* α-amylase in the three other α-amylases are also highlighted in pink. Note that there is no glutamate in the *Bacillus cereus* oligo-1,6-glucosidase corresponding to Glu225 from *X. campestris* enzyme. Sources of the enzymes: BceOGL: *Bacillus cereus* oligo-1,6-glucosidase (Watanabe et al., 1990); XcaAMY: *Xanthomonas campestris* α-amylase (Abe et al., 1996); BmeAMY: *Bacillus megaterium* α-amylase (Metz et al., 1988); DthAMY: *Dictyoglomus thermophilum* AmyC α-amylase (Horinouchi et al., 1988); TmaAMY: *Thermotoga maritima* α-amylase (Liebl et al., 1997).

gor, 1993; Svensson, 1994; Janeček, 1997a; Kuriki & Imanaka, 1999). Nevertheless, as more sequence data became available, it was recognised that certain amino-acid residues closely preceding and/or succeeding the four conserved sequence regions may also be significant with regard to the enzyme specificity (e.g., Janeček et al., 1995) – in other words the longer conserved sequence regions would be more meaningful (MacGregor et al., 2001). This might be especially true for the conserved sequence region III (Fig. 1) that was established as a stretch of 4 amino acids (230_EVLD in Taka-amylase A; Nakajima et al., 1986) and that, if considered as a 4-residue-long region only, can easily be misidentified in a sequence (Fig. 2a). It is worth mentioning that this region is, in fact, not conserved at all despite the fact that it contains the invariant β5-strand catalytic glutamate (cf. Fig. 1). Its identification in a new sequence should thus be carried out with special care in conjunction with the other conserved sequence regions of various enzyme specificities, especially if there

is limited experimental support (Fig. 2b). ABE et al. (1999) solved the problem by site-directed mutagenesis on isoamylase, which identified the correct glutamic acid. A similar strategy was applied by, e.g., DEVULAPALLE et al. (1997) to glucosyltransferase, SARCABAL et al. (2000) to amylosucrase and NAKAYAMA et al. (2001) to glycogen debranching enzyme.

Importance of the additional conserved sequence regions

The fifth conserved sequence region – in domain B
The fifth conserved sequence region (173_LPDLD in Taka-amylase A; Fig. 1) was proposed first for α-amylases (JANEČEK, 1992) and later was identified also in the other members of the α-amylase family (JANEČEK, 1995a). It is positioned near the C-terminus of domain B around the aspartate involved in binding the calcium ion, Asp175 in Taka-amylase A (e.g. MATSUURA et al., 1984; QIAN et al., 1993; KADZIOLA et al., 1994; MACHIUS et al., 1995). In most cases there are 28–30 residues between this Ca^{2+}-binding aspartate and the β4-strand catalytic aspartate, i.e. Asp206 in Taka-amylase A (JANEČEK, 1992, 1995a).

Although this region seems to be not fully conserved or not easily identifiable in some family members (JANEČEK, 1997a), in several cases this stretch is very characteristic for a special enzyme specificity (Fig. 3). The most striking example is the oligo-1,6-glucosidase with the template sequence QPDLN. This region in the form of QxDLN (x = P, A, I, W) is found also in α-glucosidase, dextran-glucosidase, trehalose-6-phosphate hydrolase, amylosucrase, sucrose phosphorylase, isomaltulose synthase and trehalose synthase (e.g., JANEČEK, 1997a, 2000a). There is a characteristic change from QPDLN (the oligo-1,6-glucosidase group) to MPKLN (the neopullulanase group) via the MPDLN belonging to a group of odd members of the α-amylase family (Fig. 3). It has already been pointed out (JANEČEK et al., 1997) that structurally the domain B of these members should resemble very closely the domain B of oligo-1,6-glucosidase from *Bacillus cereus* (KIZAKI et al., 1993). Remarkably, the mammalian proteins associated with transport of dibasic and neutral amino acids across cell membranes also unambiguously contain this conserved sequence region (Fig. 3) as well as an entire segment that corresponds perfectly (almost 50% identity) with the whole domain B of the enzymes from the α-amylase family (JANEČEK, 2000b). The sequence QxDLN as the fifth conserved sequence region could thus be used as a fingerprint that defines the proposed "oligo-1,6-glucosidase group" within the frame of the α-amylase family (OSLANCOVÁ & JANEČEK, 2002).

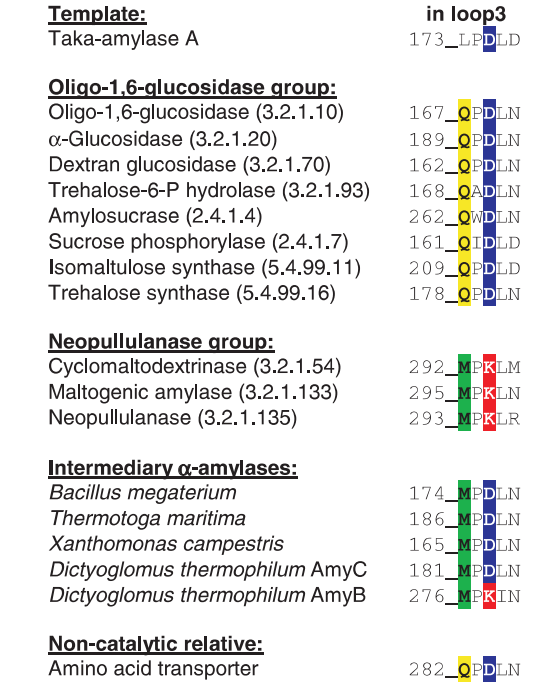

Fig. 3. The fifth conserved sequence region of the α-amylase family enzymes. The figure focuses on the so-called oligo-1,6-glucosidase and neopullulanase groups of the family. Sources of the enzymes: Taka-amylase A (TODA et al., 1982); oligo-1,6-glucosidase (WATANABE et al., 1990); α-glucosidase (JANDA et al., 2000); dextran glucosidase (RUSSELL & FERRETTI, 1990); trehalose-6-phosphate hydrolase (RIMMELE & BOOS, 1994); amylosucrase (POTOCKI DE MONTALK et al., 1999); sucrose phosphorylase (FERRETTI et al., 1988); sucrose isomerase (ZHANG et al., 2002); trehalose synthase (TSUSAKI et al., 1996); cyclomaltodextrinase (PODKOVYROV & ZEIKUS, 1992); maltogenic amylase (KIM et al., 1999b); neopullulanase (TONOZUKA et al., 1993); *Bacillus megaterium* (METZ et al., 1988); *Thermotoga maritima* (LIEBL et al., 1997); *Xanthomonas campestris* (ABE et al., 1996); *Dictyoglomus thermophilum* AmyC (HORINOUCHI et al., 1988); *Dictyoglomus thermophilum* AmyB (HORINOUCHI et al., 1988). For illustration also the corresponding segment is shown of the human protein responsible for the transport of dibasic and neutral amino acids across cell membranes (amino-acid transporter; BERTRAN et al., 1993) containing the oligo-1,6-glucosidase-type of domain B (JANEČEK et al., 1997). Colour code: Ca^{2+}-binding aspartate – blue; lysine playing the role of a calcium – red; oligo-1,6-glucosidase-like glutamine – yellow; neopullulanase-like methionine – green.

The members of the family with characteristic sequences in this region MPKLN (or MPKIN and MPKLR; Figs 1,3), such as cyclomaltodextrinases, maltogenic amylases and neopullulanases, have a shorter domain B than the majority and presumably do not bind a calcium ion due to the Asp→Lys substitution (JANEČEK et al., 1997). Enzymes with these specificities could also form a group of closely-related members of the α-amylase family (MATZKE et al., 2000; PARK et al., 2000; MACGREGOR et al., 2001; OH et al., 2001), the neopullulanase group (OSLANCOVÁ & JANEČEK, 2002) (Fig. 3). Indeed, it has recently been shown that the longer side-chain of lysine, corresponding in neopullulanase to the Ca^{2+}-binding aspartate of Taka-amylase A, may play the stabilising role of a Ca^{2+} ion by occupying the equivalent position and fixing the relative orientation and location of the catalytic $(\beta/\alpha)_8$-barrel and domain B (KAMITORI et al., 1999). In isoamylase, for instance, this region is 342_GANFN (Fig. 1) with the side-chain of Asn344 occupying the position of the Ca^{2+} ion conserved in other members of the α-amylase family, while isoamylase does not contain this Ca^{2+}, but has Ca^{2+} at a different position (KATSUYA et al., 1998).

Finally, the members of the α-amylase family containing the intermediary sequence MPDLN, especially some particular α-amylases (Fig. 3), may exhibit mixed substrate specificity of α-amylases, cyclomaltodextrinases, and neopullulanases (BRUMM et al. 1991; ABE et al., 1996; JANEČEK, 1997a). In this respect, the two α-amylases from *Dictyoglomus thermophilum*, designated as AmyB and AmyC (HORINOUCHI et al. 1988), containing MPKIN and MPDLN, respectively (Fig. 3), should be of great interest (JANEČEK, 2000a).

The sixth conserved sequence region – strand β2

As far as the conserved sequence region covering strand β2 of the catalytic $(\beta/\alpha)_8$-barrel is concerned (Fig. 4), it has been shown to be very helpful in discriminating, for example, the CGTases from α-amylases (JANEČEK, 1994a, JANEČEK et al., 1995). It was defined first for α-amylases (JANEČEK, 1994b) and the efforts were made to show that this β2-strand stretch, flanked in loops by a conserved glycine and proline, could also be evolutionarily important for the $(\beta/\alpha)_8$-barrel fold (JANEČEK, 1995c, 1996a,b, 1997b). The conserved sequence region at strand β2 is thus characterised by a glycine residue followed by seven or eight amino-acid residues and then a conserved proline, although in several cases the Gly and Pro are replaced by other residues (not included in Fig. 4). In order to distinguish CGTases from α-amylases the length of this region has to be taken into account, together with special sequence features (Fig. 4). The CGTases usually have eight residues between the Gly and Pro, while the α-amylases have seven (JANEČEK et al., 1995). This is, however, not always the case, and several exo-α-amylases, such as maltotriohydrolases and maltotetraohydrolases, also have eight residues between Gly and Pro (KOBAYASHI et al., 1994; FUJITA et al., 1989), while the two known archaeal CGTases appear to have seven (YAMAMOTO et al., 2000; RASHID et al., 2002).

In a wider sense this conserved sequence region may be of help in distinguishing between the true α-amylases, true CGTases, and oligosaccharide-producing amylases (JANEČEK, 1995b), such as maltotriohydrolase (Fig. 4). This is possible due to the presence of a glutamine residue preceding the invariant proline at the C-terminus of this region, a feature highly characteristic of CGTases (JANEČEK & TÓTH, 1994; JANEČEK et al., 1995). The α-amylases have a gap in the position corresponding to that glutamine (Gln78 in both *Bacillus circulans* CGTases) (Fig. 4). The other feature that seems to be characteristic of CGTases is the tryptophan (Trp75 in both *Bacillus circulans* CGTases) but it is found also in certain α-amylases. The residue in that position in many α-amylases is be glutamine (Fig. 4), especially in the so-called animal group and/or animal-like α-amylases (JANEČEK, 1994b; DA LAGE et al., 1998; SUMITANI et al., 1998; D'AMICO et al., 2000; PUJADAS & PALAU, 2001). The two archaeal *Thermococcus* CGTases with shorter β2-strand regions (YAMAMOTO et al., 2000; RASHID et al., 2002) seem to contain more α-amylase-like sequence features also in other conserved sequence regions (Š. JANEČEK, unpublished results). Remarkably, the recently-sequenced acarviose transferase with specificity very close to a CGTase (HEMKER et al., 2001) has also the shorter version of this region and lacks the glutamine characteristic of CGTase. The predictions concerning the exact enzyme specificity should thus be made carefully in conjunction with experimental biochemical data (WIND et al., 1995; DEL-RIO et al., 1997).

The seventh conserved sequence region – strand β8

This conserved sequence region (Fig. 5) was identified together with the one covering strand β2 (JANEČEK, 1994b). Although it does not contain invariant residues conserved throughout all the α-amylase family members (JANEČEK, 1997a;

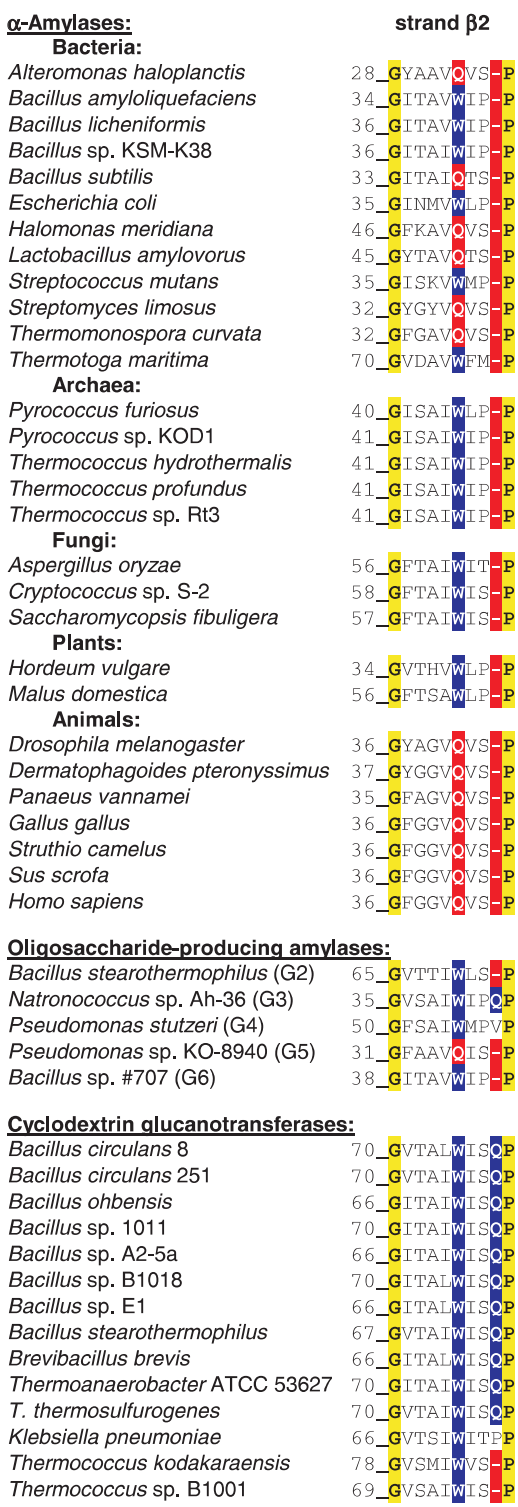

Fig. 4. The sixth conserved sequence region of the α-amylase family enzymes. The figure focuses on the α-amylases and cyclodextrin glucanotransferases from the family. Sources of the enzymes: α-amylases: *Alteromonas haloplanctis* (Feller et al., 1992), *Bacillus amyloliquefaciens* (Takkinen et al., 1983), *Bacillus licheniformis* (Yuuki et al., 1985), *Bacillus* sp. KSM-K38 (Hagihara et al., 2001), *Bacillus subtilis* (Yang et al., 1983), *Escherichia coli* (Raha et al., 1992), *Halomonas meridiana* (Coronado et al., 2000), *Lactobacillus amylovorus* (Giraud & Cuny, 1997), *Streptococcus mutans* (Simpson & Russell, 1998), *Streptomyces limosus* (Long et al., 1987), *Thermomonospora curvata* (Petříček et al., 1992), *Thermotoga maritima* (Liebl et al., 1997), *Pyrococcus furiosus* (Dong et al., 1997a), *Pyrococcus* sp. KOD1 (Tachibana et al., 1996), *Thermococcus hydrothermalis* (Lévêque et al., 2000), *Thermococcus profundus* (Lee et al., 1996), *Thermococcus* sp. Rt3 (Jones et al., 1999), *Aspergillus oryzae* (Toda et al., 1982), *Cryptococcus* sp. S-2 (Iefuji et al., 1996), *Saccharomycopsis fibuligera* (Itoh et al., 1987), *Hordeum vulgare* (Rogers & Milliman, 1983), *Malus domestica* (Wegrzyn et al., 2000), *Drosophila melanogaster* (Boer & Hickey, 1986), *Dermatophagoides pteronyssimus* (Mills et al., 1999), *Panaeus vannamei* (Van Wormhoudt & Sellos, 1996), *Gallus gallus* (Benkel et al., 1997), *Struthio camelus* (Kabuto et al., 2000), *Sus scrofa* (Pasero et al., 1986), *Homo sapiens* (Nishide et al., 1986); oligosaccharide-producing amylases: *Bacillus stearothermophilus* (G2) (Diderichsen & Christiansen, 1988), *Natronococcus* sp. Ah-36 (G3) (Kobayashi et al., 1994), *Pseudomonas stutzeri* (G4) (Fujita et al., 1989), *Pseudomonas* sp. KO-8940 (G5) (Shida et al., 1992), *Bacillus* sp. #707 (G6) (Tsukamoto et al., 1988); CGTases: *Bacillus circulans* 8 (Nitschke et al., 1990), *Bacillus circulans* 251 (Lawson et al., 1994), *Bacillus ohbensis* (Sin et al., 1991), *Bacillus* sp. 1011 (Kimura et al., 1987), *Bacillus* sp. A2-5a (Ohdan et al., 2000), *Bacillus* sp. B1018 (Itkor et al., 1990), *Bacillus* sp. E1 (Yong et al., 1996), *Bacillus stearothermophilus* (Fujiwara et al., 1992), *Brevibacillus brevis* (Kim et al., 1998), *Thermoanaerobacter* sp. ATCC 53627 (Jørgensen et al., 1997), *Thermoanaerobacterium thermosulfurigenes* (Bahl et al., 1991), *Klebsiella pneumoniae* (Binder et al., 1986), *Thermococcus kodakaraensis* (Rashid et al., 2002), *Thermococcus* sp. B1001 (Yamamoto et al., 2000). Colour code: conserved glycine and proline at the C- and N-terminal sides of the region – yellow; α-amylase-like features (glutamine and a gap in the positions $i-4$ and $i-1$, respectively, in respect to the conserved proline) – red; CGTase-like features (tryptophane and glutamine in the positions $i-4$ and $i-1$, respectively, with respect to the conserved proline) – blue.

2000a), it usually starts with a very well-conserved glycine (Gly323 in Taka-amylase A) followed by a proline in the $i+2$ position (Fig. 1).

For some specificities it is not easy to identify this region without information derived from the three-dimensional structure (cf. JANEČEK, 2000a; PRZYLAS et al., 2000).

Nevertheless, it has already been demonstrated that this region may contain the features characteristic of archaeal α-amylases (JANEČEK et al., 1999). Another example is shown in Figure 5 highlighting the features typical for either the "oligo-1,6-glucosidase" or the "neopullulanase" groups that were also discussed above with regard to the fifth conserved sequence region (Fig. 3). These features are mainly the tyrosine and aspartate (Tyr454 and Asp456 in neopullulanase in Fig. 5) at the C-terminal side of the region. They could be taken as characteristic residues of the neopullulanase group, while the oligo-1,6-glucosidase-like enzymes should contain glutamine and glutamate, respectively, in the corresponding positions (Fig. 5). Concerning the β8-strand region, amylosucrase, sucrose phosphorylase, and trehalose synthase seem to be rather more like neopullulanase members, although they seem to be clear oligo-1,6-glucosidase-like enzymes with regard to the fifth conserved sequence region (cf. Figs 3,5). This is not surprising at all if the parts of their sequences between strands β7 and β8 are taken into account; these are different from, and longer than, the equivalent segments of oligo-1,6-glucosidase, α-glucosidase, dextran glucosidase and trehalose-6-phosphate hydrolase (OSLANCOVÁ & JANEČEK, 2002). This observation may also be related to the fact that the former members are transferases while the later ones are hydrolases (MACGREGOR et al., 2001). Finally one can imagine that many possible combinations of the various characteristic motifs will actually be possible and may occur in nature.

The non-catalytic proteins, related to the α-amylase family from the structural and evolutionary points of view (JANEČEK et al., 1997; BRÖER & WAGNER, 2002), have the segment covering strand β8 of the enzymatic members of the family perfectly conserved in the corresponding positions (Fig. 5). This fact should be of interest because three of the four best-conserved sequence regions with catalytic residues (strands β4, β5 and β7) are less well conserved (JANEČEK, 2000b).

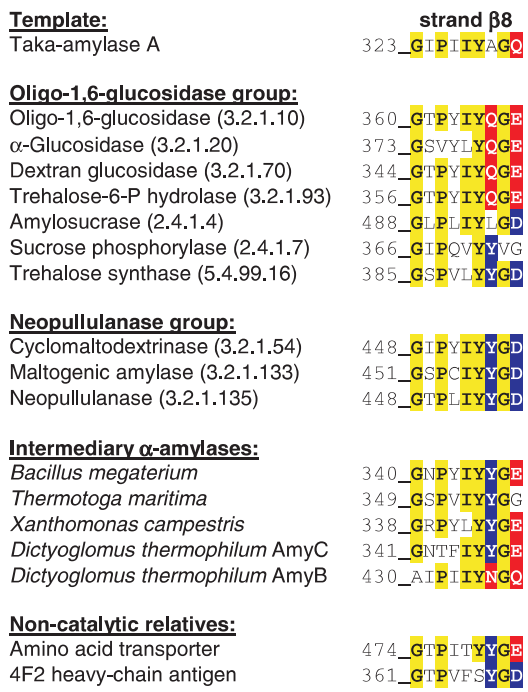

Fig. 5. The seventh conserved sequence region of the α-amylase family enzymes. The figure focuses on the so-called oligo-1,6-glucosidase and neopullulanase groups of the family. Sources of the enzymes are the same as in Figure 3, the one for the 4F2 heavy-chain antigen (GOTTESDIENER et al., 1988) being added. Colour code: conserved residues – yellow; oligo-1,6-glucosidase-like glutamine and glutamate – red; neopullulanase-like tyrosine and aspartate – blue.

Conclusion

According to our present knowledge, the α-amylase family consisting of families 13, 70 and 77 of glycoside hydrolases (i.e. clan GH-H) covers hydrolases, transferases, and isomerases (i.e. the enzyme classes 3, 2 and 5). The amino acid transporters and 4F2 heavy-chain antigens (i.e. non-enzymes) are also recognised as relatives of the family. It is possible that the main α-amylase family (clan GH-H) is related to the amylolytic family 57 of glycoside hydrolases, although the hierarchy will very probably be at a higher level (i.e. less strict common criteria) than a clan in terms of its definition. The members of the α-amylase enzyme family act mainly on α-1,4- and α-1,6-glucosidic linkages; however, they also operate on α-1,1-, α-1,2-, α-1,3- and α-1,5-glucosidic bonds. Despite the presence of several conserved sequence regions the invariant amino-acid residues are extremely few, i.e. three catalytic residues in strands β4, β5 and β7 plus the arginine in position $i-2$ with regard to the β4-strand catalytic aspartate.

Concerning the number of conserved sequence regions it could be postulated that: (i) each

enzyme specificity and/or even a distinct taxonomic group from the α-amylase family may have its own conserved sequence regions and/or characteristic sequence similarities located in the catalytic $(\beta/\alpha)_8$-barrel domain and/or outside this domain; (ii) these regions may discriminate between the α-amylase family members from the point of view of their specificity, structure and evolution; and (iii) all conserved sequence regions may not be common for the entire α-amylase family and they may indicate closer evolutionary relatedness with functionally different enzymes and proteins.

If conserved sequence regions are considered, it can be concluded that the α-amylase family enzymes should be characterised by 4, 5 and in some cases 7 conserved sequence regions, i.e. by as many as possible. So if we go back to the question raised in the title of this review: "How many conserved sequence regions are there in the α-amylase family?", the exact answer is less important than the practical use of the information characteristic of specific regions.

Acknowledgements

This work was supported in part by the VEGA Grant No. 2/2057/22 from the Slovak Grant Agency.

References

Abe, J. I., Shibata, Y., Fujisue, M. & Hizukuri, S. 1996. Microbiology **142:** 1505–1512.

Abe, J. I., Ushijima, C. & Hizukuri, S. 1999. Appl. Environ. Microbiol. **65:** 4163–4170.

Amemura, A., Chakraborty, R., Fujita, M., Noumi, T. & Futai, M. 1988. J. Biol. Chem. **263:** 9271–9275.

Bahl, H., Burchhardt, G., Spreinat, A., Haeckel, K., Wienecke, A., Schmidt, B. & Antranikian, G. 1991. Appl. Environ. Microbiol. **57:** 1554–1559.

Beier, L., Svendsen, A., Andersen, C., Frandsen, T. P., Borchert, T. V. & Cherry, J. R. 2000. Protein Eng. **13:** 509–513.

Benkel, B. F., Nguyen, T., Ahluwalia, N., Benkel, I. & Hickey, D. A. 1997. Gene **192:** 261–270.

Bertran, J., Werner, A., Chillaron, J., Nunes, V., Biber, J., Testar, X., Zorzano, A., Estivill, X., Murer, H. & Palacin, M. 1993. J. Biol. Chem. **268:** 14842–14849.

Binder, F., Huber, O. & Böck, A. 1986. Gene **47:** 269–277.

Binderup, K. & Preiss, J. 1998. Biochemistry **37:** 9033–9037.

Boer, P. H. & Hickey, D. A. 1986. Nucleic Acids Res. **14:** 8399–8411.

Bourne, Y. & Henrissat, B. 2001. Curr. Opin. Struct. Biol. **11:** 593–600.

Bröer, S. & Wagner, C. A. 2002. Cell Biochem. Biophys. **36:** 155–168.

Brumm, P. J., Hebeda, R. E. & Teague, W. M. 1991. Starch **43:** 315–319.

Candussio, A., Schmid, G. & Bock, A. 1990. Eur. J. Biochem. **191:** 177–185.

Coronado, M. J., Vargas, C., Mellado, E., Tegos, G., Drainas, C., Nieto, J. J. & Ventosa, A. 2000. Microbiology **146:** 861–868.

Coutinho, P. M. & Henrissat, B. 1999. Carbohydrate-Active Enzymes server at URL: http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html; accessed March 22, 2002.

Da Lage, J. L., Renard, E., Chartois, F., Lemeunier, F. & Cariou, M. L. 1998. Proc. Natl. Acad. Sci. USA **95:** 6848–6853.

D'Amico, S., Gerday, C. & Feller, G. 2000. Gene **253:** 95–105.

Dauter, Z., Dauter, M., Brzozowski, A. M., Christensen, S., Borchert, T. V., Beier, L., Wilson, K. S. & Davies, G. J. 1999. Biochemistry **38:** 8385–8392.

del-Rio, G., Morett, E. & Soberon, X. 1997. FEBS Lett. **416:** 221–224.

Devulapalle, K. S., Goodman, S. D., Gao, Q. Hemsley, A. & Mooser, G. 1997. Protein Sci. **6:** 2489–2493.

Diderichsen, B. & Christiansen, L. 1988. FEMS Microbiol. Lett. **56:** 53–60.

Dong, G., Vieille, C., Savchenko, A. & Zeikus, J. G. 1997a. Appl. Environ. Microbiol. **63:** 3569–3576.

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

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

Feller, G., Lonhienne, T., Deroanne, C., Libioulle, C., Van Beeumen, J. & Gerday, C. 1992. J. Biol. Chem. **267:** 5217–5221.

Ferretti, J. J., Huang, T. T. & Russell, R. R. B. 1988. Infect. Immun. **56:** 1585–1588.

Friedberg, F. 1983. FEBS Lett. **152:** 139–140.

Frillingos, S., Linden, A., Niehaus, F., Vargas, C., Nieto, J. J., Ventosa, A., Antranikian, G. & Drainas, C. 2000. J. Appl. Microbiol. **88:** 495–503.

Fujita, M., Torigoe, K., Nakada, T., Tsusaki, K., Kubota, M., Sakai, S. & Tsujisaka, Y. 1989. J. Bacteriol. **171:** 1333–1339.

Fujiwara, S., Kakihara, H., Woo, K. B., Lejeune, A., Kanemoto, M., Sakaguchi, K. & Imanaka, T. 1992. Appl. Environ. Microbiol. **58:** 4016–4025.

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

Giraud, E. & Cuny, G. 1997. Gene **198:** 149–157.

Gottesdiener, K. M., Karpinski, B. A., Lindsten, T., Strominger, J. L., Jones, N. H.,

Thompson, C. B. & Leiden, J. M. 1988. Mol. Cell. Biol. **8:** 3809–3819.
Gottschalk, T. E., Tull, D., Aghajari, N., Haser, R. & Svensson, B. 2001. Biochemistry **40:** 12844–12854
Hagihara, H., Hazashi, Y., Endo, K., Igarashi, K., Ozawa, T., Kawai, S., Ozaki, K. & Ito, S. 2001. Eur. J. Biochem. **268:** 3974–3982.
Hemker, M., Stratmann, A., Goeke, K., Schröder, W., Lenz, J., Piepersberg, W. & Pape, H. 2001. J. Bacteriol. **183:** 4484–4492.
Henrissat, B. 1991. Biochem. J. **280:** 309–316.
Henrissat, B. & Bairoch, A. 1996. Biochem. J. **316:** 695–696.
Hong, S. H. & Marmur, J. 1986. Gene **41:** 75–84.
Horinouchi, S., Fukusumi, S., Ohshima, T. & Beppu, T. 1988. Eur. J. Biochem. **176:** 243–253.
Ibuka, A., Tonozuka, T., Matsuzawa, H. & Sakai, H. 1998. J. Biochem. **123:** 275–282.
Iefuji, H., Chino, M., Kato, M. & Iimura, Y. 1996. Biochem. J. **318:** 989–996.
Imamura, H., Fushinobu, S., Jeon, B. S., Wakagi, T. & Matsuzawa, H. 2001a. Biochemistry **40:** 12400–12406.
Imamura, H., Fushinobu, S., Yamamoto, M., Kumasaka, T., Wakagi, T. & Matsuzawa, H. 2001b. J. Appl. Glycosci. **48:** 171–175.
Itkor, P., Tsukagoshi, N. & Udaka, S. 1990. Biochem. Biophys. Res. Commun. **166:** 630–636.
Itoh, T., Yamashita, Y. & Fukui, S. 1987. FEBS Lett. **219:** 339–342.
Janda, L., Damborský, J., Petříček, M., Spížek, J. & Tichý, P. 2000. J. Appl. Microbiol. **88:** 773–783.
Janeček, Š. 1992. Biochem. J. **288:** 1069–1070.
Janeček, Š. 1994a. FEBS Lett. **353:** 119–123.
Janeček, Š. 1994b. Eur. J. Biochem. **224:** 519–524.
Janeček, Š. 1995a. FEBS Lett. **377:** 6–8.
Janeček, Š. 1995b. Biologia, Bratislava **50:** 515–522.
Janeček, Š. 1995c. Protein Sci. **4:** 1239–1242.
Janeček, Š. 1996a. Protein Sci. **5:** 1136–1143.
Janeček, Š. 1996b. Biochem. J. **319:** 1005–1006.
Janeček, Š. 1997a. Prog. Biophys. Mol. Biol. **67:** 67–97.
Janeček, Š. 1997b. Int. J. Biol. Macromol. **21:** 277–280.
Janeček, Š. 1998. Folia Microbiol. **43:** 123–128.
Janeček, Š. 2000a. In: Ohnishi, M., Hayashi, T., Ishijima, S. & Kuriki, T. (eds) Glycoenzymes, Japan Academic Societies Center, Tokyo, pp. 19–54.
Janeček, Š. 2000b. Trends Glycosci. Glycotechnol. **12:** 363–371.
Janeček, Š., Lévêque, E., Belarbi, A. & Haye, B. 1999. J. Mol. Evol. **48:** 421–426.
Janeček, Š., MacGregor, E. A. & Svensson, B. 1995. Biochem. J. **305:** 685–686.
Janeček, Š., Svensson, B. & Henrissat, B. 1997. J. Mol. Evol. **45:** 322–331.
Janeček, Š. & Tóth, D. 1994. Biologia, Bratislava **49:** 301–306.
Janse, B. J. H., Steyn, A. J. C. & Pretorius, I. S. 1993. Curr. Genet. **24:** 400–407.
Jeon, B. S., Taguchi, H., Sakai, H., Ohshima, T., Wakagi, T. & Matsuzawa,H. 1997. Eur. J. Biochem. **248:** 171–178.
Jespersen, H. M., MacGregor, E. A., Henrissat, B., Sierks, M. R. & Svensson, B. 1993. J. Protein Chem. **12:** 791–805.
Jespersen, H. M., MacGregor, E. A., Sierks, M. R. & Svensson, B. 1991. Biochem. J. **80:** 51–55.
Jones, R. A., Jermiin, L. S., Easteal, S., Patel, B. K. & Beachman, I. R. 1999. J. Appl. Microbiol. **86:** 93–107.
Jørgensen, S. T., Tangney, M., Starnes, R. L., Amemiya, K. & Jørgensen, P. L. 1997. Biotechnol. Lett. **19:** 1027–1031.
Kabuto, S., Ogawa, T., Muramoto, K., Oosthuizen ,V. & Naude, R. J. 2000. Comp. Biochem. Physiol. **127B:** 481–490.
Kadziola, A., Abe, J. I., Svensson, B. & Haser, R. 1994. J. Mol. Biol. **239:** 104–121.
Kamitori, S., Kondo, S., Okuyama, K., Yokota, T., Shimura, Y., Tonozuka, T. & Sakano, Y. 1999. J. Mol. Biol. 287: 907–921.
Katsuya, Y., Mezaki, Y., Kubota, M. & Matsuura, Y. 1998. J. Mol. Biol. 281: 885–897.
Kim, J. S., Cha, S. S., Kim, H. J., Kim, T. J., Ha, N. C., Oh, S. T., Cho, H. S., Cho, M. J., Kim, M. J., Lee, H. S., Kim, J. W., Choi, K. Y., Park, K. H. & Oh, B. H. 1999a. J. Biol. Chem. **274:** 26279–26286.
Kim, M. H., Sohn, C. B. & Oh, T. K. 1998. FEMS Microbiol. Lett. **164:** 411–418.
Kim, T. J., Kim, M. J., Kim, B. C., Kim, J. C., Cheong, T. K., Kim, J. W. & Park, K. H. 1999b. Appl. Environ. Microbiol. **65:** 1644–1951.
Kim, T. J., Nguyen, V. D., Lee, H. S., Kim, M. J., Cho, H. Y., Kim, Y. W., Moon, T. W., Park, C. S., Kim, J. W., Oh, B. H., Lee, S. B., Svensson, B. & Park, K. H. 2001. Biochemistry **40:** 14182–14190.
Kimura, K., Kataoka, S., Ishii, Y., Takano, T. & Yamane, K. 1987. J. Bacteriol. **169:** 4399–4402.
Kizaki, H., Hata, Y., Watanabe, K., Katsube, Y. & Suzuki, Y. 1993. J. Biochem. 113: 646–649.
Klein, C. & Schulz, G. E. 1991. J. Mol. Biol. **217:** 737–750.
Kobayashi, K., Kato, M., Miura, Y., Kettoku, M., Komeda, T. & Iwamatsu, A. 1996. Biosci. Biotech. Biochem. **60:** 1882–1885.
Kobayashi, T., Kanai, H., Aono, R., Horokoshi, K. & Kudo, T. 1994. J. Bacteriol. **176:** 5131–5134.
Kuriki, T. 1992. Trends Glycosci. Glycotechnol. **4:** 567–572.
Kuriki, T. & Imanaka, T. 1989. J. Gen. Microbiol. **135:** 1521–1528.
Kuriki, T. & Imanaka, T. 1999. J. Biosci. Bioeng. **87:** 557–565.
Kuriki, T., Kaneko, H., Yanase, M., Takata, H., Shimada, J., Handa, S., Takada, T., Umeyama,

H. & Okada, S. 1996. J. Biol. Chem. **271:** 17321–17329.

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

Lawson, C. L., Van Montfort, R., Strokopytov, B., Rozeboom, H. J., Kalk, K. H., De Vries, G. E., Penninga, D., Dijkhuizen, L. & Dijkstra, B. W. 1994. J. Mol. Biol. **236:** 590–600.

Lee, J. T., Kanai, H., Kobayashi, T., Akiba, T. & Kudo, T. 1996. J. Ferment. Bioeng. **82:** 432–438.

Leemhuis, H., Dijkstra, B. W. & Dijkhuizen, L. 2002. FEBS Lett. **514:** 189–192.

Lévêque, E., Haye, B. & Belarbi, A. 2000. FEMS Microbiol. Lett. **186:** 67–71.

Liebl, W., Stemplinger, I. & Ruile, P. 1997. J. Bacteriol. **179:** 941–948.

Long, C. M., Virolle, M. J., Chang, S. Y., Chang, S. & Bibb, M. J. 1987. J. Bacteriol. **169:** 5745–5754.

MacGregor, E. A. 1988. J. Protein Chem. **7:** 399–415.

MacGregor, E. A. 1993. Starch **7:** 232–237.

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

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

MacGregor, E. A. & Svensson, B. 1989. Biochem. J. **259:** 145–152.

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

Matsui, I. & Svensson, B. 1997. J. Biol. Chem. **272:** 22456–22463.

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

Matzke, J., Herrmann, A., Schneider, E. & Bakker, E. P. 2000. FEMS Microbiol. Lett. **183:** 55–61.

Meissner, H. & Liebl, W. 1998. Eur. J. Biochem. **250:** 1050–1058.

Metz, R. J., Allen, L. N., Cao, T. M. & Zeman, N. W. 1988. Nucelic Acids Res. **16:** 5203.

Mills, K. L., Hart, B. J., Lynch, N. R., Thomas, W. R. & Smith, W. 1999. Int. Arch. Allergy Immunol. **120:** 100–107.

Monchois, V., Lakey, J. H. & Russell, R. R. B. 1999. FEMS Microbiol. Lett. **177:** 243–248.

Monchois, V., Vignon, M., Escalier, P. C., Svensson, B. & Russell, R. R. B. 2000. Eur. J. Biochem. **267:** 4127–4136.

Mori, H., Bak-Jensen, K. S., Gottschalk, T. E., Motawia, M. S., Damager, I., Moller, B. L. & Svensson, B. 2001. Eur. J. Biochem. **268:** 6545–6558.

Morishita, Y., Hasegawa, K., Matsuura, Y., Katsube, Y., Kubota, M. & Sakai, S. 1997. J. Mol. Biol. **267:** 661–672.

Nakajima, R., Imanaka, T. & Aiba, S. 1986. Appl. Microbiol. Biotechnol. **23:** 355–360.

Nakayama, A., Yamamoto, K. & Tabata, S. 2001. J. Biol. Chem. **276:** 28824–28828.

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

Nishide, T., Nakamura, Y., Emi, M., Yamamoto, T., Ogawa, M., Mori, T. & Matsubara, K. 1986. Gene **41:** 299–304.

Nitschke, L., Heeger, K., Bender, H. & Schulz, G. E. 1990. Appl. Microbiol. Biotechnol. **33:** 542–546.

Oh, B. H., Kim, J. S., Cho, H. S., Kim, T. J., Lee, H. S. & Park, K. H. 2001. In: The First Symposium on the Alpha-Amylase Family, Smolenice Castle, Slovakia, Sep 30 – Oct 4, 2001, p. 45.

Ohdan, K., Kuriki, T., Takata, H. & Okada, S. 2000. Appl. Microbiol. Biotechnol. **53:** 430–434.

Oslancová, A. & Janeček, Š. 2002. Cell. Mol. Life Sci. (in press).

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

Pasero, L., Mazzéi-Pierron, Y., Abadie, B., Chicheportiche, Y. & Marchis-Mouren, G. 1986. Biochim. Biophys. Acta **869:** 147–157.

Petříček, M., Tichý, P. & Kuncová, M. 1992. Gene **112:** 77–83.

Podkovyrov, S. M. & Zeikus, J. G. 1992. J. Bacteriol. **174:** 5400–5405.

Potocki De Montalk, G., Remaud-Simeon, M., Willemot, R.M., Planchot, V. & Monsan, P. 1999. J. Bacteriol. **181:** 375–381.

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.

Qian, M., Haser, R. & Payan, F. 1993. J. Mol. Biol. **231:** 785–799.

Raha, M., Kawagishi, I., Mueller, V., Kihara, M., Macnab, R. M. 1992. J. Bacteriol. **174:** 6644–6652.

Rashid, N., Cornista, J., Ezaki, S., Fukui, T., Atomi, H. & Imanaka, T. 2002. J. Bacteriol. **184:** 777–784.

Rimmele, M. & Boos, W. 1994. J. Bacteriol. **176:** 5654–5664.

Rogers, J. C. 1985. Biochem. Biophys. Res. Commun. **128:** 470–476.

Rogers, J. C. & Milliman, C. 1983. J. Biol. Chem. **258:** 8169–8174.

Romeo, T., Kumar, A. & Preiss, J. 1988. Gene **70:** 363–376.

Roujeinikova, A., Raasch, C., Burke, J., Baker, P. J., Liebl, W & Rice, D. W. 2001. J. Mol. Biol. **312:** 119–131.

Russell, R. R. B. & Ferretti, J. J. 1990. J. Gen. Microbiol. **136:** 803–810.

Saab-Rincón, G., del-Río, G., Santamaría, R. I., López-Munguía, A. & Soberón, X. 1999. FEBS Lett. **453:** 100–106.

Sarçabal, P., Remaud-Simeon, M., Willemot, R., Potocki de Montalk, G., Svensson, B. & Monsan, P. 2000. FEBS Lett. **474:** 33–37.

Shida, O., Takano, T., Takagi, H., Kadowaki, K. & Kobayashi, S. 1992. Biosci. Biotech. Biochem. **56:** 76–80.
Shirokizawa, O., Akiba, T. & Horikoshi, K. 1990. FEMS Microbiol. Lett. 70: 131–136.
Simpson, C. L. & Russell, R. R. B. 1998. J. Bacteriol. **180:** 4711–4717.
Sin, K. A., Nakamura, A., Kobayashi, K., Masaki, H. & Uozumi, T. 1991. Appl. Microbiol. Biotechnol. **35:** 600–605.
Skov, L. K., Mirza, O., Henriksen, A., De Montalk, G. P., Remaud-Simeon, M., Sarçabal, P., Willemot, R. M., Monsan, P. & Gajhede, M. 2001. J. Biol. Chem. **276:** 25273–25278.
Sumitani, J. I., Nagae, H., Kawaguchi, T. & Arai, M. 1998. J. Ferment. Bioeng. **85:** 428–432.
Svensson, B. 1988. FEBS Lett. 230: 72–76.
Svensson, B. 1994. Plant Mol. Biol. **25:** 141–157.
Svensson, B., Tovborg Jensen, M., Mori, H., Sass Bak-Jensen, K., Bønsager, B., Nielsen, P. K., Kramhøft, B., Prætorius-Ibba, M., Nøhr, J., Juge, N., Greffe, L., Williamson, G. & Driguez, H. 2002. Biologia, Bratislava **57 (Suppl. 11):** 5–19.
Tachibana, Y., Mendez Leclere, M., Fujiwara, S., Takagi, M. & Imanaka, T. 1996. J. Ferment. Bioeng. **82:** 224–232.
Takata, H., Kuriki, T., Okada, S., Takesada, Y., Iizuka, M., Minamiura, N. & Imanaka, T. 1992. J. Biol. Chem. **267:** 18447–18452.
Takkinen, K., Pettersson, R. F., Kalkkinen, N., Palva, I., Söderlund, H. & Kääriäinen, L. 1983. J. Biol. Chem. **258:** 1007–1013.
Terada, Y., Fujii, K., Takaha, T. & Okada, S. 1999. Appl. Environ. Microbiol. **65:** 910–915.
Toda, H., Kondo, K., & Narita, K. 1982. Proc. Japan Acad. **B58:** 208–212.
Tonozuka, T., Ohtsuka, M., Mogi, S. I., Sakai, H., Ohta, T. & Sakano, Y. 1993. Biosci. Biotech. Biochem. **57:** 395–401.
Tsai, Y.-W., Chia, J.-S., Shiau, Y.-Y., Chou, H.-C., Liaw, Y.-C., Lou, K.-L. 2000. FEMS Microbiol. Lett. **188:** 75–79.
Tsukamoto, A., Kimura, K., Ishii, Y., Takano, T. & Yamane, Y. 1988. Biochem. Biophys. Res. Commun. **151:** 25–31.
Tsusaki, K., Nishimoto, T., Nakada, T., Kubota, M., Chaen, H., Fukuda, S., Sugimoto, T. & Kurimoto, M. 1997. Biochim. Biophys. Acta **1334:** 28–32.
Tsusaki, K., Nishimoto, T., Nakada, T., Kubota, M., Chaen, H., Sugimoto, T. & Kurimoto, M. 1996. Biochim. Biophys. Acta **1290:** 1–3.
Uitdehaag, J. C. M., van der Veen, B. A., Dijkhuizen, L. & Dijkstra, B. W. 2002. Enzyme Microb. Technol. **30:** 295–304.
van der Maarel, M. J., van der Veen, B., Uitdehaag, J. C., Leemhuis, H. & Dijkhuizen, L. 2002. J. Biotechnol. **94:** 137–155.
van der Veen, B. A., Uitdehaag, J. C., Dijkstra, B. W. & Dijkhuizen, L. 2000. Biochim. Biophys. Acta **1543:** 336–360.
van Lieshout, J. F. T., Verhees, C. H., van der Oost, J. & de Vos, W. M. 2001. In: The First Symposium on the Alpha-Amylase Family, Smolenice Castle, Slovakia, Sep 30 – Oct 4, 2001, p. 26.
Van Wormhoudt, A., & Sellos, D. 1996. J. Mol. Evol. **42:** 543–551.
Vihinen, M. & Mäntsälä, P. 1989. Crit. Rev. Biochem. Mol. Biol. **24:** 329–418.
Watanabe, K., Kitamura, K., Iha, H. & Suzuki, Y. 1990. Eur. J. Biochem. **192:** 609–620.
Wegrzyn, T., Reilly, K., Cipriani, G., Murphy, P., Newcomb, R., Gardner, R. & MacRae, E. 2000. Eur. J. Biochem. **267:** 1313–1322.
Wind, R. D., Buitelaar, R. M. & Dijkhuizen, L. 1998. Eur. J. Biochem. **253:** 598–605.
Wind, R, D., Liebl, W., Buitelaar, R. M., Penninga, D., Spreinat, A., Dijkhuizen, L., Bahl, H. 1995. Appl. Environ. Microbiol. **61:** 1257–1265.
Yamamoto, T., Fujiwara, S., Tachibana, Y., Takagi, M., Fukui, K. & Imanaka, T. 2000. J. Biosci. Bioeng. **89:** 206–209.
Yang, M., Galizzi, A. & Henner, D. 1983. Nucl. Acids Res. **11:** 237–249.
Yong, J., Choi, J. N., Park, S. S., Park, C. S., Park, K. H. & Choi, Y. D. 1996. Biotechnol. Lett. **18:** 1223–1228.
Yuuki, T., Nomura, T., Tezuka, H., Tsuboi, A., Yamagata, H., Tsukagoshi, N. & Udaka, S. 1985. J. Biochem. **98:** 1147–1156.
Zhang, D., Li, X & Zhang, L. H. 2002. Appl. Environ. Microbiol. **68:** 2676–2682.
Zhou, J., Baba, T., Takano, T., Kobayashi, S. & Arai, Y. 1989. FEBS Lett. **255:** 37–41.

Received May 7, 2002
Accepted May 31, 2002