J Mol Evol (1999) 48:421–426

JOURNAL OF MOLECULAR EVOLUTION

Close Evolutionary Relatedness of α-Amylases from Archaea and Plants

Štefan Janeček,[1] Emmanuel Lévêque,[2,3] Abdel Belarbi,[2] Bernard Haye[3]

[1] Institute of Microbiology, Slovak Academy of Sciences, Štefánikova 3, SK-81434 Bratislava, Slovakia
[2] Laboratoire de Microbiologie Générale et Moléculaire, EUROPOL'AGRO, UFR Sciences de Reims, BP 1039, F-51687 Reims Cedex 2, France
[3] Laboratoire de Biochimie, EUROPOL'AGRO, UFR Sciences de Reims, BP 1039, F-51687 Reims Cedex 2, France

Received: 3 June 1998 / Accepted: 20 August 1998

Abstract. The amino acid sequences of 22 α-amylases from family 13 of glycosyl hydrolases were analyzed with the aim of revealing the evolutionary relationships between the archaeal α-amylases and their eubacterial and eukaryotic counterparts. Two evolutionary distance trees were constructed: (i) the first one based on the alignment of extracted best-conserved sequence regions (58 residues) comprising β2, β3, β4, β5, β7, and β8 strand segments of the catalytic $(\alpha/\beta)_8$-barrel and a short conserved stretch in domain B protruding out of the barrel in the β3 → α3 loop, and (ii) the second one based on the alignment of the substantial continuous part of the $(\alpha/\beta)_8$-barrel involving the entire domain B (consensus length: 386 residues). With regard to archaeal α-amylases, both trees compared brought, in fact, the same results; i.e., all family 13 α-amylases from domain Archaea were clustered with barley pI isozymes, which represent all plant α-amylases. The enzymes from *Bacillus licheniformis* and *Escherichia coli*, representing liquefying and cytoplasmic α-amylases, respectively, seem to be the further closest relatives to archaeal α-amylases. This evolutionary relatedness clearly reflects the discussed similarities in the amino acid sequences of these α-amylases, especially in the best-conserved sequence regions. Since the results for α-amylases belonging to all three domains (Eucarya, Eubacteria, Archaea) offered by both evolutionary trees are very similar, it is proposed that the investigated conserved sequence regions may indeed constitute the ''sequence fingerprints'' of a given α-amylase.

Key words: α-Amylase — *Thermococcus hydrothermalis* — Archaeons — Eubacteria — Eukaryotes — Conserved sequence regions — Evolutionary relationships

Introduction

α-Amylases (EC 3.2.1.1) are widely found enzymes capable of hydrolyzing the α-1,4-glucosidic bonds in starch. These enzymes constitute two sequence families (13 and 57) in the sequence-based classification of glycosyl hydrolases (Henrissat and Bairoch 1996). Although it has recently been shown that both the extant α-amylase families, 13 and 57, may share a common distant ancestor (Janeček 1998), the two families still have to be elucidated separately. Family 57 of α-amylases comprises a few amylolytic enzymes predominantly from hyperthermophiles (Bauer et al. 1998). The present study deals with family 13 α-amylases (for a recent review, see Janeček 1997), which involves more than one hundred α-amylases belonging to all the three domains of life: Eucarya, Eubacteria, and Archaea.

We have recently cloned and sequenced a gene coding for a family 13 α-amylase from the hyperthermophilic archaeon *Thermococcus hydrothermalis*. The full nucleotide and amino acid sequences of this α-amylase will be presented elsewhere, together with the description of its biochemical properties (E Lévêque, M Ned-

Correspondence to: Š. Janeček; *e-mail:* sjanecek@ue.savba.sk

Table 1. α-Amylases used in the present study

Domain and abbreviation	Source	Accession no.[a]	Reference
Bacteria			
Aerhy	*Aeromonas hydrophila* strain JMP636	P22630	Gobius and Pemberton 1988
Altha	*Alteromonas haloplanctis* strain A23	X58627	Feller et al. 1992
Bacli	*Bacillus licheniformis* strain ATCC 27811	P06278	Yuuki et al. 1985
Bacsu	*Bacillus subtilis*	P00691	Yang et al. 1983
Ecoli	*Escherichia coli* strain JA11	P26612	Raha et al. 1992
Lacam	*Lactobacillus amylovorus* strain CIP 102989	U62096	Giraud and Cuny 1997
Stral	*Streptomyces albidoflavus*	P09794	Long et al. 1987
Thtma	*Thermotoga maritima* strain DSM 3109	Y11359	Liebl et al. 1997
Archaea			
Pyrfu	*Pyrococcus furiosus* strain DSM 3638	U96622[b]	Jørgensen et al. 1997
Pyrsp	*Pyrococcus* sp. strain KOD1	D83793	Tachibana et al. 1996
Thchy	*Thermococcus hydrothermalis* strain AL662	AF068255	Lévêque et al. 1998
Thcpr	*Thermococcus profundus* strain JCM 9378	—[c]	Lee et al. 1996
Thcsp	*Thermococcus* sp. strain Rt3	AF017454	Jones et al. 1998
Eucarya			
Aspor	*Aspergillus oryzae*	P10529	Toda et al. 1982
Crysp	*Cryptococcus* sp. strain S-2	D83540	Iefuji et al. 1996
BarHIG	*Hordeum vulgare* (barley, high-p*I* isozyme)	P04063	Rogers 1985
BarLOW	*Hordeum vulgare* (barley, low-p*I* isozyme)	P00693	Rogers and Milliman 1983
Drome	*Drosophila melanogaster* (fruit fly)	P08144	Boer and Hickey 1986
Chicken	*Gallus gallus* (chicken, pancreas)	U63411	Benkel et al. 1997
HumanS	*Homo sapiens* (human, saliva)	P04745	Nishide et al. 1986
PigP	*Sus scroffa* (pig, pancreas)	P00690	Pasero et al. 1986
Shrimp	*Panaeus vannamei* (shrimp, hepatopancreas)	X77318	Van Wormhoudt and Sellos 1996

[a] Accession numbers from the SWISS-PROT protein [all start with ''P'' (Bairoch and Apweiler 1998)] and GenBank DNA (Benson et al. 1998) sequence databases.
[b] An identical sequence from the same strain was determined independently by Dong et al. 1997 (GenBank: AF001268).
[c] This sequence is not available from a sequence database.

jma, B Haye, A Belarbi, submitted). Here we focus on the identification of the best-conserved sequence regions of the *T. hydrothermalis* α-amylase, which allows us to predict roughly the structural features of this archaeal α-amylase, mainly the most important β-strands of the parallel $(\alpha/\beta)_8$-barrel fold. Since the most detailed evolutionary study published so far dealt only with microbial (eubacterial and fungal), plant, and animal α-amylases (Janeček 1994), the main goal of the present work was to reveal the evolutionary relationships of archaeal α-amylases to their eubacterial and eukaryotic counterparts.

Materials and Methods

All α-amylases used in the present study are listed in Table 1. All known archaeal α-amylases with determined sequence were used. As far as the α-amylases from Eubacteria and Eucarya are concerned, respectively, they were chosen as representatives of previously identified evolutionary related groups (Raimbaud et al. 1989; Janeček 1994, 1995, 1997; Jespersen et al. 1993; Janeček et al. 1997). Thus the studied α-amylases represent a wide spectrum of taxonomically different species. There are representatives for liquefying (*Bacillus licheniformis*) as well as saccharifying (*Bacillus subtilis* and *Lactobacillus amylovorus*) bacterial α-amylases, intracellular α-amylases (*Escherichia coli*), and α-amylases from actinomycetes (*Streptomyces albidoflavus*), which, along with the α-amylase from *Alteromonas haloplanctis* (psychrophile), form one closely related group, with animal α-amylases represented by the enzymes from *Drosophila melanogaster* (insects), shrimp (sea animals), chicken (birds), and human saliva and pig pancreas (isozymes of mammals). The eubacterial set is further completed by two α-amylases, those from the facultative anaerobe *Aeromonas hydrophila* and the extremal thermophile *Thermotoga maritima*. As for the remaining eukaryotic α-amylases, the two α-amylase pI isozymes from barley cover all plant α-amylases, and the α-amylases from *Aspergillus oryzae* and *Cryptococcus* sp. cover all fungal and yeast enzymes, respectively.

In each α-amylase amino acid sequence the conserved sequence regions were identified to give a set of aligned sequence stretches that then served for calculation of an evolutionary distance tree. Another similar tree was constructed for all studied α-amylases based on the sequence alignment of the large segment cut out from each α-amylase, which corresponded with the part of their $(\alpha/\beta)_8$-barrel domain from the start of strand β2 to the end of strand β8 (involving the domain B positioned in the loop3). All sequence alignments were performed using the program CLUSTAL W (Thompson et al. 1994) and then manually tuned where applicable. The method used for building the trees in both cases was the neighbor-joining method (Saitou and Nei 1987). The Phylip format tree output was applied using the bootstrapping procedure (Felsenstein 1985); the number of bootstrap trials used was 1000. The trees were drawn with the program Tree View (Page 1997).

Results and Discussion

The extracted conserved sequence regions of the studied archaeal, eubacterial, and eukaryotic α-amylases are

Source	β2	β3	loop3	β4	β5	β7	β8
Bacteria							
Aerhy	25 GYKQVLISP	75 GIAVYADVVLNH	163 LPDLD	190 GFRVDAVKH	217 HVFGEVIT	288 FAITHD	323 GSPLVYSDH
Altha	28 GYAAVQVSP	78 GVDIYVDTLINH	142 LADLD	170 GFRFDASKH	196 VVFQEVID	259 FVDNHD	297 GYPKVMSSY
Bacli	36 GHTAVWIPP	94 DINVYGDVVINH	198 YADID	227 GFRLDAVKH	257 FTVAEYWQ	323 FVDNHD	357 GYPQVFYGD
Bacsu	33 GYTAIQTSP	91 GIKVIVDAVINH	144 LYDWN	172 GFRFDAAKH	204 FQYGEILD	264 WVESHD	299 STPLFFSRP
Ecoli	35 GHNMVWLPP	94 DIAVLLDVVVNH	202 GENID	231 GFRIDAVKH	261 FIVAEYWS	327 LVANHD	361 GVPSVFYPD
Lacam	45 GYTAVQTSP	103 NIRIIVDATLND	156 FYDWN	184 GFRYDAATH	218 FQYGEVLQ	278 WVESHD	313 SVPLFFDRP
Stral	32 GYGYVQVSP	82 GVKVVADSVINH	145 LADLD	173 GFRIDAAKH	200 YWKQEAIH	263 FVDNHD	297 GSPDVHSGY
Thtma	70 GVDAVWFMP	117 GIKVIMDLVINH	186 MPDLN	214 GFRIDAAKH	254 ILVGEVFS	305 FLENHD	349 GSPVIYYGG
Archaea							
Pyrfu	40 GISAIWLPP	99 GIKVIADVVINH	161 FPDIC	193 GWRFDYVKG	217 WAVGEYWD	283 FVANHD	307 GQPVIFYRD
Pyrsp	41 GISAIWIPP	100 GIKVIADIVINH	162 FPDIA	194 AWRFDYVKG	218 WAVGEYWD	284 FVANHD	308 GQPVIFYRD
Thchy	41 GISAIWIPP	100 NMKVIADIVINH	162 YPDIC	194 AWRFDYVKG	218 WAVGEYWD	284 FVANHD	308 GQPAIFYRD
Thcpr	41 GISAIWIPP	100 GIKVIADIVINH	162 FPDIA	194 AWRFDYVKG	218 WAVGEYWD	284 FVANHD	308 GQPVIFYRD
Thcsp	41 GISAIWIPP	100 GIKVIADIVINH	162 FPDIA	194 AWRFDYVKG	217 LAVGEYWD	283 FVANHD	307 GQPVIFYRD
Eucarya							
Aspor	56 GFTAIWITP	111 GMYLMVDVVANH	173 LPDLD	202 GLRIDTVKH	226 YCIGEVLD	292 FVENHD	323 GIPIIYAGQ
Crysp	58 GFTAIWISP	118 GMYLMVDVVVNH	186 LVDLR	215 GLRIDSLQQ	240 YMVGEVFN	307 FLENQD	338 GIPITYYGQ
BarLOW	34 GVTHVWLPP	82 GVQAIADIVINH	147 APDID	176 AWRLDFARG	201 LAVAEVWD	286 FVDNHD	318 GIPCIFYDH
BarHIG	33 GHTHVWLPP	81 GVKAIADIVINH	146 APDID	175 GWRFDFAKG	200 FAVAEIWT	284 FVDNHD	316 GTPCIFYDH
Drome	36 GYAGVQVSP	88 GVRTYVDVVFNH	154 LRDLN	182 GFRVDAAKH	219 YIVQEVID	283 FVDNHD	322 GTPRVMSSF
Chicken	36 GFGGVQVSP	90 GVRIYVDAVVNH	165 LLDLA	193 GFRIDAAKH	229 FIYQEVID	295 FVDNHD	334 GFTRVMSSY
HumanS	36 GFGGVQVSP	90 GVRIYVDAVINH	165 LLDLA	193 GFRIDASKH	229 FIYQEVID	295 FVDNHD	334 GFTRVMSSY
PigP	36 GFGGVQVSP	90 GVRIYVDAVINH	165 LLDLA	193 GFRLDASKH	229 FIFQEVID	295 FVDNHD	334 GFTRVMSSY
Shrimp	35 GFAGVQVSP	91 GVRIYVDAVINH	165 LNDLN	193 GFRIDASKH	229 FIFQEVID	293 FIDNHD	332 GYTRVMSSY
				*	*	*	
consensus	G avw P	g D viNh	d	gfR Da kh	v Ev d	fv nhD	g p

Fig. 1. The conserved sequence regions of α-amylases. The abbreviations of enzyme sources are given in Table 1. The best-conserved parts of an α-amylase sequence comprise strands β2, β3, β4, β5, β7, and β8. There is also a short conserved sequence stretch located near the C terminus of the longest loop connecting strand β3 amd helix α3 (loop 3). The three proposed catalytic residues (Asp, Glu, and Asp in strands β4, β5, and β7, respectively) are marked by *asterisks.* The sequence features highly characteristic of archaeal α-amylases are *boxed.* A residue is written in the consensus sequence (*consensus*) if it is present in more than half of the α-amylases.

shown in Fig. 1. These regions comprise most of the β-strand segments forming the catalytic $(\alpha/\beta)_8$-barrel domain as well as the short pentapeptide stretch located near the C terminus of the very long loop3 (in the β3 → α3 connection), which forms, in fact, a small distinct domain, called domain B (Janeček et al. 1997). They may be taken as the ''sequence fingerprints'' of a given α-amylase since they involve most of the common amino acid residues important from both the functional and the structural points of view and often contain the residues highly characteristic of an individual α-amylase (Janeček 1997).

Figure 1 offers the opportunity to trace the sequence features exclusive for the α-amylases from archaeons (unless otherwise specified, all amino acid numbering throughout the text corresponds to mature α-amylase from *Thermococcus hydrothermalis*): Ile42, Pro48–Pro49, Ile107, Ile165, Ala194–Trp195, Tyr199, Gly202, Ala219, Tyr223–Trp224, Ala286, Gln309, Ile312–Phe313–Tyr314, and Asp316 (compare the boxed residues in Fig. 1). From sequence comparison of the conserved sequence regions of *T. hydrothermalis* α-amylase with the rest of family 13 α-amylases (Fig. 1), it is evident that the α-amylase from *T. hydrothermalis* (i) clearly ranks among the α-amylase family 13; (ii) adopts an $(\alpha/\beta)_8$-barrel structure, with domain B protruding out of the barrel between strand β3 and helix α3; and (iii) contains the three amino acid residues playing the catalytic role, which are the Asp198, Glu222, and Asp289 in strands β4, β5, and β7, respectively.

The presence of the residues highly characteristic of all the archaeal α-amylases in the enzymes from other kingdoms indicates the candidates for closest relatives to archaeal α-amylases. As is easily seen from Fig. 1, the plant α-amylases contain most of the archaeal sequence features, especially those connected with strand β4 around the catalytic aspartate residue (Asp198). It should be pointed out that the tryptophan equivalent with Trp195 of *T. hydrothermalis* α-amylase is present in 16 out of 17 plant α-amylases available in the SWISS-PROT (Barioch and Apweiler 1998) and GenBank (Benson et al. 1998) databases, and the glycine corresponding to Gly202 of the archaeal α-amylase is found at the end of this region of all plant α-amylases. While no special role for tryptophan equivalent to Trp195 has been assigned in the structure of the barley α-amylase–acarbose complex, the glycine corresponding to Gly202 provides in the plant enzyme a specific ligand for calcium ion (Kadziola et al. 1998). It is worth mentioning that no other α-amylase from more than 100 available in the sequence databases contains either the tryptophan or the glycine in this region except for the α-amylase from *Dictyoglomus thermophilum* AmyB (Horinouchi et al. 1988) containing the glycine at the end of the β4-strand region (Š. Janeček, unpublished results). As far as the tryptophan residue in position $i + 2$ from the β5-strand catalytic glutamic acid is concerned (Trp224), its equivalent forms a stacking interaction with one of the acarbose rings bound in the active site of barley α-amylase (Kadziola et al. 1998). This residue is again perfectly con-

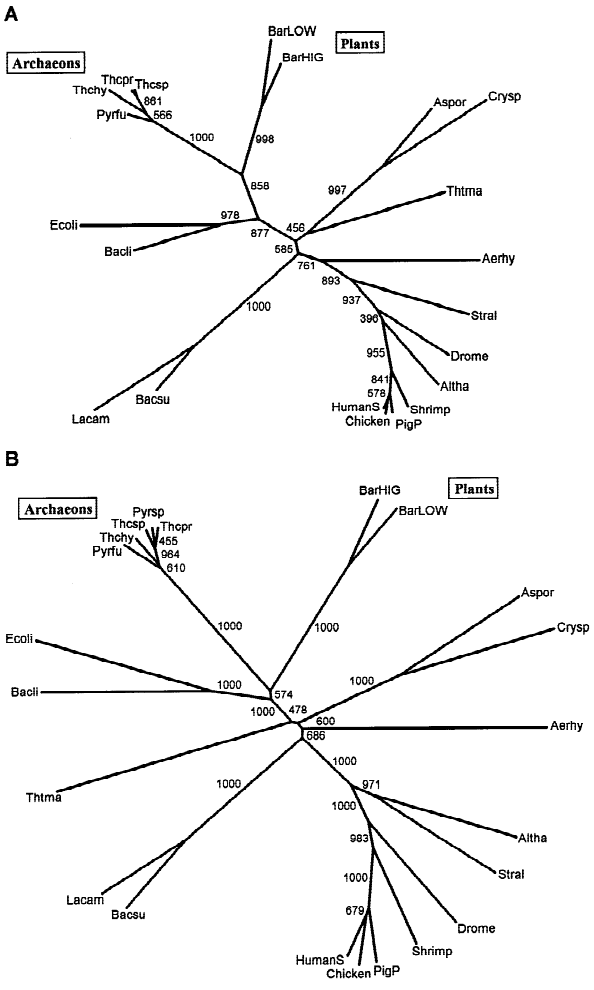

Fig. 2. Evolutionary trees of α-amylases. The abbreviations of enzyme sources are given in Table 1. The trees are based on the alignment of the conserved sequence regions shown in Fig. 1 (58 residues) **A** and on the alignment of a substantial part of the $(\alpha/\beta)_8$-barrel involving the domain B protruding out of the barrel in the $\beta3 \rightarrow \alpha3$ loop (consensus length: 386 residues) **B.** The branch lengths are proportional to the sequence divergence. *Numbers along branches* are bootstrap values (1000 replicates). In tree A the α-amylase from *Pyrococcus* sp. occupies the same position as the α-amylase from *T. profundus.*

served in and highly characteristic of both plant and archaeal α-amylases and is, moreover, present in the α-amylases from *B. licheniformis* and *E. coli* (Fig. 1).

In order to draw relevant evolutionary conclusions, an evolutionary tree (Fig. 2A) was constructed based on the alignment shown in Fig. 1. For comparison, a similar tree (Fig. 2B) was calculated that was based on the alignment of a continuous segment starting with the β2-strand and ending with the β8-strand and involving the entire domain B in the $\beta3 \rightarrow \alpha3$ loop connection (the alignment is not shown). Both the trees equally clearly demonstrate the evolutionary relatedness of archaeal and plant α-amylases. It is evident that these revealed evolutionary relationships are dictated by the similarities in the extracted best-conserved sequence regions (Fig. 1) as well as in the remaining less-conserved parts of their sequences (not shown). Interestingly, the archaeal *Sulfolobus solfataricus* α-glucosidase from family 31 of glycosyl hydrolases was also clustered together with eukaryotic (plant, fungal, and mammalian) counterparts (Rolfsmeier et al. 1998). This observation, along with the results of this study, may support the view that the archaeal and eukaryotic pathways of carbohydrate metabolism could have a common evolutionary origin. At least the partial explanation of the close evolutionary relatedness between archaeal and plant α-amylases described here can be provided by analogy with enzymes of central carbon metabolism, such as transketolase (Martin and Schnarrenberger 1997), whose mitochondrial, cyanobacterial (i.e., plant ancestral), and archaeal (i.e., eukaryote ancestral) genes might have persisted to the present in nuclear genomes.

Despite the fact that the archaeal and plant α-amylases are placed on the adjacent branches of a larger common cluster, they still retain their own evolutionary originality (documented by the long archaeal as well as plant branches). In these terms the position of the α-amylase from *B. licheniformis*, close to the archaeal and plant enzymes in both trees (Fig. 2), should be of interest. This α-amylase represents the liquefying α-amylases. When taking into account the fact that the archaeal α-amylase from *T. hydrothermalis* is also a liquefying α-amylase (E. Lévêque, unpublished results), it seems reasonable to point out that the evolutionary trees may reflect the mode of action of α-amylases, too. If so, then the residues common to *B. licheniformis* and archaeal α-amylases highlighted in Fig. 1 might become the useful candidates for mutagenesis studies aimed at elucidating the residues responsible for liquefying properties of α-amylases. This could be of special importance since the residues responsible for the hyperthermostability of *B. licheniformis* α-amylase are known, together with its three-dimensional structure (Declerck et al. 1995; Machius et al. 1995).

The similar results of both trees (Fig. 2) indicate that, when drawing the main features of evolutionary relations among α-amylases, it is not absolutely necessary to work with the entire sequences; i.e., it is satisfactory to use only the best conserved sequence regions (Fig. 1). These regions thus may be considered to be the ''sequence fingerprints'' of a given α-amylase. On the other hand, it should be pointed out that using only the four well-established conserved sequence regions covering strands β3, β4, β5, and β7 (e.g., Nakajima et al. 1986) may be risky since it is evident that some residues characteristic of a given α-amylase can be found just in strands β2 and β8 or even in the short conserved sequence region positioned near the C terminus of domain B (Janeček et al. 1997). The observation of this phenomenon can be extended to the entire α-amylase family (Janeček 1997).

Acknowledgments. Emmanuel Lévêque is a fellow from EUROPOL'AGRO. This work was supported in part by VEGA Grant 2/3013/98 to Štefan Janeček.

References

Bairoch A, Apweiler R (1998) The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1998. Nucleic Acids Res 26:38–42

Bauer MW, Driskill LE, Kelly RM (1998) Glycosyl hydrolases from hyperthermophilic microorganisms. Curr Opin Biotechnol 9:141–145

Benkel BF, Nguyen T, Ahluwalia N, Benkel I, Hickey DA (1997) Cloning and expression of a chicken α-amylase gene. Gene 192: 261–270

Benson DA, Boguski MS, Lipman DJ, Ostell J, Ouellette BFF (1998) GenBank. Nucleic Acids Res 26:1–7

Boer PH, Hickey DA (1986) The α-amylase gene in *Drosophila melanogaster:* nucleotide sequence, gene structure and expression motifs. Nucleic Acids Res 14:8399–8411

Declerck N, Joyet P, Trosset J-Y, Garnier J, Gaillardin C (1995) Hyperthermostable mutants of *Bacillus licheniformis* α-amylase: multiple amino acid replacements and molecular modelling. Protein Eng 8:1029–1037

Dong G, Vieille C, Savchenko A, Zeikus JG (1997) Cloning, sequencing, and expression of the gene encoding extracellular α-amylase from *Pyrococcus furiosus* and biochemical characterization of the recombinant enzyme. Appl Environ Microbiol 63:3569–3576

Feller G, Lonhienne T, Deroanne C, Libioulle C, Van Beeumen J, Gerday C (1992) Purification, characterization, and nucleotide sequence of the thermolabile α-amylase from the antarctic psychotroph *Alteromonas haloplanctis* A23. J Biol Chem 267:5217–5221

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Giraud E, Cuny G (1997) Molecular characterization of the α-amylase genes of *Lactobacillus plantarum* A6 and *Lactobacillus amylovorus* reveals an unusual 3′ end structure with direct tandem repeats and suggests a common evolutionary origin. Gene 198:149–157

Gobius KS, Pemberton JM (1988) Molecular cloning, characterization, and nucleotide sequence of an extracellular amylase gene from *Aeromonas hydrophila.* J Bacteriol 170:1325–1332

Henrissat B, Bairoch A (1996) Updating the sequence-based classification of glycosyl hydrolases. Biochem J 316:695–696

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

Iefuji H, Chino M, Kato M, Iimura Y (1996) Raw-starch-digesting and thermostable α-amylase from the yeast *Cryptococcus* sp. S-2: purification, characterization, cloning, and sequencing. Biochem J 318:989–996

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

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

Janeček Š (1997) α-Amylase family: Molecular biology and evolution. Prog Biophys Mol Biol 67:67–97

Janeček Š (1998) Sequence of archaeal *Methanococcus jannaschii* α-amylase contains features of families 13 and 57 of glycosyl hydrolases: a trace of their common ancestor? Folia Microbiol 43: 123–128

Janeček Š, Svensson B, Henrissat B (1997) Domain evolution in the α-amylase family. J Mol Evol 45:322–331

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

Jones RA, Patel B, Beacham IR (1998) Amylase gene from a hyperthermophilic archaebacterium (Thermococcus). Unpublished. GenBank: AF017454.

Jørgensen S, Vorgias CE, Antranikian G (1997) Cloning, sequencing, characterization, and expression of an extracellular α-amylase from the hyperthermophilic archaeon *Pyrococcus furiosus* in *Escherichia coli* and *Bacillus subtilis.* J Biol Chem 272:16335–16342

Kadziola A, Søgaard M, Svensson B, Haser R (1998) Molecular structure of a barley α-amylase-inhibitor complex: implications for starch binding and catalysis. J Mol Biol 278:205–217

Lee JT, Kanai H, Kobayashi T, Akiba T, Kudo T (1996) Cloning, nucleotide sequence, and hyperexpression of α-amylase gene from an archaeon, *Thermococcus profundus.* J Ferment Bioeng 82:432–438

Lévêque E, Nedjma M, Haye B, Belarbi A (1998) Cloning and expression of an α-amylase gene from extremely thermophilic archaebacterium *Thermococcus hydrothermalis* in *Escherichia coli.* Unpublished. GenBank: AF068255.

Liebl W, Stemplinger I, Ruile P (1997) Properties and gene structure of the *Thermotoga maritima* α-amylase AmyA, a putative lipoprotein of a hyperthermophilic bacterium. J Bacteriol 179:941–948

Long CM, Virolle M-J, Chang S-Y, Chang S, Bibb MJ (1987) α-Amylase gene of *Streptomyces limosus:* nucleotide sequence, expression motifs, and amino acid sequence homology to mammalian and invertebrate α-amylases. J Bacteriol 169:5745–5754

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

Martin W, Schnarrenberger C (1997) The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr Genet 32:1–18

Nakajima R, Imanaka T, Aiba S (1986) Comparison of amino acid sequences of eleven different α-amylases. Appl Microbiol Biotechnol 23:355–360

Nishide T, Nakamura Y, Emi M, Yamamoto T, Ogawa M, Mori T, Matsubara K (1986) Primary structure of human salivary α-amylase gene. Gene 41:299–304

Page RDM (1997) TreeView (Win16), Version 1.40. (http://taxonomy.zoology.gla.ac.uk/rod/rod.html)

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

Raha M, Kawagishi I, Mueller V, Kihara M, Macnab RM (1992) *Escherichia coli* produces a cytoplasmic α-amylase, *AmyA.* J Bacteriol 174:6644–6652

Raimbaud E, Buleon A, Perez S, Henrissat B (1989) Hydrophobic cluster analysis of the primary sequences of α-amylases. Int J Biol Macromol 11:217–225

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

Rogers JC, Milliman C (1983) Isolation and sequence analysis of a barley α-amylase cDNA clone. J Biol Chem 258:8169–8174

Rolfsmeier M, Haseltine C, Bini E, Clark A, Blum P (1998) Molecular characterization of the α-glucosidase gene (*malA*) from the hyperthermophilic archaeon *Sulfolobus solfataricus.* J Bacteriol 180: 1287–1295

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Tachibana Y, Mendez Leclere M, Fujiwara S, Takagi M, Imanaka T (1996) Cloning and expression of the α-amylase gene from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1, and characterization of the enzyme. J Ferment Bioeng 82:224–232

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment

through sequence weighting, positions specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

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

Van Wormhoudt A, Sellos D (1996) Cloning and sequencing analysis of three amylase cDNAs in the shrimp *Penaeus vannamei* (Crustacea decapoda): evolutionary aspects. J Mol Evol 42:543–551

Yang M, Galizzi A, Henner D (1983) Nucleotide sequence of the amylase gene from *Bacillus subtilis.* Nucleic Acids Res 11:237–249

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