Biologia, Bratislava, **57/Suppl. 11:** 181—189, 2002 **Review**

Diversity and evolution of the α-amylase genes in animals

Jean-Luc Da Lage[1]*, Alain Van Wormhoudt[2] & Marie-Louise Cariou[1]

[1] *UPR 9034 Populations, génétique et évolution, CNRS, 91198 Gif sur Yvette cedex, France; tel: ++ 33 1 69 82 37 27, fax: ++ 33 1 69 07 04 21, e-mail: jldl@pge.cnrs-gif.fr*
[2] *Station de Biologie marine, Collège de France et Muséum National d'Histoire Naturelle, 29900 Concarneau, France*

Da Lage, J. L., Van Wormhoudt, A. & Cariou, M. L., Diversity and evolution of the α-amylase genes in animals. Biologia, Bratislava, **57/Suppl. 11:** 181—189, 2002; ISSN 0006-3088.

Many living organisms show a tendency to have multiple amylase genes. It is also true in animals. In Drosophilids, the number of *Amy* copies varies from two to at least seven, including the divergent paralog *Amyrel*. Comparisons between copies show that the divergence, the gene arrangement and the intron/exon structure are variable among copies within species. In addition, there are often differential expressions of the different copies. A survey of a number of animal *Amy* sequences shows a high level of protein variability. The inferred protein sequences suggest some evolutionary events which could have adaptative or functional significance, such as the loss or gain of some amino acid stretches: a motif of nine aminoacids has been found in some species but not in others, independently from the phylogenetic relationships. Also, additional cysteines may create new disulfide bridges in some amylases.

Key words: intron, disulfide bridge, gene duplication, flexible loop.

Introduction

Evolutionary biologists concerned in molecular evolution have several favorite model genes and organisms. Among nuclear genes, alpha-amylase has been one of these models over the last forty years. The reasons are several, one of these being the ease of demonstrating amylase activity on electrophoresis gels, and another the low cost. When studies on enzymatic polymorphism developed, amylase was shown to exhibit an interesting variability, and also gene duplications in a number of organisms such as *Drosophila* (Bahn, 1967) or rodents (Meisler et al., 1986; Nielsen, 1977). The evolutionary significance of these independent duplications has remained a puzzling problem. Fortunately, amylase function is clear, and is connected to the environment of the organism, which should, in theory, be helpful to the understanding of this evolutionary process.

It is known that gene duplications can lead to several, not mutually-exclusive situations. Increasing the number of copies, without changing the sequence, may yield a high amount of protein. Moderate changes in coding sequences, and consequently, in the protein, may modify, sometimes subtly, the biochemical properties of the enzyme. These changes may be accompanied by modifications in regulatory regions, and hence modify the expression pattern (time- or tissue-specificity).

* Corresponding author

Table 1. PCR primers used to investigate the region of interest.

Primer name	Sequence	Strand	Protein motif [a]
WWERYQP	GGTGGGARCGTTAYCARCC	+	58–64
KHM	GAYGCNNNNAARCAYATGTGGC	+	197–204
DNHD	GAYAAYCAYGAYAAYCARCG	+	297–303
CEHREV	TGNCKCCANCGRTGYTCRCA	–	384–390
REV1400+	CCNGADATNAYRTCRCARTA	–	449–455

[a] The protein motif matching the residues (pig numbering).

Lastly, more drastic changes should lead eventually to a significantly different function, involving other substrates, cofactors, etc. In *Drosophila ananassae*, these various evolutionary pathways coexist in the *Amy* family. This species illustrates and summarizes well what may happen to duplicated genes, except, to our knowledge, becoming pseudogenes (Da Lage et al., 2000).

In this article, we show that in many animal species, amylase genes and proteins have undergone more or less drastic sequence evolution, with occurrence of some convergent molecular events. In particular, the presence of two amino acid motifs has been screened by PCR. This method has provided evidence for several cases of gene duplication. The extent of *Amy* duplication/diversification emphasizes the potential, but still enigmatic benefits in terms of adaptation.

New α-amylase sequences

Genomic DNAs from *Asterias rubens* (Echinodermata, Asteriidae), *Osmia cornuta* (Hymenoptera, Megachilidae), *Ceratitis capitata* (Diptera, Tephritidae), *Megaselia scalaris* (Diptera, Phoridae), *Blaps mucronata* (Coleoptera, Tenebrionidae), *Lithobius forficatus* (Myriapoda, Chilopoda), *Corbicula fluminea* (Bivalvia, Corbiculidae), *Patella vulgata* (Gastropoda, Patellidae), were prepared from fresh or frozen samples by two alternative methods: (1) Tris-EDTA-SDS cell disruption followed by potassium-acetate precipitation of proteins and ethanol precipitation of nucleic acids; (2) Tris-EDTA-NaCl-CTAB-polyvinylpyrollidone cell disruption in the presence of proteinase K, followed by chloroform extraction and ethanol precipitation of nucleic acids. Degenerate PCR primers were designed to match best the conserved regions of the amylase protein (Tab. 1). The combinations of (+) primers and (-) primers all encompassed the regions focused on, which correspond to the sequence between Asp297 and Trp388 in the pig alpha-amylase numbering. The amplification parameters were adapted to each case, but the annealing temperature was generally ∼53 ℃. The new sequences were deposited at GenBank with accession numbers AF286345, AF467103, AF146757, AF146758, AF467104, AF462603, AF468013-AF468016.

The *Drosophila ananassae* model

About 3000 Drosophila species have been described to date. Among those species that serve commonly as models in the laboratory, *D. ananassae* shows the most complex situation regarding the *Amy* family. Seven different copies have been cloned and sequenced from a single strain that expressed four electrophoretic variants AMY1,2,3,4. It was shown that each variant was encoded by a specific gene. These four genes are organized as two genetically independent clusters (Da Lage et al., 1992; 2000). The proteins encoded by the remaining three copies were not detected *in vivo*, but the most divergent one has been identified as the paralog *Amyrel*, which is also present in a number of Drosophila species (Da Lage et al., 1998). *Amyrel* is transcribed in larvae but not in adults, unlike the classical genes. Its function remains unknown, but the expression of the gene product in cultured Drosophila S2 cells demonstrated its ability to hydrolyze starch *in vitro* (unpublished data). However, the striking differences in *Amyrel* compared with classical *Amy* at the protein level (see below) suggest an enzymatic function somewhat different *in vivo*. Another gene of *D. ananassae*, *Amy c1*, is more enigmatic. No expression was detected until now, and the electric charge of the putative protein, computed as (Arg+Lys)-(Asp+Glu) is null. In this respect, it is reminiscent of that of fungi-feeding flies whose amylase is positively charged (Prigent et al., 1998). In addition, this gene has an unusually low codon bias, compared to other *Amy* genes in Drosophila. Figure 1 summarizes some features of the *Amy* family in *D. ananassae*.

The regulation of the *Amy* genes in *D. ananassae* is not well understood. We have shown

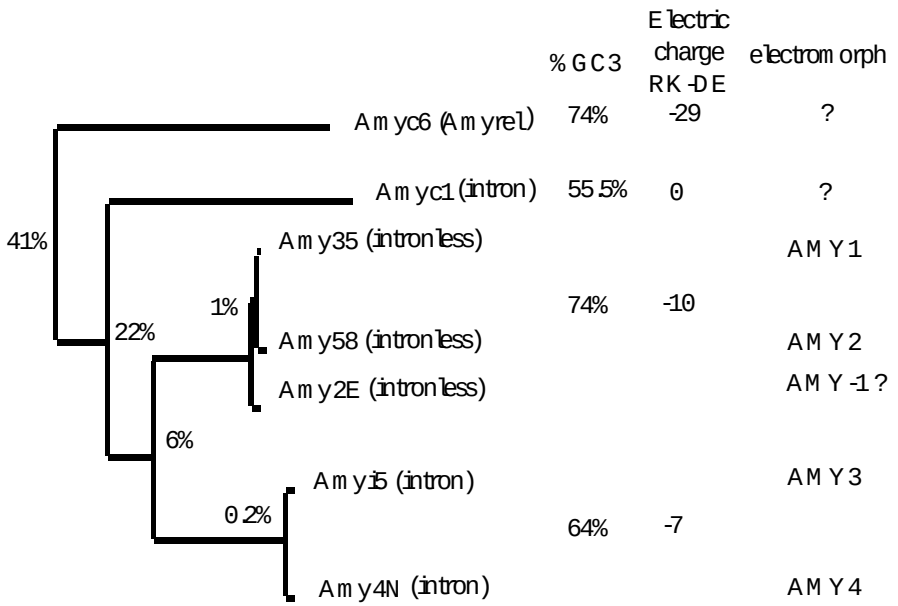

Fig. 1. Summary of the characteristics of the *Amy* gene family in *Drosophila ananassae*. The tree was made with CLUSTALW (THOMPSON et al., 1994). Numbers along the branches of the tree are percentages of divergence in amino acids. %GC3 is the G+C content at the third positions of codons. Electric charges of mature proteins were computed as (Arg+Lys)-(Asp+Glu).

(DA LAGE et al., 1996) that the cluster that contains the genes coding for AMY1 and AMY2 is active preferentially in larvae while AMY3 and AMY4 are more steadily produced, or at the adult stage. Tissue-specific expression along the midgut is also gene-cluster dependent. Nucleotide changes in the regulatory regions may be involved in this functional differentiation.

In summary, *D. ananassae* exhibits a range of examples of what a duplicated gene may become through evolutionary processes.

An overview on Drosophilids: independent duplications

Other Drosophila species were investigated by several researchers. From their work, it appeared clear that the duplications often reported had occurred independently (see references in Table 2). Except *Amyrel*, which seems present in every Drosophila species, the number of reported *Amy* genes varies from one in *D. virilis* (and possibly in many species of the *Drosophila* subgenus (TADLAOUI-OUAFI, 1993) to seven in *D. ananassae*. The copy number may be subject to variation within species. This has been reported in *D. pseudoobscura* (POPADIC et al., 1996) and suspected in *D. ananassae* (DA LAGE et al., 2000). Differential expression has been observed in *D. kikkawai* (INOMATA & YAMAZAKI, 2000) as in *D. ananassae*. Molecular events such as intron loss in some copies also occurred independently in closely related taxa (e.g. in the *montium* subgroup; ZHANG et al., in press). The relative orientation of the copies in tandemly arranged clusters also varies between species.

These results on Drosophila provide a lot of molecular data on gene and protein diversity and evolution after duplication, but they do not answer the central questions: what is beneficial, or why is it beneficial? We have undertaken an extension of this study to other animals, in order to find clues, i.e. can we link some molecular features to environment or feeding preferences? Is it possible to classify amylase proteins according to these features? And if so, is it a consequence of common ancestry or of convergent evolution?

An overview of animal α-amylases: similarities and divergences

Size of the Amy family

Data from the literature as well as from databases show numerous cases of *Amy* gene duplications in various phyla (Tab. 2). Our experiments were not targeted to detect duplications. However, PCR often results in heterogeneous products which are

Table 2. Number of amylase gene copies detected in various animal species.

Taxon	Species	Number of known *Amy* copies	References
Diptera	*Drosophila virilis*	1 + *Amyrel*	Hickey, unpubl.; Tadlaoui-Ouafi, 1993
	D. melanogaster	2 + *Amyrel*	Bahn, 1967
	D. pseudoobscura	3 + *Amyrel*	Brown et al., 1990
	D. kikkawai	4 + *Amyrel*	Inomata & Yamazaki, 2000
	D. ananassae	≥ 6 +*Amyrel*	Da Lage et al., 2000
	Ceratitis capitata	1 + *Amyrel*	this study
	Aedes aegypti	2	Grossman et al., 1997
Hymenoptera	*Apis mellifera*	1	Ohashi et al., 1999; this study
Lepidoptera	*Spodoptera frugiperda*	2	Da Lage, unpubl.
	Bombyx mori	2	Kikkawa, 1953
Coleoptera	*Tribolium castaneum*	2 very similar	Hickey, unpubl.
	Blaps mucronata	2 very divergent	this study
	Sitophilus oryzae	2	Baker et al., 1990
Myriapoda	*Lithobius forficatus*	2	Da Lage, unpubl.
Crustacea	*Penaeus vannamei*	3	Van Wormhoudt & Sellos, 1996
	Asellus aquaticus	6	Oxford, 1986
	Sphaeroma serratum	2	Laulier, 1988
Mollusca	*Crassostrea gigas*	2	Moal et al., 2000
Echinodermata	*Asterias rubens*	1	this study
Vertebrata	*Mus musculus*	4 + 1 pseudogene	www.ncbi.nlm.nih.gov
	Homo sapiens	5 or more + pseudogene(s)	www.ncbi.nlm.nih.gov; Groot et al., 1989

amplified from duplicate genes. If divergence is significant, it is evidence for a multicopy state. When divergence is low, the status of allele vs. paralog is debatable. In our assays, it happened several times that two sequences were obtained, sufficiently different to be undoubtedly duplicates (see Table 2). Another characteristic of PCR is that some extant genes may have been missed, because one primer does not match the sequence perfectly. In this respect, when the number of known *Amy* genes is one for some species, it may be just that some copies have not been detected yet. However, for the sea star *Asterias rubens*, Southern blotting with a homologous PCR fragment as a probe suggested that there was only one gene, except perhaps very divergent paralogs unable to stick to the probe (not shown). The same evidence leads to the same conclusion for *Drosophila virilis* and other species of the *Drosophila* subgenus (Tadlaoui-Ouafi, 1993). We must also point out that a number of *Amy* genes were cloned from cDNAs, i.e. from the most active copies. In *Aedes aegypti*, the very strange *Amy1* was cloned first, because it was expressed at a high level, although it is so different from other insect amylases. The much more classical *Amy2* was isolated later, and its expression remained questionable (Grossman et al., 1997).

Intron-exon structure

The intron-exon variability between duplicates, which was evident within *D. ananassae*, was also found in other animals, such as the beetle *Blaps mucronata* (not shown). Between taxa, the structure of *Amy* genes is highly variable. As an example, it seems that Diptera amylases have few introns, while Lepidoptera have a lot of them. This example is restricted to insects, but numerous cases of intron loss-or-gain may have occurred in animals (Sellos & Van Wormhoud, 2002).

Sequence divergence

Still regarding *Amy* duplicates at the intraspecific level, we observed pairs of almost identical genes (*Tribolium*, human...). This is generally expected if the gene copies are physically close to each other, through "concerted evolution". But several cases of high levels of divergence have been found between copies. A 40% value in amino acids was found in the beetle *Blaps mucronata*; a similar level of change was observed in the moth *Spodoptera frugiperda* (unpublished). However, in most cases, α-amylase proteins translated from the DNA sequences may be aligned. A tree can then be drawn. First of all, it must be said that α-amylase is a difficult material for phylogenetic studies, simply because of its multigene struc-

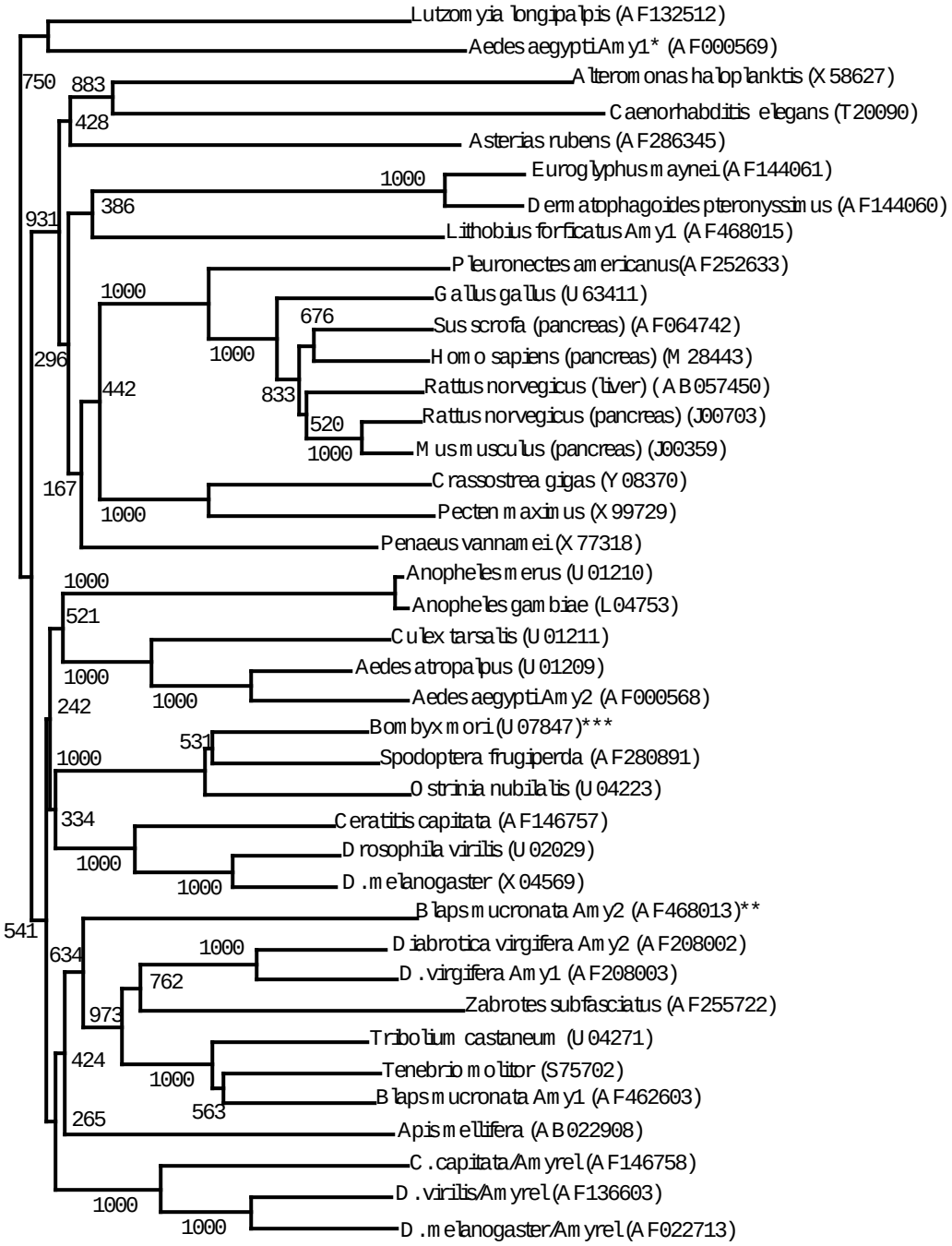

Fig. 2. Unrooted neighbor-joining tree of mature animal α-amylase protein sequences aligned with CLUSTALW. Numbers along branches are bootstrap values (1000 replicates). *: the N-terminal extension in *Aedes aegypti Amy1* was removed. Accordingly, C-terminal tails of some sequences were removed.**: sequences are partial, but do not alter the topology. ***: the only available sequence for *Bombyx mori* has a long deletion inside the coding sequence. *Alteromonas haloplanktis*, a bacterium with an α-amylase of animal type, was included (see discussion in text). The accession numbers are given into brackets. The tree was drawn with NJ PLOT by M. Gouy (`http://pbil.univ-lyon1.fr`).

ture. In contrast, discrepancies between the α-amylase tree and a reliable phylogenetic tree are good indications of peculiar evolutionary histories in the relevant lineages, or divergence between paralogs within species (Fig. 2). Incorrect branchings are easily visible: the crustacea *Penaeus*, and bivalves, are strongly associated to vertebrates (VAN WORMHOUDT & SELLOS, 1996). The huge difference between the two α-amylases of *Aedes aegypti* is striking. The tree also suggests that the gene cloned from the phlebotome (sand fly) *Lutzomyia* is orthologous to *Amy1* of *Aedes*. Also, the position of *Amyrel* shows its divergence from the classical gene, but as long as its origin is unknown, it is not clear from the tree whether it is related to any other known α-amylase, although its shares some molecular features with various *Amy* sequences.

Protein sequence: some similarities between unrelated species

Animal α-amylases share a number of protein motifs, some of which are not shared by other organisms; e.g. WWERYQPISYKL (position 58-69 in pig), WVCEHRW (382-388). We have focused our attention on two motifs that are "optional", in that they are present in some species (or in paralog proteins) and absent in others. The first one is a short motif of 3-4 amino acids GHG or GHGA (position 304-307 in pig), that belongs to the so-called "flexible loop", possibly involved in a trap-release mechanism for the substrate hydrolysis (QUIAN et al., 1997; STROBL et al., 1998). The absence of this motif in some α-amylases has already been noticed (STROBL et al., 1998). It is widespread in insects, but not in all of them. More over, both types are present in Drosophila: the motif is absent in *Amyrel*. It is also missing in *C. elegans* (nematode) (Fig. 3A). This motif, according to its phylogenetic distribution, may be ancestral. But the deletions should then have occurred independently, by a mechanism that remains unknown. The selective pressure and constraints needed in this evolutionary process should have been exerted repeatedly and strongly. One possible, strong environmental factor may be the need for circumventing α-amylase inhibitors that are abundant in plant seeds, a favorite food of many insects (PEREIRA et al., 1999). The benefit for the animal of overcoming inhibitors is obvious (MORTON et al., 2000). Crystallographical and biochemical studies of interactions between α-amylases and their inhibitors enable us progressively to understand the involvement of the flexible loop in the inhibition.

The second motif we focused on in this study is made of nine residues, in a region named the variable loop. When present, it is a well-conserved stretch of amino acids (Fig. 3B). The 9 amino-acid motif is present in vertebrates but not in a sea star; it is present in some arthropods but not in insects. It is present in some bivalves but in not all of them; and it is absent in *C. elegans*. As in the case of the "GHGA" motif, its phylogenetic distribution can be explained best by common ancestry and subsequent independent losses. The mechanism of this deletion is still enigmatic. We must explain how the same block exactly (except maybe in *Pecten*) has been removed in all occurrences. We suggest that cryptic splicing sites might have been activated in the neighborhood of this sequence, thus removing a piece of coding sequence during an event of intron loss in this area. We should remark that mechanisms for intron loss are themselves not fully understood, but this event frequently occurs in an independent manner at the same intron position (DA LAGE et al., 1996; GOTOH, 1998; ROBERTSON, 1998). Interestingly, there is an intron within the 9 amino-acid motif in mammals. However there are no clear cryptic splicing sites. Alternatively, this optional motif could be a former intron sequence, fixed as a coding sequence in a number of lineages, and lost in others. But this "intron" would have been very short and this hypothesis is less likely. Unlike the "GHGA" motif, the 9 amino-acid motif has not been found yet to coexist with genes lacking the 9 amino acids within a genome. In addition, the function of this stretch of amino acids is unknown.

It is interesting to look at cysteines in the α-amylase protein. A number of insect α-amylases have four disulfide bridges. But in several species, or even in paralogous genes such as *Amyrel*, a fifth putative bridge is likely, between Cys419 and Cys442 (numbering from *Penaeus*). Table 3 shows the distribution of this bridge that seems to lock the structure of domain C (the "greek key"). Another pair of cysteins (Cys6-Cys38) was found in Bivalves, as already mentioned by D'AMICO et al. (2000). Again, the phylogenetic distribution of the Cys419-Cys442 is enigmatic. In *Ceratitis capitata* (Diptera, Tephritidae), both classical and *Amyrel* genes have this pair of cysteines. But in Drosophila, this was lost in the classical gene and conserved in *Amyrel*. Moreover, it is present in more remote Diptera such as *Megaselia* (Phoridae) and in mosquitoes, except *Anopheles*. Thus, we can assume that this bond is ancestral in flies. The loss of this bond would thus be independent. This is suggested by the distribution in Diptera. However, although we observed this bond in Hymenoptera and Lepidoptera, it is absent from bee-

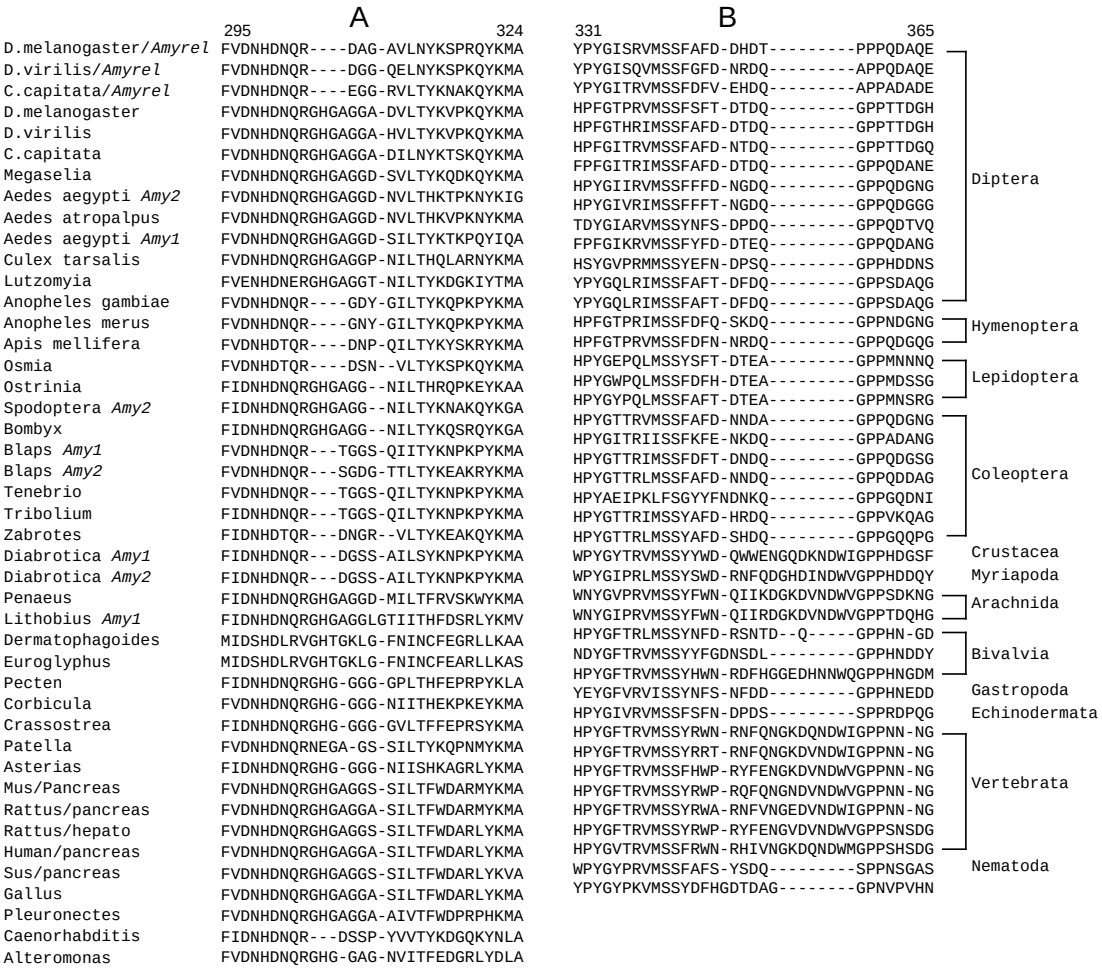

Fig. 3. Alignments of α-amylase sequences in the regions with "optional motifs". A: region of the flexible loop; B: region of the variable loop. Numbering is that of the mature porcine α-amylase.

tles, except in the divergent paralog *Amy2* of *Blaps mucronata*. We also know that these cysteines are present in the shrimp *Penaeus*, but it is the only amylase known in Crustacea. In other animal taxa investigated, this bond is absent. Is this a reacquisition, or a relic of the ancestral state, is it really limited to Arthropods? Further data are needed to provide an answer. Indeed, the questions of its origin and function remain open. The adaptive consequences of such a structure might be important.

Other single substitutions may change some important characteristics of the protein. Chloride-dependent α-amylases have been reviewed (D'Amico et al., 2000). An arginine residue at position 337 is needed for chloride binding. This arginine is present in most animal α-amylases known so far. There are some exceptions in Lepidoptera, which have a glutamine residue instead. Similarly, a glutamine is found in *Amyrel* of *Drosophila virilis* and other species from the *Drosophila* subgenus, i.e. *Amyrel* should be chloride dependent in some Drosophilids and independent in some others. The evolutionary significance of these substitutions is mysterious.

Very divergent amylases

The tree in Figure 2 shows that some amylases are at the tip of very long branches, which illustrates their high level of overall divergence. It is the case in two dust mites, *Dermatophagoides pteronyssinus* and *Euroglyphus maynei*, and also in the sea star, the nematode and some mosquito amylases. The case of *A. aegypti (Amy1)* is striking because the protein has an N-terminal extension

Table 3. Distribution of the additional disulfide bridge in animals.

Taxon	Species (gene)	SS bond
Diptera	*Drosophila melanogaster*	–
	D. melanogaster (*Amyrel*)	+
	Ceratitis capitata	+
	C. capitata (*Amyrel*)	+
	Megaselia scalaris	+
	Aedes aegypti (*Amy1*)	+
	Aedes aegypti (*Amy2*)	+
	Aedes atropalpus	+
	Culex tarsalis	+
	Anopheles gambiae	–
	Anopheles merus	–
Hymenoptera	*Apis mellifera*	+
	Osmia cornuta	+
Lepidoptera	*Ostrinia nubilalis*	+
	Spodoptera (*Amy2*)	+
	Bombyx mori	+
Coleoptera	*Blaps* (*Amy1*)	–
	Blaps (*Amy2*)	+
	Diabrotica (*Amy1*)	–
	Diabrotica (*Amy2*)	–
	Tenebrio molitor	–
	Tribolium castaneum	–
	Zabrotes subfasciatus	–
Crustacea	*Penaeus vannamei*	+
Myriapoda	*Lithobius* (*Amy1*)	–
Arachnida	*Dermatophagoides*	–
	Euroglyphus	–
Bivalvia	*Pecten maximus*	–
	Crassostrea gigas	–
Nematodes	*Caenorhabditis elegans*	–
Echinodermata	*Asterias rubens*	–
Vertebrates	*Mus musculus* (pancreas)	–
	Human (pancreas)	–
	Gallus gallus	–
	Pleuronectes	–

with a length of 237 amino acids (GROSSMAN & JAMES, 1993). This is, at present, unique in animal α-amylases. The putative ortholog in *Lutzomyia longipalpis* lacks this stretch, whose function is unknown. Interestingly, *Amy1* is expressed in salivary glands, but not at all (or very little) in the midgut, unlike *Drosophila* (ABRAHAM & DOANE, 1978), *Tenebrio molitor* (CRISTOFOLETTI et al., 2001) and *Spodoptera frugiperda* (BOLOGNESI et al., 2001) for instance, where midgut expression occurs. The evolution of the protein might be linked to tissue-specificity. More generally, it is probable that specific ecological constraints (diet, parasitic way of life) may lead to non-classical amylase sequences.

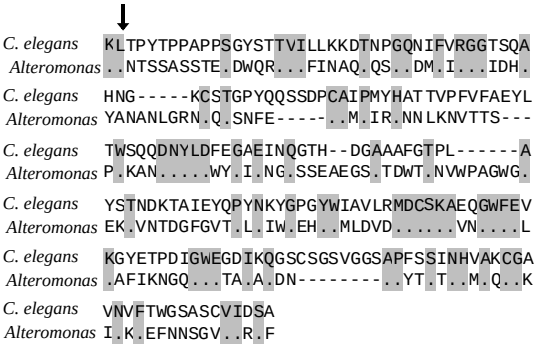

Fig. 4. Alignment of the C-terminal extensions of α-amylases from *Caenorhabditis elegans* (GenBank: T20090) and *Alteromonas haloplanktis* (X58627). Identical residues are shaded. The arrow indicates the beginning of the tail, after the most frequently found end position.

Animal-like bacterial α-amylases

Several workers have reported that some Eubacteria possess amylase genes/proteins with significant similarity to animal α-amylases. This has been found in *Alteromonas haloplanktis* (FELLER et al., 1992), several *Streptomyces* (e.g., BAHRI & WARD (1993) or *Thermomonospora* (PETRICEK et al., 1992); reviewed in JANECEK (1994). As for true animal α-amylases, we can look at the presence of the GHGA and 9 amino-acid motifs: the GHGA motif is present in *Alteromonas*, but not in *Streptomyces limosus* (GenBank: M18244), and the 9 amino-acid motif is absent in both. It is possible that those bacteria which express animal-like α-amylases also have bacterial-type *Amy* genes (which are themselves highly divergent from each other (JANECEK, 1994)). However, the activity of animal-like α-amylases should be predominant over bacterial types, since only the former have been discovered in these species. Although the animal-like sequences of these α-amylases have been noticed some years ago, no convincing explanation has been proposed. Two major hypotheses may be mentioned: first, ancestry of several amylase types, one of which would have remained in animals, along with a few, unrelated bacteria; second, lateral gene transfer between metazoa and bacteria, probably from animals towards prokaryotes. Although the second hypothesis seems more likely, evidence is difficult to obtain. In this respect, the similarity found in the C-terminal tails of *C. elegans* (animal) and *A. haloplanktis* (bacteria) (Fig. 4) seems to us a powerful clue towards the elucidation of this enigma.

Conclusion

The diversity of α-amylase proteins, and of the corresponding multigene family shows many occurrences of convergent events, motif loss, disulfide bonds in proteins, gene duplications, intron losses, etc., in genes, which make this model very attractive for the study of molecular adaptive evolution. However, before being able to tell the story without too many nebulous chapters, a large amount of data is still required from a broad range of animal taxa, especially the primitive diploblastic organisms. At a smaller scale, it is clear, as we have shown, that short-time quick divergence is frequent in amylases, so that a better sampling in insects and other arthropods, for example, will be a valuable source of information.

Acknowledgements

We thank A. Adoutte, C. H. Langley, S. D'Amico and S. Janecek for their ideas and knowledge.

References

Abraham, I. & Doane, W. W. 1978. Proc. Natl. Acad. Sci. USA. **75:** 4446–4450.

Bahn, E. 1967. Hereditas **58:** 1–12.

Bahri, S. M. & Ward, J. M. 1993. Gene **127:** 133–137.

Baker, J. E., Halliday, W. R. & Lum, P. T. M. 1990. J. Stored Prod. Res. **26:** 7–10.

Bolognesi, R., Ribeiro, A. F., Terra, W. R. & Ferreira, C. 2001. Arch. Insect Biochem. Physiol. **47:**62–75.

Brown, C. J., Aquadro, C. F. & Anderson, W. W. 1990. Genetics **126:** 131–138.

Cristofoletti, P. T., Ribeiro, A. F. & Terra, W. R. 2001. J. Insect Physiol. **47:** 143–155.

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

Da Lage, J.-L., Lemeunier, F., Cariou, M.-L. & David, J. R. 1992. Gen. Res. Camb. **59:** 85–92.

Da Lage, J.-L., Klarenberg, A. & Cariou, M.-L. 1996. Heredity **76:** 9–18.

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

Da Lage, J.-L., Maczkowiak, F. & Cariou, M.-L. 2000. J. Mol. Evol. **51:** 391–403.

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

Gotoh, O. 1998. Mol. Biol. Evol. **15:** 1447–1459.

Groot, P. C., Bleeker, M. J., Pronk, J. C., Arwert, F., Mager, W. H., Planta, R. J., Eriksson, A. W. & Frants, R. W. 1989. Genomics **5:** 29–42.

Grossman, G. L., Campos, Y., Severson, D. W. & James, A. A. 1997. Insect Biochem. Mol. Biol. **27:** 769–781.

Grossman, G. L. & James, A. A. 1993. Insect Mol. Biol. **1:** 223–232.

Inomata, N. & Yamazaki, T. 2000. Mol. Biol. Evol. **17:** 601–615.

Janecek, S. 1994. Eur. J. Biochem. **224:** 519–524.

Kikkawa, H. 1953. Adv. Genet. **5:** 107–140.

Laulier, M. 1988. Genet. Sel. Evol. **20:** 63–74.

Meisler, M. H., Antonucci, T. K., Treisman, L. O., Gumucio, D. L. & Samuelson, L. C. 1986. Genetics **113:** 713–722.

Moal, J., Daniel, J. Y., Sellos, D., Van Wormhoudt, A. & Samain, J. F. 2000. J. Comp. Physiol. **170:** 21–26.

Morton, R. L., Schroeder, H. E., Bateman, K. S., Chrispeels, M. J., Armstrong, E. & Higgins, T. J. V. 2000. Proc. Natl. Acad. Sci. USA. **97:** 3820–3825.

Nielsen, J. T. 1977. Genetics **85:** 155–169.

Ohashi, K., Natori, S. & Kubo, T. 1999. Eur. J. Biochem. **265:** 127–133.

Oxford, G. S. 1986. Heredity **56:** 105–110.

Pereira, P. J., Lozanov, V., Patthy, A., Huber, R., Bode, W., Pongor, S. & Strobl, S. 1999. Structure **7:** 1079–1088.

Petricek, M., Tichy, P. & Kuncova, M. 1992. Gene **112:** 77–83.

Popadic, A., Norman, R. A., Doane, W. W. & Anderson, W. W. 1996. Mol. Biol. Evol. **13:** 883–888.

Prigent, S., Matoub, M., Rouland, C. & Cariou, M.-L. 1998. Comp. Biochem. Physiol. **119B:** 407–412.

Quian, M., Spinelli, S., Driguez, H. & Payan, F. 1997. Protein Sci. **6:** 2285–2296.

Robertson, H. M. 1998. Genome Res. **8:** 449–463.

Sellos, D. Y & Van Wormhoudt, A. 2002. Biologia, Bratislava **57 (Suppl. 11):** in press.

Strobl, S., Maskos, K., Betz, M., Wiegand, G., Huber, R., Gomis-Ruth, F. X. & Glockshuber, R. 1998. J. Mol. Biol. **278:** 617–628.

Tadlaoui-Ouafi, A. 1993. Thesis, Université Pierre et Marie Curie, Paris.

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

Van Wormhoudt, A. & Sellos, D. 1996. J. Mol. Evol. **42:** 543–551.

Received October 4, 2001
Accepted February 12, 2002