Acta Biochim Biophys Sin 2010, 42: 863–872 | © The Author 2010. Published by ABBS Editorial Office in association with Oxford University Press on behalf of the Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. DOI: 10.1093/abbs/gmq103.

Original Article

Comparative molecular analysis of evolutionarily distant glyceraldehyde-3-phosphate dehydrogenase from *Sardina pilchardus* and *Octopus vulgaris*

Tarik Baibai[1*†], Laila Oukhattar[1,2†], Driss Mountassif[3], Omar Assobhei[2], Aurelio Serrano[4], and Abdelaziz Soukri[1*]

[1]Laboratoire de Physiologie et Génétique Moléculaire, Faculté des Sciences Aïn Chock, Université Hassan II, Casablanca B.P. 5366, Morrocco
[2]Laboratoire de Biotechnologies Marines et Environnement, Faculté des Sciences, Université Chouaib Doukkali, El Jadida B.P. 24000, Morrocco
[3]Laboratoire de Biochimie et Biologie Moléculaire, Faculté des Sciences Aïn Chock, Université Hassan II, Casablanca 5366, Morocco
[4]Instituto de Bioquímica Vegetal y Fotosíntesis (CSIC-University of Seville), Seville 41092, Spain
[†]These authors contributed equally to this work.
*Correspondence address. Tel: +212-522-230680/84; Fax: +212-522-230674; E-mail: a_soukri@hotmail.com (A.S.)/baiman21@yahoo.fr (T.B.)

The NAD$^+$-dependent cytosolic glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.12), which is recognized as a key to central carbon metabolism in glycolysis and gluconeogenesis and as an important allozymic polymorphic biomarker, was purified from muscles of two marine species: the skeletal muscle of *Sardina pilchardus Walbaum* (Teleost, Clupeida) and the incompressible arm muscle of *Octopus vulgaris* (Mollusca, Cephalopoda). Comparative biochemical studies have revealed that they differ in their subunit molecular masses and in p*I* values. Partial cDNA sequences corresponding to an internal region of the *GapC* genes from *Sardina* and *Octopus* were obtained by polymerase chain reaction using degenerate primers designed from highly conserved protein motifs. Alignments of the deduced amino acid sequences were used to establish the 3D structures of the active site of two enzymes as well as the phylogenetic relationships of the sardine and octopus enzymes. These two enzymes are the first two GAPDHs characterized so far from teleost fish and cephalopod, respectively. Interestingly, phylogenetic analyses indicated that the *sardina* GAPDH is in a cluster with the archetypical enzymes from other vertebrates, while the *octopus* GAPDH comes together with other molluscan sequences in a distant basal assembly closer to bacterial and fungal orthologs, thus suggesting their different evolutionary scenarios.

Keywords glyceraldehyde-3-phosphate dehydrogenase; cDNA; *GapC* gene; RT-PCR; molecular phylogeny; *Octopus vulgaris*; *Sardina pilchardus*

Received: May 16, 2010 Accepted: September 6, 2010

Introduction

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.12), one of the enzymes studied most in the glycolytic pathway, reversibly catalyses the oxidative phosphorylation of D-glyceraldehyde-3-phosphate to form 1,3-diphosphoglycerate in the presence of NAD$^+$ and inorganic phosphate [1]. This enzyme is widely distributed in nature in a variety of species ranging from bacteria to humans [2]. It is found mainly in the cytosol and in some organelles (i.e. chloroplasts). Organellar GAPDHs are expressed as precursor polypeptides and then post-translationally imported into the organelles [3]. In addition to its well-characterized glycolytic activity, a housekeeping function essential for the normal metabolism of all cells, GAPDH also plays a pivotal role in the Embden-Meyerhoff pathway in gluconeogenesis [2]. This enzyme has been well characterized not only because of its key role in the central metabolism, but also because of its abundance, easy preparation and remarkable conservation during evolution. It is a multimeric protein with a native molecular mass in the range of 140–200 kDa and composed of four identical subunits of approximately 35–50 kDa [2,4]. The ubiquity and evolutionary conservation of this enzyme implicate a highly important physiological function. There is now accumulating evidence that this protein is implicated in a large spectrum of cellular functions unrelated to its glycolytic function [5]. These include its roles in membrane fusion, phosphotransferase activity, DNA replication and DNA repair [6], specific binding to 3′ and 5′ regions of mRNA by direct influence of the translation or/and replication of RNA [7], nuclear RNA export [8] and neuronal apoptosis [9], as well as in neurodegenerative diseases [10] and several types of cancer

[11]. These novel activities may be related to the subcellular localization and oligomeric structure of GAPDH *in vivo*.

The glycolytic pathway is particularly suitable for testing the processes of enzyme evolution and the involvement of possible gene/genome duplications and/or horizontal gene transfer events. This central metabolic route is highly conserved and ancient; it is therefore possible to compare the enzymes included in this pathway from phylogenetically distant organisms. GAPDH is one of the most highly conserved glycolytic enzymes, for instance the rate of evolution of the catalytic domain is only 3% per 100 million years [2]. Thus, the catalytic domains in eukaryotic and eubacterial enzymes are >60% identical. Therefore, GAPDH genes are often used as phylogenetic markers for 'deep' phylogenies [12–14], as a prototype or controls for studies of genetic organization, expression and regulation and to describe the taxonomic positions of several species at different levels [15–20]. In this respect, it should be noted that, in general, GAPDH phylogeny is congruent with taxonomic data obtained for prokaryotic and eukaryotic species using other reported molecular markers [12]. The apparent incongruences found in some cases have been explained either by functional specificities (i.e. the highly divergent cytosolic and chloroplastic GAPDH isoforms found in photosynthetic eukaryotes [13,17]) or by intra-/inter-domains horizontal gene transfers (i.e. eukaryotic-type GAPDHs found in some bacteria [15]).

No information is available to date on the molecular properties of GAPDHs and the genes encoding this dehydrogenase in clupeida and cephalopoda, or on their phylogenetic relationship with GAPDHs from other species (prokaryotics and eukaryotics). In this study, we compared the kinetic and biomolecular parameters of the GAPDH purified from the skeletal muscle of *Sardina pilchardus* and the incompressible arm muscle of *Octopus vulgaris*, and tried to find some distinguishing characteristics of other GAPDHs studied previously. cDNA fragments of the corresponding *GapC* genes were obtained by polymerase chain reaction (PCR), sequenced and identified as the internal region of these genes containing the catalytic site. The phylogenetic relationship of the sardine and octopus GAPDHs with the orthologs from other vertebrate and invertebrate species was analyzed and discussed.

Materials and Methods

Biological material

Sardina pilchardus. The European pilchard *S. pilchardus* (Walbaum, 1792) is one of the most important species of small pelagic fishes of the northwest coast of Africa [21] and represents the most abundant species, particularly in the central and southern regions of Morocco. The constant exploitation of this resource for several decades places Morocco at the place of sardine–producer countries. The sardine distribution in this region extends from Cap Blanc (21°N) to Cap Spartel (35°45′N). Sardines are originated from the continental platform of the Atlantic Ocean to the Moroccan coastline. Samples used in this study were purchased from fishermen operating small fishing vessels, and were immediately frozen at −20°C until use.

Octopus vulgaris. Among the world's inshore fisheries, *O. vulgaris* is one of the most commercially valuable mollusc species fished in the world, especially in Morocco. In general terms, octopus is the most abundant and valuable species in the cephalopod fisheries in the Moroccan coast, accounting for 70–80% of total landings. Adult octopuses, *O. vulgaris*, were directly purchased from fishermen operating by day trip with small fishing vessels in the coastline of Casablanca region (33°N). The biological material was then transported to laboratory in boxes with dry ice in an isothermal truck within 6–8 h after capture.

Purification and characterization of sardine and octopus GAPDHs

GAPDH was purified to electrophoretic homogeneity from a soluble protein fraction of *S. pilchardus* skeletal muscle and *O. vulgaris* arm muscle, using a simple procedure involving a fractioned precipitation in the 60–88% (W/V) saturation range of ammonium sulfate and only one column chromatography step, namely dye-affinity chromatography on Cibacron Blue-Sepharose 4B (Amersham Pharmacia Biotech, Piscataway, NJ, USA) [22,23].

Determination of native molecular weight was carried out by electrophoresis on non-denaturing polyacrylamide slab gels (Bio-Rad, Hercules, USA) using the following protein standards: amylase, 200 kDa; alcohol dehydrogenase, 150 kDa; dimeric BSA, 132 kDa; monomeric BSA, 66 kDa; ovalbumin, 45 kDa; and lactalbumin, 14.2 kDa. As described by Hedrick and Smith [24], a calibration curve can be calculated from the relative mobility of standard proteins on non-denaturing polyacrylamide gels with different acrylamide concentrations (6%, 8%, 10%, and 12%, W/V). By constructing the Ferguson plot [Log (Rf × 100) versus the concentration of polyacrylamide gels (%)], the resulting slopes versus the known molecular weights of standard native proteins allowed the determination of the native molecular weight of the purified GAPDH.

Isoelectric focusing was carried out in 5% polyacrylamide slab gels holding ampholite-generated pH gradients in the range of 3.5–10 (Pharmacia Biotech, Uppsala, Sweden), using 25 mM NaOH and 20 mM acetic acids as cathode and anode solutions, respectively. The standard protein marker (Bio-Rad) was applied to estimate the p*I* of the purified enzymes.

Kinetics parameters and influence of pH and temperature on purified GAPDHs activity

For kinetic studies, initial velocities of the enzymatic reaction were determined by varying the concentration of the substrates, D-G3P (from 0.04 mM to 10 mM) or NAD^+ (from 0.02 to 2 mM). Values of the Michaelis constants (K_m) and dissociation constants (K_D) were obtained by mathematical calculation according to the method of Cleland [25], and the K_{cat} was calculated based on maximal velocity value (V_{max}). One unit of enzymatic activity was defined as the amount of enzyme that catalyses the reduction of 1 μmol of NAD^+ per min under the conditions used.

To determine the optimal pH, enzymatic activity was measured over a wide range of pH values (4–11) using a mixture of different buffers with different pK_a (Tris, MES, HEPES, potassium phosphate at 50 mM, and sodium acetate at 180 mM) adjusted to the same ionic strength as the standard reaction mixture. Temperature effects were characterized by activation and denaturation processes. For activation, the Tricine–NaOH buffer (50 mM, pH 8.5) was incubated for 10 min at temperatures from 15°C to 85°C using a thermostated cuvette holder connected with a refrigerated bath circulator. Then 2 mM NAD^+, 200 mM sodium arsenate, and 10 μg purified GAPDH were added to the mixture. The reaction was started immediately by the addition of 10 mM D-G3P. For denaturation, 10 μg purified GAPDH was incubated at temperatures from 15°C to 85°C for 10 min in the 50 mM Tricine–NaOH buffer. Then 2 mM NAD^+ and 200 mM sodium arsenate were added. The enzymatic activities were measured after 2 min incubation at 25°C immediately started by the addition of 10 mM D-G3P.

RNA isolation, RT-PCR methodology and cDNA sequencing of sardine and octopus GAPDH genes

Total RNA was isolated from sardina skeletal muscle and octopus arm muscle using the method of Chomczynski and Sacchi [26]. First-strand cDNA was generated by reverse transcription (RT) of total RNA (3 μg), using 200 units of reverse transcriptase-moloney murine leukemia virus (MMLV) transcriptase (Promega, Madison, USA) and the reverse primer named Gap 2: 5′-CCSCAYTCRTTRTCRT ACCA-3′ in a reaction mixture containing 50 mM Tris–HCl buffer (pH 8), 3 mM $MgCl_2$, 10 mM dithiothreitol and 0.2 mM of each deoxynucleoside triphosphate for 1 h at 42°C. An aliquot from this template (1/10 of the reaction volume) was used in a subsequent PCR using 1.25 units of Go Taq DNA polymerase (Promega), 0.04 μM of reverse (Gap2) and forward primer Gap 1: 5′-GCYWSYTGYACS ACSAAYTG-3′. Gap 1 and Gap 2 are degenerated oligonucleotides constructed from conserved regions (ASCTTNC, WYDNEW(C)G) present in all GAPDHs so far studied [2]. Amplifications of cDNA fragments (ca. 0.5 kb) corresponding to internal regions (ca. 70% of the full ORF) of *GapC* genes from sardine and octopus were carried out by PCR using the same degenerate primers. Amplification conditions for both species were 35 cycles of 92°C for 1 min, 45°C for 1 min and 72°C for 1 min. The PCR-amplified cDNA fragments from sardine and octopus were visualised on 2% (w/v) agarose gels, and purified using a phenol/chloroform protocol [27]. DNA sequencing was performed using the PCR-amplified cDNA, Gap 1 or Gap 2 primers, and the BigDye 3.1 sequencing kit (Applied Biosystems, California, USA). Extension products were purified with DyeEx-96 (Qiagen) and electrophoretically separated and detected in an ABI PRISM 377XL automated sequencer (Applied Biosystems). Electropherograms were revised and sequences were edited using BioEdit 7.0.1 [28].

Nucleotide sequence accession numbers

The nucleotide sequences from *S. pilchardus* and *O. vulgaris* reported in this paper have been deposited in the GenBank sequence database under the accession numbers EF621524 and EF634059, respectively.

Protein sequences alignment, phylogenetic analyses and active-site modeling of sardine and octopus GAPDHs

Multiple sequences alignment of GAPDH protein regions corresponding to the cDNA fragments of *sardina* and *octopus GapC* genes was done with the CLUSTAL X v.1.8 program [29]. No internal gaps were established. Through this alignment, phylogenetic trees were constructed using the distance neighbor-joining algorithm (Kimura distance calculations) and the minimum evolution method, as well as the maximum parsimony method, with the programs MEGA4 [30], TREE-PUZZLE v.5.2 [31], and PROTPARS v.3.573c (PHYLIP package v.3.5c w1993x Felsenstein, J., Department of Genetics, University of Washington, Seattle, USA) [32]. Bootstrap analyses (values being presented on a percentage basis) were computed with 1000 or 10,000 replicates without significant changes in tree topologies and bootstrap values, and all positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (pairwise deletion option) for the distance trees.

For maximum parsimony method, there was a total of 175 positions in the final data set, out of which 44 were parsimony informative. For maximum likelihood analysis, estimations of support (also expressed as percentages) were assigned to each internal branch by the algorithm quartet puzzling. This method applies maximum likelihood tree reconstruction to all possible quartets that can be formed from *n* sequence [31].

BLAST searches were made employing the National Center for Biotechnology Information Website facilities (http://www.ncbi.nlm.nih.gov/). Published amino acid sequences of animal GAPDHs used for the alignment were from Mammalia (*Mus musculus*, P16858; *Jaculus orientalis*, P80534; *Oryctolagus cuniculus*, P46406; *Sus scrofa*, P00355; *Bos taurus*, P10096; and *Homo sapiens*, P00354); Avian (*Gallus gallus*, P00356 and *Columba livia*, AAB88869); Amphibia (*Pleurodeles waltz*, AF343978); Teleost fishes (*Onchorhyncus mykiss*, AAB82747; *Sparus aurata*, ABG23666, *Dicentrarchus labrax*, AAW56452; *Tribolodon brandtii*, AB266388; *Oplegnathus fasciatus* GAPDH isoforms 1, ACF35052, and 2, ACF35053); Mollusca (*Crassostrea gigas*, CAD67717; *Marisa* sp., AAS02316 and *Pinctda fucata*, BAD90588; *Leptochiton* sp. strain SJB-2006, ABM97664 and *Haliotis discus*, ABO26632) and other major invertebrate clades (Acoelomata: *Fasciola hepatica*, AAG23287; Arthopoda: *Daphnia pulex*, CAB94909, and *Bombyx mori*, BAE96011; Cnidaria: *Hydra magnipapillata*, XP_0021655; Echinodermata: *Asterias rubens*, ABM97661; Nematoda: *Caenorhabditis briggsae*, CAP22176; Nemertea: *Cerebratulus* sp., ABM97662; and Sipuncula: *Phascolion strombus*, ABM97666). The GAPDH encoded by the enterobacterial *Escherichia coli* G*apA* gene (accession number P06977) and a fungal ortholog encoded by the *Saccharomyces cerevisiae GAPDH1* gene (P00360) were also included as outgroups.

The 3D structures of *S. ilchardus* and *O. vulgaris* GAPDHs are unknown, although the structures of several GAPDHs have been reported [3,33]. As GAPDHs are among the most conserved proteins sharing high degrees of primary sequence similarity [2], we subsequently generated homology models of both sardine and octopus GAPDHs active-site domains based upon known 3D structures of GAPDHs from other species, using Swiss Model [34] and the program Deep view/SWISS PDB Viewer v3.7 that provides an interface allowing to analyze several proteins, to superimpose them in order to deduce structural alignments and compare their active sites [35]. These models were subjected to spatial motif searches with the aim of identifying possible distinctive features as well as amino acids in the active site of this enzyme conserved during evolution.

Results

Kinetic properties of GAPDH and influence of pH and temperature on purified GAPDH activity

GAPDH has been purified to apparent electrophoretic homogeneity from muscle soluble protein fraction of the European pilchard, *S. pilchardus* and common octopus, *O. vulgaris*. The purification of the enzyme was performed by a straightforward procedure involving ammonium sulfate precipitation and only one chromatography step, namely dye-affinity chromatography. **Table 1** summarizes a representative purification protocol for the two species. A value of approximately 36 U/mg of protein was obtained for the specific activity of purified sardine GAPDH with a yield of 25% and a purification factor of approximately 78 fold, while for the octopus enzyme a specific activity value of 9.2 U/mg of protein and a purification factor of about 26 fold were obtained. SDS-PAGE analysis of the final enzyme preparations, using the dye-affinity columns, showed a 37 kDa and 36 kDa protein bands corresponding to the GAPDH subunit for sardine and octopus, respectively (**Fig. 1**). Non-denaturing PAGE showed that the native molecular mass of the obtained proteins are approximately 155 kDa and 153 kDa for sardine and octopus, respectively. Isoelectric focusing of the purified proteins showed a single band but with very different p*I* values for the two species, namely 7.9 for the sardine and 6.6 for the octopus enzymes (**Fig. 2**). Therefore, the fish GAPDH is a basic protein according to previous reports on other metazoan GAPDHs [33,36–38], whereas its molluscan ortholog is unexpectedly an acidic protein-like microbial GADPH [13,15]. As GAPDH catalyzes a two-substrate reaction, the K_m values for D-G3P and NAD^+, which have been determined by varying the concentration of one substrate and keeping the concentration of the other constant, were 73.4 μM and 92 μM for sardine and 320 μM and 66 μM for octopus, respectively (**Table 2**). The V_{max} values calculated for the sardine and octopus GAPDHs

Table 1 Comparative steps of GAPDH purification from skeletal muscle of *S. pilchardus* and arm muscle of *O. vulgaris*

	Total protein (mg)		Specific activity (units/mg)		Yield (%)		Purification (fold)	
	Sardina	*Octopus*	*Sardina*	*Octopus*	*Sardina*	*Octopus*	*Sardina*	*Octopus*
Crude extract	4700	810.9	0.46	0.35	100	100	1	1
Ammonium sulfate fractionation (60%–88%)	840	110.25	1.54	2.11	60	81	3.34	5.95
Blue-sepharose CL-6B chromatography	15	5	35.7	9.2	25	16	77.71	25.91

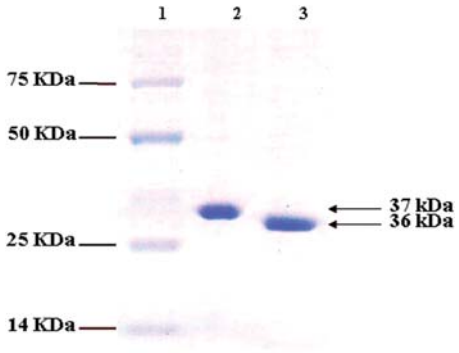

Figure 1 Comparison of GAPDHs subunit molecular masses from *Sardina pilchardus* skeletal muscle and *O. vulgaris* arm muscle Purified proteins were run on 12% SDS-PAGE and gels were stained with Coomassie Brilliant Blue. Lanes 2 and 3 show pure GAPDHs (25 μg per lane) from sardina and octopus, respectively. Lane 1 corresponds to molecular mass standards (Broad Range MW, Bio-Rad).

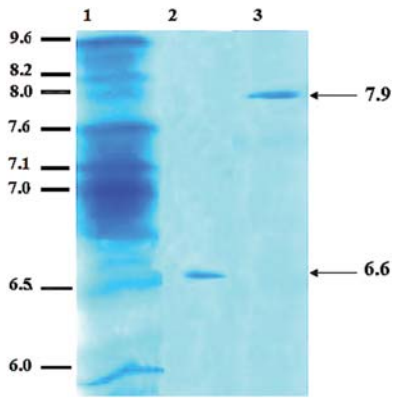

Figure 2 Isoelectric focusing of purified *S. pilchardus* and *O. vulgaris* GAPDHs Isoelectric focusing was performed on 5% (w/v) acrylamide gel holding Ampholyte-generated pH gradients (pH range, 3.5–10). Lane 1 corresponds to isoelectric focusing protein markers (p*I* range, 4.0–9.6). Lanes 2 and 3 correspond to pure GAPDH from octopus (p*I* 6.6) and sardina (p*I* 7.9), respectively.

were 37.6 U/mg and 21.8 U/mg, respectively. The catalytic efficiencies, expressed in terms of V_{max}/K_m ratios, indicate that the octopus enzyme is less efficient with both substrates than its teleostian ortholog.

Pre-incubation of both sardina and octopus GAPDHs for 10 min at temperatures varying between 15°C and 30°C did not irreversibly affect the enzyme activity. Thermal inactivation, however, occurred above 35°C and resulted in total activity loss at 60°C and 75°C, respectively for sardina enzyme and octoups GAPDH. Studies on the effect of temperature on enzymes activities revealed an optimal value at 28–32°C for sardina GAPDH, and around 35°C for octopus enzyme. The pH activity profile of purified GAPDHs was determined in a pH range from 4 to 11; the maximum of GAPDH relative enzymatic activity in octopus is observed between pH 7 and 7.5, lower than the value observed for GAPDH of sardina that occurred between pH 7.5 and 8.5.

Protein sequences alignment, phylogenetic analysis and active-site modeling

RT-PCR amplification using primers constructed from two highly conserved GAPDH regions produced for the two species a single cDNA fragment of the expected size (~0.5 kb) comprising most of the coding region of a *GapC* gene. The nucleotide sequences determined for the amplified cDNA fragments (507 bp and 525 bp for sardine and octopus *GapC* genes, respectively) were deposited in the GenBank/EMBL databases with the accession numbers EF621524 and EF634059. The cDNAs obtained from two independent RT-PCR experiments for each species were sequenced and found to be identical.

These sequences were aligned and compared with other GAPDHs selected to include species representatives of the main phyla of aquatic and terrestrial vertebrates and invertebrates, as well as two model bacterial and fungal species, by using the CLUSTAL X program [29] (**Fig. 3**). Relatively high percentage of identity (71%) and similarity (79%) were found between the amino acid sequences of GAPDHs of the two marine species.

The above-described multiple sequences alignment was used to construct phylogenetic trees obtained with the MEGA and Tree-Puzzle programs to infer the evolutionary relationship of the sardina and octopus GAPDHs using distance methods (neighbor-joining, minimum evolution), and maximum parsimony and likelihood methods [30,39,40] [**Fig. 4(A)**, **(B)** and **(C)**]. All consensus phylogenetic trees yielded similar topologies, and showed distant evolutionary positions for sardina and octopus GAPDH. The sardina protein conforms a cluster with other teleost fishes orthologs within a well-supported assembly of GAPDH of other main vertebrate groups, while the octopus protein arranges with other molluscan orthologs in a rather basal assembly near to microbial (fungal and bacterial) GAPDH.

GAPDH of other major invertebrate groups display between these vertebrates and molluscan assembling. Interestingly, a second cluster of teleost fishes GAPDH was also found, probably due to the occurrence in some species of enzyme isoforms (paralogs), as has been reported for many other vertebrates [33,36,37].

Sardina and octopus GAPDHs shared an amino acid sequence identity of 86% and 75% with human and 33 and 32% with *E. coli* GAPDH, respectively. Their active-site 3D structures were therefore based on the structure of the templates of these two well-studied enzymes, as they fulfill both the criteria of high-sequence identity and high resolution of the experimentally determined model structure. The holoforms have been determined at 1.75 Å and 1.8 Å, respectively for sardina and octopus GAPDH. The final models were visualized with the SPDB Viewer program (**Fig. 5**).

Table 2 Comparison of the kinetic parameters (K_m and V_{max}) for the oxidation reaction of the purified GADPHs from *S. pilchardus* and *O. vulgaris*

	K_m G3P (μM)	K_m NAD^+ (μM)	K_D NAD^+ (μM)	V_{max} (U/mg)	K_{cat} (min^{-1})
Sardina GAPDH	73.4 ± 8.1	92.0 ± 7.4	173.1 ± 28.2	37.6 ± 2.9	12.3 ± 0.9
Octopus GAPDH	320.0 ± 35.3	66.0 ± 7.6	103.0 ± 23.5	21.8 ± 3.5	8.5 ± 1.4

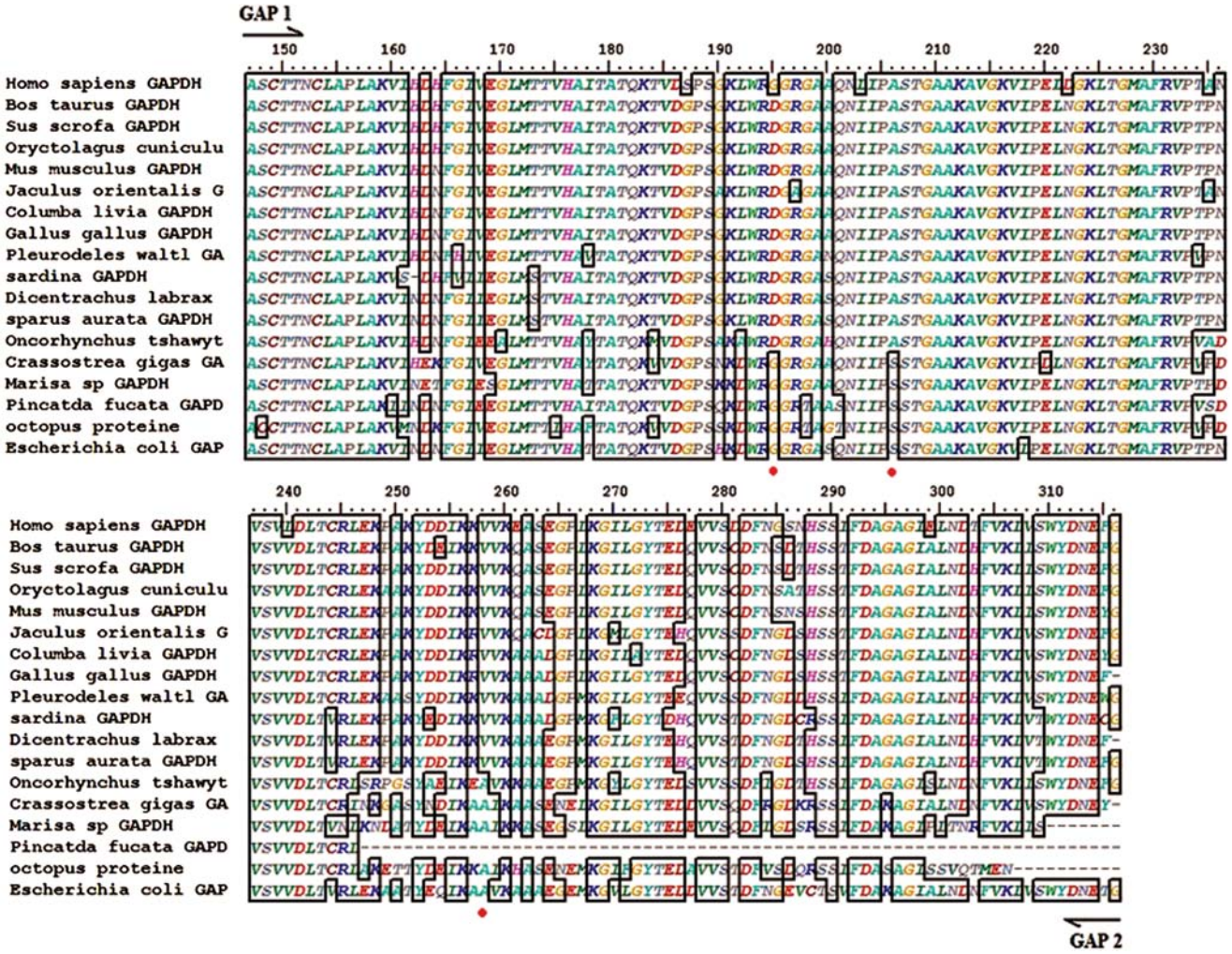

Figure 3 Multiple sequences alignment of *S. pilchardus* and *O. vulgaris* GAPDHs compared with 16 GAPDHs, selected to include representatives of the main aquatic and terrestrial vertebrate phyla and one bacterial species, by using the CLUSTAL X (v. 1.8) program. The 170 amino acid sequences corresponding to the proteins encoded by the RT-PCR-amplified cDNA fragments of the *GapC* genes from sardina and octopus correspond to a conserved internal region of GAPDH. Conserved amino acid residues throughout all the aligned sequences are farmed, and those distinctively found in molluscs and bacterial GAPDH sequences are indicated by red marks.

Discussion

GAPDH is the most highly conserved protein of all glycolytic enzymes. It plays a key role in central carbon metabolism and shows both genetic and post-translational regulations. This enzyme is responsible for the oxidative phosphorylation of G3P in the presence of NAD^+ and inorganic phosphate. In this work, GAPDH from *S. pilchardus* skeletal muscle and *O. vulgaris* arm muscle were purified to electrophoretic homogeneity from a soluble protein fraction. **Table 1** summarizes a representative purification protocol for the two species. A value of 36 U/mg of protein was obtained for the specific activity of the purified sardine GAPDH with a yield of 25% and a purification factor of 78 fold, while for the octopus GAPDH, a value of 9.2 U/mg of protein for the specific activity and a purification factor of about 26 fold were obtained. These differences in specific activity and factor of purification

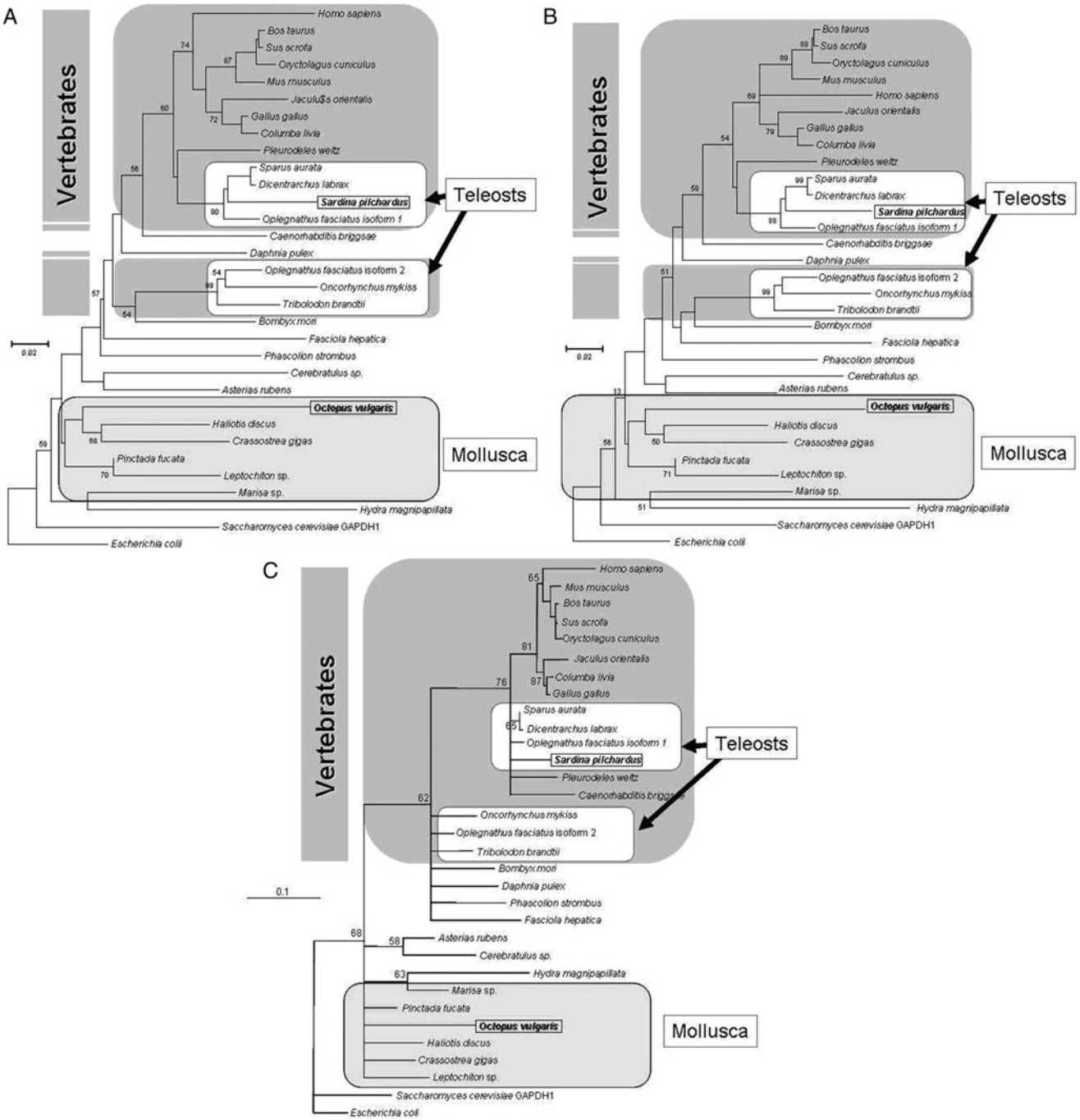

Figure 4 Evolutionary relationships of sardine and octopus GAPDHs (A) Using the neighbor-joining method. Selected GAPDH sequences representing the main vertebrate and invertebrate phylogenetic groups were used. A fungal (*S. cerevisiae*) and a bacterial (*E. coli*) GAPDHs sequences are shown as out groups. The bootstrapped consensus distance tree was inferred from 10,000 replicates. The scale bar represents 0.02 amino acid substitutions per site. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (pairwise deletion option). There were a total of 170 amino acid positions in the final data set. (B) Using the minimum evolution (ME) method. The optimal distance tree with the sum of branch length = 1.96 is shown. ME tree was searched using the close-neighbor-interchange (CNI) algorithm at a search level of 1. The neighbor-joining (NJ) algorithm was used to generate the initial tree. Other parameters are indicated as above. Phylogenetic analyses were conducted in MEGA4. (C) Using maximum likelihood analysis. Tree reconstruction was performed with the Quarter puzzling procedure and parameters estimated using Quarter sampling on an NJ tree. Bootstrap analysis of 10,000 re-samplings was performed and percent values are presented. The percentages of replicate trees in which the associated sequences clustered together in higher than 50% of the bootstrap test are shown next to the branches. Nodes with less than 50% support are collapsed. The scale bar represents 0.1 amino acid substitutions per site. Phylogenetic analyses were conducted in Tree-Puzzle v. 5.2.

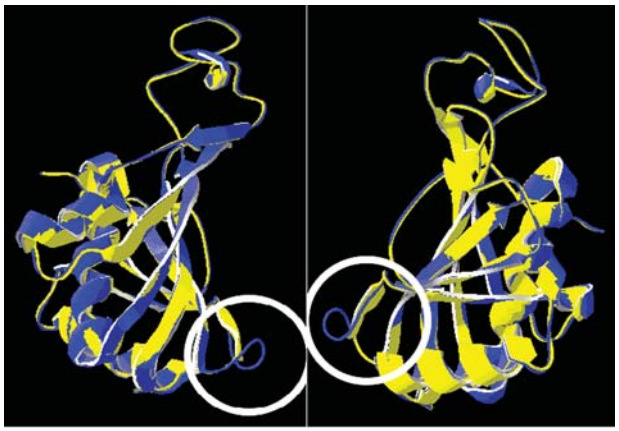

Figure 5 Stereoviews of the modeled structures of GAPDHs of *S. pilchardus* and *O. vulgaris* obtained from the secondary structure alignment The back bone of sardina and octopus GAPDHs are shown as gold and blue ribbons, respectively. Note the presence of unordered extra lop of the octopus GAPDH (white circle). The figure was created with the SWISS-PDB Viewer program.

between the two species may be explained by the shorter shelf life of octopus muscle due to its high autolytic activity, namely 25 times greater than in gadoids and 3 times greater than in squid [41]. Dye-affinity chromatography on Blue-Sepharose seems to be a very effective technique for the purification of sardine and octopus GAPDH, as previously described for the NAD^+-dependent GAPDH of other sources [22,36,42,43]. No additional purification steps were required to obtain homogeneous GAPDH preparations. As stated above, SDS-PAGE of the purified octopus sp and sardine enzymes showed a single protein band subunit of 36 kDa and 37 kDa, respectively. This result, when compared with the corresponding native molecular masses (154 kDa and 153 kDa, respectively), suggests that both of the enzymes have a homotetrameric structure like most other GAPDHs studied so far [4,33,36]. However, the sardine enzyme subunit exhibited an estimated molecular mass of 37 kDa that is somewhat higher than the one reported for bacterial and molluscan species (ca. 35 kDa), but is identical to those reported for the amphibian (pleurodeles) and mammalian (jerboa or human) species GAPDHs [33,36,38]. Moreover, the isoelectric focusing technique revealed very different values of p*I* for the single protein band present in the enzyme preparations purified from the two species, namely 7.9 and 6.6 for sardine and octopus GAPDH, respectively. This result indicates that a single enzyme isoform with quite different molecular properties occurs in the muscle tissues of the two species, and strongly suggests that single highly divergent *GapC* genes are expressed in these tissues. A single GAPDH isoform has been previously found in other animal tissues and microorganisms, both prokaryotes and eukaryotes [4,33,44,45]. However, it does not seem to be a general rule, as the presence of several GAPDH isoforms have also been reported in phylogenetically different organisms [37,38,45,46].

The kinetic parameters of octopus and sardina GAPDH (K_m values for D-G3P and NAD^+ being, respectively, in the ranges of 70 μM−300 μM and 70 μM−90 μM) (**Table 2**) are comparable to those found for cytosolic GAPDHs purified from other eukaryotes, protists (*Tetrahymena pyriformis* [43]), lower metazoa like the mollusk *Loligo vulgaris* [47], or mammalians like *Jaculus orientalis* or *Homo sapiens* [36,38]. However, on the whole the kinetic parameters analyses show that the octopus enzyme is catalytically less efficient than GAPDH of sardine. This suggests possible differences in the mechanism of the catalytic reaction. Therefore, some catalytic features of the GAPDH of the two marine species reported here differ from those previously described for GAPDH from other sources [36−38,43], reflecting protein differences between species.

The amino acid sequences derived from cDNA obtained from each species correspond to a highly conserved region around the catalytic subunit including many residues strictly conserved in GAPDH from very diverse organisms [2]. These sequences were compared and aligned with those of other selected GAPDHs representing main vertebrates and invertebrates phylogenetic groups. However, the limited information available on molluscan GAPDH allowed including sequences of a few oyster species (Bivalva) and only one gastropod (Haliotis). In fact, the octopus sequence presented here is the first GAPDH sequence reported so far of a species of the Cephalopoda clade, for which recent studies proved muscle specializations in its motor system [48]. The GAPDH sequence alignment has shown that over the analyzed region the orthologs of the two marine species studied here share a high percent of amino acid identity. The conservation is slightly higher in the catalytic domain between the amino acid residues 147 and 320, corresponding to the most conserved region for all GAPDH so far studied, especially the strictly conserved histidine 176 directly involved in the catalysis reaction (**Fig. 3**). The minor differences between the two sequences include different content and distribution of several residues like isoleucine 175, glycine 195 and threonine at positions 198 and 201, in the amino acid sequence of octopus GAPDH that are not present in the other sequences. On one hand, the sardine GAPDH sequence contains other histidine residue at position 164 that are not found in the octopus one. On the other hand, the presence of a Serine residue in position 173 seems to characterize the group of teleost fishes, but the most important remark is the distinctive conservation of three amino acid residues (aspartate, glycine and serine, respectively, at positions 192, 195 and 206) only in the sequences of octopus, other mollusks (*Crassostrea gigas*, *Marisa* sp. and *Pinctada fucata*), as well as in the sequence of *E. coli* (**Fig. 3**).

The sardine and octopus GAPDH share a relatively high-sequence identity with the *Homo sapiens* and *E. coli* GAPDHs, respectively, in accordance with the high conservation between orthologs from all phyla ranging from bacteria to vertebrate. Therefore, modeling of the crystal structures of both GAPDH was based on the structure of the templates of these two species, as they fulfill both the criteria of high percent sequence identity and high-resolution structure determination. The quality of the final models reflect the sequence identity shared between octopus and sardine GAPDH with exception in a small region, which seems not located in the catalytic site. The final models corresponding to one subunit of the holo-enzyme is shown in **Fig. 5**.

The phylogenetic trees obtained show the evolutionary relationships among teleost fishes and molluscan GAPDH sequences suggesting two divergent subfamilies, and reveal that while the sardine GAPDH is closely related to its orthologs of other teleost fishes and near to the group of amphibian enzymes the molluscan GAPDHs, including the octopus enzyme, conform a more deeply distant branching group, closely related to fungal and bacterial orthologs. Rather than a paraphyletic relationship between the molluscan and vertebrates GAPDHs these results may reflect an ancestral character of the first group of enzymes. It should be noted that these results depict the molecular phylogeny of the GAPDH protein only and they do not necessarily represent phylogenetic relationships between metazoan species. This interesting divergence is in agreement with the catalytic and structural differences between the two GAPDH orthologs reported in this paper. However, it remains to be clarified whether this scenario is due to true phylogenetic relationships among anciently diverged phylogenetic clades or to a specific specialization of GAPDH in the molluscan muscle, a tissue in which a particular glycolytic metabolism is feasible. On the other hand, the possibility of anomalous phylogenetic relationships due to horizontal gene transfer and enzyme functional substitution, as reported in other GAPDH-based phylogenies [2,3,33,43], cannot be ruled out. In any case, the comparative study presented in this work on novel eukaryotic purified proteins and gene sequences will shed new light on the functional and evolutionary relationship of the GAPDH from two representative species of marine animals, *S. pilchardus* and *O. vulgaris*, and on the possibility of using GAPDH enzyme as molecular marker for the studies of stock management belonging to the international coast. Current work is currently underway in our laboratories to go further on this research.

Conclusions

The NAD^+-dependent cytosolic GAPDH (EC 1.2.1.12) has been purified to homogeneity from muscle tissues of *S. pilchardus* and arm muscle of *O. vulgaris*. Comparative studies revealed that the two proteins differ in their subunit molecular masses, p*I* values and some kinetic parameters. Molecular phylogenetic studies using the amino acid sequences obtained from the cDNA fragments corresponding to an internal region of the *GapC* genes from sardine and octopus confirm the distant phylogenetic positions of the GAPDH of these two marine species, and suggest their possible use as molecular markers for seafood analysis and marine species traceability.

Acknowledgements

The authors gratefully thank Dr. L. Fourat for his help during the realization of the present work. Thanks are also extended to Junta de Andalucia PAIDI (group BIO-261) and FEDER EU program for support.

Funding

This work was supported by Genosphere Biotechnology and Novamics Biosciences Company Researches and the Morocco–Spain Cooperation Project (AECID, No A/8870/07).

References

1 Harris JI and Waters M. Glyceraldehyde-3-phosphate dehydrogenase. In: Boyer PD ed. The Enzymes, 3rd edn. New York: Academic Press, 1976, 1–49.
2 Fothergill-Gilmore LA and Michels PA. Evolution of glycolysis. Prog Biophys Mol Biol 1993, 59: R105–R235.
3 Cerff R. The chimeric nature of nuclear genomes and the antiquity of introns as demonstrated by GAPDH gene system. In: Go M and Schimmel P eds. Tracing Biological Evolution in Protein and Gene Structures. Amsterdam: Elsevier, 1995, 205–227.
4 Iddar A, Valverde F, Serrano A and Soukri A. Expression, purification, and characterization of recombinant nonphosphorylating NADP-dependent glyceraldehyde-3-phosphate dehydrogenase from Clostridium acetobutylicum. Protein Expr Purif 2002, 25: 519–526.
5 Sirover MA. New insights into an old protein: the functional diversity of mammalian glyceraldehyde-3-phosphate dehydrogenase. Biochim Biophys Acta 1999, 1432: 159–184.
6 Baxi MD and Vishwanatha JK. Uracil DNA glycosylaseglyceraldehyde-3 phosphate dehydrogenase is an Ap4A binding protein. Biochemistry 1995, 34: 9700–9707.
7 Schultz DE, Hardin CC and Lemon SM. Specific interaction of glyceraldehyde-3-phosphate dehydrogenase with the 59-nontranslated RNA of hepatitis A virus. J Biol Chem 1996, 271: 14134–14142.
8 Singh R and Green MR. Sequence-specific binding of transfer RNA by glyceraldehyde-3-phosphate dehydrogenase. Science 1993, 259: 365–368.
9 Saunders PA, Chen RW and Chuang DM. Nuclear translocation of glyceraldehyde-3-phosphate dehydrogenase isoforms during neuronal apoptosis. J Neurochem 1999, 72: 925–932.
10 Mazzola JL and Sirover MA. Reduction of glyceraldehyde-3-phosphate dehydrogenase activity in Alzheimer's disease and in Huntington's disease fibroblasts. J Neurochem 2001, 76: 442–449.

11 Rondinelli RH, Epner DE and Tricoli JV. Increased glyceraldehyde-3-phosphate dehydrogenase gene expression in late pathological state human prostate cancer. Prostate Cancer Prostatic Dis 1997, 1: 66–72.
12 Canback B, Andersson SG and Kurland CG. The global phylogeny of glycolytic enzymes. Proc Natl Acad Sci USA 2002, 99: 6097–6102.
13 Iddar A, Valverde F, Assobhei O, Serrano A and Soukri A. Widespread occurrence of non-phosphorylating glyceraldehyde-3-phosphate dehydrogenase among gram-positive bacteria. Int Microbiol 2005, 8: 251–258.
14 Buhay JE, Moni G, Mann N and Crandall KA. A molecular taxonomy in the dark: evolutionary history, phylogeography, and diversity of cave crayfish in the subgenus Aviticambarus, genus Cambarus. Mol Phylogenet Evol 2007, 42: 435–448.
15 Lawrence JG, Hartl DL and Ochman H. Molecular considerations in the evolution of bacterial genes. J Mol Evol 1991, 33: 241–250.
16 Schlaeper BS and Zuber H. Cloning and sequencing of the genes encoding glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase, and triose phosphate isomerise (gap operon) from mesophilic Bacillus megaterium: comparison with corresponding sequences from thermophilic Bacillus species. Gene 1992, 122: 53–63.
17 Liaud MF, Valentin C, Martin W, Bouget FY, Kloareg B and Cerff R. The evolutionary origin of red algae as deduced from the nuclear genes encoding cytosolic and chloroplast glyceraldehyde-3-phosphate dehydrogenases from Chondrus crispus. J Mol Evol 1994, 38: 319–327.
18 Brown EW, Davis RM, Gouk C and Zwet TV. Phylogenetic relationships of necrogenic Erwinia and Brenneria species as revealed by glyceraldehyde-3-phosphate dehydrogenase gene sequences. Int J Syst Evol Microbiol 2000, 50: 2057–2068.
19 Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G and Zumla A. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. Biotechniques 2004, 37: 112–119.
20 Rubie C, Kempf K, Hans J, Su T, Tilton B, Georg T and Brittner B, *et al*. Housekeeping gene variability in normal and cancerous colorectal, pancreatic, esophageal, gastric and hepatic tissues. Mol Cell Probes 2005, 19: 101–109.
21 Ettahiri O, Berraho A, Vidy G, Ramdani M and Do chi T. Observation on the spawning of Sardina and Sardinella off the south Moroccan Atlantic coast (21–26°N). Fish Res 2003, 60: 207–222.
22 Baibai T, Oukhattar L, Moutaouakkil A and Soukri A. Purification and characterization of glyceraldehyde-3-phosphate dehydrogenase from European Pilchard *Sardina pilchardus*. Acta Biochim Biophys Sin 2007, 39: 947–954.
23 Oukhattar L, Baibai T, Moutaouakkil A, Assobhei O and Soukri A. Isolation and characterization of glyceraldehyde-3-phosphate dehydrogenase from the common octopus (*Octopus vulgaris* Cuvier, 1797). Rev Fish Biol Fisheries 2007, 18: 263–271.
24 Hedrick JL and Smith AJ. Size and charge isomer separation and estimation of molecular weights of proteins by disc gel electrophoresis. Arch BiochemBiophys 1968, 126: 155–164.
25 Cleland WW. The kinetics of enzyme catalysed reaction with two or more substrates or products, nomenclature and rate equations. Biochim Biophys Acta 1963, 67: 104–137.
26 Chomczynski P and Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenolchloroform extraction. Anal Biochem 1987, 162: 156–159.
27 Sambrook J, Fritsch EF and Maniatis T. Molecular Cloning, A Laboratory Manual. New York: Cold Spring Harbor Laboratory Press, 1989, 1659.
28 Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 1999, 41: 95–98.
29 Thompson JD, Gibson TJ, Plewnisk F, Jeanmougin F and Higgins DG. The CLUSTAL X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997, 25: 4876–4882.
30 Tamura K, Dudley J, Nei M and Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 2007, 24: 1596–1599.
31 Strimmer K and von Haeseler A. Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. Mol Biol Evol 1996, 13: 964–969.
32 Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985, 39: 783–791.
33 Mounaji K, Erraiss N, Iddar A, Wegnez M, Serrano A and Soukri A. Glyceraldehyde-3-phosphate dehydrogenase from the newt Pleurodeles waltl. Protein purification and characterization of a GapC gene. Comp Biochem Physiol B Biochem Mol Biol 2002, 131: 411–421.
34 Schwede T, Kopp J, Guex N and Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 2003, 31: 3381–3385.
35 Guex N and Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 1997, 18: 2714–2723.
36 Soukri A, Valverde F, Hafid N, Elkebbaj MS and Serrano A. Characterization of muscle glyceraldehyde-3-phosphate dehydrogenase isoforms from euthermic and induced hibernating Jaculus *orientalis*. Biochim Biophys Acta 1995, 1243: 161–168.
37 Fourrat L, Iddar A and Soukri A. Purification and characterization of cytosolic glyceraldehyde-3-phosphate dehydrogenase from the dromedary camel. Acta Biochim Biophys Sin 2007, 39: 148–154.
38 Heinz F and Freimuller B. Glyceraldehyde-3-phosphate dehydrogenase from human tissues. Methods Enzymol 1982, 89: 301–305.
39 Rzhetsky A and Nei M. A simple method for estimating and testing minimum evolution trees. Mol Biol Evol 1992, 9: 945–967.
40 Nei M and Kumar S. Molecular Evolution and Phylogenetics. New York: Oxford University Press, 2000, 300 pp.
41 Hurtado JL, Borderias J, Montero P and An H. Characterization of proteolytic activity in octopus (Octopus vulgaris) arm muscle. J Food Biochem 1999, 23: 469–483.
42 Thompson ST, Cass KH and Stellwagen E. Blue dextransepharose: an affinity column for the dinucleotide fold in proteins. Proc Natl Acad Sci USA 1975, 72: 669–672.
43 Hafid N, Valverde F, Villalobo E, Elkebbaj MS, Torres A, Soukri A and Serrano A. Glyceraldehyde-3-phosphate dehydrogenase from Tetrahymena pyriformis: enzyme purification and characterization of a gap C gene with primitive eukaryotic features. Comp Biochem Physiol B Biochem Mol Biol 1998, 119: 493–503.
44 Marangos PJ and Constantinides SM. Multiple forms of flounder muscle glyceraldehyde-3-phosphate dehydrogenase. Subunit composition, properties, and tissue distribution of the forms. J Biol Chem 1974, 249: 951–958.
45 Soukri A, Hafid N, Valverde F, ElKebbaj MS and Serrano A. Evidence for a posttranslational covalent modification of liver glyceraldehyde-3-phosphate dehydrogenase in hibernating jerboa (*Jaculus orientalis*). Biochim Biophys Acta 1996, 1292: 177–187.
46 Serrano A, Mateos MI and Losada M. ATP-driven transhydrogenation and ionization of water in a reconstituted glyceraldehyde-3-phosphate dehydrogenase (phosphorylating and non-phosphorylating) model system. Biochem Biophys Res Commun 1993, 197: 1348–1356.
47 Storey KB and Hochachka PW. Squid muscle glyceraldehyde-3-phosphate dehydrogenase: control of the enzyme in a tissue with an active a-glycero-P cycle. Comp Biochem Physiol B 1975, 52: 179–182.
48 Kier WM and Schachat FH. Muscle specialization in the squid motor system. J Exp Biol 2008, 211: 164–169.