December 2010 ISSN : 0974 - 7435 Volume 4 Issue 4

Trade Science Inc.

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 4(4), 2010 [179-189]

# Comparative molecular analysis of evolutionary distant glyceraldehyde-3-phosphate dehydrogenases from two marine species of fisheries interest, *Sardina pilchardus* and *Octopus vulgaris*

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

[1]Laboratoire de Physiologie et Génétique Moléculaire, Université Hassan II, Faculté des Sciences Aïn Chock, Km 8 route d'El Jadida, BP. 5366 Mâarif, Casablanca, (MORROCCO)

[2]Laboratoire de Biotechnologies Marines et Environnement, Faculté des Sciences, Université Chouaib Doukkali, B.P. 20 El Jadida, (MORROCCO)

[3]Laboratoire de Biochimie et Biologie Moléculaire, Université Hassan II, Faculté des Sciences Aïn Chock, km 8 route d'El Jadida BP. 5366, Mâarif, Casablanca, (MOROCCO)

[4]Instituto de Bioquímica Vegetal y Fotosíntesis (CSIC-University of Seville), Seville, (SPAIN)

E-mail : a_soukri@hotmail.com; tarikbaibai@yahoo.com

*Received: 12th May, 2010 ; Accepted: 22nd May, 2010*

## ABSTRACT

The $NAD^+$ dependent cytosolic glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.12), a key enzyme of central carbon metabolism and bioenergetics, has been purified to homogeneity from muscle tissues of two distinct marine species, skeletal muscle of *Sardina pilchardus* Walbaum (Teleost, Clupeida) and arm incompressible muscle of *Octopus vulgaris* (Mollusca, Cephalopoda). Comparative biochemical studies revealed that the two proteins differ by their subunit molecular masses, p*I* values and exhibit some specific catalytic features. Partial cDNA sequences corresponding to an internal region of the *GapC* genes from Sardina and Octopus were obtained by the polymerase chain reaction using degenerate primers constructed from highly conserved protein motifs. Alignments of deduced amino acid sequences were used to establish the 3D active site structures of the two enzymes as well as the phylogenetic relationships of the sardine and octopus enzymes, which are the first GAPDHs characterized so far from a teleost fish and a cephalopod, respectively. Interestingly, phylogenetic analyses place the sardine GAPDH in a cluster with the archetypical enzymes from other vertebrates, while the octopus GAPDHs comes together with other molluscan sequences, in a distant basal assembly closer to their bacterial and fungal homologs.



## KEYWORDS

Glyceraldehyde-3-phosphate dehydrogenase;
cDNA;
*GapC* gene;
RT-PCR;
Molecular phylogeny;
*Octopus vulgaris*;
*Sardina pilchardus*.

## INTRODUCTION

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.12), is one of the most studied enzymes in the glycolytic pathway, which reversibly catalyses the oxidative phosphorylation of D-glycer-

aldehyde-3-phosphate to form 1,3-diphosphoglycerate in the presence of the $NAD^+$ and inorganic phosphate[15]. This enzyme is widely distributed in nature in a variety of species ranging from bacteria to humans[11]. It is found mainly in the cytosol, and in some organelles (i.e., chloroplasts). Organellar GAPDHs are encoded by nuclear genes as precursor polypeptides and post-translationally imported into the organelles[6]. In addition to its well characterized glycolytic activity, a housekeeping function essential for the normal metabolism of all cells, GAPDH plays a pivotal role in the Embden-Meyerhoff pathway also in gluconeogenesis[11]. This enzyme has been well characterized not only because of its central role in the central metabolism, but also because of its abundance, ease of preparation and its conservation during evolution, being a protein with a native molecular mass in the range of 140-200 kDa and composed of four identical subunits of approximately 35–50 kDa[11,19]. The ubiquity and evolutionary conservation of this enzyme indicate 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[38]. These include its roles in membrane fusion, phosphotransferase activity, replication and DNA repair[2], specific binding to 3' and 5' regions of mRNA[34], nuclear RNA export activity[37], neuronal apoptosis[32], and in neurodegenerative diseases[24] and prostate cancer[28]. These novel activities may be related to the sub-cellular localization and oligomeric structure of GAPDH in vivo.

The glycolytic pathway is particularly suitable for testing theories 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 the most highly conserved of all glycolytic enzymes, for instance the rate of evolution of the catalytic domain is only 3% per 100 million years[11]. Thus, these domains in eukaryotic and eubacterial enzymes are >60% identical. Therefore GAPDHs genes have often been used as phylogenetic markers for "deep" phylogenies[4,5,18], as a prototype for studies of genetic organization, expression and regulation and to describe the taxonomic positions of several species at different levels[21,22,33], in addition to its utilization as endogenous controls or reference in appropriate for most experimental conditions[3,8,29].

No information was available to date on the structure of the gene encoding this dehydrogenase in clupeida and cephalopoda, and on their phylogenetic relationships with other GAPDHs. In this paper we compare kinetic and biomolecular parameters of GAPDH purified from the skeletal muscle of *Sardina pilchardus* and the incompressible arm muscle of *Octopus vulgaris* seeking some distinguishing characteristics of other GAPDHs. cDNA fragments of the corresponding *GapC* genes were amplified by polymerase chain reaction techniques, sequenced and identified as the internal region of these genes containing the catalytic site. The phylogenetic relationships of the Sardine and Octopus GAPDHs with homologous dehydrogenases of other vertebrate and invertebrate species are discussed. Our results show that these two enzymes have remarkable differences; hence GAPDH may be considered a suitable molecular marker of fisheries interest useful for marine species traceability.

## MATERIALS AND METHODS

### Biological material

#### a) European pilchard: *Sardina pilchardus*

The European pilchard *Sardina pilchardus* is one of the most important species of small pelagic fishes of the northwest coast of Africa[9]. Its distribution extends from Cap Blanc (21°N) to Cap Spartel (35°45'N). The samples are originated from the continental platform of the Atlantic Ocean, near the Moroccan coastline. Animals used in this study were purchased from fishermen operating small fishing vessels, in the coast line of Casablanca (33°N) and they were immediately frozen at -20°C until use.

#### b) Octopus: *Octopus vulgaris*

Adult octopuses, *Octopus 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 our laboratory in boxes with dry ice in an isothermal truck within 6 to 8 h of capture.

### Purification and characterization of sardine and octopus GAPDHs

GAPDH has been purified to electrophoretic homogeneity from a soluble protein fraction of Sardina pilchardus skeletal muscle and arms of Octopus vulgaris using a simple procedure involving only one column chromatography step, namely dye-affinity chromatography on Cibacron blue-Sepharose 4B (Amersham Pharmacia Biotech). The specific characterization of the enzymes purified from the two species was described elsewhere[1,27]

### RNA isolation, RT-PCR methodology and DNA sequencing of the GAPDHs from *Sardina pilchardus* and *Octopus vulgaris*

Total RNA was isolated from sardina skeletal muscle and octopus arm muscle previously frozen at -20°C, using the method of Chomczynski and Sacchi[7]. First-strand cDNA was generated by reverse transcription (RT) of total RNA (3µg), using 200 units of reverse transcriptase (MMLV) Moloney Murine Leukemia Virus Transcriptase (Promega) and the reverse primer named Gap2; 5'-CCSCAYTCRTTRTCRTACCA-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 during 1 h at 42°C. An aliquot from this template (1/10 of the reaction volume) was used in a subsequent polymerase chain reaction (PCR) using 1.25U Go Taq DNA polymerase (Promega), 0.04 µM of reverse (Gap2) and forward primer Gap1; 5'-GCYWSYTGYACSACSAAYTG-3'. Amplifications of cDNA fragments (ca. 0.5 kb) corresponding to internal regions (70% of the full ORF) of *GapC* genes from sardine and octopus were carried out by PCR using the same degenerate primers constructed from two highly conserved motifs (ASCTTNC, WYDNEW(C)G) present in all GAPDHs so far studied[11]. 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 with the addition of ethidium bromide according to Sambrook[31], and purified using a phenol/chloroform protocol. They exhibited the expected size (approximately 0.5 kb) for an internal GapC gene fragment comprising approximately the half of the complete coding region.

Sequence reactions were performed by using the second PCR-amplified cDNA, primers Gap1 and Gap2 described above (forward and reverse primers), and a BigDye Terminator Cycle Sequence Ready Reaction Kit version 1.0 (Applied Biosystems, Foster City, CA). The following PCR conditions were used: an initial denaturation step at 94ºC for 5 min, followed by 30 cycles of denaturation at 94ºC for 10 s, annealing at 50ºC for 10 s, and extension at 62ºC for 4 min. After the sequence reaction, excess dye terminators were removed by gel filtration in Centri-Sep spin column (Applied Biosystems). The purified DNA samples were dried and suspended in the template suppression reagent (TSR) or a mix of formamide and Blue Dextran from Applied Biosystems. The dissolved DNA samples were heated at 95ºC, 2 min for denaturation, and then immediately cooled on ice. Sequences were analyzed with an automated DNA sequencer ABI Prism 377 from Applied Biosystems. Complete sequences were aligned, assembled, and compared by eye with sequences of GAPDHs of other species for verification using BioEdit[14].

### Nucleotide sequence accession numbers

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

### Sequence alignment, phylogenetic analysis and active-site modelling of *S. pilchardus* and *O. vulgaris* GAPDHs

Multiple sequence alignment of GAPDH protein regions corresponding to the cDNA fragments of Sardina and Octopus GapC was done with the CLUSTAL X v.1.8 program[44]. Through this alignment, phylogenetic trees were constructed using the distance Neighbor-Joining algorithm (Kimura distance calculations) and Minimum Evolution method, as well as the Maximum Parsimony method with the programs MEGA4[43], TREE-PUZZLE v.5.2[42] and PROTPARS v.3.573c (PHYLIP package v.3.5c w1993x Felsenstein, J., Dept. of Genetics, Univ. of Washington, Seattle,

USA)[10]. Bootstrap analyses (values being presented on a percentage basis) were computed with 10000 replicates, and all positions containing alignment gaps and missing data were eliminated only in pair wise sequence comparisons (Pairwise deletion option), for the distance trees. For maximum parsimony method there were a total of 175 positions in the final dataset, 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 quarter puzzling[42].

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). A bacterial GAPDH encoded by the *Escherichia coli gap1* gene (accession number P06977) and a fungal homolog encoded by the *Saccharomyces cerevisiae GAPDH1* gene (P00360) were also included as outgroups.

The 3D structures of *Sardina pilchardus* and *Octopus vulgaris* GAPDHs are unknown, although the structures of several GAPDHs have been reported up to date. As GAPDHs are among the most conserved proteins sharing high degrees of primary sequence similarity[11], we subsequently generated homology models of both Sardina and Octopus GAPDHs active sites based upon known 3D structures of GAPDHs from others species, using Swiss Model[35], and the program SPDBV37SP5. These models were subjected to spatial motif searches with the aim of identifying most amino acids conserved in the active site of this enzyme during evolution.

## RESULTS

### Kinetic properties of GAPDHs purified from muscle tissue of Sardina and Octopus

GAPDH has been purified to electrophoretic homogeneity from muscle soluble protein fraction of the european pilchard, *Sardina pilchardus* and common octopus, *Octopus 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 35 U/mg of protein was obtained for the specific activity of purified Sardina GAPDH with a yield of 25 % and a purification factor of approximately 77 folds, while for the octopus enzyme a specific activity value of approximately 9.2 U/mg of protein and a purification factor of about 26 fold were obtained. SDS-PAGE analysis of the different fractions obtained throughout the purification procedure showed a progressive enrichment in a 37 kDa and 36 kDa protein band corresponding to the GAPDH subunit for Sardina and Octopus, respectively (Figure 1). Only this protein band was seen in the electrophoretically homogeneous final enzyme preparations, thus confirming they actually correspond to the GAPDH subunit (Figure 1). Non-denaturing PAGE showed that the native molecular mass of obtained proteins are approximately 155 and 153 kDa for Sardina and Octopus, respectively. Isoelectric focusing of the purified proteins showed a single band with very different pI values for the two species, namely 7.9 for Sardina and 6.6 for Octopus (Figure 2). Therefore, the fish GAPDH is a basic protein according to previous reports on other animal GAPDHs[12,16,25,40], whereas its molluscan homolog is unexpectedly an acidic

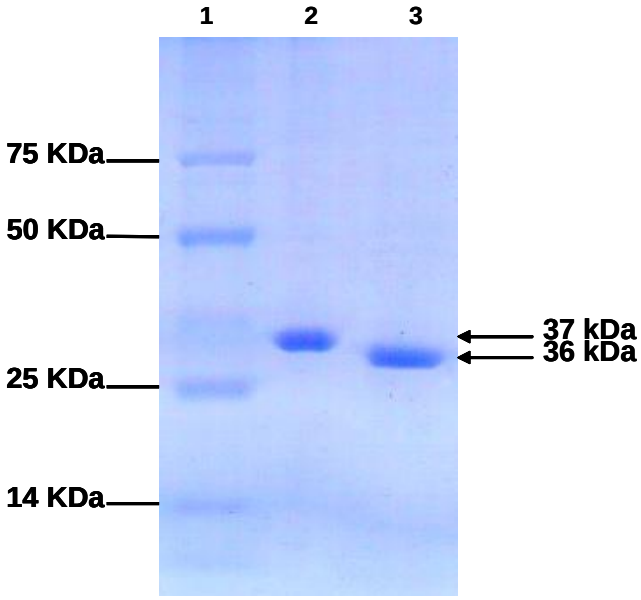


**Figure 1: Comparison of GAPDHs subunit molecular masses from *Sardina pilchardus* skeletal muscle and *Octopus vulgaris* arms. Purified proteins were run on 12% SDS-PAGE and gels were stained with Coomassie Brilliant Blue. Lanes 2 and 3 show pure GAPDHs (25 µg) from sardina and octopus, respectively.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

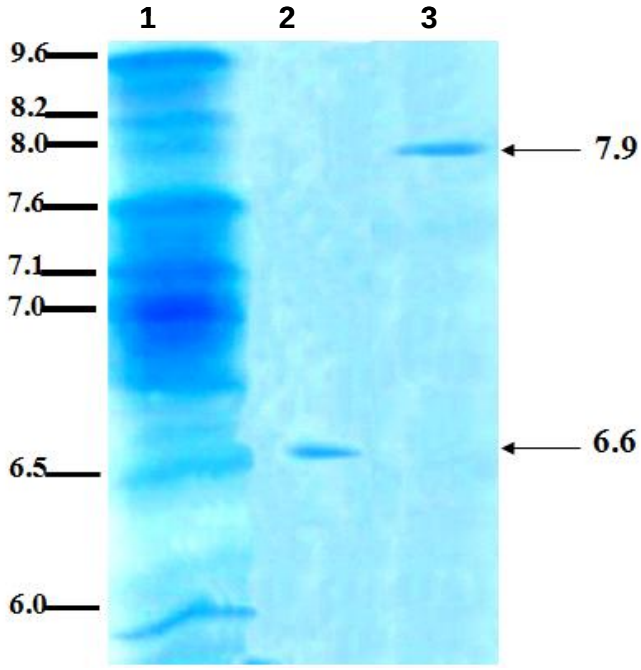


**Figure 2: Isoelectric focusing of *Sardina pilchardus* and *Octopus vulgaris* GAPDHs. Isoelectric focusing was performed on 5% (w/v) acrylamide gel holding ampholyte-generated pH gradients (pH range, 3.5 – 10). Lane 1 corresponds to isoelectric focusing protein markers (p*I* range, 4 – 9.6). Lanes 2 and 3 corresponds to pure GAPDH from octopus (p*I* 6.6) and sardina (p*I* 7.9) respectively.**

protein like most microbial GADPHs[18,21]. 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 constant concentration of the other, were respectively 73.4 and 92 µM, for Sardina, and 320 and 66µM, for Octopus (TABLE 2). The $V_{max}$ values calculated for Sardina and Octopus GAPDHs were 37,6 U/mg and 21,8 U/mg, respectively. The catalytic efficiencies, expressed in terms of $V_{max}/K_m$, indicate that the Octopus enzyme is less efficient with both substrates.

## Sequence alignment, phylogenetic analysis and active-site modelling

RT-PCR amplification using primers constructed from two highly conserved GAPDH regions produced a single cDNA fragment of the expected size (approx. 0.5 kb) comprising approximately half of the coding region of a GapC gene for the two species. The nucleotide sequence determined for the amplified cDNA fragment (507 and 525 bp respectively for Sardina and Octopus were deposited in the GenBank/EMBL databases with the accession numbers: EF621524 and EF634059). The cDNA obtained from two independent RT-PCR experiments for each species were sequenced and found to be identical. The derived amino acid sequence corresponds to a highly conserved region around the catalytic site including many residues strictly conserved in GAPDHs from very diverse organisms[11]. 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 and two model bacterial and fungal species by using the CLUSTAL X (v. 1.8) program[44] (Figure 3). Relatively high percent identity (71 %) and similarity (79%) values were found between the amino acid sequences of GAPDH proteins of the two marine species. The above described multiple sequence alignment was used to construct phylogenetic trees or cladograms obtained with the MEGA and Tree-Puzzle programs to infer the evolutionary relationship of the Sardine and Octopus GAPDHs using distance methods (Neighbor-Joining, Minimum Evolution), and Maximum Parsimony and Likelihooh methods[26,30,43], (Figure 4. A, B and C). All consensus phylogenetic trees yielded similar topologies, and show dis-

FULL PAPER

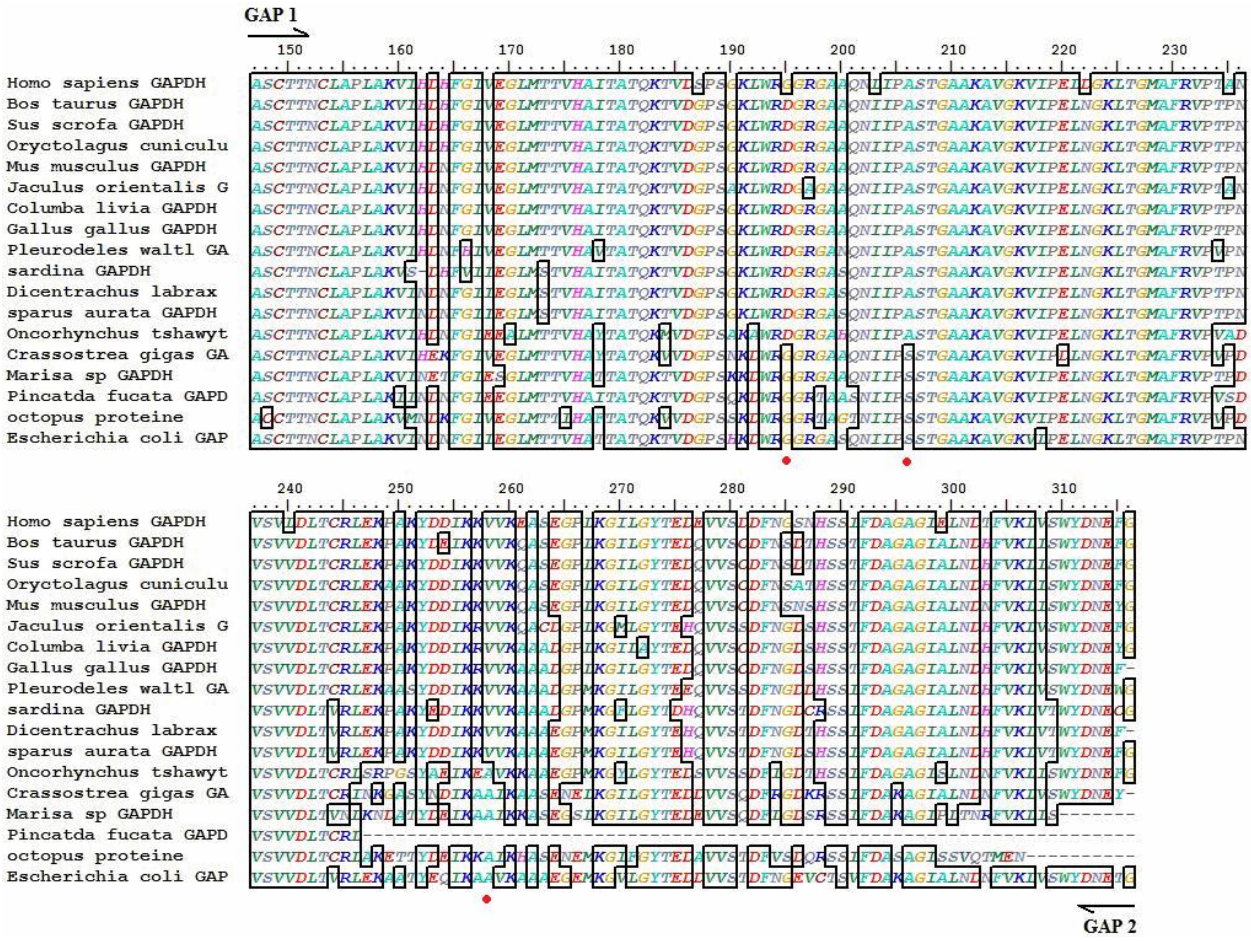


**Figure 3: Multiple sequence alignment of *Sardina pilchardus* and *Octopus vulgaris* GAPDHs compared with 16 GAPDHs, selected to include representatives of the main aquatic and terrestrial vertebrate phyla and one bacterial species, by using the CLUSTAL X (v. 1.8) program (Thompson et al., 1997).The 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.**

tant evolutionary positions for the Sardina and Octopus GAPDHs. The Sardina protein conform a cluster with other teleost fishes homologs within a well supported assembly of GAPDHs of other main vertebrate groups, while the Octopus protein arrange with other molluscan homologs in a rather basal assembly near to microbial (fungal and bacterial) GAPDHs. GAPDH paralogs of other mayor invertebrate groups display in between these vertebrates and molluscan assemblings. Interestingly, a second cluster of teleost fishes GAPDHs is also found, probably due to the ocurrence of enzyme isoforms, as was reported for many vertebrates[12,25,40].

Sardina and Octopus GAPDHs shared a sequence identity of 86% and 75% with the Homo sapiens and 33% and 32% with Escherichia coli GAPDH respectively. The active site 3D structures of both Octopus and Sardina GAPDHs were based in the structure of the templates of this two species, as they fulfil both the criteria of high percent sequence identity and high resolution of the model determined structure. The holoforms have been determined at 1, 75 and 1.8 Å, respectively for Sardina and Octopus enzymes. The final models were elaborated with the SPDBV37SP5 program. (Figure 5)

## 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

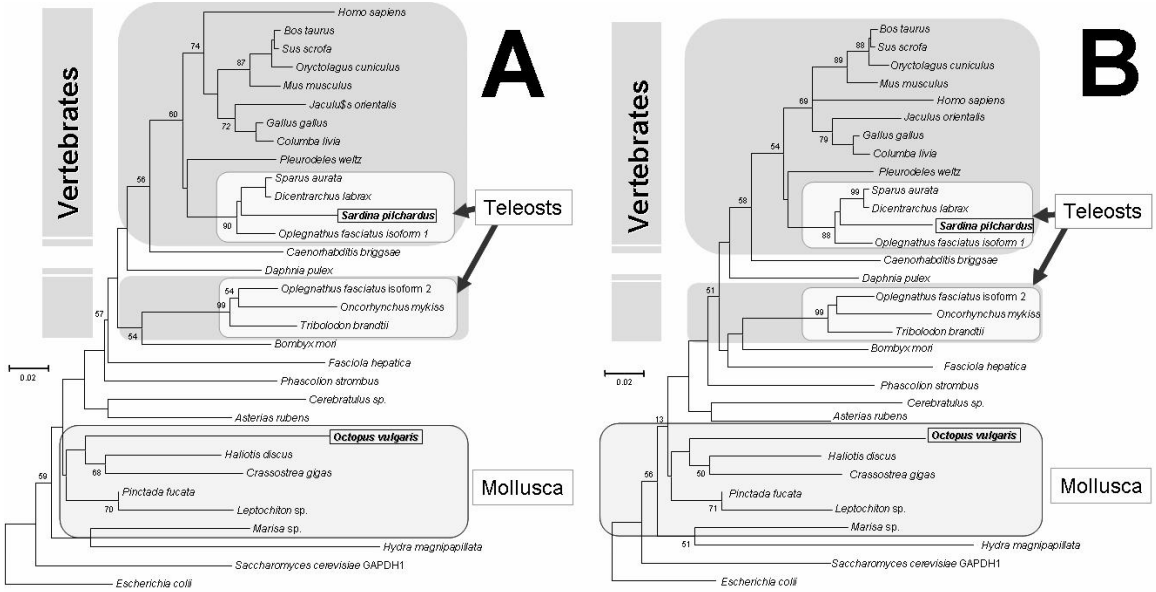


**Figure 4: A and B. A: Evolutionary relationships of sardine and octopus GAPDHs obtained using the Neighbor-Joining method. Selected GAPDH sequences representing the main vertebrate and invertebrate phylogenetic groups were used. A fungal (S. cerevisiae) and an bacterial (E. coli ) GAPDH sequences are shown as outgroups. The bootstrapped consensus distance tree was inferred from 10000 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 dataset. Phylogenetic analyses were conducted in MEGA4. B: Evolutionary relationships of sardine and octopus GAPDHs obtained using the Minimum Evolution 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 algorithm was used to generate the initial tree. Other parameters as above indicated. Phylogenetic analyses were conducted in MEGA4.**

NAD+ and inorganic phosphate. In this work, GAPDH from Sardina pilchardus skeletal muscle and arm muscle of Octopus vulgaris were purified to electrophoretic homogeneity from a soluble protein fraction. TABLE 1 summarizes a representative purification protocol for the two species. A value of approximately 35 units/mg of protein was obtained for the specific activity of Sardina purified GAPDH with a yield of 25 % and a purification factor of approximately 77 fold. While for the octopus one, a value of approximately 9.2 U/mg of protein was obtained for the specific activity and a purification factor of about 26 fold. This difference in specific activity and factor of purification between the two species may be explained by the short shelf life of Octopus muscle, due to its high autolytic activity, 25 times greater than in gadoids and 3 times greater than in squid[17]. Dye-affinity chromatography on Blue Sepharose seems to be effective for the purification of Sardina and Octopus GAPDHs, as for other NAD+ dependent GAPDH purifications[1,13,45]. No additional purification steps were required to obtain homogenous GAPDH samples. As stated above, SDS-PAGE of the purified GAPDHs showed a single protein bands corresponding to 36 and 37 kDa, in both species Octopus sp and Sardina.sp, This results, compared with the natives molecular weights (154 kDa and 153kDa), suggests that the enzymes has a homotetrameric structure like the majority of other GAPDHs so far studied[19,25,40]. However, the Sardina enzyme subunit exhibited an estimated molecular weight of 37 kDa that is some what higher than the one reported for bacterial species and mollusk species (35 or 36 kDa), but it's very near to value reported for the pleurodel specie and mammalian species like human or jerboa GAPDH[13,16].

The pI value was estimated by the isoelectric focusing technique. A single protein band was detected with different values of pI for the two species, respectively 7.9 for Sardina. sp and 6.6 for the Octopus sp, the purified GAPDHs were found to be homogeneous. This result indicates that only one isoform of the enzyme, although with quite different molecular properties, occurs in the muscle of each one of the two species, strongly suggesting that a single GapC gene encoding a protein with differences in some key amino

FULL PAPER

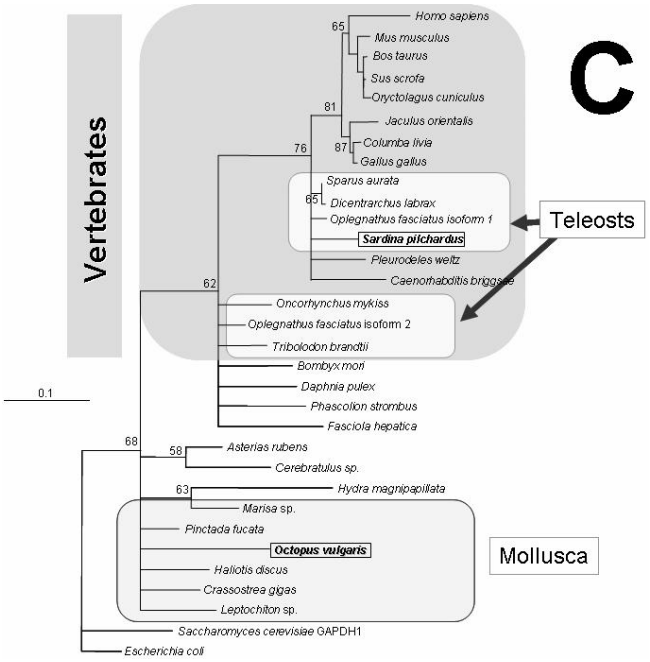


**Figure 4 C : Evolutionary relationships of sardine and octopus GAPDHs obtained by Maximum Likelihood analysis. Tree reconstruction was performed with the Quartet puzzling procedure and parameters estimated using Quartet sampling on a NJ tree. Bootstrap analysis of 10,000 re-samplings was performed 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.**

acid residues (i.e., catalytic site region) is expressed in these tissues. A single GAPDH isoform has been also found in other animal tissues and microorganisms, both prokaryotes and eukaryotes[19,23,25]. However, it does not seem to be a general rule, as the presence of several GAPDH isoforms has been reported in phylogenetically very different organisms[16,39]. Concerning the kinetics parameters of GAPDH, the Km values for D-G3P and NAD+, have been determined, being respectively 73.4 and 92 μM, for Sardina, and 320 and 66μM, for Octopus (TABLE 2). These values are comparable to those found for cytosolic GAPDHs purified from other eukaryotes, both protists (Tetrahymena pyriformis)[13], lower metazoa, like the mollusk Loligo vulgaris[41], or mammalians like Homo sapiens or Jaculus orientalis[16,40]. However, on the whole the kinetic parameter analyses show that the Octopus enzyme is catalytically less efficient. This indicates possible differences in the mechanism of the catalytic reaction. The kinetic parameter values obtained for the two marine species, Sardina and Octopus, GAPDHs differ in a number of instances from those described previously for GAPDHs from other sources[12,13,16,40], reflecting protein differences between species.

## Sequences alignment, phylogenetic analysis and active site modelling

The two cDNA obtained from independent RT-PCR experiments for each species were sequenced and the derived amino acid sequence corresponds to a highly conserved region of the catalytic subunit including many residues strictly conserved in GAPDHs from very diverse organisms[11]. These sequences were aligned and compared with other selected GAPDHs representing main vertebrates and invertebrates Phylogenetic groups. However, the limited information available on molluscan GAPDHs 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[20]. The GAPDHs sequence alignment of the two marine species, *Sardina pilchardus* and *Octopus vulgaris* sharing high percent identity. The conservation is slightly higher in the catalytic domain between 147 and 320 Amino acids, corresponding to the most conserved region for all GAPDHs so far studied especially the histidine 176 involved in the catalysis reaction (Figure 3). The minor differences between the two sequences include different content and distribution of several amino acids like leucine 175, glycine 195, threonine 198 and 201 in the amino acids sequence of octopus that's not present in the other sequences. On the other hand, the *Sardina pilchardus* GAPDH sequence comprises two other histidine residues (162 and 164) that are not found in the *Octopus vulgaris* sequence, and the presence of the amino acid serine in the position 172 seems to characterize the group of teleost fishes, but the most important remark is the identity of three amino acids (Aspartate, Glycin and Serine respectively at the positions 192, 195 and 206) not only in the sequence of *Octopus vulgaris*, and other mollusks sequences (*Crassostrea gigas*, *Marisa* sp. and *Pinctada fucata*), but also in

the sequence of *Escherichia coli*.

The phylogenetic trees obtained shows analogous evolutionary relationships among the GAPDHs of teleost fishes and molluscan sequences and reveal that while the Sardina GAPDH is closely related to homologs of other teleost fishes and near to the group of amphibian enzymes while the molluscan GAPDHs, including the Octopus enzyme conform a distant more deeply 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 reported in this paper. However, it remains to be clarified if this is actually due to true phylogenetic relationships among anciently-diverged phylogenetic clades or to 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, reported among other GAPDH-based phylogenies[6,11,13], cannot be ruled out. In any case, the comparative study presented in this work on purified proteins and novel gene sequences will shed new light on the functional and evolutionary relationship of the GAPDHs from representative species of two main groups of marine animals. Current work is currently underway in our laboratories to go further on this research.

Sardina and Octopus GAPDHs shared a relatively high sequence identity with the *Homo sapiens* and *Escherichia. coli* GAPDHs, respectively, in accordance with the high conservation between homologs from all phyla ranging from bacteria to vertebrate. The crystal structures of both Octopus and Sardina GAPDHs were based in the structure of the templates of this two species, as they fulfil both the criteria of high percent sequence identity and high resolution structure determi-

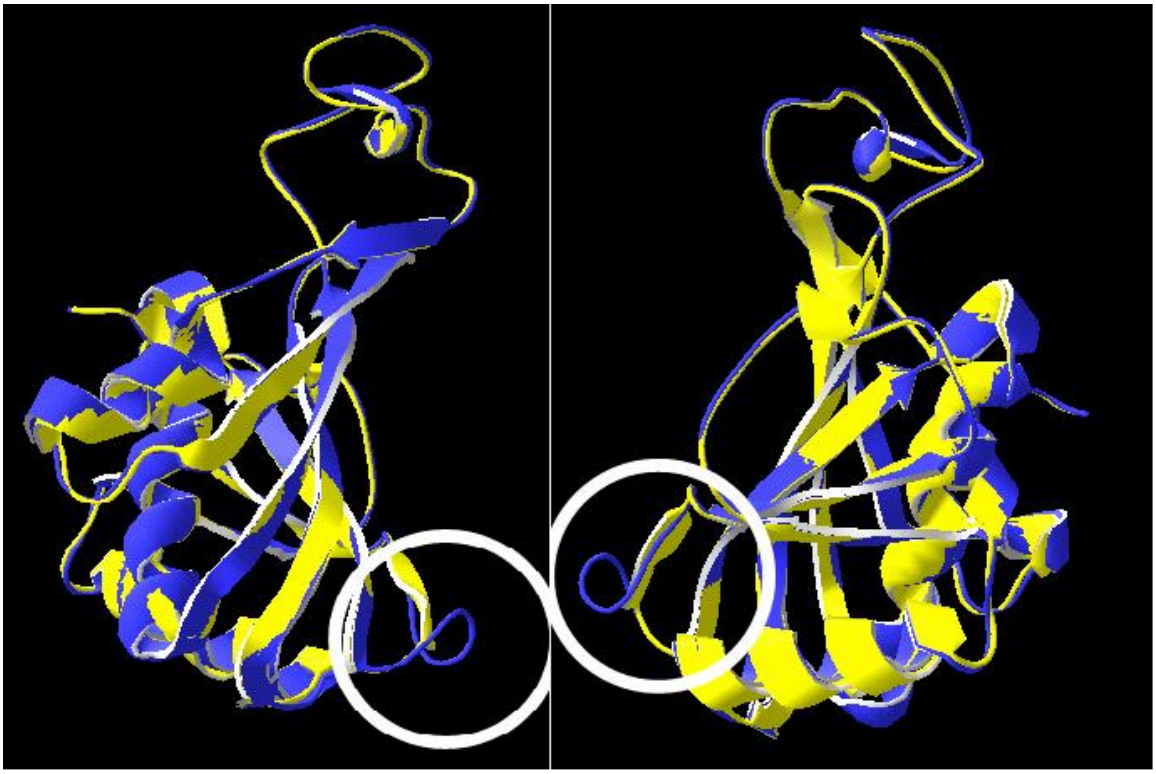

**Figure 5: Stereoviews of the modeled GAPDH structures of *Sardina pilchardus* and *Octopus vulgaris* obtained from the secondary structure alignment. The back bone of sardina and octopus GAPDHs are shown as gold and blue ribbons, respectively. Note the unordered extra lop of the octopus GAPDH (white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

nation. The quality of the final models reflects the sequence identity shared between *Octopus vulgaris* and *Sardina pilchardus* GAPDHs with exception in a small region which seems not located in the catalytic site. The final models corresponding to one subunit of the holoenzyme is shown in Figure 5.

## CONCLUSION

The $NAD^+$-dependent cytosolic glyceraldehyde-3-phosphate dehydrogenase (GAPDH, EC 1.2.1.12) has been purified to homogeneity from muscle tissues of *Sardina pilchardus* and arm muscle of *Octopus vulgaris*. Comparative studies revealed that the two proteins differ by their subunit molecular masses, p*I* values and kinetics 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 GAPDHs of these two marine species, and suggest their possible uses as molecular markers for seafood analysis and marine species traceability.

## ACKNOWLEDGMENTS

The authors thank greatfully Dr. L. Fourat for his help during the realization of the present work. This work was supported by Genosphere Biotechnology and Novamics Biosciences Company Researches and the Morocco – Spain cooperation project (AECID, N° A/8870/07). Thanks are also due to Junta de Andalucia PAIDI (group BIO-261) for support.

## REFERENCES

[1] T.Baibai, L.Oukhattar, A.Moutaouakkil, A.Soukri; Acta Biochimica et Biophysica Sinica, **39(12)**, 947-954 **(2007)**.
[2] M.D.Baxi, J.K.Vishwanatha; Biochemistry, **34**, 9700-9707 **(1995)**.
[3] E.W.Brown, R.M.Davis, C.Gouk, T.V.Zwet; International Journal of Systematic and Evolutionary Microbiology, **50**, 2057-2068 **(2000)**.
[4] J.E.Buhay, G.Moni, N.Mann, K.Crandall; Molecular Phylogenetics and Evolution, **42**, 435-448 **(2007)**.
[5] B.Canback, S.G.Andersson, C.G.Kurland; Proc. Natl.Acad.Sci.USA, **99**, 6097-6102 **(2002)**.
[6] R.Cerff; 'The Chimeric Nature of Nuclear Genomes and the Antiquity of Introns as Demonstrated by GAPDH Gene System'. In: M.Go, P.Schimmel, (Eds.), Tracing Biological Evolution in Protein and Gene Structures, Elsevier, Amsterdam, 205, 227 **(1995)**.
[7] P.Chomczynski, N.Sacchi; Anal.Biochem., **162**, 156-159 **(1987)**.
[8] K.Dheda, J.F.Huggett, S.A.Bustin, M.A.Johnson, G.Rook, A.Zumla; Biotechniques, **37**, 112-9 **(2004)**.
[9] O.Ettahiri, A.Berraho, G.Vidy, M.Ramdani, T.Do Chi; Fish Res, **60**, 207-222 **(2003)**.
[10] J.Felsenstein; Evolution, **39**, 783-791 **(1985)**.
[11] L.A.Fothergill-Gilmore, P.A.Michels; Prog.Biophys. Mol.Biol., **59**, R105-235 **(1993)**.
[12] L.Fourrat, A.Iddar, A.Soukri; Acta Biochimica et Biophysica Sinica, **39**, 148-154 **(2007)**.
[13] N.Hafid, F.Valverde, E.Villalobo, M.S.Elkebbaj, A.Torres, A.Soukri, A.Serrano; Comp.Biochem. Physiol.B, **119**, 493-503 **(1998)**.
[14] T.A.Hall; Nucl.Acids Symp.Ser., **41**, 95-98 **(1999)**.
[15] J.I.Harris, M.Waters; 'Glyceraldehyde-3-phosphate Dehydrogenase'. In: P.D.Boyer, (Ed.), The Enzymes, Third Ed., Academic Press, New York, 1-49 **(1976)**.
[16] F.Heinz, B.Freimuller; Methods Enzymol., **89**, 301-305 **(1982)**.
[17] J.L.Hurtado, J.Borderias, P.Montero, H.An; J.Food Biochem., **23**, 469-483 **(1999)**.
[18] A.Iddar, F.Valverde, O.Assobhei, A.Serrano, A.Soukri; International Microbiology, **8**, 251-258 **(2005)**.
[19] A.Iddar, F.Valverde, A.Serrano, A.Soukri; Prot.Exp.Purif., **25**, 519-526 **(2002)**.
[20] W.M.Kier, F.H.Schachat; The Journal of Experimental Biology, **211**, 164-169 **(2008)**.
[21] J.G.Lawrence, D.L.Hartl, H.Ochman; J.Molec. Evol., **33**, 241-250 **(1991)**.
[22] M.F.Liaud, C.Valentin, W.Martin, F.Y.Bouget, B.Kloareg, R.Cerff; J.Mol.Evol., **38**, 319±327 **(1994)**.
[23] P.J.Marangos, S.M.Constantinides; Biol.Chem., **249**, 951-958 **(1974)**.
[24] J.L.Mazzola, M.A.Sirover; J.Neurochem., **76**, 442-449 **(2001)**.
[25] K.Mounaji, N.Erraiss, A.Iddar, M.Wegnez, A.Serrano, A.Soukri; Comp.Biochem.Physiol.B

Biochem.Mol.Biol., **131**, 411-421 **(2002)**.

**[26]** M.Nei, S.Kumar; 'Molecular Evolution and Phylogenetics'. Oxford University Press, New York, **(2000)**.

**[27]** L.Oukhattar, T.Baibai, A.Moutaouakkil, O.Assobhei, A.Soukri; Rev.Fish.Biol.Fisheries, **(2007)**.

**[28]** R.H.Rondinelli, D.E.Epner, J.V.Tricoli; Prostate Cancer Prostatic Dis., **1**, 66-72 **(1997)**.

**[29]** C.Rubie, K.Kempf, J.Hans, T.Su, B.Tilton, T.Georg, B.Brittner, B.Ludwig, M.Schilling; Molecular, Cellular Probes., **19(2)**, 101-109 **(2005)**.

**[30]** A.Rzhetsky, M.Nei; Molecular Biology and Evolution, **9**, 945-967 **(1992)**.

**[31]** J.Sambrook, E.F.Fritsch, T.Maniatis; 'Molecular Cloning', a Laboratory Manual. Cold Spring Harbor Laboratory Press, New York, **(1989)**.

**[32]** P.A.Saunders, R.W.Chen, D.M.Chuang; J.Neurochem., **72**, 925-932 **(1999)**.

**[33]** B.S.Schlaeper, H.Zuber; Gene, **122**, 53-63 **(1992)**.

**[34]** D.E.Schultz, C.C.Hardin, S.M.Lemon; J.Biol. Chem., **271**, 14134-14142 **(1996)**.

**[35]** T.Schwede, J.Kopp, N.Guex, M.C.Peitsch; Nucleic Acids Res., **31**, 3381-3385 **(2003)**.

**[36]** A.Serrano, M.I.Mateos, M.Losada; Biochem. Biophys.Res.Comm., **197**, 1348-1356 **(1993)**.

**[37]** R.Singh, M.R.Green; Science, **259**, 365-368 **(1993)**.

**[38]** M.A.Sirover; Biochim.Biophys.Acta, **1432**, 159-184 **(1999)**.

**[39]** A.Soukri, N.Hafid, F.Valverde, M.S.Elkebbaj, A.Serrano; Biochim.Biophys.Acta, **1292**, 177-187 **(1996)**.

**[40]** A.Soukri, F.Valverde, N.Hafid, M.S.Elkebbaj, A.Serrano; Biochim.Biophys.Acta, **1243**, 161-168 **(1995)**.

**[41]** K.B.Storey, P.W.Hochachka; Comp.Biochem. Physiol.B, **52(1)**, 179-182 **(1975)**.

**[42]** K.Strimmer, A.von Haeseler; Mol.Biol.Evol., **13**, 964-969 **(1996)**.

**[43]** K.Tamura, J.Dudley, M.Nei, S.Kumar; Molecular Biology and Evolution, **24**, 1596-1599 **(2007)**.

**[44]** J.D.Thompson, T.J.Gibson, F.Plewnisk, F.Jeanmougin, D.G.Higgins; Nucleic Acid Res., **25**, 4876-4882 **(1997)**.

**[45]** S.T.Thompson, K.H.Cass, E.Stellwagen; Proc.Natl. Acad.Sci.USA, **72**, 669-672 **(1975)**.