RESEARCH ARTICLE

Gene structure of *Drosophila* diaphorase-1: diversity of transcripts in adult males and females, in different organs and at different stages of development

PAVLINA M. IVANOVA[1]*, BORIS H. DUNKOV[2] and KIRIL H. RALCHEV[1]

[1]*Sofia University St Kliment Ohridski, Faculty of Biology, Department of Genetics, 8 Dragan Tzankov Blvd., 1164 Sofia, Bulgaria*

[2]*University of Shoumen Konstantin Preslavsky, Faculty of Natural Sciences and Functional Biology, Shoumen 9700, Bulgaria*

Abstract

The gene *EG:22E5.5* or *CG4199* (accession number O77266, Q9W529) from Berkeley *Drosophila* Genome Project (BDGP) was found using the partial amino acid sequences of three tryptic peptides obtained from purified *Drosophila virilis* diaphorase-1. This gene is located on the X chromosome at position 2C9–2C10. The structure of the gene reveals three exons and two long introns. Using BDGP, we found six transcripts in this gene. The difference between these transcripts is in their 5′ ends; the 3′ ends of the six transcripts are identical. Thirty-four ESTs from different cDNA libraries were found, most of them from Schneider L2 cell culture (SH) cDNA library. The transcripts are represented at very low level in the cells of different organs and at different stages of *Drosophila* development. Using RT-PCR, we obtained five of these transcripts in cDNA samples from female adult flies. However, we could not find any of them in cDNA samples from male adult flies. Moreover, we obtained only the third transcript (CG4199-RC) in the sample of testis from adult flies and the fourth transcript (CG4199-RD) in an embryo sample. None of the other five transcripts were found in the samples of different organs and in the samples obtained at different stages of *Drosophila* development.

[Ivanova P. M., Dunkov B. H. and Ralchev K. H. 2008 Gene structure of *Drosophila* diaphorase-1: diversity of transcripts in adult males and females, in different organ and at different stages of development. *J. Genet.* 141–146]

Introduction

Eukaryotic genes display great diversity in their structure and mechanisms of regulation. Investigation of model systems can therefore make a significant contribution to the understanding of the genome as a whole. Study of the diaphorases (NAD(P)H: acceptor oxidoreductases; EC 1.6) (Enzyme nomenclature 1973) in *Drosophila*, as a gene–enzyme system may contribute in widening our knowledge in this sphere. The diaphorases are an ubiquitous class of flavin-bound enzymes that catalyze the reduction of various dyes which act as hydrogen acceptors from the reduced form of diphosphopyridine and triphosphopyridine nucleotides, i.e. NADH and NADPH (Onyenwoke and Wiegel 2007).

Previous investigations in our laboratory have shown that these enzymes in *Drosophila* are encoded by at least four structural genes, and differ in their specificity for nicotinamide cofactors. Diaphorase-1 and diaphorase-3 can oxidize NADH as well as NADPH, while diaphorase-2′ and diaphorase-2″ oxidize only NADH. Genes coding for diaphorase-1 and diaphorase-3 were localized on the X chromosome of *D. virilis* (Ralchev and Dunkov 1987). Diaphorase-1 and diaphorase-2 were isolated from two *Drosophila* species, *D. virilis* and *D. melanogaster*, and purified by gel filtration, affinity chromatography, immunoaffinity chromatography and ion-exchange chromatography. The molecular weights of both the enzymes were same in each species. The molecular weight of diaphorase-1 was same under both denaturating and nondenaturating conditions, close to 60 kD, indicating a monomeric structure (Ralchev *et al*. 1992). Diaphorase-2′ has been identified as

*For correspondence. E-mail: pavlyna@biofac.uni-sofia.bg.

Keywords. *Drosophila*; diaphorase-1; expressed sequence tags (EST); RT-PCR.

a dihydrolipoamide dehydrogenase with diaphorase activity (Ivanova *et al*. 2004). Diaphorase-2″ has been identified as a xantine dehydrogenase, also showing diaphorase activity (Dunkov *et al*. 1991) and is coded by the gene *rosy*.

In the present paper, we report on the gene structure of *CG4199* from *D. malanogaster*, the diversity of the gene transcripts in female and male adult flies, and the diversity of the transcripts in different organs collected from adult flies at different stages of *Drosophila* development. The product of *D. melanogaster* (*CG4199*) gene for diaphorase-1 is a NADH-dependent oxidoreductase containing NADH-binding domain and Rieske-type (2Fe-2S) cluster. There is a possibility that this enzyme play an important role in biological electron transfer reactions, like the other proteins containing these clusters.

Materials and methods

Drosophila culture

Adult *D. melanogaster* were maintained on a standard cornmeal/agar medium at 25°C. Embryos (2 h), larvae (third instar), pupae (light and dark) were selected from mass cultures. The different organs (ovaries, testes and heads) were collected from adult flies.

RT-PCR

The cDNA samples from male and female-adult flies, from different developmental stages, and from different organs were prepared using Oligotex mRNA Purification kit (Qiagen, Valencia, USA) and SuperScript II RT kit (Invitrogen, Carlsbad, USA).

Using SuperScript II RT (Invitrogen, Carlsbad, USA), we synthesized the first strand of cDNA from purified mRNA. We used this cDNA to perform the PCR amplification using appropriate primers for the six transcripts. The reverse primer DIA.R (5′-CGGAGCTCCGTAATGGGTGC-3′) was same for the six transcripts. We used the following forward primers: DIA.F1-2 (5′-GGCAGCACTGGCCAGACG-3′) for the first and the second transcript; DIA.F3 (5′-GCCTCGCATTGCATTGGC-3′) for the third transcript; DIA.F4-5 (5′-GCATCTCGGACACGTTTGGG-3′) for the fourth and fifth transcript; DIA.F6 (5′-GGTCTCCAGTTCATTCG-3′) for the sixth transcript.

The RT-PCR with primers (FER.3F1 and FER.3R) for ferritin small subunit was used as quantitative control for the mRNA preparation from female and male adult flies.

The PCR amplification was performed at 94°C for 5 min; 30 cycles at 94°C for 1 min, 60°C for 1 min, 72°C for 1 min; 72°C for 10 min in 50 μl mixtures containing 5 μl of template cDNA. After the amplification, the PCR products were separated on a 1% agarose gel, and the corresponding bands were excised and purified using the Sephaglass BandPrep kit (Pharmacia Biotech, Piscataway, USA). Purified PCR fragments were cloned in the PCR II vector, using the TA cloning kit (Invitrogen, Carlsbad, USA), and both strands were sequenced using vector-specific and insert-specific primers. DNA sequencing was performed by the automated sequencing facility of the Arizona Research Laboratories, Division of Biotechnology, University of Arizona, Tucson, USA.

Computer analyses

Diaphorase-1 was purified from *D. virilis* (Ralchev *et al*. 1992) and partial amino acid sequences of three tryptic peptides were obtained (Petkov P., Dunkov B. and Ralchev K. *unpublished data*). The sequences of these peptides (DIA-1 pep1 - AAFGEEIGGR; DIA-1 pep2 - VNIGHYQLAQYHGR; DIA-1 pep3 - LNTDFLANSGVR) showed similarity with the gene *EG: 22E5.5* or *CG4199* (accession number O77266, Q9W529), which was obtained from the BDGP (http://www.fruitfly.org). Amino acid sequence alignment of different proteins was done using BLAST software.

Results

The gene *EG:22E5.5* or *CG4199* (accession number O77266, Q9W529) structure reveals three exons and two long introns. Using BDGP, we found six transcripts from this gene (CG4199RF, CG4199RC, CG4199RD, CG4199RB, CG4199RE and CG4199RA). The difference between these transcripts are in their 5′ ends. The 3′ ends of the six transcripts are identical (figure 1).

Four of these transcripts (CG4199RF, CG4199RC, CG4199RD and CG4199RB) have an identical open reading frame (ORF) and encode the same protein (CG4199-PF, CG4199-PC, CG4199-PD and CG4199-PB). The other two transcripts (CG4199RE and CG4199RA) encode two different proteins (CG4199-PE and CG4199-PA).

The cDNAs encoding these transcripts were identified in the expressed sequence tag (EST) database BDGP. Twenty three of the ESTs are from Schneider L2 cell culture (SH) cDNA library, five from embryo (RE and LD) cDNA libraries, four from adult head (RH and GH) cDNA libraries, one from testes (AT) cDNA library and one from ovary (GM) cDNA library (table 1).

The information in table 1 shows that the transcripts can be found in SH and are present at a very low level in the cells of different organs and at the different stages of *Drosophila* development. Unfortunately, we could not find them using Northern blot hybridization. To solve this problem, we used RT-PCR method using samples of mRNA from female and male adult flies.

We expected to identify the first (CG4199-RE) and the second (CG4199-RF) transcripts using forward primer DIA.F1-2 and reverse primer DIA.R. Their expected size is 814 bp for CG4199-RF and 554 bp for CG4199-RE transcript. As shown in figure 2, we derived only CG4199-RE. We obtained this transcript in the sample of cDNA, prepared from female adult flies, but did not find it in cDNA sample from male adult flies.

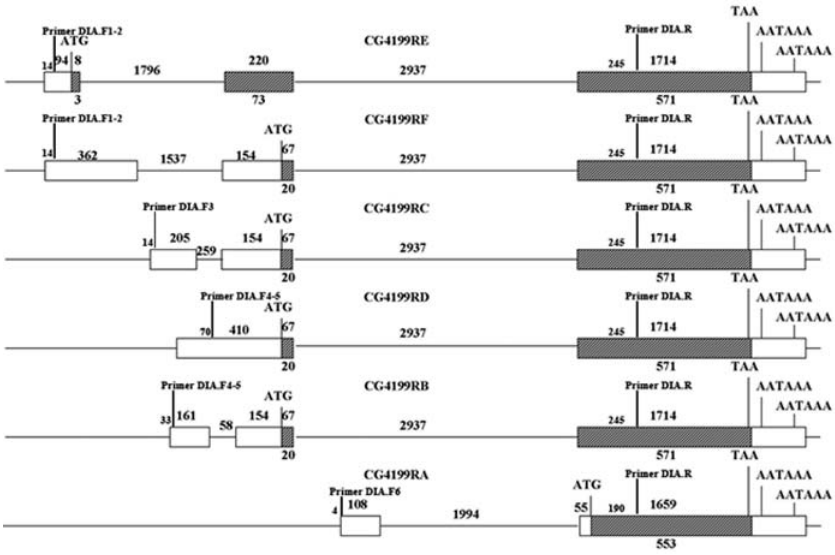

Figure 1. Diversity in the exon–intron organization of diaphorase-1 gene. The six transcripts with their differences. Exons are shown as boxes; the number of nucleotides in the exons are shown above the boxes. Protein-coding regions are marked with the numbers of the encoded amino acids indicated below them. Stop codons are indicated above the boxes. Introns are designated by thin lines with the number of nucleotides indicated above.

Table 1. *D. melanogaster* DIA-1 cDNAs. The six transcripts (CG4199RE, CG4199RF, CG4199RC, CG4199RD, CG4199RB and CG4199RA) and the number of ESTs for them identified from different cDNA libraries.

Transcript	BDGP Schneider L2 cell culture (SH)	BDGP embryo (LD)	BDGP embryo (RE)	BDGP head (RH)	BDGP head (GH)	BDGP ovary (GM)	BDGP testes (AT)
CG4199RE			53167 20109				
CG4199RF	22520 05956 22368 15644				11732		
CG4199RC	24769 05213 18368 17018 20904 02019 02121 17830 20386 19004 13127				14669		17482
CG4199RD	06560 14951		69164	22855 22056			
CG4199RB	04546 24522						
CG4199RA	23281 37636 05890 26317	36801 37636				01986	

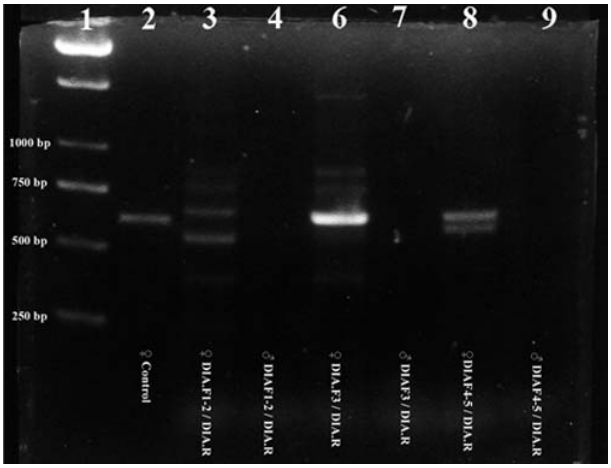

Figure 2. Amplified DNA fragments using PCR were separated on a 1% agarose gel. Lane 1, 1kb plus DNA ladder; lane 2, control sample (cDNA, prepared from female adult flies with FER3.F1 and FER3.R primers); lane 3, cDNA, prepared from female adult flies with DIA.F1-2 and DIA.R primers; lane 4, cDNA, prepared from male adult flies with DIA.F1-2 and DIA.R primers; lane 5, cDNA, prepared from female adult flies with DIA.F3 and DIA.R primers; lane 6, cDNA, prepared from male adult flies with DIA.F3 and DIA.R primers; lane 7, cDNA, prepared from female adult flies with DIA.F4-5 and DIA.R primers; lane 8, cDNA, prepared from male adult flies with DIA.F4-5 and DIA.R primers.

Using forward primer DIA.F3 and reverse primer DIA.R we expected to receive the third (CG4199-RC) transcript, the size of which is 657 bp. Again, we obtained it in the sample of cDNA, prepared from female adult flies, but we could not find it in the sample from male adult flies (figure 3).

The combination of forward primer DIA.F4-5 and reverse primer DIA.R resulted in the expected fourth (CG4199-RD) and fifth (CG4199-RB) transcripts. The size of CG499-RD and CG4199-RB transcripts was the expected one 652 bp and 594 bp, respectively. We obtained them only from the cDNA sample from female adult flies (figure 2).

Using forward primer DIA.F6 and reverse primer DIA.R, we obtained the expected sixth (CG4199-RA) transcript with a size of 349 bp. Again it was found only in the cDNA sample from female adult flies (figure 3).

These results reveal that five of the expected six transcripts are represented in female mRNA samples, but none in male mRNA samples. The possibility that quantitative differences in the mRNA samples result in the above-mentioned differences between females and males can be excluded because the intensity of the bands obtained using forward and reverse primers (FER.3.F1 and FER.3.R) for the ferritin small subunit is equal in both samples (figures 2 and 3). As described before, these transcripts are poorly represented in the cells. We used RT-PCR method to find them using samples of mRNA from ovaries, testes and heads of *Drosophila* adult flies.

Figure 4,a&b shows the result of RT-PCR using samples of mRNA from different organs. Unfortunately, we

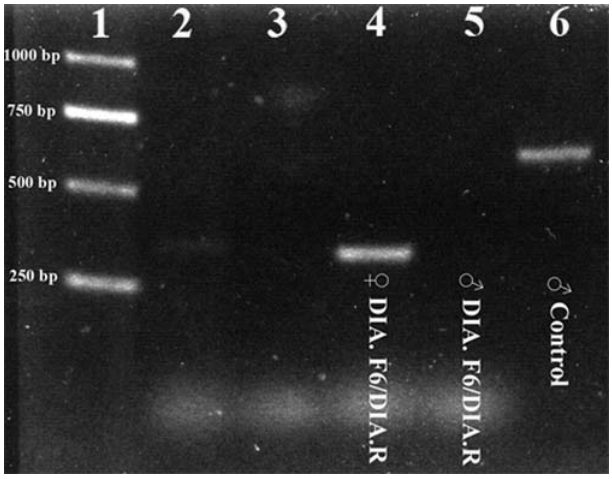

Figure 3. Amplified DNA fragments using PCR were separated on a 1% agarose gel. Lane 1, 1 kb plus DNA ladder; lane 4, cDNA, prepared from female adult flies with DIA.F6 and DIA.R primers; lane 5, cDNA, prepared from male adult flies with DIA.F6 and DIA.R primers; lane 6, control sample (cDNA, prepared from male adult flies with FER3.F1 and FER3.R primers).

obtained only CG4199-RC transcript. We got this transcript from the sample of cDNA prepared from testes. We could not find the other five transcripts in the samples of testes, ovaries and heads. These results confirm the data in table 1.

In the next set of experiments, we probed the expression of these six transcripts at different stages of *Drosophila* development. Once again we used RT-PCR method to find them in samples of mRNA from embryo, larvae, light and dark pupae.

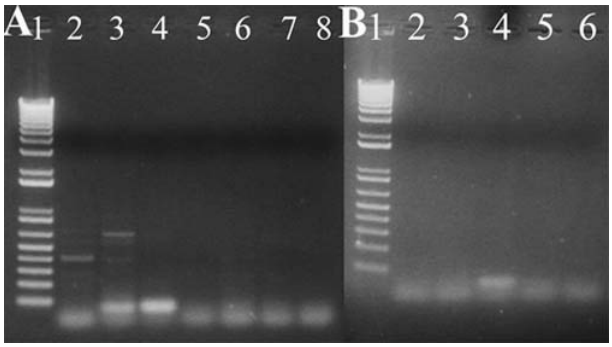

Figure 4. DNA fragments, amplified using PCR, were separated on a 1% agarose gel. (A) Lanes 1, 1 kb plus DNA ladder; 2, cDNA, prepared from testes with DIA.F1-2 and DIA.R primers; 3, cDNA, prepared from testes with DIA.F3 and DIA.R primers; 4, cDNA, prepared from testes with DIA.F4-5 and DIA.R primers; 5, cDNA, prepared from testes with DIA.F6 and DIA.R primers; 6, cDNA, prepared from ovaries with DIA.F1-2 and DIA.R primers; 7, cDNA, prepared from ovaries with DIA.F3 and DIA.R primers; 8, cDNA, prepared from ovaries with DIA.F4-5 and DIA.R primers.
(B) Lanes: 1, 1 kb Plus DNA ladder; 2, cDNA, prepared from ovaries with DIA.F6 and DIA.R primers; 3, cDNA, prepared from head with DIA.F1-2 and DIA.R primers; 4, cDNA, prepared from head with DIA.F3 and DIA.R primers; 5, cDNA, prepared from head with DIA.F4-5 and DIA.R primers; 6, cDNA, prepared from head with DIA.F6 and DIA.R primers.

Figure 5,a–c shows the result of RT-PCR with samples at different stages of *Drosophila* development. We obtained only the fourth (CG4199-RD) transcript, and it was found in samples of cDNA from embryo. We did not find the other five transcripts in the samples of embryo, larvae, light and dark pupae.

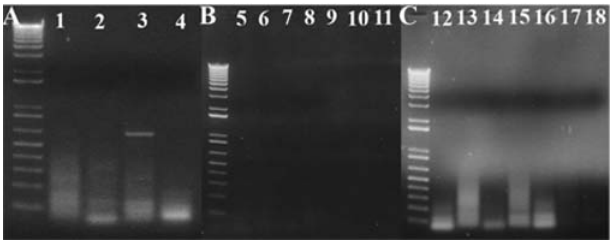

Figure 5. DNA fragments, amplified using PCR, were separated on a 1% agarose gel. (A) Embryo: lane 1, cDNA, prepared from embryos with DIA.F1-2 and DIA.R primers; lane 2, cDNA, prepared from embryos with DIA.F3 and DIA.R primers; lane 3, cDNA, prepared from embryos with DIA.F4-5 and DIA.R primers; lane 4, cDNA, prepared from embryos with DIA.F6 and DIA.R primers. (B) Larvae and light pupae: lane 5, cDNA, prepared from larvae with DIA.F1-2 and DIA.R primers; lane 6, cDNA, prepared from larvae with DIA.F3 and DIA.R primers; lane 7, cDNA, prepared from larvae with DIA.F4- 5 and DIA.R primers; lane 8, cDNA, prepared from larvae with DIA.F6 and DIA.R primers; lane 9, cDNA, prepared from light pupae with DIA.F1-2 and DIA.R primers; lane 10, cDNA, prepared from light pupae with DIA.F3 and DIA.R primers; lane 11, cDNA, prepared from light pupae with DIA.F4-5 and DIA.R primers. (C) Light pupae and deep pupae: lane 12, cDNA, prepared from deep pupae with DIA.F1- 2 and DIA.R primers; lane 13, cDNA, prepared from deep pupae with DIA.F3 and DIA.R primers; lane 14, cDNA, prepared from deep pupae with DIA.F4-5 and DIA.R primers; lane 15, cDNA, prepared from deep pupae with DIA.F6 and DIA.R primers; lane 16, cDNA, prepared from light pupae with DIA.F6 and DIA.R primers; lane 17, cDNA, prepared from light pupae with DIA.F3 and DIA.R primers; lane 18, cDNA, prepared from light pupae with DIA.F4-5 and DIA.R primers.

A database search using the software BLAST (figure 6) revealed that the gene product (CG4199-PF, which we named *DmelI*) shares similarities along the entire molecule with the ferredoxin reductase-like protein in *Xenopus laevis* (accession number D86491; amino acid identity is 38%); putative pyridine nucleotide-disulphide oxidoreductase in *Anopheles gambia*e (accession number EAA03876; amino acid identity is 55%); *D. melanogaster* feridoxin reductase, which we named as *DmelII* (gene name is CG10700; amino acid identity is 59%) and *Homo sapiens* unnamed protein product (accession number BAC04434; amino acid identity is 38%).

Discussion

The gene *EG:22E5.5* or *CG4199* (accession number O77266, Q9W529) is located on the X chromosome at position 2C9-2C10, which confirms our previous results (Ralchev and Dunkov 1987). Its structure reveals three exons and two long introns. Six transcripts that were found from this gene show difference in their 5′ ends, suggesting that these transcripts are products of different transcriptional start sites. They are poorly represented in the cells of different organs and at different stages of *Drosophila* development.

Since five of these transcripts are obtained in females, but none in males, it is likely that there are differences between male and female adult flies in the control of expression of this gene.

Amino acid sequence alignment shows that the gene product (CG4199-PF) , named *DmelI*, shares similarities along the entire molecule with the ferredoxin reductase-like protein in *X. laevis* (Hatada *et al*. 1997); putative pyridine nucleotide-disulphide oxidoreductase in *A. gambiae, D. melanogaster* feridoxin reductase, which was named *DmelII*, and *H. sapiens* unnamed protein. There is conserved sequence in the motif containing two cysteines and two histidines: Cys-X-His-15 to 17 amino acids-Cis-X-X-His (where X is any amino acid).

Analysis of these amino acid sequences reveals that the cysteine residues are conserved in two regions (Schagger *et al*. 1987). From a comparison of these conserved regions to those observed in the terminal oxygenase and the ferridoxin components of the dioxygenases (Mason and Cammack 1992) the same consensus sequence emerged.

All of the above mentioned proteins contain an iron-sulfur cluster that has Rieske type properties and we suppose that the product of *D. melanogaster CG4199* gene (diaphorase-1, named *DmelI*) has similar function. Proteins containing Rieske-type (2Fe-2S) clusters play important roles in many biological electron transfer reactions such as aerobic respiration, photosynthesis and biodegradation of various alkene and aromatic compounds (Berry *et al*. 2000). Typically, (2Fe-2S) clusters are not directly involved in the catalytic transformation of substrate, but rather supply electrons for the active site (Schmidt and Shaw 2001).

There is also conserved sequence in the motif containing three glycines: Gly-X-Gly-X-X-Gly-X-X-X-Ala-X-X-X-X-X-X-Gly (where X is any amino acid). This conserved sequence of *Drosophila* gene product involves a nucleotide-binding $\beta\alpha\beta$ unit domain including three essential glycines (Mason and Cammack 1992). This analysis bolsters our hypothesis that the *D. melanogaster* CG4199 gene product is an NADH-dependent oxidoreductase (diaphorase-1).

These results indicate that the *D. melanogaster CG4199* gene product, diaphorase-1, is likely a NADH-dependent oxidoreductase containing NADH-binding domain and Rieske-type (2Fe-2S) cluster. There is a possibility that this product plays an important role in biological electron transfer reactions, like the other proteins containing these clusters.

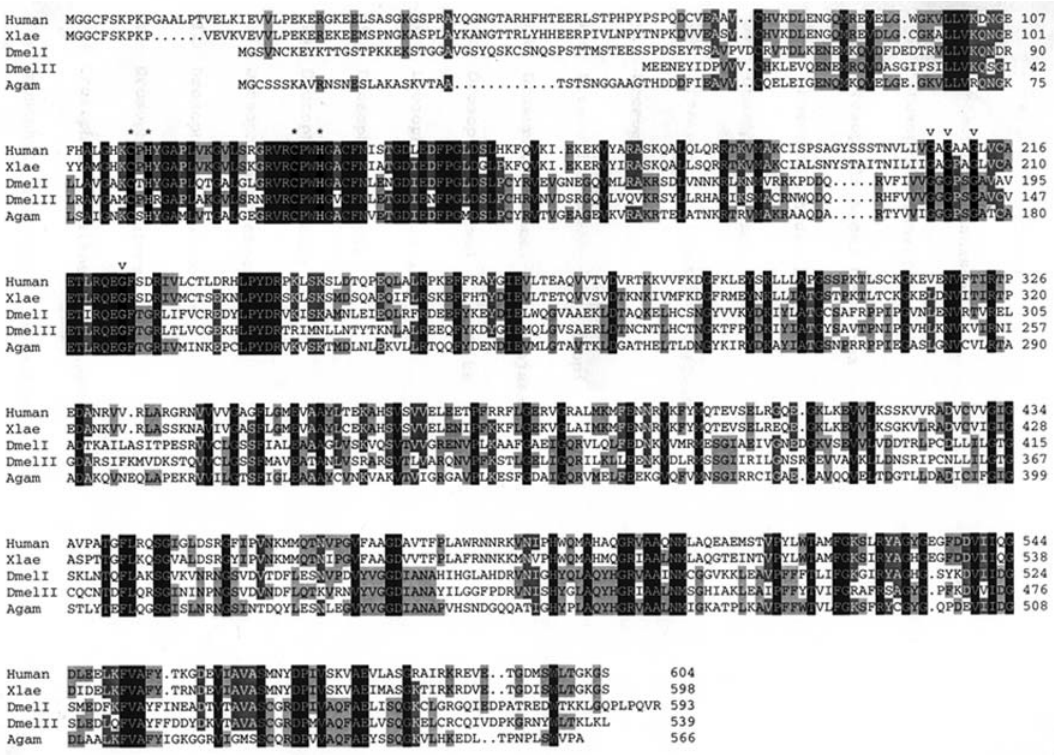

Figure 6. Amino acid sequence alignment of invertebrate and human proteins. Human, *Homo sapiens* unnamed protein product (GenBank no. BAC04434); *Xlae, X. laevis* ferredoxin reductase-like protein (GenBank no. D86491); *DmelI, D. melanogaster* unnamed protein product of the gene (GenBank no. CG4199); *DmelII, D. melanogaster* feridoxin reductase (GenBank no. CG10700); Agam, *Anopheles gambiae* putative pyridine nucleotide-disulphide oxidoreductase (GenBank no. EAA03876).

References

Berry E. A., Guergova-Kuras M., Huang I. S. and Crofts A. R. 2000 Structure and function of cytochrome bc complexes. *Annu. Rev. Biochem*. **69**, 1005–1075.

Dunkov B., Georgieva T. G. and Ralchev K. H. 1991 Diaphorase activity associated with *Drosophila* xanthine dehydrogenase. *Dros. Inform. Serv*. **70**, 62–65.

Enzyme nomenclature 1973 Recommendations of the commission on biochemical nomenclature on the classifications of enzymes together with their units and symbols of enzyme kinetics. Elsevier, Amsterdam, VI: **16**, 443–444.

Hatada S., Kinoshita M., Sakumoto H., Nishihara R., Noda M. and Asashima M. 1997 A novel gene encoding a ferredoxin reductase-like protein expressed in the neuroectoderm in *Xenopus neurula*. *Gene* **194**, 297–299.

Ivanova P. M., Harizanova N. T., Krasteva M., Dimov S. G. and Ralchev K. H. 2004 *Drosophila* dihydrolipoamide dehydrogenase. protein structure. developmental expression. *Annuaire de L'Universite de Sofia "St. Kliment Ohridski"*, *Biochem. Biophys. Biol. mol. Genet. Microbiol. Physiol*. **96**, 115–123.

Mason J. and Cammack R. 1992 The electron-transport proteins of hydroxylating bacterial dioxygenases. *Annu. Rev. Microbiol*. **46**, 277–305.

Onyenwoke R. U. and Wiegel J. 2007 Iron (III) reduction: A novel activity of the human NAD(P)H:oxidoreductase. *Biochem. Biophys. Res. Commun*. **353**, 389–393.

Ralchev K. H. and Dunkov B. H. 1987 Electrophoretic spectrum of diaphorases in *Drosophila*. *Dros. Genet. Sel*. **20**, 196–202.

Ralchev K. H., Petkov P. M. and Dunkov B. H. 1992 Purification and characterization of diaphorases from some *Drosophila* species. *Biochem. Genet*. **30**, 305–315.

Schagger H., Borchart U., Machleidt W., Link T. A. and von Jagow G. 1987 Isolation and amino acid sequence of the 'Rieske' iron sulfur protein of beef heart ubiquinol:cytochrome c reductase. *FEBS Lett*. **219**, 161–168.

Schmidt C. L. and Shaw L. 2001 A comprehensive phylogenetic analysis of Rieske and Rieske-type iron-sulfur proteins. *J. Bioenerg. Biomembr*. **33**, 9–26.

Received 20 December 2007, in final revised form 24 March 2008; accepted 31 March 2008
Published on the Web: 25 July 2008