Sorbitol Dehydrogenase of Drosophila

GENE, PROTEIN, AND EXPRESSION DATA SHOW A TWO-GENE SYSTEM*

Teresa Luque‡§, Lars Hjelmqvist¶, Gemma Marfany‡, Olle Danielsson‡, Mustafa El-Ahmad‡,
Bengt Persson‡, Hans Jörnvall***, and Roser González-Duarte‡

From the ¶Department de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, E-08071
Barcelona, Spain and the **Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77
Stockholm, Sweden

The Drosophila melanogaster sorbitol dehydrogenase (SDH) is characterized as a two-enzyme system of the medium chain dehydrogenase/reductase family (MDR). The SDH-1 enzyme has an enzymology with $K_m$ and $k_{cat}$ values an order of magnitude higher than those for the human enzyme but with a similar $k_{cat}/K_m$ ratio. It is a tetramer with identical subunits of ~38 kDa. At the genomic level, two genes, Sdh-1 and Sdh-2, have a single transcriptional start site and no functional TATA box. Expression is greater in larvae and adults than in pupae, where it is very low. At all three stages, Sdh-1 constitutes the major transcript. Sdh-1 and Sdh-2 genes were located at positions 84E-F and 86D in polytene chromosomes. The deduced amino acid sequences of the two genes show 90% residue identity. Evaluation of the sequence and modeling of the structure toward that of class I alcohol dehydrogenase (ADH) show altered loop and side chain relationships as in mammalian Sdh-1, establishes that SDH, despite gene multiplicity and larger variability than the “constant” ADH of class III, is an enzyme conserved over wide ranges.

Sorbitol dehydrogenase (SDH)1 is a zinc-containing enzyme that catalyzes the oxidation of sorbitol to fructose using NAD+ as cofactor. Together with aldose reductase, SDH is a constituent of the polyl pathway, an important bypass of glycolysis (cf. Ref. 1). Overloading of this pathway can lead to accumulation of sorbitol, which has been implicated in the pathogenesis of diabetic complications (2). However, inhibitors of sorbitol dehydrogenase also have been reported to alleviate diabetic complications (3, 4), and sorbitol accumulation is perhaps therefore not as decisive a step in the development of the complications as frequently believed.

SDH is a zinc enzyme (5) and belongs to the medium chain dehydrogenase/reductase protein family (MDR), which also includes the well known mammalian alcohol dehydrogenases (ADHs) (6, 7). Despite distinct structural and mechanistic relationships, SDH and ADH differ in zinc content, SDH has only one zinc atom per subunit, at the catalytic site, whereas ADH has two per subunit. In addition, the enzymes differ in quaternary structure, SDH is tetrameric whereas mammalian ADH is dimeric. SDH and ADH do not share substrates, SDH catalyzes the oxidation of polyols such as sorbitol, xylitol, D-mannitol, and L-iditol but has no activity toward primary alcohols.

The SDH enzyme has been characterized from a few sources, including sheep (8), human (9, 10), and Bacillus subtilis (11). The conformation has been modeled toward that of ADH (12), largely explaining the active site residue functions. However, reports on the nature of its zinc binding site (5, 12–16) have not fully agreed. The third ligand has been interpreted as Glu155 from the model structure (12) or Glu159 from mutant studies (16), while Glu149 also has seemed possible from the conservation pattern. However, recent evidence from crystal structures of homologous tetrameric ADH forms (17) again suggest Glu155 as the third ligand. The cloning of Sdh cDNA from rat (18), Bombyx mori (19), and human (20) has been reported, and the gene has been described in Bacillus subtilis (11), Saccharomyces cerevisiae (21), human (22), and B. mori (23). Two allelic variants of Sdh have been observed in fish, toads, and mouse whereas an Sdh pseudogene, which lacks exon 1 and shows a premature stop codon in exon 8, has been characterized in human (24).

Three different SDH activities of Drosophila melanogaster were described earlier: soluble NAD+-dependent SDH, a mitochondrial NADP+-dependent SDH, and a soluble NADP+-dependent SDH (25), which were assumed to be encoded by three different genes. The structural gene for the cytosolic NADP+-dependent Sdh was mapped at 91B9-93F on the third chromosome of D. melanogaster. To date, only another member of the MDR family, class III ADH, has been characterized in Drosophila at the protein (26) and gene level (27). Here, we report protein and gene relationships for the highly different SDH member of the MDR family. We find two similar genes, Sdh-1 and Sdh-2 (80% nucleotide identity in the coding regions), and have characterized them and their transcription products. The major enzyme purified was found to be SDH-1, and this form was enzymatically characterized and conformationally modeled, allowing further interpretations of the active site relationships.

EXPERIMENTAL PROCEDURES

Screening of Genomic and cDNA Libraries—D. melanogaster genomic DNA (1 μg) was purified according to a modification of the guanidine thiocyanate method (28) and amplified using two degenerate

* This work was supported by Comisión Interministerial de Ciencia y Tecnología Grant 96-0329, Swedish Medical Research Council Projects 13X-3532 and 01P-11312, European Community Contract BIO4-CT97-2123, and the Foundations of Magn. Bergvall, Fredrik och Ingrid Thuring, Åke Wiberg, and in memory of Lars Hierta. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Recipient of a fellowship from Berth von Kantzow’s Foundation.

¶ Recipient of a Formacio´n del Personal Investigador fellowship from the Ministerio de Educación y Ciencia.

** To whom correspondence should be addressed: Dept. of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden. Tel.: 46-8-7287702; Fax.: 46-8-337462; E-mail: Hans.Jornvall@mbb.ki.se.

§ Recipient of a fellowship from the Ministerio de Educación y Ciencia.

** This work was supported by Comisión Interministerial de Ciencia y Tecnología Grant 96-0329, Swedish Medical Research Council Projects 13X-3532 and 01P-11312, European Community Contract BIO4-CT97-2123, and the Foundations of Magn. Bergvall, Fredrik och Ingrid Thuring, Åke Wiberg, and in memory of Lars Hierta. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of a fellowship from Berth von Kantzow’s Foundation.

§ Recipient of a Formacio´n del Personal Investigador fellowship from the Ministerio de Educación y Ciencia.

** To whom correspondence should be addressed: Dept. of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden. Tel.: 46-8-7287702; Fax.: 46-8-337462; E-mail: Hans.Jornvall@mbb.ki.se.

1 The abbreviations used are: SDH, sorbitol dehydrogenase; MDR, medium chain dehydrogenase/reductase family; ADH, alcohol dehydrogenase; RACE, rapid amplification of cDNA ends; PCR, polymerase chain reaction; FPLC, fast protein liquid chromatography; HPLC, high performance liquid chromatography; bp, base pair(s); kb, kilobase pair(s).
Sorbitol Dehydrogenase of D. melanogaster oligonucleotides, designed from species-conserved SDH regions (22) (5'-GSGHATYTGCGHWSCGA-3' and 5'-GGCRTATRGCASGRCRTG-3'), corresponding to positions 57–43 and 91–85 of the Drosophila SDH amino acid sequence). The reaction conditions were as follows. First, a denaturing step of 1 min 30 s at 94 °C and then 35 cycles of 1 s at 94 °C and 30 s at 55 °C. The product of this amplification was labeled with [α-32P]dCTP (NEN Life Science Products) using a random primed DNA labeling kit (Boehringer Mannheim). It was then used as a probe to screen 1.5–2 × 10^6 phages of a dASH-2 genomic library of D. melanogaster (27). The same probe was used to screen 3–4 × 10^6 phages of a tag10 cDNA library from D. melanogaster (29). DNA from positive recombiant phages was isolated according to Ref. 30.

DNA Mapping and Sequencing—Positive clones were characterized by restriction mapping (enzymes from Boehringer Mannheim and Promega). Regions homologous to Sdh were identified after Southern blot analysis and subcloned into the Bluescript phagemid (Strategene). DNA sequencing was performed on double-stranded DNA by the dideoxy method (31) with [35S]dATP (NEN Life Science Products) and a MegaBACE sequencer (Amersham Pharmacia Biotech) or with ABIPRISM dye terminator cycle sequencing ready reaction kit (Perkin-Elmer). A set of internal primers was used, and each nucleotide was sequenced at least three times.

Identification of the Transcriptional Start Site—Total cellular RNA and poly(A) RNA from adults obtained as described (27). The ends of the transcripts were characterized using the 5'–3' RACE (rapid amplification of cDNA ends) Kit (CLONTECH), using the primers 5'-TCATGGGCTTAGTGAGCAC-3' (position 61–56 of the amino acid sequence) for cDNA synthesis and 5'-CCTGATGC-GTTGGAGAACG-3' (position 13–7 of the amino acid sequence) for nested PCR. RACE products were subcloned using the SureClone ligation kit (Amersham Pharmacia Biotech) and sequenced with the M13 universal primer.

Northern Blot Analysis and Expression Studies—Total cellular RNA from larvae, pupae, and adult D. melanogaster was isolated by the guanidinium thiocyanate method (32), separated by 1.2% agarose/formaldehyde gel electrophoresis, and transferred to nylon membranes (Amersham Pharmacia Biotech) for hybridization with the same probe as that used for the library screening. A control hybridization was performed with a d melanogaster actin gene probe.

Sdh expression was studied by reverse transcriptase-PCR from larval, pupal, and adult RNA. PCR was performed with internal Sdh oligonucleotides, and gfd/odh was amplified as a control.

In Situ Hybridization on Polytenic Chromosomes—Polytenic chromosomes were prepared from D. melanogaster larval salivary glands using standard techniques. They were screened with a gene fragment amplified from genomic DNA clones using the oligonucleotides corresponding to positions 45–50 and 350–344 of the amino acid sequence. The PCR product was random primed and labeled with digoxigenin-dUTP (Boehringer Mannheim). The hybridized probe was detected with antidigoxigenin-peroxidase and 3,3'-diaminobenzidine as substrate (Boehringer Mannheim). Polytenic chromosomes were stained with 5% Giemsa in 50 mM sodium phosphate and then with 2% orcein in 60% acetic acid.

Genomic DNA Southern Blot Analysis—Genomic DNA, 4 to 8 μg, purified from adults of D. melanogaster, was digested with EcoRI, BamHI, and HindIII, resolved in 0.6% agarose gel, and transferred by Southern blotting (30) to a nylon membrane (Hybond N, Amersham Pharmacia Biotech). The probe used and the hybridization conditions were the same as those for library screening.

Protein Purification—D. melanogaster Adh* adults were bred and harvested following conventional procedures. After homogenization (about 40 g of flies) in 2 ml Tris/HCl, pH 7.5 (about 150 ml), and centrifugation, proteamine sulfate (2%) and ammonium sulfate fractionations (0–35% and 35–75%) were carried out. Sephadex G-25 exclusion chromatography was used to desalt the sample. SDH activity was detected by a standard assay (33). The enzyme was purified to get rid of short chain Drosophila ADH by affinity chromatography on Blue-Sepharose equilibrated with 20 mM Tris/HCl, pH 7.5 (about 150 ml), and centrifugation, proteamine sulfate (2%) and ammonium sulfate fractionations (0–35% and 35–75%) were carried out. Sephadex G-25 exclusion chromatography was used to desalt the sample. SDH activity was detected by a standard assay (33). The enzyme was purified to get rid of short chain Drosophila ADH by affinity chromatography on Blue-Sepharose equilibrated with 20 mM Tris/HCl, pH 7.5, and washed with 0.15 M NaCl in 50 mM sodium phosphate buffer, pH 7.5, and with 1 mM NaCl in 50 mM sodium phosphate buffer, pH 7.5. The mature enzyme was then subjected to fast protein liquid chromatography (FPLC) on Mono S in 10 mM sodium phosphate, pH 6.2, with a gradient of 0–1 mM NaCl. The SDH activity eluted at 0.2 M and was submitted to exclusion chromatography on Superdex-200 (FPLC). Total protein during purification was determined by the Bradford method (34). SDS-polycrylamide gel electrophoresis was carried through with a Phast system (Amersham Pharmacia Biotech) and staining with Coomassie Brilliant Blue. The pl was determined by isoelectric focusing on Phast IEF 3–9 gels and staining with silver nitrate. Activity staining was performed with nitro blue tetrazolium/phenazine methosulfate (26).

For structural analysis, the pure protein was reduced and carboxymethylated in 10 mM Tris base, pH 8.0, followed by limited proteolytic digestion with Lys-C protease, peptide purification by HPLC, sequence analysis, and peptide mass determinations, as described (35).

Enzymatic Measurements—Activity with polyols, isopropyl alcohol, and fructose was tested in solution and via activity staining after isoelectric focusing. K_m and k_cat values were determined with alcohols at pH 10.0 in 0.1 M glycine/NaOH with 10 mM NAD^+ and with fructose at pH 7.1 in 0.05 M sodium phosphate buffer with 1 mM NAD^+. K_m for NAD^+ was analyzed with 50 mM sorbitol. Enzymatic activities were determined at 25 °C by monitoring the change in absorbance at 340 nm with a Beckman DU 68 or 64 spectrophotometer.

Sequence Evaluations and Modeling—Primary structures were aligned using the program CLUSTAL W (36), and alignments were visualized using the program PRALIN. Phylogenetic trees were constructed with the program CLUSTAL W. Positions with gaps were excluded, and correction was made for multiple substitutions. Bootstrap analysis (37) was used for confidence evaluation.

Three-dimensional models of Drosophila and human SDH were obtained by adopting their amino acid sequences into the known fold of the human class I β ADH subunit (38) using the program ICM (version 2.7, Molsoft LLC, Metuchen, NJ). With ICM docking of sorbitol to the enzyme model was performed using a non-rigid procedure, allowing free movement of the substrate, the rotatable bonds of the substrate, and the χ angles of the ADH molecule and with an additional distance restraint of 2.0–2.4 Å between the sorbitol hydroxyl oxygen (O2) and the catalytic zinc. For comparisons the sheep liver SDH model (12) also was used.

RESULTS

Isolation and Analysis of Sdh Genes—A genomic DNA fragment was amplified with degenerate oligonucleotides. Its size, 164 bp, corresponds to the expected length of an intronless genomic region coding for the selected part of the SDH enzyme. The amino acid sequence, deduced from the sequenced DNA, agrees with that expected from reports of the enzyme from other species. The fragment obtained was used to screen a D. melanogaster genomic library. Several positive clones were isolated and analyzed. Restriction analysis of more than 10 positive clones allowed us to distinguish two different patterns (Fig. 1), named Sdh-1 and Sdh-2. By genomic restriction mapping, the Sdh-2 gene was located adjacent to the gfd/odh gene, the first member of the medium chain ADH family reported in Drosophila (27). Despite the close juxtaposition, the two genes, Sdh-2 and gfd/odh, have an opposite transcriptional orientation.

Screening of a cDNA library yielded several positive clones that correspond to the Sdh-1 and/or Sdh-2 genes.

The nucleotide sequences of the coding region of Sdh-1 and Sdh-2 exhibit 80% identities and their deduced amino acid sequences 90% identities. The genomic sequence of Sdh-1 is deposited under the GenBank™ accession number AF002212 (AF002212) of Sdh-2 under the GenBank™ accession number AF002213. Both genes consist of 6 exons and 5 introns, which are located at equivalent positions. The cDNA sequences further corroborate the intron locations. Introns follow the GT-AG rule, and their lengths differ between the intronated genes (Fig. 2). Compared with the average Drosophila intron length (60–70 bp), the size of the second intron of both genes (386 bp in Sdh-1 and 306 bp in Sdh-2) and that of the third intron of Sdh-1 (with 264 bp) is large.

Regulatory signals were identified in the 5' and 3' untranslated regions of both genes. Sdh-1 shows a conventional TATA box at position 98, and Sdh-2 shows one at position 250. Other consensuses sequences for different transcription factors were detected in both genes (Fig. 3). The cDNA sequences are long enough to allow us to determine the length of the trailer (43 bp in Sdh-1 and 21 bp in Sdh-2). A polyadenylation signal with 1 nucleotide mismatch in relation to the consensus is present in
both of the two Sdh genes (Fig. 3).

**Transcriptional Start Site**—A single RACE extension product using adult RNA was detected for each Sdh gene. Cloning and sequencing of the RACE fragments confirmed that the transcriptional start site was 67 and 71 bp upstream from the ATG initiator codon in Sdh-1 and Sdh-2, respectively (Fig. 3). Thus, the TATA boxes detected in the Sdh-1 and Sdh-2 genomic sequences appear too far away from the transcriptional start sites to be functional.

**Sdh Expression**—Northern blot analysis showed a single transcript, whose length, 1.3 kb, was in agreement with the Sdh coding region. Sdh transcripts were observed in the larval and the adult stage but not in pupae (Fig. 4). Since transcript lengths for Sdh-1 and Sdh-2 have almost the same size, it was not possible to assess whether the single band resolved on the Northern blot corresponded to one or two genes. Then, reverse transcriptase-PCR analyses were carried out with larval, pupal, and adult RNA, using specific oligonucleotides for Sdh-1 and Sdh-2 genes. Both transcriptional products were amplified at the three stages analyzed, even in pupae where no expression had been detected by Northern blot analysis. At all stages, the Sdh-1 transcript was more abundant than that of Sdh-2, in
agreement with results from the protein purification (below). The gfd/odh transcriptional product was amplified and used as control.

In Situ Analysis on Polytene Chromosomes—Two different hybridization signals were observed (Fig. 5) on the R arm of chromosome 3 of D. melanogaster when hybridized with an Sdh PCR probe. One gene signal mapped at division 486D, at the same site known to harbor the gfd/odh gene. From the genomic restriction map analysis, we could assign the Sdh-2 gene to this position. The second signal mapped at division 84E-F, and it corresponds to Sdh-1.

Genomic DNA Southern Blot Analysis—DNA of D. melanogaster was digested with EcoRI, BamHI, and HindIII and hybridized with an Sdh probe, which had no restriction site for these enzymes. Two bands were obtained in all digests in agreement with the two Sdh genes (Fig. 6).

Protein—SDH was purified in batches from 40 g of whole flies by a 5-step method utilizing two precipitation steps and
three chromatographies. The preparation obtained had a specific activity of 35 units/mg after a minimally 500-fold to some 1000-fold purification (non-SDH activity in initial steps makes purification-fold difficult to estimate). SDS-polyacrylamide gel electrophoresis demonstrated an apparently homogeneous preparation with the presence of a single band positioned identically to that of the sheep liver enzyme (Fig. 7) and corresponding to a subunit molecular mass of about 38 kDa. Isoelectric focusing showed a band at pI 9, positive with sorbitol in the activity test with nitro blue tetrazolium and phenazine methosulfate.

Tests with sorbitol at pH 5–11 demonstrated an activity optimum at pH 10.0 where the activity with sorbitol was about 1.5-fold that with xylitol, which in turn was about 2-fold that with ribitol. With the most active preparation, values at pH 10.0 for sorbitol were: $K_m$ 23 mM, $k_{cat}$ 6700 min$^{-1}$, and $k_{cat}/K_m$ 290 min$^{-1}$ 1 min$^{-1}$. Corresponding values for the human enzyme at pH 10 are $K_m$ 0.67 mM, $k_{cat}$ 312 min$^{-1}$, and $k_{cat}/K_m$ 466 min$^{-1}$ 1 min$^{-1}$ (9). The $K_m$ value for NAD$^+$ was found to vary considerably with pH and was about 20-fold lower at pH 7.1 (0.064 mM) than at 10.0 (1.5 mM).

To identify the protein prepared versus the two isozyme structures deduced from the DNA, the protein was carboxymethylated and digested with Lys-C protease. Peptides were separated by HPLC and submitted to MALDI-TOF mass analysis or sequencer analysis. Six pure peptides were analyzed and found to correspond to the DNA-deduced primary structure at positions 83–102 (experimental mass 2358 versus calculated mass 2357), 106–129 (2902 versus 2903), 138–161 (2603 versus 2604), 210–220 (1195 versus 1195), 221–234 (1587 versus 1588), and 316–325 (1266 versus 1266). They correspond to isozyme-unique regions which all identified the preparation as SDH-1.

To determine the quaternary structure, the native enzyme was submitted to FPLC exclusion chromatography on Superdex-200, which was calibrated with ovalbumin (43 kDa), bovine serum albumin (67 kDa), aldolase (158 kDa), and catalase (232 kDa). Correlation with the elution position of the SDH preparation gave a value of $\theta$ 145 kDa, corresponding to a tetrameric structure with subunits of $\theta$ 38 kDa.

Combined, these results identify the protein as a sorbitol dehydrogenase with an activity profile and quaternary structure like those of the mammalian enzymes and with constituent peptide regions corresponding to $Sdh$-1 of the two genes identified.

Conformational Modeling—Since previous conformational and zinc ligand assignments of mammalian SDHs versus the related ADHs have involved shifts and reinterpretations in a gap-rich central segment, we decided to model also the Drosophila SDH toward the crystallographically analyzed human class I $\beta_1$ ADH (38) of the same MDR protein family. We tested an alignment in this central region based on Glu149/Glu150 (numbering system of the human enzyme) as the third active site zinc ligand in agreement with results from site-directed mutagenesis (13, 16), and one based on Glu150 (as that ligand in agreement with a previous modeling (12). The tests showed that Drosophila SDH could be modeled into the crystallographically determined human $\beta_1$ ADH subunit with both alignments. However, the alignment centering on Glu149-Glu150 as the third zinc ligand gave three larger differences...
toward the mammalian SDH than the previous alignment. One is a shortcut in SDH (Drosophila positions 109/112) down chain from the structural zinc in ADH, the second an insertion at residues 158–162 increasing a loop, and the third a gap corresponding to Phe116 in the sheep SDH model. This gap would involve the wall of the substrate binding site (12) and make the substrate pocket slightly larger. This new alignment was therefore not further pursued since the enzymology now suggests an unaltered $k_{cat}/K_m$ ratio compatible with a lack of large differences. Also, recent crystallography of homologous tetrameric ADH forms appear compatible with Glu155 as the third ligand in SDH (17). Hence we used that alignment for the Drosophila SDH modeling. We also modeled in the same alignment the human SDH subunit, such that the human/Drosophila SDH relationships can be evaluated both in relation to enzymology and to differences with ADH.

The results show that the SDH sequences can be modeled into an ADH-related fold (Fig. 9), giving structures overall similar to that obtained for sheep SDH (12). The Drosophila SDH structure is found to have all the Cys residues (Drosophila FIG. 8. $D. melanogaster$ SDH amino acid sequences aligned with other SDH sequences. Conserved positions are shown against a black background. Alignment was performed using the CLUSTAL W program (36). Arrows indicate functional residues discussed under “Results.” Residue numbers are shown for the Drosophila and human enzymes, in both cases not counting the initiator Met, since this is known for the human enzyme (10) to be removed during post-translational processing to form the mature enzyme.

FIG. 9. Stereo views showing the monomer backbone (A) and active site (B) of the Drosophila sorbitol dehydrogenase model. In A, the monomer backbone of the Drosophila sorbitol dehydrogenase model (blue) is compared with those of the human sorbitol dehydrogenase model (red; Ref. 38). Alignment in a central region with low similarity and alternative possibilities (cf. text) uses Glu155 in SDH toward Cys174 in ADH, in agreement with the alignment in sheep SDH analysis (12). All three structures show overall similarities supporting a conserved fold. Differences are noticeable in loop regions down chain of the structural zinc segment of alcohol dehydrogenase. In B, sorbitol (green) is docked into the active site. The polypeptide backbone is colored black, residues close to the active site red, residues close to the coenzyme blue, coenzyme magenta, and the catalytic zinc atom gray. Since the modeling is performed on a single subunit, the final positions of residues at the outer part of the active site may be different in the oligomer because of changes upon subunit interactions.
positions 95, 98, 101, 109; Fig. 8) that correspond to the Cys ligands at the second zinc atom of ADH (ADH positions 97, 100, 103, 111). This is also a feature with the B. mori, Saccharomyces pombe, B. subtilis, and Escherichia coli SDHs, but with no other SDH. Nevertheless, the SDHs with these Cys residues still appear not to have the second zinc atom, as proven for the B. subtilis enzyme (11), and therefore they resemble the one-zinc (5) mammalian SDHs.

The results also show that the human and Drosophila SDH models differ very little. Of 11 substrate-interacting positions as deduced for the mammalian enzyme (12) and now compatible with docking to the Drosophila model (Fig. 9B), 8 are identical between the human and Drosophila SDHs (Cys to Tyr, Ser to Met, Pro to Ser, Phe to Leu, His to Tyr, and Glu to Gln). This overall similarity is compatible with the largely conserved enzymatic $k_{cat}/K_m$ ratio between the human and Drosophila enzymes. Only three of the substrate-adjacent positions differ between human and Drosophila SDH (Tyr to Arg of Drosophila versus Asp of the human enzyme, Met to Leu, and Arg to Lys) and appear likely to contribute explanations to the higher values of both $k_{cat}$ and $K_m$ in Drosophila versus human SDH. A proper fit of sorbitol into the active site of the modeled Drosophila SDH could be demonstrated by substrate docking (Fig. 9B), supporting the conclusions on the substrate interactions.

**DISCUSSION**

Sorbitol dehydrogenase from D. melanogaster has been characterized at the protein and genomic levels, allowing comparisons with other genes and analysis of the enzyme properties. The genomic and cDNA sequences reveal the presence of an Sdh duplication, with highly similar Sdh-1 and Sdh-2 components. Duplications are quite common in the MDR family, also of recent origin visible only in some lines (35). We now see that this applies also to SDH and this time in the insect line. Different rates of divergence is a characteristic feature of MDR enzymes, and they have been traced in the ADH branches (26), where conservation differs 5-fold between the variable class II (41) and constant class III (26). We now see that SDH is fairly constant but more variable than class III ADH, although still traceable through much of the same life forms (Fig. 10). In Drosophila, this enzyme activity is needed to exploit and detoxify the high polyol concentrations in fermenting fruits of the insect feeding and breeding sites.

By in situ hybridization on polytene chromosomes the two Sdh genes were mapped on chromosome 3R, Sdh-1 at position 84E-F and Sdh-2 at position 86D. The latter is adjacent to the

---

**Fig. 10. Unrooted phylogenetic tree relating the present enzymes to other sorbitol dehydrogenases and class III alcohol dehydrogenases.** Relationships shown are those obtained with the program CLUSTAL W (36), with distances corrected for multiple substitutions and positions with gaps excluded. Numbers show results from Bootstrap analysis (37).
Sorbitol Dehydrogenase of D. melanogaster

gfd/odh gene encoding class III ADH but transcriptionally inverted. It is not infrequent that clustered related genes belong to multigene families as now is the case with Sdh-2 and gfd/odh. The two Sdh genes are expressed and encode two isozymes. We have no evidence of a close pseudogene copy, as has been reported for Sdh in humans (24). Neither Sdh-1 nor Sdh-2 mapped at 91B-93F, the position suggested by Bischoff (42) for the NAD+-dependent SDH. Moreover, the presence of Asp199 argues in favor of the conclusion that SDH-1 and SDH-2 are NAD+-dependent (12) and correspond to the NAD+-dependent SDH observed earlier. Hence, the previous mapping at 91B-93F is unsupported by recent data. Similarly, by genomic screening we could not identify a gene responsible for the NAD+-dependent form reported earlier (25).

The exon/intron distribution is shared by the two Sdh genes and consists of 6 exons and 5 introns. Comparisons with the S. cerevisiae, B. mori, and human SDH genes show that in this system the number of introns increases with genome complexity. This is also the case for the class III Adh genes (27). Introns could have been gradually acquired with the appearance of eukaryotic organisms (43) or, if ancestrally present, could have been forced to disappear (44). No introns were reported in S. cerevisiae Sdh, while there are 7 in B. mori and 8 in human. Interestingly, the two first and the two last intron positions are shared by the D. melanogaster, B. mori, and human genes whereas the third Drosophila intron lies between the fifth and sixth human introns. This conservation of intron positions at both ends of the coding sequence does not occur in class III Adh and Drosophila Adh, where intron loss seems to have occurred directionally, from 3′ to 5′. In addition, B. mori introns 4 and 5 are at the same positions as human introns 4 and 6 (23). As B. mori is more closely related to D. melanogaster than to humans, these two introns may have been lost in the lineage of D. melanogaster. Hence, the SDH ancestral gene should have had a minimum of 6 introns, and subsequent intron losses/gains have occurred in the different organisms.

Sdh transcripts (1.3 kb) were present in larvae and adults, and a very low expression was detected in pupae. At all these stages, Sdh-1 was more abundantly expressed than Sdh-2, suggesting differential transcriptional regulation. Although both genes have a TATA-less promoter, a single transcription start site was identified for each gene. In contrast, in S. cerevisiae and in B. mori, several TATA boxes have been characterized (21, 23), while no TATA element was identified in the human genes, where two transcription initiation sites were described (22). The consensus TATA box present in the two Drosophila Sdh genes is probably non-functional, as in each case it is too far from the transcription start site.

Several transcription factor binding sites were identified in the two Drosophila Sdh genes. Since Sdh-2 and gfd/odh genes have opposite transcriptional orientations, they could share the 5′-non-coding region. Thus, this upstream gfd/odh region was also searched for transcription factor binding sites. Nevertheless, the two Sdh genes and gfd/odh showed different expression patterns. While gfd/odh was expressed at all developmental stages, Sdh expression was practically undetectable in pupae. A TFIID binding site was located downstream to both of the two Sdh transcription start sites, as frequently reported for Drosophila TATA-less genes (45), but was absent in gfd/odh. All three genes showed Sp1 binding sites.

Although a duplication of the Sdh gene was suggested earlier for salmonid fish (46) and an Sdh pseudogene was recently described in humans (24), isozyme-related genes have not been characterized before. The Drosophila Sdh genes encode very similar SDH proteins. Their deduced amino acid sequences show 90% residue identity, and both proteins share zinc binding residues (12), the Rossmann fold for NAD+ binding (47), and the sequence motif Pro-Xaa-Ile/Val-Xaa-Gly-His-Glu-Xaa-Gly, common to all medium chain alcohol dehydrogenases (48). In addition, the different modeling now tested with Drosophila SDH show that the remaining differences are compatible, although with three localized conformational consequences, with each of the two different alignments regarding the third active site zinc ligand in SDH. In the absence of major improvements with the new model, we have kept to the alignment with Glu155 assigning a liganding role. The biological significance of the Sdh duplication remains unknown. When a complete model can either retain function, evolve into a new gene, or become functionless. Some reports suggest that a duplicated locus is much more likely to become a pseudogene than to evolve a new gene function (49), but this development appears not to have been the case for Sdh-1 or Sdh-2. The two genes exhibit similar divergence rates (Fig. 10). Moreover, genes from the alcohol dehydrogenase family share a common ancestor, and gene duplications constitute one of the mechanisms for new functional appearances (50). Comparisons between functional genes show much higher synonymous than nonsynonymous substitution rates (Ks and Kn, respectively) and a high G + C content in the third codon position. Ks and Kn values were calculated to estimate the evolutionary pattern of Drosophila Sdh genes. As expected, much higher Kn values (1.195 ± 0.147) than Kn (0.056 ± 0.009) values were obtained. In addition, both Sdh genes show a high G + C percentage in the third codon position (67% for Sdh-1 and 73% for Sdh-2). These results further support functionality for these two genes.

Acknowledgments—We thank Berta Alsina and Florenci Serras for their valuable help in the in situ hybridization experiments and the Serveis Cientifico-Tecnis, Universitat de Barcelona, for technical support in automated DNA sequencing.

REFERENCES

1. Jeffery, J., and Jornvall, H. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 901–905
2. Gabbay, K. H. (1973) N. Engl. J. Med. 288, 831–836
3. Geisen, K., Uts, R., Grotsh, H., Lang, H. J., and Nimmesgern, H. (1994) Arzneim. Forsch. Drug Res. 44, 1032–1043
4. Tilton, B. G., Chang, K., Nyegaard, J. R., Van den Enden, M., Ide, Y., and Williamson, J. B. (1995) Diabetes 44, 254–262
5. Jeffery, J., Chesters, J., Mills, C., Sadler, P. J., and Jornvall, H. (1984) EMBO J. 3, 357–360
6. Jovell, H., von Bahr-Lindstrom, H., and Jeffery, J. (1984) Eur. J. Biochem. 140, 17–23
7. Jovell, H., Danielsson, O., Eklund, H., Hjelmqvist, L., Hoog, J.-O., Peres, P., and Shadquat, J. (1993) in Enzymology and Molecular Biology of Carbohydr Metabolism (Weiner, H., Crabb, D. W., and Flynn, T. G., eds) Vol. 4, pp. 533–544, Plenum Press, New York
8. Jeffery, J., Cederlund, E., and Jornvall, H. (1984) Eur. J. Biochem. 140, 7–16
9. Maret, W., and Auld, D. S. (1988) Biochemistry 27, 1622–1628
10. Karlsson, C., Maret, W., Auld, D. S., Hoog, J.-O., and Jornvall, H. (1989) Eur. J. Biochem. 166, 453–500
11. Ng, K., Ye, R., Hu, X-C., and Wong, S-L. (1992) J. Biol. Chem. 267, 24989–24994
12. Eklund, H., Horjales, E., Jornvall, H., Branden, C.-I., and Jeffery, J. (1985) Biochemistry 24, 8005–8012
13. Karlsson, C., and Hoog, J.-O. (1993) Eur. J. Biochem. 216, 103–107
14. Feiters, M. C., and Jeffery, J. (1989) Biochemistry 28, 7257–7262
15. Maret, W. (1989) Biochemistry 28, 9944–9949
16. Karlsson, C., Jornvall, H., and Hoog, J.-O. (1995) in Enzymology and Molecular Biology of Carbohydr Metabolism (Weiner, H., Holmes, R. S., and Wernmuth, B., eds) Vol. 5, pp. 397–406, Plenum Press, New York
17. Korkhim, Y., Kahle-Giboa, A. J., Perez, M., Bogan, O., Burstein, Y., and Frolove, P. (1993) J. Mol. Biol. 276, 967–981
18. Karlsson, C., Jornvall, H., and Hoog, J.-O. (1991) Eur. J. Biochem. 198, 761–765
19. Niimi, T., Yamashita, O., and Yaguma, T. (1993) Eur. J. Biochem. 213, 1125–1131
20. Lee, F. K., Cheung, M. C., and Chang, S. (1994) Genomics 21, 354–358
21. Sarthy, A. V., Schopp, C., and Idler, K. B. (1994) Gene (Amst.) 140, 121–126
22. Iwata, T., Tocques, N., Zimonjic, D. B., Karlsson, C., Hoog, J.-O., Vaca, G., Rodriguez, I. R., and Carper, D. (1995) Genomics 26, 55–62
23. Niimi, T., Yamaishi, O., and Yaguma, T. (1996) Insect Mol. Biol. 5, 269–280
24. Carr, I. M., Markham, A. F., and Coletta, P. L. (1997) Eur. J. Biochem. 245, 760–767
25. Bischoff, W. L. (1976) Biochem. Genet. 14, 1019–1039
26. Danielsson, O., Atrian, S., Luque, T., Hjelmqvist, L., Gonzalez-Duarte, R., and...
Sorbitol Dehydrogenase of *Drosophila*: GENE, PROTEIN, AND EXPRESSION
DATA SHOW A TWO-GENE SYSTEM
Teresa Luque, Lars Hjelmqvist, Gemma Marfany, Olle Danielsson, Mustafa El-Ahmad,
Bengt Persson, Hans Jörnvall and Roser Gonzàlez-Duarte

*J. Biol. Chem.* 1998, 273:34293-34301.
doi: 10.1074/jbc.273.51.34293

Access the most updated version of this article at [http://www.jbc.org/content/273/51/34293](http://www.jbc.org/content/273/51/34293)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 48 references, 10 of which can be accessed free at [http://www.jbc.org/content/273/51/34293.full.html#ref-list-1](http://www.jbc.org/content/273/51/34293.full.html#ref-list-1)