Amino acid sequence homology among the 2-hydroxy acid dehydrogenases: Mitochondrial and cytoplasmic malate dehydrogenases form a homologous system with lactate dehydrogenase

(evolution/active site/enzyme)

JENS J. BIRKTOFT, ROSS T. FERNLEY*, RALPH A. BRADSHAW, AND LEONARD J. BANASZAK

Department of Biological Chemistry, Washington University School of Medicine, St. Louis, Missouri 63110.

Communicated by Robert L. Hill, July 15, 1982

The amino acid sequence of porcine heart mi-ABSTRACT tochondrial malate dehydrogenase (mMDH; L-malate:NAD oxidoreductase, EC 1.1.1.37) has been compared with the sequences of six different lactate dehydrogenases (LDH; Llactate:NAD⁺ oxidoreductase, EC 1.1.1.27) and with the "x-ray" sequence of cytoplasmic malate dehydrogenase (sMDH). The main points are that (i) all three enzymes are homologous; (ii) invariant residues in the catalytic center of these enzymes include a histidine and an internally located aspartate that function as a proton relay system; (iii) numerous residues important to coenzyme binding are conserved, including several glycines and charged residues; and (iv) amino acid side chains present in the subunit interface common to the MDHs and LDHs appear to be better conserved than those in the protein interior. It is concluded that LDH, sMDH, and mMDH are derived from a common ancestral gene and probably have similar catalytic mechanisms.

A large number of enzymes occur in multiple forms (1). In some cases, the different forms are compartmentalized in unique cellular locations, such as in the mitochondrion and cytoplasm. For the aspartate aminotransferases, which show such a distribution, it has been shown that the mitochondrial and cytoplasmic forms are homologous proteins as judged by comparisons of their amino acid sequences and their molecular structures (2). The cytoplasmic (or soluble) and the mitochondrial forms of malate dehydrogenase, sMDH and mMDH (L-malate:NAD⁺ oxidoreductase, EC 1.1.1.37), respectively, represent a comparable pair of enzymes.

The MDHs are the products of different nuclear genes and both are synthesized in the cytoplasm (3). sMDH remains in the cytosol after synthesis whereas mMDH is translocated into the mitochondrial matrix, a process that may be facilitated by the presence of a leader sequence at the amino terminus of the polypeptide chain (4). The compartmentalization of sMDH and mMDH has physiological significance; sMDH functions in the malate-aspartate shuttle whereas mMDH is a key enzyme in the same shuttle and in the citric acid cycle. Both sMDH and mMDH are dimers of identical subunits having subunit molecular weights of 36,000 and 34,000, respectively (5). Most of their catalytic properties are similar, with the exception of the different inhibitory effects of high concentrations of the substrate oxaloacetate and the substrate analogs mono- and difluorooxaloacetate (6). The activity of sMDH appears less affected than that of mMDH by these compounds. Both sMDH and mMDH have been crystallized and studied by x-ray diffraction methods. However, while the x-ray studies on mMDH are still in their

preliminary stages, the molecular structure of sMDH is known at high resolution (7). The molecular structure of porcine heart sMDH is very similar to that of another NAD-dependent dehydrogenase, lactate dehydrogenase (LDH; L-lactate:NAD⁺ oxidoreductase, EC 1.1.1.27) (8).

Unlike the MDHs, which are cellular organelle specific, the lactate dehydrogenase isozymes are found in the cytoplasm of different cell types. Three different forms of LDH have been described: the A or muscle form (LDHA), the B or heart form (LDHB), and the C or testes form (LDHC) (1). Although the A and B forms occur throughout the respective organism, their relative abundance varies from tissue to tissue. LDHA is more abundant than LDHB in liver and skeletal muscle while LDHB is present at higher levels in heart and kidney tissue (1). The LDHC gene is expressed only in adult testes (3). The threedimensional structures and the amino acid sequences of the three LDHs are very similar despite the fact that they are the products of three distinct genes (9-11). The structural homology among the LDHs extends to the level of subunit-subunit interactions in that it is possible to form heterotetramers of the three forms of LDH, as has been observed both in vivo and in vitro (12).

The recent determination of the primary structure of pig heart mMDH (13) has made it possible to compare the amino acid sequence of this enzyme with those of the LDHs. A preliminary "x-ray sequence" for porcine heart sMDH has also been determined (unpublished data). Thus, it is possible to compare the amino acid sequences and structures of LDH, mMDH, and sMDH in detail.

MATERIAL AND METHODS

In the comparison studies, the amino acid sequence of mMDH is that reported by Fernley *et al.* (13). The amino acid sequences for the LDHAs and LDHBs and their alignments are taken from Eventoff *et al.* (14) and that of mouse LDHC is as reported by Pan *et al.* (15). Atomic coordinates for the apo form and ternary complex of dogfish LDHA and for the apo form of mouse LDHC were obtained from the protein data bank at Brookhaven National Laboratory (16). The x-ray sequence for sMDH has been derived by correlating the available information on the amino acid sequence of a number of peptides (unpublished observations) with electron density maps based on an x-ray structure

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact:

Abbreviations: sMDH and mMDH, cytoplasmic and mitochondrial malate dehydrogenase, respectively; LDH, lactate dehydrogenase; LDHA, muscle LDH; LDHB, heart LDH; LDHC, testes LDH. * Present address: The Howard Florey Institute of Experimental Phys-

^{*} Present address: The Howard Florey Institute of Experimental Physiology and Medicine, University of Melbourne, Parkville, Victoria 3052, Australia.

partially refined by restrained least-squares methods (17, 18). The refinement has been continued, and at this time (June 1982) the conventional crystallographic *R* factor for all diffraction data to 2.5-Å resolution is 33%. A listing of the "x-ray sequence" based on the partially refined x-ray coordinates can be obtained from the authors upon request. Considering that the resolution of the x-ray diffraction data of sMDH is only 2.5 Å and that only 92% of all nonhydrogen atoms, as based on the amino acid composition (5), were used in the initial stages of the refinement, the amino acid sequence of sMDH presented here undoubtedly contains some incorrect assignments. The structural homology between sMDH and LDH is described elsewhere (18) and forms the basis of the alignment of sMDH and LDH shown in Fig. 1.

The amino acid sequences of mMDH and several LDHs were initially compared by using the procedure described by Mc-Lachlan with spans of 11 residues and weighting factors of 1, 2, 2, 3, 3, 3, 3, 3, 2, 2, 1 (19). The effect of the individual amino acid substitutions on the structure, function, and properties of mMDH was analyzed by using the known structure of dogfish apo-LDHA as a starting point and making appropriate changes in the model. This analysis was done on a MMS-X interactive graphics system (20) using the programs BUILD (21) and NEWNIP (22).

RESULTS AND DISCUSSION

Overall Similarity of Structures. The alignment of the amino acid sequence of mMDH, the sequences of six different LDHs, and the x-ray sequence of sMDH is shown in Fig. 1. In this alignment, 31% of the 314 amino acids in mMDH have an identical counterpart in one of the six LDH sequences. When the comparison is done with the individual LDH sequences, the homology index varies between 22% and 24% as shown in Table 1. While this value is low, it is by no means outside the range observed for related protein systems (23). The alignment shown in Fig. 1 introduced about 14 gaps in the sequences or 4.4 gaps per 100 residues. This value, although high, is comparable with that obtained with the bacterial serine proteases (23, 24). These gaps are, with one exception, located on the molecular surface. Only at the junction between the nucleotide binding and the catalytic domain, residue 146,[†] does an internally located deletion occur. The overall sequence homology appears to be slightly better in the NAD-binding domain than in the catalytic domain, although the differences, as given in Table 1, may not be significant. The region where the greatest difficulty was encountered in aligning the sequences involved residues 182-209, which includes the β -sheet strands β H and β J and helix α 1G. As used in Fig. 1, α and β indicate helical and β secondary structure, respectively, and the roman letter is assigned relative to the sequence order beginning at the amino terminus. In both LDH and sMDH, β H, β J, and α 1G are segments of the polypeptide chain that follow a meandering pathway across the protein surface and make several hairpin loops. In this region, the mMDH sequence is 9 residues shorter than the dogfish LDH sequence and 12 residues shorter than the other listed LDH sequences. It is noteworthy that, in this part of the structure, rather large conformational differences are observed between LDH and sMDH (18) and among the three known LDH structures (10, 11)

Amino acid residues in the interior of these dehydrogenases appear to be less susceptible to evolutionary changes than those that are in contact with the surrounding solvent medium, as can be seen from Fig. 1. If one defines as internal those residues that, in the apo form of a sMDH subunit, do not appear to be exposed to solvent, visual inspection of the hypothetical mMDH model indicates that 35 out of 94 (37%) internal residues in mMDH have an identical counterpart in one of the six LDHs. The remainder are for the most part replaced by residues of a similar chemical nature. In a few instances, a nonpolar residue is replaced by a residue with a polar side chain or vice versa—for example, leucine to threonine at position 37. But in no instance are any charged side chains introduced into the protein interior.

In this type of sequence comparison, one of the most troublesome problems is the replacement of a nonproline residue by a proline. In some instances, this can necessitate a conformational change in the homologous protein. No difficulties were encountered in incorporating most of the prolines found in the sequence into the putative mMDH structure, although three unusual substitutions did occur-i.e., prolines-15, -125, and -297—in the middle of α -helices. These replacements are expected to interrupt the helical structure because of the disruption of the helical hydrogen bonding pattern and because of steric problems caused by the pyrrolidine ring. However, the presence of a proline residue in the middle of a stretch of helical structure is not without precedent. In all such instances, the proline introduces a bend in the otherwise linear helical axis. Typically, in the case of lamprey hemoglobin, proline-25 is found in the A-helix and the bend is about 20° (25). In LDH, helix αD is separated from helix αE by a proline and the axes of these two helices make an angle of about 30° with each other. In sMDH, no proline is found in the corresponding position and helices αD and αE do, in fact, form a continuous and straight helix (7). Prolines-15 and -125 are located in internal positions in helices αB and $\alpha 1F$, respectively, both of which are part of the structure of the NAD-binding domain. Helix α B would also be part of the Q-axis subunit interface in mMDH as will be discussed below. Proline-297 occurs in the middle of helix α H, which is stretched across the surface of the molecule and is near the COOH-terminus. This proline would be located on the side of the helix that is exposed to solvent and a helical kink, similar to those observed in lamprey hemoglobin and LDH, could be accomodated in the putative mMDH structure without causing major structural disruption. Considering that the overall homologies adjacent to these proline residues are reasonably good, it is possible that no major structural rearrangements have occurred.

Subunit-Subunit Interactions. Both mMDH and sMDH are found as dimeric proteins (5), whereas LDH usually occurs in a tetrameric form (26). Comparison of the structure of dogfish LDH and pig heart sMDH showed that these two oligomeric enzymes share a common twofold symmetry axis, the Q axis according to the nomenclature of Rossmann et al. (27). The two additional symmetry elements present in LDH are called the P and R axes, respectively. Inspection of the mMDH sequence showed that those residues located in the Q-axis interface are better preserved than those participating in the two other types of subunit-subunit interactions. These results are summarized in Table 2. It should also be noted that, in LDH, a larger number of residues participate in interactions in the Q-axis interface than in the other two and that the residues around the Q axis are better preserved among the LDHs than those in the P- and R-axis interfaces (13). Overall, the data given in Table 2 suggest that the mMDH dimer is formed around the Q axis and thus resembles the sMDH dimer.

At low pH, mMDH undergoes reversible dissociation into monomers, such that, at pH 5.0, the enzyme at low concentrations is exclusively in the monomeric form (28). Several charged

[†] Unless stated otherwise, the numbering system used is based on the mMDH sequence (13). The corresponding numbering for LDH can be obtained from Fig. 1.

NNDH No: LDH No: MNDH (P) LDHB (P) LDHB (C) LDHA (P) LDHA (C) LDHA (D) LDHA (D) LDHC (M) SMDH (P)	0	s	P] 22 N N N N Z 1	K K K K K K R	2 V I X I I V		5 V 1 V V V V V V V V V V V T			G G X G X G A	1(S () Z 1 Z 1 Z 1 Z 1 Z 1 Z 1 Z 1 Z 1 Z 1 Z 1		I M M M M M F	S G A A A A T	1 QCCCCCCL	5 P 1 A	2 S 1 S 1 S 1 S 1 S 1 S 1 S 1 S 1 S	40 L I I I I I I		L G G M M M L B	K		S F	25 45 1 1 1 1 1 G				S D D D B D D I				Y V V V V V		35 1 / 7 1 7 1 7 1 7 1 7 1 7 1 7 1 7 1	55 		E E E Z E T Z	- D D D B D B Z					A M M M M M M M M M M	6 A M M M M M M L M	5 D 1 D 1 D 1 D 1 D 1 D 1 D 1 D 1 D 1 D 1		
				R		I	[]	L I N	L I I N		N	NI	, ^I			I	1 0	2 9	I	I	Q				I				R		ı″	1	N	N	N 1	N	H		Q	Q M	1	¢	ີ	C		Q	Q	Q
MMDH No: LDH No: MMDH (P) LDHB (P) LDHB (C) LDHA (C) LDHA (C) LDHA (D) LDHC (M) SMDH (P)	1 6 6 6 6 6 7 -	70 E S S S S S S -	50 T L L L L L L L	D A F L L F L L F L L - I	- QQRKHSP	75 T T T T T T T T T				80 L A S S F F	G N B G G G G G	60 P 1 K 1 K 1 K 1 K 1 K 1 K 1	85 Q Y Y Y Y Y Y Y Y S S	L	P S A N S S N G	6 V V V V V B	5 C 1 T 1 T 1 T 1 S 1 S 1 Y 1		90 GNNNHGNS	C S S S S S S Z	O D K K R K K B	V I L L L V		5 V I V V V I I V I I I I V I	75 P T T T T T L		10 G 1 G 1 G 1 G 1 G 1 G 1 G 1 G 1	P R R R R R R S	80 R Q Q Q Q Q M ▲	1 K Q Q Q Q Q Q V K	05 P (E		85 5 5 5 5 7 -] R R R R R R R R R R R R	10 D L L L L L L L	D 1 N 1 N 1 N 1 N 1 D 1 K 1	90 F V V V V L V L V L	N O O O O O O O O	I 1 5 T R R R R R R R R		93 A 1 V 1 V 1 V 1 V 1 V 1 V 1	5 E 1 H 1 H 1 H 1 H 1 K 1	12 I V V P I P I P I P I P	O A K K K K K K K K K	1 F F F F F C	00 L I I I I I Z	12 T / I I I I I I I I G /	SAQZPN PNPGA
	Q			I	:	L R		β I	C							1				1	I	1	β1 1	D I N	I N	N	N	N									1	•	1	1		1	a [1 N)		N I		•
MMDH No: LDH No: MMDH (P) LDHB (P) LDHB (C) LDHA (C) LDHA (C) LDHA (D) LDHC (M) SMDH (P)	C 1 1 1 1 1 1 1 1	105 A V V V V V V B	1 Q K X K K K K K	30 4 C 7 S 7 S 7 S 1 S 8 S 7 S	P P X P P P K	1: - 1 - 1 - 1 - 1 - 1 - 1 K 9			13 1 1 2 1 2 2 1 1 2 1 2 1 2 1 2 1 2 1 2	5 C I X L X L I I I	1 V V V V V V V V V V		14 5 N 5 N 5 N 5 N 5 N 5 N 5 N 5 N 5 N 5 N	0 P P P P X P P P	V V V V X V X V A	NDDDXDDT	14 S 1 I 1 I 1 X 1 V 1 I 1 B 1		12 P Y Y Y X Y L	5 I V V V V V V V T	1 T A T A A V A	50 W 1 W 1 W 1 W 1 W 1 W 1 W 1 W 1	13 K 1 K 1 K 1 K 1 K 1 K 1	30 V F L S L S I S L S L S B S	K G G G G G A	K		5 V 	15 Y L F F L F L	5 N P P P P P P P P B	14 P 1 K 1 K 1 K 1 G 7 K 1	40 N K H J H J H J H J K K K		0 F I I I I I V	1 GGGGGGG B	45 V 1 S (0 S (0 S (0 S (0 S (0 S (0 S (0)		5 T H N N N N N N N N N N N N N N N N N N	L L L L L L L L	1 D D D D D D D D B	50 1 5 5 5 5 5 1	70 V 1 A	R I I I I I I I I I I I I I I I I I I I	N R R R R R R R R R R R R R R R R R R R	15 A Y Y H Y S	5 F L L L L L L L L L M	V N N N N I L	A A E E E E E Z
	1	a	E	1	1	1.	1	1	/: 1	} E 1	I	I] N 2	I N	1	1	1	1 1	. 1	a 1	F I	1			1		1				R	i	r-	1	βı	1	- 1 : N	1 1	-	I N	1		Q	2 21 Q 1 S	, 7 1		 I	1	Q Q
MMDH No: LDH No: MMDH (P) LDHB (P) LDHA (C) LDHA (C) LDHA (C) LDHA (C) LDHA (C) SMDH (P)	16 L R R R R R K K	0 180 K L L L L L L L		DH HH HH N S	16 P P P P S P P	85 A S S S S S S S S		S H H H H H H H H H	17 G G G G B	0 190 W W W W W V			0000000	1751 G E E Q E E Z	95 H H H H H H H	A G G G G G G G G	G I	18 - S - S - S - S - S - S	I S S S S V S Z	20(1 V V V P V F	P A A P P S P G	1 V V V V I T	85 I 9 W 9 W 9 U 1 I 1 	20 S Q S G S G S G S G S G S G S G S G S G S G	5 C V V X M V L	1 T N N X W N Z	90 20 P 1 V 1 V 1 X 1 N 1 L 1 L 1	9 V G G X G G X G G X G G X G G X G G X G G X G G X G G X G G X G G X G G X G G X G G X G G X G G X G G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X G X	D V V K	1 FVSSX- TZ	95 2 P - - - - -		21 Q Q Q Q Q Q C N C N C S C S C S C S C S		200 S N H H H B R	2: T · · P ! P ! P ! P ! P !	15 E H E H E L D H E L S K		T T T T T T B	2 - D B D B N B Z	20 	- DBDBBW	S 1 S 1 K 1 K 2 S (22 W W W W	5 KKKKK KKK S	- E E A E K N I	H V H V H V H L H V H V H
			F	P	Q	Q P	I			1	1	[]	1		s			I	I	I	I	1	ı ′′	, P	I	1	1	9				P	P	ī											u	10	1	
MMDH No: LDH No: MMDH (P) LDHB (P) LDHB (C) LDHA (P) LDHA (C) LDHA (D) LDHA (D) SMDH (P)	230 G K K K K K K K K K K	205 R M Q E Q D G V a	1 Q V V V V V V V V V V I Z 1G	E E D D B Z	235 S S S S S S R I	10 G T A Y A Y A Y A Y A Y G V	E E E E Z V	VVVVVVVVVVVVVVVVVVVVVVVVVVVV	240 V I I I I V G	215) K K K K K K K K C Z		(C C C C C C C A	2 2 0 2 4 5 7 Y Y Y Y Y Y B	GTTTTTB	S N N N S S S			O T G G G G G G G G G G G G G G G G G G	222 L L L L L L F	S S S S S S S S S S	2 M V V V L	55 A 1 A 1 A 1 A 1 B 1 	23 Y A D L Z L D L D L D L B L J L G	0 G I C A A A V	2 A E Z E E E E R K	60 R 1 S 1 T 1 S 1 T 1 S 1 H 1	23 7 V H L H L H L H I H I H I L L V	5 F K K K K K K	S N N N G 	2	65 L \ L \ L \ L \ L \ L \ L \ L \ S \ 2	24 7 DR R RR R RR R RR R RR R R	0 A I V V V V V L	2 M H H H H H H A	70 N P S P P P Z	24 G K V S I S I S V S V T M G		G M L M A L I	279 V V I V V V A		50 E G G G G G D G G	C	S F		25 K M T L M F F K	S Y Y Y H Y H Y H	80 Q E G I G I G I G I G I A A
			1 1 S			I	Q	Q	I		Q	2	Q	N	N	1	Q () N	I	Q	Q	I	10	Q I	I				R	R		RI	? P	P	1	I	1 1	1		I								
MMDH No: LDH No: MMDH (P) LDHB (P) LDHB (C) LDHA (P) LDHA (C) LDHA (D) LDHC (M) SMDH (P)	T E Q K K K K S	260 N 1 D 1 D 1 D 1 D 1 D 1 P 1	28 C P D V D V V V V V V V K I -	5 F F F F F F F		65 SILSL SV SV SL SV SL SC	29 P P P P P P P V		L V V I V V R		29 G I N S G G D G D G D I G D I G D I G D I G D I G D I S G D I S G D I S G D I S G D I S G D I S G D I S G D I S G D S G D S S G D S G D S S S S S S S S S S S S S S S S S S S) S K K S S C S K S S C S K S S S S S S S	G G G G G T	275 I L I X I X /3 M	3 E T T S X S T K	01 K S D X N B I 	N 1 1 - 1 1 1 - 1 1 1 - 1 1 1 - 1 1 1 - 1 1 1 - 1 1 - 1 1 - 1 1 - 1 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	28 / I / I / V / V / V / V / V	0 INNKXK SG	30 G Q V X M V L	K K K T I K B P	2 I I I I I I I I I	85 K 1 K 1 K 1 K 1 K 1 K 1 K 1 K 1	31 P F D D D D P D P D P N D F	O E E E E E E Z S	2 V V E E L R	90 K 1 A 1 E 0 A 1 E 0 A 1 E 1		5 A K K K K Q K N	2 B N K K K K K K L	95 A S S S S S S T 	32 I 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1			300 K W W W W W W	3 G S G D B Z	25 S 1 I Q I Q I Q Z Z		30 K D E E D D T	5 GLLLLL E	30 E K K Q Q K E F	E D D	3 F L F F F F L E				314 	

FIG. 1. Sequence homology among 2-hydroxy acid dehydrogenases. (P), (C), (D), and (M) refer to pig, chicken, dogfish, and mouse, respectively, as the enzyme source. α and β indicate secondary structural elements according to the usual nomenclature for the dehydrogenases (7). Various features of individual residues are indicated below the structural element designations: I, internally located; C, participating in catalysis; N, in contact with NAD; S, involved in substrate binding; Q, part of the Q-axis interactions; P and R, involved in P-axis or R-axis interactions, respectively (LDH only). The upper line of the numbering system refers to the linear mMDH sequence (13), and the lower line refers to the crystal structure of dogfish LDHA (14). The P and R assignments are based on table 1 in ref. 10, and the I, C, N, S, and Q assignments are from analysis of the sMDH structure (17, 18). Amino acids are indicated by the single-letter code; the NH₂ terminus of sMDH is acetylated and is indicated by an O.

Table 1.	Summary of	homologies	among	the 2-hyd	roxy
acid dehy	drogenases				

	Identities,	
Comparison	no./no.	% homology
mMDH vs. any of six LDH sequences	97/314	31
mMDH vs. LDHA (pig)	69/314	22
mMDH vs. LDHA (dogfish)	76/314	24
mMDH vs. LDHB (pig)	68/314	22
mMDH vs. LDHC (mouse)	68/314	22
Dinucleotide binding domain		
mMDH vs. any of six LDH sequences	50/146	34
mMDH vs. LDHB (pig)	34/146	23
Catalytic domain		
mMDH vs. any of six LDH sequence	47/168	28
mMDH vs. LDHB (pig)	34/168	20

Identities were not tabulated for the chicken LDH enzymes because of the long stretches for which the sequences have not been determined.

side chains are located in the putative Q-axis interface in mMDH (aspartic acids-43 and -164, histidine-46, glutamic acid-212, lysine-215). Protonation of a single amino acid has been shown to be responsible for this pH-dependent dissociation (29). Either one of the acidic amino acids or possibly the histidine could be responsible for this dissociation. Except for aspartic acid-164, these residues are also conserved in LDH and this enzyme has been shown to undergo reversible dissociation into dimers and monomers at pH 3.5-4.5 (30).

Active Site and Substrate Binding. The principal difference in the enzymatic properties of LDH and MDH is found in their substrate specificities although both enzymes utilize 2-hydroxy (or the corresponding 2-keto) acids as substrates. LDH shows a preference for straight-chain alkyl 2-hydroxy acids, L-lactate (pyruvate) being optimal (31), whereas MDH favors 2-hydroxy dicarboxylic acids with L-malate (oxaloacetate) being the preferred substrate(s) (32). In addition, the dissociation constants for the cofactors NAD and NADH and the pH dependence of their binding are similar for both the MDHs and LDHs (26, 33, 34).

The catalytic sites in both LDH and sMDH are believed to contain a histidine-aspartate pair linked by a hydrogen bond (17) and both of these residues are preserved in mMDH as histidine-176 and aspartic acid-149. This aspartate is nearly completely shielded from the solvent by surrounding protein, and many of the amino acid side chains that form this shield are either invariant (residues 148, 152, and 176), or replaced by similar residues in mMDH (residues 146, 151, 172, 180, and 206). Thus, the function of this histidine-aspartate pair should be the same in mMDH as has been proposed for LDH and sMDH (18).

It has previously been shown that the binding sites for the co-factor NAD (NADH) are structurally very similar in LDH and sMDH (8, 17). The adenine ring is located in a shallow crev-

Table 2. Conservation of residues in subunit-subunit interfaces of the 2-hydroxy acid dehydrogenases

			Identities									
	Resi- dues.*	mM vs. ar six L	DH ny of DHs	mM vs. Ll (pi	DH DHA g ⁾	mM vs. Ll (pi	DH DHB g ⁾	Six LDHs				
Axis	no.	No.	%	No.	%	No.	%	No.	%			
Q	22	15	68	13	59	13	59	15	68			
Ρ	14	3	21	1	7	3	21	6	43			
R	12	2	17	2	17	1	8	4	33			

* Based on table 2 of ref. 13.

ice lined by predominantly hydrophobic residues. Four out of nine residues contained in this crevice are conserved in mMDH—i.e., residues 33, 74, 76, and 96—and the remainder consist of fairly conservative substitutions—i.e., residues 6, 32, 34, 61, and 100. Therefore, the hydrophobic nature of the adenine binding site is also conserved. The residue at position 61 is a tyrosine in LDH and is hydrogen bonded to a nitrogen on the adenine ring. This residue is replaced in mMDH by a glutamine so the possibility for maintaining this particular hydrogen bond still exists. The serine assigned to this position in sMDH does not appear to be in contact with NAD (17).

The interactions between the enzyme and the ribose/pyrophosphate/ribose moieties of NAD are largely polar in nature, with nearly every oxygen atom forming a hydrogen bond with the enzyme (17). Most of the hydrogen bonds described for the coenzyme complexes of sMDH (17) and LDH (9) appear to exist in mMDH. The most noticeable difference is the replacement of arginine by proline at position 79 in mMDH. In LDH, this residue is thought to be involved in an ionic interaction with the negatively charged phosphate oxygens of the coenzyme. However, the succeeding two residues in the mMDH sequence are both basic, arginine-80 and lysine-81, and one of these may serve the same function as the arginine in LDH. In fact, modification of lysine-81 in mMDH causes inactivation and this reaction is inhibited by NAD, ADP, and AMP (35), strongly suggesting that lysine-81 in mMDH is involved in interaction with the pyrophosphate in NAD. In sMDH, residue 79 is thought to be serine and this side chain appears to be in contact with NAD (17).

All of the basic residues involved in NAD binding are located in the "loop" region, which in mMDH includes residues 77–90. In both LDH and sMDH, this loop undergoes a substantial conformational change on formation of the enzyme–NAD complex (9, 36). Several other residues in this part of the polypeptide chain are also conserved, including glycines at positions 77 and 83. It is possible that these two residues function as "hinges" for the conformational change accompanying coenzyme binding as observed in LDH and sMDH.

Two additional amino acids in LDH are involved in ionic interaction with NAD and are found in mMDH as well. Aspartic acid-33 is hydrogen bonded to the adenosine ribose O-2 position and is found in all NAD-binding proteins (27). Lysine 58 (in the LDH numbering system) is located nearby and probably is involved in an electrostatic interaction with the pyrophosphate moiety. It is conservatively replaced by histidine in mMDH.

In the dehydrogenases, the contacts between the nicotinamide ring of the coenzyme and the enzyme are predominantly hydrophobic in nature. Most of these residues are either invariant—i.e., residues 117, 118, and 148—or conservatively replaced—i.e., residues 10, 116, 221, and 223—by chemically similar side chains. Asparagine-118 is located close to the nitrogen of the nicotinamide ring.

When bound to LDH, the carboxamide group of NAD is probably hydrogen bonded to the carbonyl group of residue 117 and to the side chain hydroxyl group of a serine at position 145 (9, 11). These two hydrogen bonds help in the orientation of the nicotinamide ring, such that the proper stereochemistry is provided for hydride transfer. In mMDH, serine-117 is present, but unexpectedly the serine at position 145 is replaced by valine. However, the two succeeding residues (positions 146 and 147) in the mMDH sequence are both threonine. mMDH, like sMDH, has one residue less than LDH in this region (Fig. 1). At this place, a minor structural difference is observed between LDH and sMDH and may also be found between mMDH and LDH. Therefore, although the overall chemical nature of the nicotinamide binding subsite is similar in all of these enzymes, some differences may occur near the carboxamide group.

The substrate for the MDHs differs from that of LDH by the presence of an additional carboxyl group. In both LDH and sMDH, the end of the substrate opposite from the 1-carboxyl group is pointed toward a solvent-filled cavity. Only a few contacts appear to exist between enzyme and substrate at this point, explaining the lack of strict specificity of LDH. In a modelbuilding experiment, an L-malate ion was placed in the putative substrate binding site of mMDH such that hydrogen bonds were formed between arginine-152 and the 1-carboxyl group and between histidine-176 and the 2-hydroxyl group of the substrate. When this is done, the only polar side chain in mMDH that could possibly interact with the 4-carboxyl group of L-malate appears to be that of glutamine-207. The corresponding residue in LDH is valine and, in sMDH, is currently estimated to be glutamine or a glutamate residue. This side chain could account for the difference in substrate specificity between the MDHs and the LDHs.

Finally, it has been proposed that a so-called ring of negative charges near the active site of LDH may be of some functional significance (14). This group of negative charges includes residues 84, 178, 208, and 212, but only the last two are conserved in mMDH. This lack of homology with the MDHs suggests that the clustering of negative charges in this region of LDH probably has little functional significance.

CONCLUSION

Comparison of the amino acid sequence of mMDH with the sequences of several LDHs has shown that these two enzymes are homologous. Because it has previously been shown that sMDH is also structurally similar to LDH, it can now be inferred that mMDH and sMDH are homologous enzymes. Comparison of the tentative sMDH sequence derived from the electron density map with those of mMDH and LDH indicates that sMDH will show some similarities at the amino acid sequence level as well. This means that the mitochondrial and cytoplasmic forms of MDH must have a common evolutionary origin as has been established for the aspartate aminotransferases (2).

Those amino acids that are directly involved in the mechanism of action of the 2-hydroxy acid dehydrogenases are extensively conserved. Only with respect to the substrate binding site are significant differences between MDH and LDH apparent. Formation of the oligomeric forms of sMDH and LDHs is dominated by the dimerization around the so-called Q axis, where a high degree of conservation of amino acid residues is also found in mMDH. Thus, it is predicted that protein-protein interactions in mMDH are probably the same as in sMDH. The similarities among the 2-hydroxy acid dehydrogenases as exemplified by MDH and LDH suggest that they also originated from a common ancestral gene.

We are grateful to Dr. A. L. McLachlan for a copy of his sequence comparison program and to Ms. Solveig Storvick-Pollei for the preparation of this manuscript. This work was supported by National Science Foundation Grant PCM-7921864 (to L.J.B.), U.S. Public Health Service Research Grants GM-13925 (to L.J.B.) and AM-13362 (to R.A.B), and a National Institutes of Health biomedical research support grant to Washington University School of Medicine (J.J.B.).

- 1. Markert, C. L., ed. (1977) Isozymes (Academic, New York), Vols. L-3.
- Ford, G. C., Eichele, G. & Jansonius, J. N. (1980) Proc. Natl. Acad. Sci. USA 77, 2559–2563. 2
- Edwards, Y. M. & Hopkinson, D. A. (1977) in Isozymes: Current 3 Topics in Biological and Medical Research, eds. Rattazz, M. C.,

Scandalius, J. G. & Whitt, G. S. (Liss, New York), Vol. 1, pp. 19-78.

- Aziz, L. E., Chien, S.-M., Patel, H. V. & Freeman, K. B. (1981) 4. FEBS Lett. 133, 127-129.
- Banaszak, L. J. & Bradshaw, R. A. (1975) in The Enzymes, 3rd 5. Ed., ed. Boyer, P. D. (Academic, New York), Vol. 11A, pp. 369 - 396
- Kun, E. & Volfin, P. (1966) Biochem. Biophys. Res. Commun. 22, 6. 187–193.
- Hill, E. J., Tsernoglou, D., Webb, L. & Banaszak, L. J. (1972) 7. J. Mol. Biol. 72, 577–591.
- Webb, L. E., Hill, E. J. & Banaszak, L. J. (1973) Biochemistry 8 12, 5101-5108.
- Holbrook, J. J., Liljas, A., Steindel, S. J. & Rossmann, M. G. (1975) in *The Enzymes*, 3rd Ed., ed., Boyer, P. D. (Academic, 9. New York), Vol. 11A, pp. 191–292.
- 10. Musick, W. D. L. & Rossmann, M. G. (1979) J. Biol. Chem. 254, 7621-7623.
- Grau, U. M., Trommer, W. E. & Rossmann, M. G. (1981) J. Mol. 11. Biol. 151, 289-307
- Ansari, A. A. (1981) Biochem. J. 199, 75-79. 12.
- Fernley, R. T., Glatthaar, B. E., Sutton, M. R. & Bradshaw, R. 13.
- A. (1982) J. Biol. Chem., in press. Eventoff, W., Rossmann, M. G., Taylor, S. S., Turff, H.-J., 14. Meyer, H., Keil, W. & Kiltz, H.-H. (1977) Proc. Natl. Acad. Sci. USA 74, 2677-2681.
- Pan, Y.-C., Huang, S., Marciniszyn, J. P., Jr., Lee, C.-Y. & Li, 15. S. S.-L. (1980) Hoppe-Seyler's Z. Physiol. Chem. 361, 795-799.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. 16. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* 112, 535–542. Birktoft, J. J., Fernley, R. T., Bradshaw, R. A. & Banaszak, L. K.
- 17. (1982) in Molecular Structure and Biological Activity, eds. Griffin, J. F. & Duax, W. L. (Elsevier/North-Holland, New York), pp. 37-55
- Birktoft, J. J. & Banaszak, L. J. (1982) J. Biol. Chem., in press. 18.
- 19 McLachlan, A. D. (1972) J. Mol. Biol. 64, 417-437.
- Barry, C. D., Bosshard, H. E., Ellis, R. A. & Marshall, G. R., 20. (1974) Fed. Proc. Fed. Am. Soc. Exp. Biol. 33, 2368–2372.
- Miller, J. R., Abdel-Meguid, S. S., Rossmann, M. G. & Ander-21. son, D. C. (1981) J. Appl. Cryst. 14, 94-100. Lederer, F., Glatigny, A., Bethge, P. H., Bellamy, H. D. & Ma-
- 22 thews, F. S. (1981) *J. Mol. Biol.* 148, 427–448. Doolittle, R. F. (1981) *Science* 214, 149–159. James, M. N. G., Delbaere, L. T. J. & Brayer, G. D. (1978) *Can.*
- 23
- 24. . Biochem. 56, 396–402.
- 25. Hendrickson, W. A. & Love, W. E. (1971) Nature (London) 232, 197-203.
- 26. Everse, J. & Kaplan, N. O. (1973) Adv. Enzymol. 37, 61-133.
- 27. Rossmann, M. G., Liljas, A., Bränden, C.-I. & Banaszak, L. J (1975) in The Enzymes, ed. Boyer, P. D. (Academic, New York), 3rd Ed., Vol. 11A, pp. 61-102
- 28. Bleile, D. M., Schulz, R. A., Harrison, J. H. & Gregory, E. M. (1977) J. Biol. Chem. 252, 755-758.
- Wood, D. C., Hodges, C. T., Howell, S. M., Clary, L. G. & Harrison, J. H. (1981) J. Biol. Chem. 256, 9895–9900. 29.
- 30. Rudolph, R., Heider, I., Westhoff, E. & Jaenicke, R. (1977) Biochemistry 16, 3384–3390. Schwert, G. W. & Winer, A. (1963) in The Enzymes, eds. Boyer,
- 31. P. D., Lardy, H. & Myrbäck, K. (Academic, New York), 2nd
- Ed., Vol. 7, pp. 127–148. Kun, E. (1963) in *The Enzymes*, eds. Boyer, P. D., Lardy, H. 32. & Myrbäck, K. (Academic, New York), 2nd Ed., Vol. 7, pp. 149-160
- Holbrook, J. J. & Wolfe, R. G. (1972) Biochemistry 11, 2499-2502. 33.
- Rupley, J. A., Forster, L. S., Torikata, T., Johnson, R. E. & O'Neal, C. C., Jr. (1980) Biochem. Biophys. Res. Commun. 93, 34. 654-660.
- Wimmer, M. J., Mo, T., Sawyers, D. L. & Harrison, J. H. (1975) J. Biol. Chem. 250, 710-715. 35.
- Weininger, M., Birktoft, J. J. & Banaszak, L. J. (1977) in Pyridine 36. Dependent Dehydrogenases, ed. Sund, H. (de Gruyter, Berlin), pp. 87-100.