Refined Crystal Structure of Mitochondrial Malate Dehydrogenase from Porcine Heart and the Consensus Structure for Dicarboxylic Acid Oxidoreductases[†,‡]

William B. Gleason,[§] Zhuji Fu,[∥] Jens Birktoft,[⊥] and Leonard Banaszak[*,#]

Biomedical Engineering Center, Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota 55455, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, Fujian, Peoples Republic of China, Roche Research Center, Hoffmann-LaRoche, Nutley, New Jersey 07710, and Department of Biochemistry, University of Minnesota, Minneapolis, Minnesota 55455

Received August 31, 1993[⊛]

ABSTRACT: The crystal structure of mitochondrial malate dehydrogenase from porcine heart contains four identical subunits in the asymmetric unit of a monoclinic cell. Although the molecule functions as a dimer in solution, it exists as a tetramer with 222 point symmetry in the crystal. The crystallographic refinement was facilitated in the early stages by using weak symmetry restraints and molecular dynamics. The *R*-factor including X-ray data to 1.83-Å resolution was 21.1%. The final root mean square deviation from canonical values is 0.015 Å for bond lengths and 3.2° for bond angles. The resulting model of the tetramer includes independent coordinates for each of the four subunits allowing an internal check on the accuracy of the model. The crystalline mitochondrial malate dehydrogenase tetramer has been analyzed to determine the surface areas lost at different subunit–subunit interfaces. The results show that the interface with the largest surface area is the same one found in cytosolic malate dehydrogenase. Each of the subunits contains a bound citrate molecule in the active site permitting the elaboration of a model for substrate binding which agrees with that found for the crystalline enzyme from *Escherichia coli*. The environment of the N-terminal region of the crystallographic model has been studied because the functional protein is produced from a precursor. This precursor form has an additional 24 residues which are involved in mitochondrial targeting and, possibly, translocation. The crystallographic model of mitochondrial malate dehydrogenase has been compared with its cytosolic counterpart from porcine heart and two prokaryotic enzymes. Small but significant differences have been found in the polar versus nonpolar accessible surface areas between the mitochondrial and cytosolic enzymes. Using least squares methods, four different malate dehydrogenases have been superimposed and their consensus structure has been determined. An amino acid sequence alignment based on the crystallographic structures describes all the conserved positions. The consensus active site of these dicarboxylic acid dehydrogenases is derived from the least squares comparison.

The centralmost aerobic energy producing pathway in nearly all organisms is the citric acid cycle. In eukaryotic cells, the enzymes participating in this pathway are located within the inner membrane matrix of mitochondria. The citric acid cycle involves nine reactions beginning with the formation of citrate (six carbons) from oxaloacetate (four carbons) and acetyl coenzyme A (two carbons). The six-carbon tricarboxylic acid, citrate, is isomerized and then changed first to a five-carbon intermediate and then to the four-carbon forms including fumarate, L-malate, and oxaloacetate. Malate dehydrogenase catalyzes the interconversion of L-malate and oxaloacetate or the last step before additional citrate is formed by the participation of another acetyl coenzyme A. In general terms, the similarity of the citric acid cycle intermediates is striking; they are all either di- or tricarboxylic acids. As will be shown here, mitochondrial malate dehydrogenase appears to bind citrate as well as the substrates L-malate and oxaloacetate. The binding of citrate to the active site has also been observed in the crystal structure of the enzyme isolated from *Escherichia coli* (Hall *et al.*, 1992).

As noted above, in eukaryotic cells the citric acid cycle chemistry takes place within the mitochondria. However, some of the enzymes have forms which exist in other cellular compartments. Such is the case with the malate dehydrogenases. Mitochondrial malate dehydrogenase (abbreviated mMDH), like its cytosolic counterpart (abbreviated cMDH), uses NAD/NADH as a cofactor, both are homodimers in solution, and they have similar Michaelis constants. However, there is less than 20% amino acid sequence identity between mMDH and cMDH.

Why did two forms of the MDHs evolve? One reason is that the properties of the cytosol and mitochondrion are different. The inner membrane matrix of a mitochondrion from which mMDH derives is also the site of fatty acid oxidation as well as a number of other metabolic pathways. A consequence of the multitude of enzymatic reactions found in the inner membrane matrix is the very high protein concentration, suggested to approach that found in a protein crystal (Srere, 1981).

The consequence of this high protein concentration is that metabolic intermediates are present in a milieu in which diffusion may be significantly reduced. It has been suggested that, to facilitate movement of metabolites through their metabolic pathways, enzymes adjacent to each other in a pathway may have evolved with a propensity for forming

[†] Supported by a grant from the National Science Foundation: DMB 8941746.

[‡] The coordinates of all four subunits have been desposited in the Protein Data Bank under the code 1MLD.

[*] Author to whom correspondence should be addressed.

[§] Biomedical Engineering Center, Department of Laboratory Medicine and Pathology, University of Minnesota.

[∥] Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences.

[⊥] Roche Research Center.

[#] Department of Biochemistry, University of Minnesota.

[⊛] Abstract published in *Advance ACS Abstracts,* February 1, 1994.

0006-2960/94/0433-2078$04.50/0

enzyme–enzyme complexes within the mitochondrial matrix (Srere, 1990). In the case of mMDH, the adjacent citric acid cycle enzymes would be fumarase and citrate synthase, and aspartate amino transferase which is needed for the malate shuttle. Structural studies of mMDH, complemented with those of the adjacent enzymes, may well provide information useful in answering questions about the possibility of forming these proposed enzyme–enzyme complexes.

The enzyme mMDH and the other citric acid cycle enzymes are coded for by nuclear genes, and precursor forms of these proteins are synthesized in the cytoplasm. The precursor must first be targeted to the mitochondrion and then translocated through the membranes into the matrix. This occurs by an incompletely understood mechanism. However, recognition of proteins for transportation across the membrane appears to involve a segment of approximately 25 amino acids on the N-terminal of the newly synthesized matrix protein. In the case of mMDH, targeting requires the presence of an N-terminal translocation peptide (Grant *et al.*, 1987). Although the exact mechanism of translocation remains unknown, surface side chains on these proteins must play some role. In addition, the precursor form is proteolytically cleaved during the uptake process with loss of the translocation signal sequence. Peptide bonds in the precursor form must, therefore, be sterically available to the appropriate protease for processing. The refined crystal structure of mMDH presented here describes the location and steric accessibility of the newly formed N-terminal in the mature folded enzyme.

As noted above, amino acid sequence comparisons indicate that mMDH is not very similar to cMDH. The mMDH sequence is, in fact, more closely related to the prokaryotic *E. coli* enzyme (abbreviated eMDH). This finding is in accord with the generally accepted idea that the evolution of the mitochondrion in eukaryotic cells resulted from a symbiotic coexistence of two prokaryotes. However, any generalization that prokaryotic MDH structures are similar to the mitochondrial form of eukaryotes will require more structural examples, because malate dehydrogenase from the prokaryote *Thermus flavus* (abbreviated tMDH) is more similar to cMDH than it is to eMDH (Nishiyama *et al.*, 1993; Kelly *et al.*, 1993). With the completion of the crystal structure determination of mMDH, comparisons can now be made at the structural level. Despite the relatively large difference in amino acid sequence between cMDH and mMDH, the structural similarity is still very high.

The data and discussion which follow describe the crystal structure of mMDH refined at high resolution. Monoclinic crystals, obtained from citrate buffer in the presence of polyethylene glycol, contained four subunits in the asymmetric unit (Roderick Banaszak, 1986). The preliminary chain trace showed that about one-half of the mMDH molecule consisted of a nucleotide binding domain similar to other NAD binding proteins such as cMDH and lactate dehydrogenase (abbreviated LDH) (Roderick Banaszak, 1986). Some differences between mMDH and cMDH were known to be present from the C_α model. Most of them occurred in surface loops and were subsequently documented in the structural studies of eMDH (Hall *et al.*, 1992). The similarity between mMDH and eMDH was recognized before the refinement described in this report was complete. In fact, a partially refined model of mMDH from the studies described here was used to solve the crystal structure of eMDH by molecular replacement techniques (Hall *et al.*, 1992).

Previously we reported the quaternary structure of crystalline mMDH and a C_α model (Roderick Banaszak, 1986). This earlier work noted the fact that the four subunits of crystalline mMDH found in the asymmetric unit were arranged in a tetramer with point group symmetry 222. As previously noted, mMDH, eMDH, tMDH, and cMDH are all dimers in solution. The presence of four mMDH subunits in the asymmetric unit was surprising, and the size of the problem (approximately 1200 independent amino acids) made refinement difficult. The crystalline mMDH tetramer includes a tight interface conventionally described in the dehydrogenase literature as the *Q*-axis interface (Rossmann *et al.*, 1973). This interface is also found in cMDH, eMDH, tMDH, and lactate dehydrogenase.

As will be shown below, there is a relatively large surface involved in subunit–subunit contacts in mMDH. Overall they are very similar to those in cMDH, and this same interface is present in eMDH, tMDH, and the tetrameric LDHs. In spite of the similarities, the dimer interface in mMDH has special significance. At low pH, the dimer will dissociate to a monomeric form (Wood *et al.*, 1981) of significantly reduced enzymatic activity. Within the mMDH tetramer, the other subunit–subunit interfaces are unlike those found in tetrameric LDH (Grau et al., 1981). Furthermore, the extent of other subunit–subunit contacts in the crystalline mMDH tetramer is relatively small.

MATERIALS AND METHODS

X-ray Diffraction Data. Monoclinic crystals of porcine heart mMDH have cell dimensions $a = 72.75$ Å, $b = 146.76$ Å, $c = 67.58$ Å, and $\beta = 108.16°$ and belong to the space group $P2_1$ with four subunits in the asymmetric unit. The X-ray data were derived from two sources. A native data set to about 2.5-Å resolution was collected on a CAD4 diffractometer using a nickel filter and partial ω step scans. This data included 45 655 reflections and was part of the diffraction data used to produce the initial chain trace to 3-Å resolution described earlier (Roderick Banaszak, 1986). The heavy atom data was also originally collected using the CAD4 diffractometer. Despite the large unit cell, the crystals diffracted well and additional data to 1.83-Å resolution were collected and used as described below.

The 1.83-Å-resolution X-ray data were obtained on a Nicolet area detector system at Argonne National Laboratory during two different sessions using a four-circle goniostat. Most of the mMDH crystals were slabs elongated in the b^* direction which tended to align along the capillary axis. This meant that b^* was along the ω axis at $\chi = 0°$, and ω scans at this orientation were used to collect most of the data. Additional ω scans were made at a series of varying χ values. Finally, additional data were collected from two crystals mounted with β^* perpendicular to the capillary axis (and the ϕ axis). The area detector data, about 91 000 independent reflections, were processed using XENGEN (Howard *et al.*, 1985). R_{sym} for different runs ranged from 0.05 to 0.10. The two area detector data sets and the diffractometer data were scaled and merged using the local scaling option in the ROCKS program package (Reeke, 1984), and the R_{sym} based on I_{hkl} was 0.145 for all data.

The Initial Model. The previously published model contained only C_α positions which were subsequently expanded into polyalanine coordinates. Using prominent features of the 2.8-Å map and the amino acid sequence as reported by Birktoft *et al.* (1982), a preliminary sequence assignment was then made. Using the 2.5-Å data, a $2|F_o| - |F_c|$ map was calculated and the initial sequence assignment verified and modified as indicated by the map. This model then formed

the starting point of TNT refinement (Tronrud *et al.*, 1986). After 49 cycles, new electron density maps ($2|F_o| - |F_c|$ and $|F_o| - |F_c|$) were calculated, each of the four mMDH subunits in the asymmetric unit was checked, and model adjustments were made.

The rebuilt model was then used as the starting point for 71 additional cycles of TNT refinement. Both $|F_o| - |F_c|$ and $2|F_o| - |F_c|$ maps were calculated, and the model was again checked. The starting *R*-factor was 43.9%, and the root mean square value for bond lengths was 0.042 Å and for bond angles was 6.1°. The *R*-factor near the conclusion of the initial TNT refinement was 36.5%, and the root mean square value for bond lengths was 0.024 Å and for bond angles was 4.4°.

While no restraints were imposed on the inter-subunit relationships at these early stages of refinement, the four subunits were frequently compared using least squares methods in order to identify regions where one subunit deviated significantly from one or more of the other subunits. Such regions were given particular attention during the model rebuilding in order to verify that the structures were indeed different in such regions.

Refinement Strategy Including the Use of Noncrystallographic Symmetry. Because of the large size of the asymmetric unit, initial attempts to refine the structure as four independent subunits resulted in extremely slow progress. Since the protein is tetrameric with approximate 222 point symmetry, it appeared to be an excellent candidate for the application of noncrystallographic symmetry (NCS). Procedures for the application of NCS are conveniently implemented in the X-PLOR crystallographic package (Brunger, 1990) which was used for further refinement.

An initial simulated annealing run (2000 K) was made with four independent and unrestrained subunits using data from 6.0- to 2.5-Å resolution. This resulted in a structure with an *R*-factor of 0.309 and root mean square deviation in bond lengths of 0.034 Å and in bond angles of 6.1°. Ramachandran plots as well as the electron density map indicated that the model contained numerous errors, and the prospect of correcting each subunit individually was discouraging. Application of NCS would allow initial work on a single subunit and make the application of major corrections more efficient.

A strategy evolved which did not involve electron density averaging. Rather, the four independent subunits were rotated to a common reference frame and the coordinates averaged. The subunit in best agreement with the average was taken as the "best" monomer for the subsequent application of NCS. Three rounds of positional refinement, manual readjustment, and simulated annealing were then performed, resulting in a *R*-factor of 0.339.

At this point the restraints on the subunit relationships were gradually relaxed, and finally the subunits were treated independently. Some significant errors were located by comparing the model and electron density map with the previously reported crystal structure of the cytoplasmic enzyme (Birktoft *et al.*, 1989a). The resolution was gradually increased to include all of the available data which extended to approximately 1.9 Å. At the end of this phase of the refinement the *R*-factor was 0.247. The model was good enough at this point that an averaged subunit was used to solve the structure of eMDH by molecular replacement (Hall *et al.*, 1992).

Completion of the refinement involved the addition of 521 water molecules and the correction of some errors in loop regions which was expedited by the availability of the refined eMDH structure. The final *R*-factor for 83 454 reflections in the resolution range 6.00–1.83 Å was 21.1%. All available data was used without removing any weak reflections. The *R*-factor was 17.7% if the 45 647 reflections to 2.55-Å resolution were used. The final root mean square deviation from canonical values is 0.015 Å for bond lengths and 3.2° for bond angles. After the refinement had been completed, the program system PROCHECK (Laskowski *et al.*, 1993) was used to evaluate the four sets of coordinates. A particularly useful feature of this suite of programs is the Ramachandran plot evaluation which applies empirical rules based on 118 structures of resolution of at least 2.0 Å and *R*-factors no greater than 20%. According to this analysis, a good quality model should have at least 90% of non-glycine residues in the most favored regions (Morris *et al.*, 1992). For our model the percent in the most favored region is 89.3% (A subunit),[1] 91.2% (B subunit), 91.6% (C subunit), and 90.5% (D subunit). Valine 192 in the C subunit is the only residue in a disallowed region. However, valine 192 in all the other subunits is in a region classified as generously allowed. Because the electron density in the region of residues 190–193 is discontinuous in all of the four subunits, these results are not surprising. We were also unable to locate the terminal amino acid, K314, in any of the subunits. The coordinates of all four subunits have been deposited in the Protein Data Bank under the code 1MLD.

Comparing Models of the MDHs. For comparison purposes, the B subunit of mMDH was used because this subunit had the smallest root mean square deviation from an averaged subunit derived from the four independent mMDH subunits. There is only one subunit in the asymmetric unit of *E. coli* MDH, and the coordinates of Hall *et al.* (1992) were used. Rather large differences were found between the two independent subunits in cMDH (Birktoft *et al.*, 1989a). Kelly *et al.* (1993) noted, however, that the B subunit of cMDH was more similar to the enzyme from *T. flavus* than was the A subunit and argued that the B subunit was probably the better structural representation. For this reason we chose to use the B subunit of cMDH for comparison. Because the two crystallographically independent subunits in tMDH are very similar, we arbitrarily chose the A subunit for comparison.

Model coordinates from the different crystal structures of the MDHs were compared using least squares methods in a pairwise fashion. The rotated coordinates were then compared visually and by using plots of the root mean square difference versus the residue number. Positions where the root mean square value exceeded 2 Å were omitted from the next cycle of fitting. In the end a reduced coordinate set was used to optimize the superposition of the two models. A comparison of the structures in terms of similar and dissimilar positions in the complete crystallographic model will be given in the Discussion section.

RESULTS

The Four Subunits in Crystalline mMDH. Taken together, the structural studies of the MDHs have resulted in the molecular symmetry of the normally dimeric MDH molecules being presented in four different crystallographic forms. Since the monomeric units of the crystalline eMDH dimer are related by crystallographic symmetry, they are of course identical (Hall *et al.*, 1992). Similar results might have been expected for the eukaryotic forms: mMDH, cMDH, and tMDH. For cMDH, the two crystallographically independent subunits are

[1] The four independent subunits in the crystallographic asymmetric unit are labeled A–D.

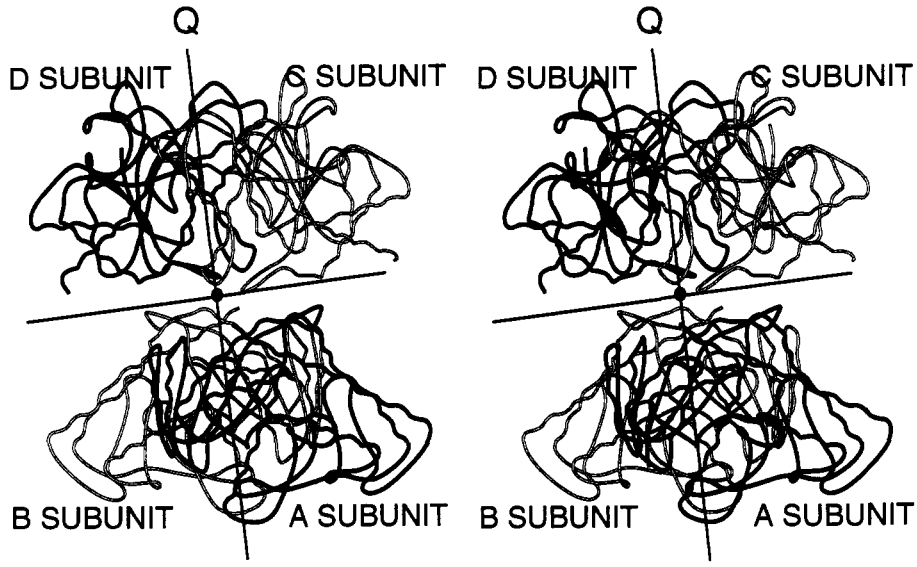

FIGURE 1: The crystalline tetramer of pig heart mMDH. The MOLSCRIPT (Kraulis, 1991) stereodrawing shows the four subunits of mMDH found in the crystalline state in the form of a main-chain coil. Although the enzyme exists as a dimer in solution, it very closely approximates a 222 point symmetrical tetramer in the crystalline state. The solid lines depict the approximate positions of the three mutually perpendicular 2-fold rotation axes. The four independent subunits are labeled A–D. For ease of viewing, the B and C subunits are depicted by hollow lines. Strong subunit–subunit contacts between A and B and between C and D are thought to be characteristic of the solution dimer and are marked by the letter "Q". The two opposed loops found in the A–C and B–D interface are those that are believed to undergo a conformational change upon binding substrate or coenzyme as described in the text.

related by an approximate 2-fold rotation axis (172°), and the root mean square difference between the two subunits is 1.62 Å for backbone atoms and 2.17 Å for all atoms (Birktoft *et al.*, 1989a). The dimeric tMDH crystallized with a dimer in the asymmetric unit, and the two monomers are related by noncrystallographic 2-fold rotational symmetry (Kelly *et al.*, 1993).

The enzyme mMDH crystallizes as a tetramer with no crystallographic symmetry relating the monomeric units which are, nonetheless, remarkably similar. The root mean square deviation of the backbone atoms for the four subunits of mMDH from an averaged subunit is as follows: A, 0.171 Å; B, 0.161 Å; C, 0.167 Å; and D, 0.190 Å. If all non-hydrogen atoms are included in the comparison, then the root mean square deviation from the average subunit is as follows: A, 0.392 Å; B, 0.413 Å; C, 0.438 Å; and D, 0.415 Å. The conformation of the citric acid molecule is also similar in all subunits. The root mean square deviation from the average value is as follows: A, 0.124 Å; B, 0.088 Å; C, 0.086 Å; and D, 0.084 Å. The rotation angles for the symmetry operators relating the subunits are as follows: (1) subunit A to B, 179.86°; (2) subunit A to C, 175.71°; and (3) subunit A to D, 179.97°. These values coupled with the close agreement in coordinates make it clear that crystalline mMDH is essentially a 222 point symmetric molecule.

In the initial crystallographic studies at 3-Å resolution, two of the subunits shared extensive protein–protein contacts. These strong dimeric interactions within the tetramer identified the solution dimer. The higher resolution X-ray structure described here confirms the original selection. A coil model of the crystalline mMDH tetramer is shown in Figure 1. If viewed with stereo, the 222 point symmetry with 2-fold rotational symmetry around axes which are nearly vertical, horizontal, and perpendicular to the plane of the drawing is apparent. The most important contacts are difficult to observe in this orientation. Note that the vertical 2-fold axis would relate subunit A to B as well as C to D. The dimer pairs A–B and C–D clearly have extensive surface contacts and are proposed to represent the solution dimer.

Table 1: Surface Area Comparisons for Oligomers of mMDH

oligomer	surface area lost (Å^2)	% lost[a]
mMDH tetramer	8373	17.0
A–B dimer[b]	3197	13.0
A–C dimer	997	4.0
A–D dimer	17	0.0

[a] Based on the surface area of a constituent monomer of 12 305 Å without bound water or the substrate analogue, citrate. [b] The different dimeric forms can be visualized in Figure 1.

Supporting evidence for these assignments comes from crystallographic studies of homologous proteins. In cMDH, the same subunit interface is present in the crystallographic dimer (Birktoft *et al.*, 1989a) and is referred to as the *Q*-axis interface in the lactate dehydrogenase tetramer (Rossmann *et al.*, 1973). Very small areas of contact are present between the A and C subunits and the equivalent B and D subunits. As will be shown below, there are essentially no contacts between the C and B and the A and D subunits.

One way to assess the relative importance of the mMDH subunit–subunit contacts is to tabulate the loss of surface area when monomer units are incorporated into subsets of the mMDH tetramer. To do this we used the surface area calculation available in X-PLOR (Brunger, 1990) which implements the algorithm of Lee and Richards (1971). The probe radius was 1.6 Å, and the results are summarized in Table 1. The formation of the crystalline tetramer results in a loss of nearly 20% of the total surface area of the component monomers. However, the data in Table 1 also indicates that nearly all of this loss results from the formation of the A–B (and C–D) subunit interface. Only about 4% of the loss occurs due to interactions in the B–D (and A–C) interface.

The surface area calculations also provide a simple way of identifying residues participating in the subunit–subunit interfaces. To do this, the residue by residue accessibilities are compared for monomer and the different dimer interfaces.

Subunit–Subunit Interactions and pH Sensitive Dissociation. The dimer interface believed to exist in solution would be formed between A and B (and C and D) in Figure 1. The

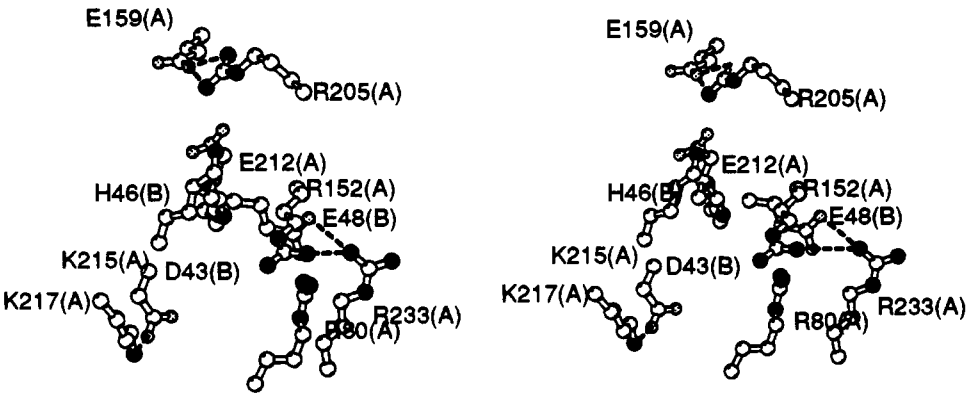

FIGURE 2: Site of pH-related subunit–subunit dissociation. The stereodrawing shows the amino acid side chains from two subunits which are believed to be related to the pH sensitive dissociation of the mMDH dimer into monomers. Taken from the crystal structure, the following side chains are shown: R80(A), K215(A), K217(A), R205(A), E159(A), R233(A), R152(A) and D43(B), H46(B), E48(B). The symbols "(A)" and "(B)" define the appropriate subunit as in Figure 1. Because of the symmetry of the mMDH dimer, these interface residues are present twice in each molecule, but only one set is shown. Nitrogen atoms are depicted by filled spheres, carbon atoms by open circles and oxygen atoms by shaded spheres. The dashed lines connect heteroatoms less than 3.2 Å apart indicating possible hydrogen bonds.

interaction between these two polypeptide chains is much more extensive than the A–C (and B–D) interface. Several regions of the polypeptide chain are involved including L19, N22, H36, A42, S45 through R50, V151 through F156, Q164 through A166, and side chains in the structure from V213 to Y229. The surface complementarity between the subunits at the A–B interface is extensive, and the convoluted contours of the individual units fit together with essentially no empty space. Since mMDH undergoes a pH sensitive dissociation with a pK of about 5.5 (Bleile *et al.*, 1977; Wood *et al.*, 1981), we have examined the interface carefully in order to identify potential triggering sites for the dissociation.

Although much of the interface consists of hydrophobic residues, a small volume of the interface includes mostly ionic side chains as shown in Figure 2. The interactions include a total of four arginines, two lysines, three glutamics, one aspartic acid, and one histidine. Using Figure 2, one can see that three of the side chains come from the B subunit (D43, E48, H46) and the rest from the A subunit (K215, K217, R80, R152, R205, E159, E212). The bottom-most segment of the ionic region contains the ion pair of R233(A) and E48(B). Above that is a cluster of three basic side chains (R152(A), R205-(A), H46(B)) and two acidic residues (E159(A), D43(B)). However, D43 is closest to K217(A) in the lower part of Figure 2. The pair K215(A) and E121(A) also appear to form a unique ion pair.

Eliminating arginines and lysines, the most plausible candidates serving as triggers in the pH sensitive dissociation would be H46, D43, E159, E48, E212, or any combination thereof. Based only on proximity, protonation of D43 and H46 would place two positive groups relatively close to each other (H46 and R152) without the benefit of a counterion, and dissociation could be favored. Note that R152 and R80 are also involved in substrate binding which could link the quaternary structure to the enzymatic activity.

Recent work by Steffan and McAlister-Henn (1991) on mutants of mMDH from yeast has shown that the mutation H46L dramatically affects the pH optimum of the enzyme from 7.5 in the native enzyme to 5.5 in the mutant. The pH shift also led to increased stability of the dimeric form of the mutant enzyme, and it was concluded that H46 is the residue responsible for the pH-induced dissociation of the dimer to a monomeric species with reduced activity. Another mutant which was studied was D43N. This change produced an mMDH with little or no activity. However, the resulting yeast strains grew on acetate, suggesting metabolism through the citric acid cycle (Steffan & McAlister-Henn, 1991). Apparently the deleterious effects of the mutation were compensated for by a high concentration of enzyme within the mitochondria (Steffan & McAlister-Henn, 1991). The site mutations mentioned above for yeast mMDH by homology correspond to amino acid side chains observed in the pig heart mMDH crystal structure.

In crystalline mMDH, the two dimers make weaker contacts through the loop region leading to the observed tetramer. This loop includes the turn between residues 80 through 84. For example, the only accessibility changes for the A–C interface occur in the residues between V78 and T92. The two closest side chains in this interface are M84(A)–M84(C) and D88-(A)–D88(C).[2] No obvious ionic interactions are present, but several residues are involved in intermolecular hydrogen bonds including side chains of D88, N91, and T92. In addition, the presence of a few water molecules in the A–C interface serve to link main-chain carbonyl oxygens. Thus the weak interface of the crystalline tetramer involves only four hydrogen bonds (two pairs) and the burying of one hydrophobic side chain from each subunit: M84.

What is especially interesting about the A–C and B–D interfaces of the crystalline tetramer is the fact that it is formed by the "loop region" which in LDH is believed (Rossmann *et al.*, 1975) to undergo a conformational change during a catalytic cycle. Since the loop is in the down position in crystalline mMDH, the binding of citrate may have been a factor in the formation of the crystal tetramer. As discussed in more detail subsequently, two arginines near this segment, R80 and R86, are also involved in substrate binding.

N-Terminal Region and Mitochondrial Import. The derived crystal model is that of the mature protein. The precursor form of this protein is 24 residues longer, and the N-terminal residue was formed by proteolysis during mitochondrial uptake. The actual amino acid sequence of the mMDH precursor peptide from pig heart is unknown, but overall, the mature protein has 96% sequence identity with rat mMDH (Grant *et al.*, 1987). Furthermore, the amino acid sequences of three known precursor peptides for mitochondrial MDHs are very similar, as is shown in Table 2. Note the high sequence identity between rat and mouse precursor (96%). All this suggests that the precursor peptide for the pig enzyme is also very similar.

[2] The subunit identification code is placed within parentheses to prevent mistaken identification with the one-letter amino acid code.

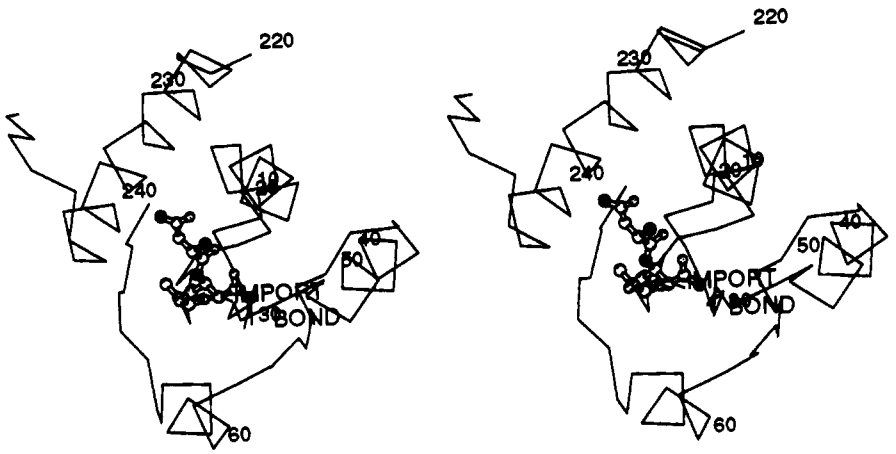

FIGURE 3: Proteolytically cleaved bond of the precursor form of mMDH. The stereodrawing contains a portion of a single subunit surrounding the position of the N-terminal which is formed upon cleavage of the precursor segment. The location of nearby residues is depicted by a C_α model which includes residues 1–75 and 220–250. All atoms are shown for the amino acid sequence -N-N-A- where -A- would correspond to residue 1 in the crystal structure. Atoms are shaded as described in Figure 2. The two penultimate residues of the precursor form are -N-N-, and the hypothetical positions of the -N-N- sequence are forced by steric principles to extend outward from the main body of the enzyme. The bond cleaved during translocation is labeled with the words "IMPORT BOND"; the cleavage occurs between -N-A-, giving an amino terminal alanine.

Table 2: The Amino Acid Sequences of Three Precursor Peptides of mMDH[a]

	−15	−5	1
yeast	MLSRVAKRAF	---------SST	VANP *YK
mouse	MLSALARPAG	AALRRSFSTS	AQNN *AK
rat	MLSALARPVG	AALRRSFSTS	AQNN *AK

[a] Yeast (Thompson *et al.*, 1988), mouse (Takeshima *et al.*, 1988), rat (Grant *et al.*, 1987). The asterisk (*) marks the beginning of the mature enzyme (after translocation and proteolytic cleavage).

Table 3: Surface Properties of mMDH[a]

A. Ionizable Amino Acids in the MDHs from Pig Heart

protein	His	Lys	Arg	Asp	Glu	total	%
mMDH	5	24	8	12	16	65/313	21
cMDH	4	31	20	25	18	88/333	26

B. Accessible Surface Area (Å^2) and Percent of Total Accessible Surface Area

protein	basic residues	acidic residues	% of total	uncharged polar residues	% of total	nonpolar residues	% of total
mMDH	1982	1187	14.1	7184	32	11 632	53
cMDH	3699	2363	22.9	8388	31	12 214	46

[a] The following definitions were used for tabulating the accessible surface areas. (b) Basic: NE, CZ, NH1, NH2 atoms of arginine; NZ atom of lysine; ND1, CE1, NE2 atoms of histidine (2) Acidic: CD, OE1, OE2 atoms of glutamic acid; CG, OD1, OD2 atoms of aspartic acid (3) polar: N, O, C from the main chain; OG1 of threonine; OG of serine; CG, OD1, ND2 of asparagine; CD, OE1, NE2 of glutamine; OH of tyrosine; SG of cysteine. (4) Nonpolar: All other carbon atoms. (5) Not included in any of the above categories was the NE1 nitrogen atom of tryptophan. The enzyme MMDH has no tryptophan, and the accessibility of NE1 in the tryptophans of cMDH was negligible.

For the mammalian precursor sequences, two asparagines precede the N-terminal of the mature enzyme. We have added these two amino acids to the crystal structure model of pig heart mMDH and placed the new atoms in a position in which they are not in van der Waals contact with other enzyme atoms. The results are shown in stereo in Figure 3. Note that the bond that is proteolytically cleaved is on or near the surface of the subunit. Furthermore, it is not near the subunit–subunit interface which exists in the mature mMDH nor any of the additional subunit–subunit interfaces observed in the crystallographic tetramer.

Based on the structural data, one would like to be able to infer something about the conformation of the precursor protein when the proteolytic processing occurs. Must it be unfolded? Or can it be folded resembling the mature enzyme? If the protease recognizes side chains on the N-terminal side of the scissile peptide bond, the folded structure shown in Figure 3 would appear to allow for such recognition. Both asparagines appear close enough to the surface and have no other obvious location except for this extended conformation. The N-terminal region is well away from any subunit interface, and therefore the oligomeric state is also unlikely to interfere with proteolytic processing. Clearly, proteolysis is always possible in an unfolded state. The crystal structure of mMDH suggests that proteolysis is also possible in a folded state.

One of the more interesting comparative properties of the eukaryotic MDHs is the fact that mMDH is a basic protein while cMDH is acidic. Ordinarily one might expect this to mean that mMDH has more basic amino acids. However, as can be seen in the top part of Table 3, this is not the case. Although more basic overall, mMDH has fewer basic side chains, 37 compared to 45 in cMDH. However, the ratio of basic to acidic amino acids is greater for mMDH despite the fact that the total number is smaller, and this leads to the higher isoelectric point.

The number of electrostatically charged side chains as well as their distribution on the surface of mMDH could be a factor in mitochondrial import. Any charged side chain might contribute to the energy barrier involved in moving the precursor of mMDH through the inner membrane of a mitochondrion. Therefore the X-ray crystallographic coordinates for both proteins have been used to calculate the accessible surface areas for all atoms in both mMDH and cMDH (Lee & Richards, 1971). For the purpose of making comparisons of individual atom types, the program AREA in CCP4 (SERC Daresbury Laboratory, 1986) was used. Values from each atom were then put into one of three categories: ionizable (basic, acidic), polar, and nonpolar according to the footnote definitions in Table 3. The results are shown in part B of Table 3.

Although both molecules have a significant surface charge density, the accessible charged surface of mMDH is reduced by nearly 50% when compared to cMDH. As can be seen in part B of Table 3, in both molecules about one-third of the

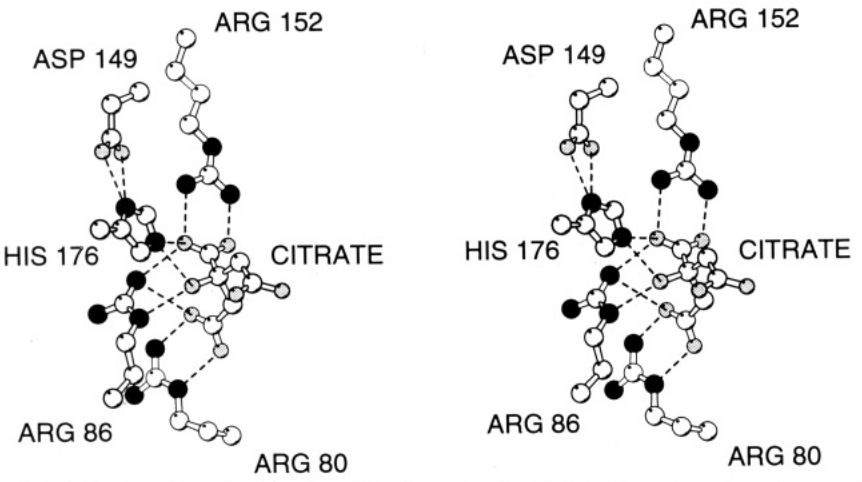

FIGURE 4: The active site of mMDH with a bound citrate ion. The stereodrawing depicts the protein atoms interacting with the bound citrate in crystalline mMDH. Five side chains are shown: R80, R86, R152, H176, and D149. The nitrogen atoms are shown as filled circles, and hydrogen bonds are given as dashed lines. The removal of one carbon and the carboxylate without any hydrogen bonds would give a malate ion, which is the normal substrate. Note the network of hydrogen bonds which connect the ion to three arginines and the histidine–aspartic acid pair.

surface is convered by polar (but uncharged) atoms. Therefore, the lower ionic area of mMDH is compensated for by an increase in the nonpolar accessible area. Again in Table 3, mMDH has 53% of the accessible surface in the form of nonpolar atoms while cMDH has 46%. A more exhaustive comparison of mitochondrial versus cytosolic enzymes will need to be done before it is possible to relate this to mitochondrial import. Nonetheless, the structural data suggests that two homologous MDHs have molecular surfaces which appear to be significantly different.

Bound Citrate and the Active Site. Portions of the mMDH electron density map contained areas which could not be accounted for by protein. These were at locations corresponding to the active site in homologous MDHs. In fact, similar electron density was observed in the X-ray crystallographic studies of eMDH which had also been crystallized in the presence of citrate (Hall *et al.*, 1992). The shape and extent of this density strongly suggested that mMDH also bound a single citrate molecule in each of the subunits in a manner similar to the prokaryotic enzyme, eMDH. The binding of this citrate ion in the active site is illustrated in stereo in Figure 4. In addition, intermolecular distances for C, N, and O atoms involved in hydrogen bonds between citrate and mMDH are given in Table 4. A word of caution: a similar table in the crystallographic report on eMDH labeled the citrate carbon atoms in a different way, but the binding interactions are essentially the same (Hall *et al.*, 1992).

In the eMDH study, it was assumed that the citrate ion was mimicking the binding of substrate. The same binding mode is maintained in mMDH, which further supports the analogy to the binding interactions for the substrate, L-malate. As was the case with eMDH, one of the three carboxylate moieties does not interact with any protein atoms but instead points toward the position of the NAD binding site.

As can be seen in Figure 4, two carboxyl groups belonging to citrate are analogous to those found in malate and are bound to R152 and R80. R152 interacts with the carboxyl group using the terminal nitrogens of the guanidino group, while R80 uses one terminal and one nonterminal guanidino nitrogen (NE atom in PDB nomenclature) in an "edge-on" interaction. Arginine 86 also appears important in properly orienting the substrate. It interacts with two of the carboxyl groups of the bound citrate. Furthermore, the NE atom of R86 is hydrogen bonded to the hydroxyl oxygen at C2. Both R152 and R80 are close to what is believed to be the pH sensitive region of the subunit–subunit interface and were shown in Figure 2.

Table 4: Hydrogen Bonds Involved in Citrate Binding to mMDH

		distance (Å)	
		mMDH	eMDH[a]
83 ARG NE[b]	CIT O2	2.96(0.09)[c]	2.88
176 HIS NE2	CIT O2	2.74(0.04)	2.88
86 ARG NH1	CIT O11	2.76(0.04)	3.19
152 ARG NH2	CIT O11	2.94(0.08)	2.81
176 HIS NE2	CIT O11	2.94(0.10)	2.99
152 ARG NH1	CIT O12	2.71(0.01)	2.85
80 ARG NH1	CIT O41	2.76(0.03)	2.78
86 ARG NH1	CIT O41	2.98(0.15)	3.19
80 ARG NE	CIT O42	2.87(0.02)	2.78
149 ASP OD1	176 HIS ND1	2.82(0.05)	2.75
149 ASP OD2	176 HIS ND1	2.88(0.05)	2.98
CIT O12	W409	2.91(0.05)	
CIT O42	W435	2.80(0.11)	
CIT O61	W407	2.75(0.07)	
CIT O61	W418	2.96(0.11)	
CIT O62	W418	2.82(0.07)	
W418	W419	2.76(0.19)	
W407	W408	2.70(0.09)	
W435	80 ARG N	2.87(0.06)	
W409	223 ALA N	2.88(0.04)	
W409	210 GLY O	2.90(0.11)	

[a] Hall *et al.*, 1992. [b] Atoms designated as residue number, residue type, and atom designator, except CIT, which designates the citrate molecule, and W*nnn*, which designates a water molecule. [c] Mean and standard deviation of equivalent distances in the four subunits.

Part of the network stabilizing the two carboxylates of the bound substrate includes residues thought to participate in catalysis. As can be seen in Figure 4, H176 is oriented for interaction with the hydroxyl at C-2. This histidine is thought to be responsible for the removal of the hydrogen of the OH in the form of a proton during a catalytic cycle. D149 is hydrogen bonded to H176 to form the histidine–aspartate proton relay system found in cMDH (Birktoft & Banaszak, 1982), later in LDH (Clarke *et al.*, 1988), and recently in eMDH (Hall *et al.*, 1992).

DISCUSSION

Four independent subunits of mMDH have been refined to convergence with high-resolution X-ray diffraction data.

```
          ---βA---        ------αB-------          -----βB-----      --
m-pig    ..AKVAVLG[7]   ASGGIGQPLS[17]  LLLKNSp...[24]  l.v.sRLTLY[32]  DIa....hTP[38]
e-MDH    ..MKVAVLG[7]   AAGGIGQALA[17]  LLLKTQlp..[25]  s.g.sELSLY[33]  DIap...vTP[40]
c-pig   sepIRVLVTG[10]  AAGQIAYSLL[20]  YSIGNGsvfg[30]  kdqpiILVLL[40]  DItpmmgvLD[50]
t-MDH  mkapVRVAVTG[11]  AAGQIGYSLL[21]  FRIAAGemlg[31]  kdqpvILQLL[41]  EIpqamkaLE[51]

         ---αC----        ---βC--      --αC'--        --βD--
m-pig   GVAADLShie[48]  tratvkgylg[58]  peqlpdcLKG[68]  CDVVVIPAgv[78]  prkpgmtrdd[88]
e-MDH   GVAVDLShip[50]  tavkikgfsg[60]  ed.atpaLEG[69]  ADVVLISAgv[79]  arkpgmdrsd[89]
c-pig   GVLMELQdca[60]  lpllkdviat[70]  dk.eeiaFKD[79]  LDVAILVGsm[89]  prrdgmerkd[99]
t-MDH   GVVMELEdca[61]  fpllagleat[71]  dd.pdvaFKD[80]  ADYALLVGaa[90]  prkagmerrd[100]

         -----αD----------αE-----       --βE--       ---------α1F--------
m-pig   lfntnatIVA[98]  TLTAACAQHC[108] pd.AMICIIS[117] NPVNSTIPIT[127] AEVFkkhgvy[137]
e-MDH   lfnvnagIVK[99]  NLVQQVAKTC[109] pk.ACIGIIT[118] NPVNTTVAIA[128] AEVLkkagvy[138]
c-pig   llkanvkIFK[109] CQGAALDKYA[119] kksVKVIVVG[129] NPANTNCLTA[139] SKSA...psi[146]
t-MDH   llqvngkIFT[110] EQGRALAEVA[120] kkdVKVLVVG[130] NPANTNALIA[140] YKNA...pgl[147]

            -βF-- --------α2F----------                 ----------βG-βH-βJ-----
m-pig   nPNKIFGVTT[147] LDIVRANAFV[157] AELKGLNPAR[167] Vs.VPVIGGH[176] AGKtIIPLIS[186]
e-MDH   dKNKLFGVTT[148] LDIIRSNTFV[158] AELKGKQPGE[168] Ve.VPVIGGH[177] SGVtILPLLS[187]
c-pig   pKENFSCLTR[156] LDHNRAKAQI[166] ALKLGVTSDD[176] VknVIIWGNH[186] SST.QYPDVN[195]
t-MDH   nPRNFTAMTR[157] LDHNRAKAQL[167] AKKTGTGVDR[177] IrrMTVWGNH[187] SST.MFPDLF[196]

         ----βG-βH-βJ---- ---------------α1G--α2G-----------------
m-pig   QCtpk.....[191] .....vdfpq[196] .dqlstLTGR[205] IQEAgtevvk[215] akagagsaTL[225]
e-MDH   QV.pg.....[191] .....vsfte[196] .qevadLTKR[205] IQNAgtevve[214] akagggsaTL[225]
c-pig   HAkvklqake[205] vgvyeavkdd[215] swlkgeFITT[225] VQQRgaavik[235] arkl..ssAM[243]
t-MDH   HAevd....g[202] rpalelv.dm[211] ewyekvFIPT[221] VAQRgaaiiq[231] arga..ssAA[239]

         --------α3G--------   -----------------βK--βL--βM-----------------
m-pig   SMAYAGARFV[235] FSLVdamngk[245] egvvecSFVK[255] sqetdcp...[262] ..yfSTPLll[270]
e-MDH   SMGQAAARFG[235] LSLVralqge[245] qgvvecAYVE[255] gdgqyar...[262] ..ffSQPLll[270]
c-pig   SAAKAICDHV[253] RDIWfgtpe.[262] .gefvsMGII[271] sdgnsygvpd[281] dllySFPV.t[290]
t-MDH   SAANAAIEHI[249] RDWAlgtpe.[258] .gdwvsMAVP[267] sqge.ygipe[276] givySFPV.t[285]

         ----βK-βL-βM------ --------------------αH-----------------
m-pig   gkkgieKNLg[280] iGKISPFEEK[290] MIAEAIPELK[300] ASIKKgeefv[310] knmk[314]
e-MDH   gkngveERKs[280] iGTLSAFEQN[290] ALEGMLDTLK[300] KDIALgqefv[310] nk[312]
c-pig   ikdktwKIVe[300] gLPINDFSRE[310] KMDLTAKELA[320] EEKETafefl[330] ssa[333]
t-MDH   akdgayRVVe[295] gLEINEFARK[305] RMEITAQELL[315] DEMEQvkalg[325] li[327]
```

FIGURE 5: Structural alignment of the amino acid sequences of the malate dehydrogenases. The structural alignment of the four dimeric MDHs was done by least squares methods, and the results are presented in terms of their amino acid sequences. The amino acids indicated by capital letters are those taken from the crystal coordinate list and used for the superposition. At the top of each segment are headings describing the secondary structure using the generally accepted nomenclature. Every tenth residue is numbered with a superscript for reference to the PDB coordinates. In most cases the presence of insertions/deletions was easy to identify. However, in a few cases, relatively large conformational differences are present which are described in the accompanying text.

Comparisons of the coordinates of the four independent subunits give root mean square differences on the order of 0.2–0.4 Å. Such values are in agreement with error estimates obtained from Luzzatti plots and are an appraisal of the expected variation in the structure. The four subunits found in the crystalline state have the properties of a tetrameric enzyme; that is, the subunits are arranged with 222 point symmetry. The significance of this symmetrical aggregation in the crystalline state to the solution properties is presently unknown.

The refined crystal structure of mMDH presented here completes a series of studies which includes the two eukaryotic forms (mitochondrial and cytosolic) and a prokaryotic enzyme, eMDH. The crystal structure of a second prokaryotic enzyme from the bacterium *T. flavus*, tMDH (Kelly *et al.*, 1993), has also been determined.

The similarity between the crystal structures of the mitochondrial and *E. coli* enzymes is striking as is the similarity between cMDH and tMDH. We have made a careful alignment of all four structures using least squares methods. This was done in a pairwise fashion using mMDH as the standard. If the least squares fit is done in a cyclical manner each time removing atoms belonging to clearly different conformations, eventually only homologous regions are used to obtain the transformations. The optimized structural superpositions can then be used to prepare the optimal amino acid sequence alignment as is shown in Figure 5. The residues shown in capital letters were used for the structural alignment and represent the conserved structural nucleus.

Root mean square differences between each of the structures for the conserved residues only are summarized in Table 5. As can be seen in Figure 5, the MDH nucleus based on the

Table 5: Structurally Conserved MDH Residues[a]

mMDH	eMDH	cMDH	tMDH
2–22	2–22	5–25	6–26
28–34	29–35	36–42	37–43
37–45	39–47	49–57	50–58
66–76	67–77	77–87	78–88
96–108	97–109	107–119	108–120
111–131	112–132	123–143	124–144
138–168	139–169	147–177	148–178
170–179	171–180	180–189	181–190
181–188	182–189	190–197	191–198
202–209	202–209	222–229	218–225
224–239	224–239	242–257	238–253
252–255	252–255	268–271	264–267
265–268	265–268	286–289	281–284
277–279	277–279	297–299	292–294
282–305	282–305	302–325	297–320
rms, Å	0.526	1.446	1.431

[a] The data shows the amino acids which are homologous in the crystal structures of the four MDHs. The root mean square differences between the backbone atoms of the homologous residues, after fitting the four crystallographic models together using the method of least squares, are shown in the bottom-most entry. The reference coordinate system was mMDH.

structural comparison consists of 191 structurally conserved positions. Of these side chains, only 27 are identical. Lack of structural homology occurs at unexpected locations. For example, the region characterized by the secondary structural elements βD–loop–αD is conformationally different. This may not be a true conformational variation. Rather the structural differences in this segment may represent different states of the enzymes. In LDH, the same loop is believed to change conformation during the binding of NAD (Waldman *et al.*, 1988). This argument becomes even more acceptable when looking at Figure 5. Note that there is no insertion/deletion in this region (mMDH, 70–90).

On the basis of the crystal structures, the N-terminal half of the MDHs has the least conformational variation. The six-stranded nucleotide binding domains, βA through βF in Figure 5, are conformationally similar in all the MDHs, but βC is an exception. In this strand there is little conformational similarity and there are no identities. βC is one of the edge β-strands of the six-membered parallel sheet and is not involved in binding the coenzyme or in the subunit–subunit interface.

Some of the largest conformational differences occur in the antiparallel β structure found in the C-terminal half of the molecule. Note for example, in Figure 5, the twisted antiparallel β-strands labeled βK-βL-βM. Nearly 25 residues are distinctly different in the two classes. Since cMDH and tMDH are about 20 residues larger than mMDH and eMDH, several insertions present unique differences.

The insertions are not randomly distributed but rather appear in five locations. Starting from the N-terminal, c- and tMDH have extra residues beginning at position 35 in mMDH. Next m- and eMDH have three extra residues which are part of the C-terminal end of the helix labeled α1F. One of the largest conformational differences occurs between residues 190 and 200 in mMDH. This segment is one of the antiparallel loops which we have labeled βG-βH-βJ. In c- and tMDH extra amino acids extend the loop into the solvent and a more exposed hairpin turn is formed.

The last insertion/deletion segment is found near residue 264 in mMDH (see Figure 5). It is near the middle of a twisted antiparallel β sheet element labeled βK-βL-βM. The insertion occurs near the middle of this segment, but the conformation at residues 261–265 of c- and tMDH is very different from that of m- and eMDH. Taken together, the major conformational differences between the mMDH versus the cMDH family occur in the C-terminal half of the subunit. This is the portion assigned to the "catalytic" domain (Rossmann *et al.*, 1975). There is no obvious structural explanation for these differences or, for that matter, the 27 identical amino acids describing this family. However, the data presented in Figure 5 now makes it possible to study further the conserved side chains and the conserved structural elements with an eye toward defining their functional significance.

The aforementioned structural similarities suggest that mitochondria are derived from the *E. coli* limb of the bacterial tree and the cytosolic enzyme from the limb containing the microorganism *T. flavus*. While the evolutionary relationships between the prokaryotic and eukaryotic enzymes seem clear, still unresolved are questions regarding the fact that many citric acid cycle enzymes are not represented in both the cytosol and mitochondria. The loss of active forms of some of these enzymes must also have occurred during the evolution of eukaryotic cells.

The evolutionary point of view can be extended to other aspects of the mMDH crystal structure. One can hypothesize that once the genetic information for mMDH was incorporated into the chromosome of the eukarotic cell, a mechanism had to evolve for the translocation of the precursor of mMDH into a mitochondrion. The enzyme mMDH synthesized in the cytosol must be translocated through both the outer and inner membranes of a mitochondrion. Targeting of matrix proteins appears to be due mainly to an extra 20 or so amino acids present on the N-terminal of the precursors and the properties of the protein itself. Three aspects of the mMDH structure are relevant to these processes. One of them is the location of the N-terminal in the 3D structure since it is formed by proteolysis during the translocation process. The N-terminal of the mature enzyme appears to be reasonably accessible in the crystal structure.

The second factor is the number and distribution of electrostatic charges and, finally, the location and ratio of polar to nonpolar residues. These properties must affect the energy needed by the cell for translocation. The effect of the surface properties on protein translocation should be independent of whether the mechanism involves folded or unfolded forms of the protein. The mitochondrial form of MDH clearly has a less polar surface than cytosolic MDH. More specifically, when analyzed in terms of the mMDH crystal structure, mMDH has a reduced accessible surface area for charged or ionizable side chains. While the fraction of surface assignable to uncharged polar moieties is the same in both proteins, mMDH has a higher fraction of nonpolar atoms on the surface. Overall mMDH appears to be a more hydrophobic protein, and this could be a factor in mitochondrial translocation. We are presently trying to prepare quantities of precursor mMDH suitable for crystallization.

Finally mMDH and eMDH were both crystallized from citrate-containing buffers. This has resulted in the binding of a citrate molecule at what we believe to be the position of malate/oxaloacetate at the active site. The mode of substrate binding for both mMDH and eMDH is essentially identical. The substrate binding site is located in proximity to three arginine residues which not only appear to fix the substrate in the correct orientation but also form a hydrogen bond network with the catalytic histidine/aspartic acid pair known to participate in a catalytic cycle. Two of these arginines are also close to what may be a pH sensitive region in the subunit–subunit interface.

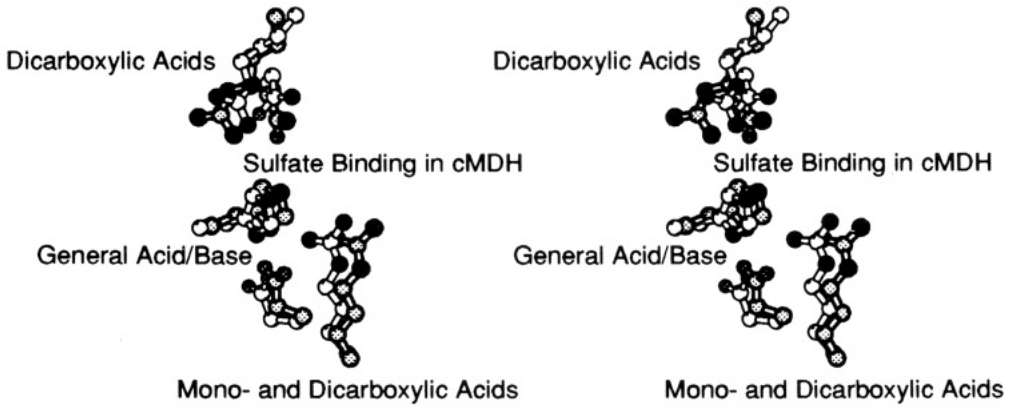

FIGURE 6: The consensus active site of the MDHs. The stereodiagram was prepared from the superimposed coordinates described in Figure 5 and in the text. The side chain carbon atoms represented by unshaded balls are taken from mMDH. Those with shaded carbon atoms belong to cMDH. Only two structures are depicted since cMDH is very similar to mMDH, and tMDH is highly homologous with cMDH. The two arginine side chains labeled mono- and dicarboxylic acids" are R152 from mMDH and R161 from cMDH. The side chains labeled "general acid/base" show the locations of D149 and H176 from mMDH, and D158 and H186 of cMDH. Two arginines labeled "dicarboxylic acids" belong to R80 of mMDH and R91 of cMDH. Behind these two arginines are asparagines: N118 in mMDH and N130 in cMDH. The role of these side chains is described more fully in the accompanying text.

The consensus active site residues required for substrate binding and part of the catalytic process are shown in Figure 6. Only mMDH and cMDH are shown since eMDH and tMDH are structurally very similar: e- to mMDH and t- to cMDH. The histidine–aspartic acid system thought to participate in the catalytic process is labeled "general acid/base" in Figure 6. Note how closely these residues overlap after the superposition of the coordinates. The His–Asp pair is believed to act as a general base in the malate to oxaloacetate direction, removing a proton from the OH of the substrate.

The two arginine side chains labeled "mono- and dicarboxylic acids" are responsible for interacting with one of the carboxylates of the four-carbon dicarboxylic acid substrate. Lactic acid is a monocarboxylic acid, and the same arginine side chain is believed to interact with it in the LDHs. The second arginine side chain, at the top of Figure 6, generates the specificity for malate as opposed to lactate. It is labeled "dicarboxylic acids" in Figure 6. This arginine side chain is shifted a small amount in cMDH but is nonetheless within easy reach of the other carboxylate of malate/oxaloacetate. Mutational changes at this position have been shown to alter the specificity of an LDH to a preference for malate (Wilks *et al.*, 1988). Unlabeled but present in Figure 6 is an asparagine side chain. It is behind the arginines that are required for binding dicarboxylic acids. The asparagine is conserved in all four MDHs (N118 in mMDH and N131 in cMDH) and undoubtedly participates in substrate binding since it is hydrogen bonded to a sulfate anion in cMDH.

No crystal data is yet available on the binding of the coenzyme NAD to mMDH, but with the similarity to eMDH, this portion of the active sites is also likely to be the same. Note, however, the variation in the dinucleotide binding consensus sequence. In all the MDHs it is -G-X-X-G-X-X-G- (Figure 5, mMDH 7–13) except for cMDH, which has an alanine in place of the last glycine. This trademark of the dehydrogenases is only preserved for a relatively short distance since the first insertions can be seen to appear at the end of the αB helix beginning at residue 24 in mMDH (see Figure 5).

At least two further points should be considered. One of them is how to use the static crystallographic model to try to understand how binding and the resulting conformational change of the NAD molecule and the surrounding enzyme activate the C4 position of the nicotinamide ring. This activation step is essential for the movement of a hydride ion back and forth between the substrate and the coenzyme. A second issue is the problem of the loop region. It is this segment of the crystal structure which is conformationally different in the MDHs yet is known to participate in substrate binding. Here the crystallographic model of mMDH may be less useful since it is this loop segment (residues 77–87) which participates in the formation of the pseudotetramer in the crystalline enzyme.

A number of extensions of this structural study remain to be done. The enzyme mMDH is thought to form complexes with enzymes adjacent to it in the citric acid cycle (Srere, 1990). Using the refined crystallographic mMDH coordinates presented here, it should be possible to search for complementarity in the molecular surfaces between mMDH and citrate synthase. If such relationships are found, they may offer additional support for the idea that multienzyme complexes serve an important function within a mitochondrion.

ACKNOWLEDGMENT

The authors wish to acknowledge the effort of Dr. Gale Rhodes, Department of Chemistry at the University of Southern Maine, in the interpretation of early electron density maps. We also gratefully acknowledge support from the Minnesota Supercomputer Institute, which was used extensively during the refinement stages. We thank Michael Hall for early access to eMDH coordinates and Dave Schuller and Ed Hoeffner for their help in maintenance of our local computer facilities. We acknowledge the guidance of Mary Westbrook in the early area detector experiments.

REFERENCES

Birktoft, J. J., & Banaszak, L. (1982) *J. Biol. Chem. 258*, 472–482.

Birktoft, J. J., Fernley, R. T., Bradshaw, R. A., & Banaszak, L. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6166–6170.

Birktoft, J. J., Rhodes, G., & Banaszak, L. (1989a) *Biochemistry 28*, 6065–6081.

Birktoft, J. J., Fu, Z., Carnahan, G. E., Rhodes, G., Roderick, S. L., & Banaszak, L. J. (1989b) *Biochem. Soc. Trans. 17*, 301–304.

Bleile, D., Schulz, R., & Harrison, J. (1977) *J. Biol. Chem. 252*, 755–758.

Brunger, A. T. (1990) *X-PLOR Manual, Version 2.1*, Yale University Press, New Haven.

Clarke, A., Wilks, H., Barstow, D., Atkinson, T., Chia, W., & Holbrook, J. (1988) *Biochemistry 27*, 1617–1622.
Grant, P., Roderick, S., Grant, G., Banaszak, L., & Strauss, A. (1987) *Biochemistry 26*, 128–134.
Grau, U., Trommer, W., & Rossmann, M. (1981) *J. Mol. Biol. 151*, 289–307.
Hall, M., Levitt, D., & Banaszak, L. (1992) *J. Mol. Biol. 226*, 867–882.
Howard, A., Nielsen, C., & Xuong, Ng. H. (1985) *Methods Enzymol. 114*, 452–471.
Kelly, C., Nishiyama, M., Ohnishi, Y., Beppu, T., & Birktoft, J. (1993) *Biochemistry 32*, 3913–3922.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283–291.
Lee, B., & Richards, F. (1971) *J. Mol. Biol. 55*, 379–400.
Morris, A. L., MacArthur, M. W., Hutchinson, E. G., & Thornton, J. M. (1992) *Proteins 12*, 345–364.
Nishiyama, M., Birktoft, J., & Beppu, T. (1993) *J. Biol. Chem. 268*, 4656–4660.
Reeke, G. (1984) *J. Appl. Crystallogr. 17*, 125–130.
Roderick, S., & Banaszak, L. (1986) *J. Biol. Chem. 261*, 9461–9464.
Rossmann, M., Adams, M., Buehner, M., Ford, G., Hackert, M., Liljas, A., Rao, S., Banaszak, L., Hill, E., Tsernoglou, D., & Webb, L. (1973) *J. Mol. Biol. 76*, 533–537.
Rossmann, M., Liljas, A., Branden, C.-I., & Banaszak, L. (1975) *The Enzymes*, Vol. 11a, pp 61–102, Academic Press, New York.
SERC Daresbury Laboratory (1986) *CCP4 Program Suite*, SERC Daresbury Laboratory, Warrington, England.
Srere, P. A. (1981) *Trends Biochem. Sci. 6*, 4–7.
Srere, P. A. (1990) *Trends Biochem. Sci. 15*, 411–412.
Steffan, J. S., & McAlister-Henn, L. (1991) *Arch. Biochem. Biophys. 287*, 276–282.
Takeshima, H., Joh, T., Tsuzuki, T., Shimade, K., & Matsukado, Y. (1988) *J. Mol. Biol. 200*, 1–11.
Thompson, L., Sutherland, P., Steffan, J., & McAlister-Henn, L. (1988) *Biochemistry 27*, 8393–8400.
Tronrud, D., Ten Eyck, L., & Matthews, B. (1986) *Acta Crystallogr. A43*, 489–501.
Waldman, A., Hart, K., Clarke, A., Wigley, D., Barstow, D., Atkinson, T., Chia, W., & Holbrook, J. (1988) *Biochim. Biophys. Acta 150*, 752–758.
Wilks, H., Hart, K., Feeney, R., Dunn, K., Muirhead, H., Chia, W., Barstow, D., Atkinson, T., Clarke, A., & Holbrook, J. (1988) *Science 242*, 1541–1544.
Wood, D., Jurgensen, S., Geesin, J., & Harrison, J. (1981) *J. Biol. Chem. 256*, 2377–2382.