

# Structural Basis for Thermophilic Protein Stability: Structures of Thermophilic and Mesophilic Malate Dehydrogenases

Bjørn Dalhus<sup>1</sup>, Markku Saarinen<sup>2</sup>, Uwe H. Sauer<sup>2</sup>, Pär Eklund<sup>2</sup>  
Kenth Johansson<sup>2</sup>, Andreas Karlsson<sup>2</sup>, S. Ramaswamy<sup>2</sup>  
Alexandra Bjørk<sup>3</sup>, Bjørnar Synstad<sup>3</sup>, Kristine Naterstad<sup>3</sup>  
Reidun Sirevåg<sup>3</sup> and Hans Eklund<sup>2\*</sup>

<sup>1</sup>Department of Chemistry  
University of Oslo  
Box 1033, Blindern  
N-0316 Oslo, Norway

<sup>2</sup>Department of Molecular  
Biology, Swedish University  
of Agricultural Sciences  
P.O. Box 590, Biomedical  
Center S-751 24 Uppsala  
Sweden

<sup>3</sup>Department of Biology  
University of Oslo  
Box 1066, Blindern  
N-0316 Oslo, Norway

The three-dimensional structure of four malate dehydrogenases (MDH) from thermophilic and mesophilic phototropic bacteria have been determined by X-ray crystallography and the corresponding structures compared. In contrast to the dimeric quaternary structure of most MDHs, these MDHs are tetramers and are structurally related to tetrameric malate dehydrogenases from Archaea and to lactate dehydrogenases. The tetramers are dimers of dimers, where the structures of each subunit and the dimers are similar to the dimeric malate dehydrogenases. The difference in optimal growth temperature of the corresponding organisms is relatively small, ranging from 32 to 55 °C. Nevertheless, on the basis of the four crystal structures, a number of factors that are likely to contribute to the relative thermostability in the present series have been identified. It appears from the results obtained, that the difference in thermostability between MDH from the mesophilic *Chlorobium vibrioforme* on one hand and from the moderate thermophile *Chlorobium tepidum* on the other hand is mainly due to the presence of polar residues that form additional hydrogen bonds within each subunit. Furthermore, for the even more thermostable *Chloroflexus aurantiacus* MDH, the use of charged residues to form additional ionic interactions across the dimer–dimer interface is favored. This enzyme has a favorable intercalation of His-Trp as well as additional aromatic contacts at the monomer–monomer interface in each dimer. A structural alignment of tetrameric and dimeric prokaryotic MDHs reveal that structural elements that differ among dimeric and tetrameric MDHs are located in a few loop regions.

© 2002 Elsevier Science Ltd. All rights reserved

**Keywords:** thermostability; crystal structure; tetrameric malate dehydrogenases; NAD

\*Corresponding author

Present addresses: U. H. Sauer, Umeå Center for Molecular Pathogenesis, UCMP, Umeå University, S-90187 Umeå, Sweden; S. Ramaswamy, Department of Biochemistry, Bowen Science Building, University of Iowa, Iowa City, IA 52242-1109, USA; B. Synstad, Department of Chemistry and Biochemistry, Agricultural University of Norway, Box 5040, N-1432 Ås, Norway; K. Naterstad, Matforsk, Osloveien 1, N-1430 Ås, Norway.

Abbreviations used: MDH, malate dehydrogenase; LDH, lactate dehydrogenase.

E-mail address of the corresponding author: [hasse@xray.bmc.uu.se](mailto:hasse@xray.bmc.uu.se)

## Introduction

The green gliding bacteria and the green sulfur bacteria represent two distinct groups of phototropic bacteria, which are found relatively far apart in the current 16 S rRNA phylogenetic tree.<sup>1</sup> However, the two groups share important characteristics, such as the types of chlorophyll and the structure of the antenna systems (chlorosomes). In order to obtain information about structure and thermostability, we have performed comparative studies on malate dehydrogenase (MDH; malate-NAD<sup>+</sup> oxidoreductase; EC 1.1.1.37) and its corresponding gene, *mdh*, from several green bacteria:

the green gliding bacterium *Chloroflexus aurantiacus* (*ca*-MDH), a thermophile, the green sulfur bacteria *Chlorobium tepidum* (*ct*-MDH), a moderate thermophile, and *Chlorobium vibrioforme* (*cv*-MDH), a mesophile. In addition, a hybrid *mdh* constructed from *C. vibrioforme* and *C. tepidum* has been over-expressed and the corresponding hybrid protein analyzed.<sup>2,3</sup> This *hybrid*-MDH contains residues 1–71 from *ct*-MDH and residues 72–310 from *cv*-MDH.

MDHs from a variety of organisms have been studied extensively. In most species so far examined, the native enzyme consists of identical subunits with a molecular mass of about 35 kDa. In most cases, including all eukaryotes, MDH is a dimer. However, in some prokaryotes, i.e. some *Bacillus* species,<sup>4,5</sup> and in phototrophic bacteria,<sup>6,7</sup> MDH is tetrameric in its native state. The significance of the oligomeric state of MDH is not known with certainty, but the dimeric structure of MDH from *Escherichia coli* is critical for enzyme activity, since the activity of monomeric mutants, unable to form dimers was decreased dramatically.<sup>8</sup> Whether the tetrameric form is important for the thermostability of the present MDHs is not known, but it has been suggested that subunits in thermostable enzymes in general are more tightly bound than in less stable homologous species.<sup>9</sup>

Comparison of amino acid sequences of a wide range of MDHs has revealed that tetrameric MDHs, including those from phototrophic bacteria, cluster together with lactate dehydrogenases (LDHs) in a dendrogram, and that their overall sequence similarity to LDHs is significantly greater than to other MDHs.<sup>2,3</sup> A high degree of similarity to LDH is also the case for the tetrameric MDH from *Bacillus* strain BI<sup>10</sup> and for MDH from the Archaea *Haloarcula marismortui*, *Archaeoglobus fulgidus* and *Methanococcus fervidus*.<sup>11</sup> On the basis of these findings, the term LDH-like MDH has been suggested for a protein with LDH-like structure and MDH enzymatic activity.<sup>3,12</sup>

The available data clearly indicate that the intrinsic thermal stability of proteins cannot be explained by a single mechanism.<sup>13</sup> On the contrary, comparative studies of the structure of homologous proteins from thermophiles and mesophiles suggest that several factors, such as tighter hydrophobic packing of the protein core,<sup>14–16</sup> increased hydrophobicity of the protein surface,<sup>17</sup> and electrostatic interactions as well as hydrogen bonding are responsible for a more rigid and therefore more stable protein. Thus, it appears that in thermostable proteins there is an abundance of ion pairs that form bonding networks that play major roles in thermal stability.<sup>18–20</sup> Reduction of the number and volume of cavities have been shown to play a role in stability<sup>14,15,21</sup> in addition to improved inter-subunit contacts within oligomeric proteins.<sup>22</sup>

Comparison of the three-dimensional structure between mesophilic and thermophilic proteins has

been performed for a number of enzyme families, including several dehydrogenases: glutamate dehydrogenase,<sup>20,23</sup> glyceraldehyde-3-phosphate dehydrogenase,<sup>24</sup> lactate dehydrogenase,<sup>25</sup> 3-isopropylmalate dehydrogenase<sup>26</sup> and malate dehydrogenase.<sup>27</sup> In the latter case, only MDHs from the thermophilic *Thermus flavus*<sup>28</sup> and the hyperthermophilic *Methanococcus jannaschii* have been compared with mesophilic MDHs. The optimum temperatures of the organisms in the present study are 32 °C for *C. vibrioforme*, 47 °C for *C. tepidum* and 55 °C for *C. aurantiacus*. However, despite the relatively small differences in the optimal growth temperature of these organisms, it is possible to explain the observed differences in MDH thermostability by variations in the protein structures.

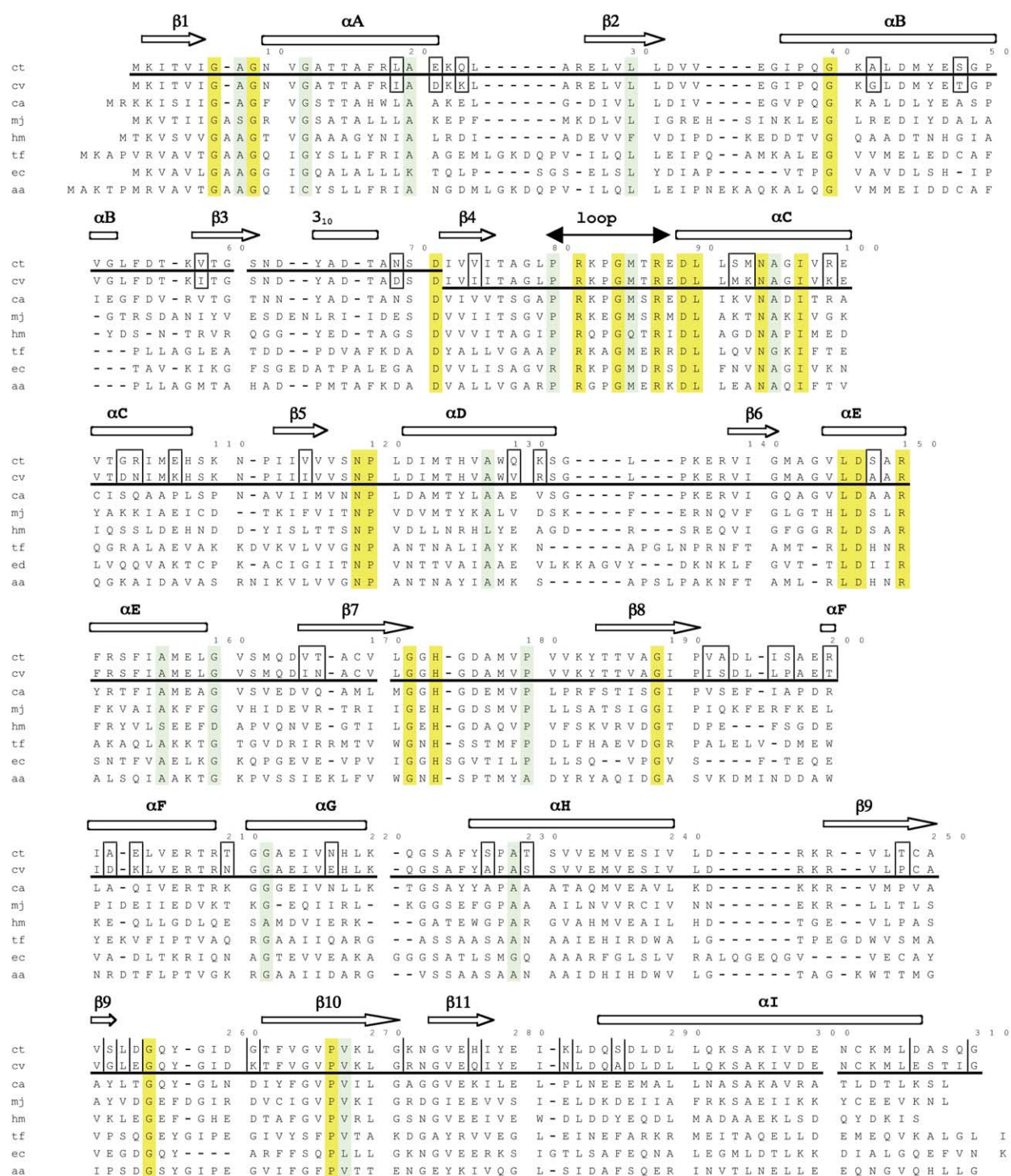
To date, the three-dimensional structures have been determined for several dimeric MDHs; those from porcine heart mitochondria<sup>29,30</sup> and cytoplasm,<sup>31,32</sup> from *E. coli*,<sup>33</sup> *T. flavus*<sup>28</sup> and *Aquaspirillum arcticum*,<sup>34</sup> and from the chloroplasts of *Sorghum vulgare*<sup>35</sup> and *Flaveria bidentis*.<sup>36</sup> By contrast, there is considerably less information available about tetrameric MDHs. The only three-dimensional structures of tetrameric MDHs reported so far are those from the two Archaea, the halophilic *H. marismortui*<sup>37,38</sup> at 2.9 Å resolution and the hyperthermophilic *M. jannaschii*<sup>27</sup> at 2.8 Å resolution.

Previously, dimeric MDHs from highly thermophilic organisms have been compared with mesophilic forms of the enzyme in order to rationalize structural factors that govern increased dimeric MDH thermostability.<sup>28</sup> In the present investigation, we have solved and analyzed the crystal structure of a series of tetrameric MDHs from closely related bacterial species with different optimal growth temperature; the thermophilic *C. aurantiacus* (*ca*-MDH), the moderate thermophilic *C. tepidum* (*ct*-MDH), the mesophilic *C. vibrioforme* (*cv*-MDH) and a *hybrid*-MDH between MDH from the latter two (*hybrid*-MDH). The present crystal structures are the first examples of tetrameric MDHs from bacteria, and MDH from *C. aurantiacus* is the first thermophilic tetrameric MDH structure to be solved by X-ray crystallography. The fact that the three-dimensional structure has now been determined for these four MDHs and that the corresponding genes in all cases are isolated and characterized, will allow us to analyze in detail the contribution of individual amino acid residues to the tetrameric thermostability through site-directed mutagenesis. In the case of *ca*-MDH, this kind of analysis is in progress in our laboratory.

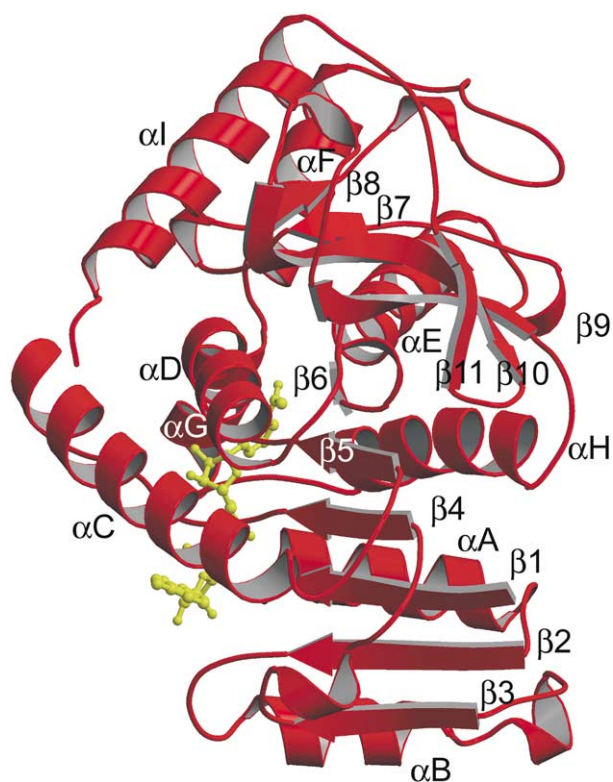
## Results and discussion

### Structure

A structure-based multiple sequence alignment and secondary structure assignment for the four



**Figure 1.** Structural alignment of the amino acid sequences of all prokaryotic MDHs with known three-dimensional structure. The ct is from *Chlorobium tepidum*, cv from *Chlorobium vibrioforme*, ca from *Chloroflexus aurantiacus*, mj from *Methanococcus jannaschii* (2.8 Å, PDB-access code 1HYG), hm from *Haloarcula marismortui* (2.9 Å, PDB-access code 1D3A), tf from *Thermus flavus* (1.9 Å, PDB-access code 1BMD), ec from *Escherichia coli* (1.9 Å, PDB-access code 1EMD) and aa from *Aquaspirillum aquaticus* (1.9 Å, PDB-access code 1B8P). The sequences used to construct the hybrid-MDH are underlined, and the sequence numbering refers to the cv and ct-MDHs. Boxed residues show sequence differences between the ct and cv MDHs. Secondary structure elements in ca-MDH as assigned by PROCHECK<sup>64</sup> are indicated above each alignment row. Fully conserved residues are marked in yellow; residues conserved in all enzymes except one are marked in green.



**Figure 2.** Ribbon diagram of a single MDH subunit with bound NAD cofactor. Strands  $\beta 1$ – $\beta 6$  and helices  $\alpha A$ – $\alpha D$  form the NAD-binding domain. NAD is in yellow. All Figures except for Figure 1 were prepared using MOLSCRIPT,<sup>65</sup> Bobscript<sup>66</sup> and Raster3D<sup>67</sup> and AdobePhotoshop 6.0 (Adobe Systems Inc, 2001).

MDHs presented in this study are shown in Figure 1 together with all other prokaryotic MDHs with known three-dimensional structure. There are 309 residues in *ca*-MDH, and 310 in each of *ct*-, *hybrid*- and *cv*-MDH. Compared to the latter three MDHs, *ca*-MDH has two more residues at the N terminus, two residues less in the C terminus as well as a one residue deletion in the turn between  $\alpha A$  and  $\beta 2$ . To simplify the discussion of specific residues in the four structures, the residue numbering for *ct*-MDH, *cv*-MDH, and the *hybrid*-MDH, which contains residues 1–71 from *ct*-MDH and residues 72–310 from *cv*-MDH, is underlined in Figure 1.

The four MDH structures are determined within the resolution range 2.0–2.5 Å with good stereochemistry (see Table 4). Approximately 90% of the residues have torsion angles that correspond to the most favored conformations for the protein backbone.<sup>39</sup> The outliers are a few lysine residues at the border of the  $\alpha_1$  region in the Ramachandran plot. As for most MDHs, the loop of residues 80–90 close to the substrate-binding site is flexible and not well defined in any of the structures.

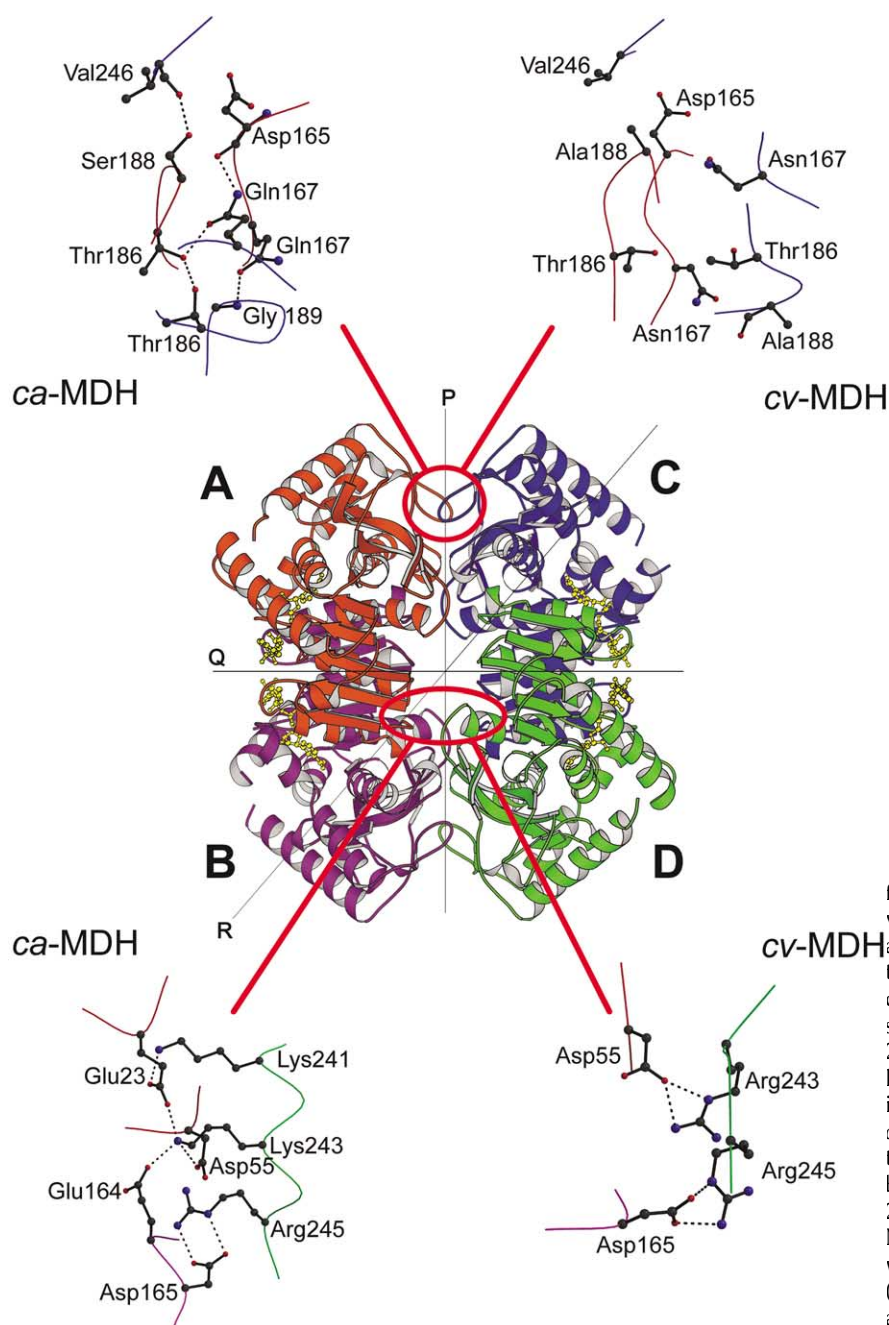
The protein monomer contains nine well-defined helices and 11 strands, of which strands  $\beta 1$ – $\beta 6$

form a parallel  $\beta$ -sheet that, together with helices  $\alpha A$ – $\alpha D$ , constitute the classical coenzyme binding domain<sup>40</sup> (Figure 2). Strands  $\beta 6$  and  $\beta 9$ – $\beta 11$  participate in a more complex anti-parallel twisted sheet, whereas strands  $\beta 7$  and  $\beta 8$  form a small, separate anti-parallel sheet between helix  $\alpha E$  and helix  $\alpha F$ . The general structure of the four MDHs is very similar. If the flexible loop 80–90 is excluded, the subunits can be superimposed on each other with root-mean-square deviations (rmsd) of  $C^\alpha$  atoms of about 0.4 Å between *ct*-MDH and the hybrid, about 0.6 Å between *cv*-MDH and *ct*-MDH, and about 1.0 Å between the *ca*-MDH and the others.

### Quaternary structure

All four enzymes are tetramers in the crystals and elute as tetramers when analyzed by gel-filtration. Sedimentation centrifugation and non-denaturing PAGE of the thermostable *ca*-MDH were also consistent with a tetramer.<sup>41</sup> On the other hand, based on results from non-denaturing PAGE, *ct*-MDH and *cv*-MDH have been reported to be dimers,<sup>42</sup> indicating that the dimer–dimer interactions are weaker in these two MDHs than in *ca*-MDH. This is supported by the results from the analysis in the following sections. Based on the characteristics of the different monomer–monomer interfaces, all four enzymes are best described as a dimer of dimers where the subunits A + B and C + D constitute each of the two dimers, which in turn interact back-to-back to form the tetramer, as depicted in Figure 3. At the dimer–dimer interface, regions 188–192 and 241–246 play important roles and fit well into crevices on the opposite dimer. The monomeric arrangement in each dimer resembles that in other MDHs where the native active enzyme is a dimer.

In general, the subunit structure of MDHs is similar to that of LDHs,<sup>40</sup> which is supported by the fact that it is possible to redesign LDH by mutagenesis in the active site to become a highly active MDH.<sup>43</sup> The tetrameric organization of the four MDHs presented in this study shows a high degree of similarity to that of the tetrameric LDHs. Thus, the quaternary structure of these MDHs can be described according to the PQR-axes nomenclature proposed for LDH<sup>44</sup> (see Figure 3). The Q-axis relates the two subunits in the dimeric MDHs whereas the P and R-axes relate the two dimers, which comprise the full tetramer. Each of the two dimers is formed by a large number of direct and water-mediated contacts across the Q-axis with most of the residues participating in these contacts located in the helices A and B in one monomer and in the helices E, F and H in the other. The dimer–dimer interactions across the PR-plane particularly involve residues located in strands 8, 9 and 10. These interactions can be either parallel with the Q-axis (from A to C or from B to D) or diagonal to it (from A to D or from B to C).



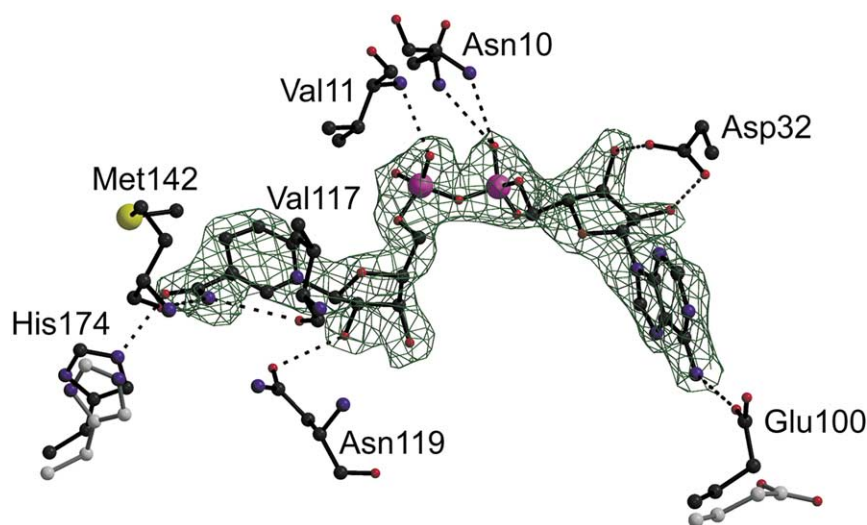
**Figure 3.** Ribbon diagram of the full MDH tetramer with close-up views of two regions in *ca*-MDH and *cv*-MDH containing residues that interact across the dimer-dimer interface. The upper panel shows polar interactions across the 2-fold *P*-axis (A–C) while the lower panel illustrates differences in ionic interactions. The Glu23 (in *ca*-MDH only) and Asp55 form contacts to residues 241 and 243 between monomers related by the 2-fold *R*-axis (A–D). Glu164 (in *ca*-MDH only) and Asp165 interact with residues across the *P*-axis (B–D). All these interactions occur also on symmetry-related interfaces.

In *ca*-MDH and *ct*-MDH, the asymmetric unit contains a dimer in which the two monomers are related by a non-crystallographic 2-fold axis, the *Q*-axis in Figure 3. The two monomers are almost identical, with an rmsd between equivalent C $\alpha$  atoms of 0.53 and 0.34 Å in *ca*-MDH and *ct*-MDH, respectively. The full tetramer is generated by proper crystallographic 2-fold symmetry along the *P* and *R*-axes (Figure 3). In *cv*-MDH and *hybrid*-MDH, the crystallographic asymmetric unit contains a full tetramer, and there are similar relationships between the monomers, except that all axes are non-crystallographic. The rmsd values between corresponding C $\alpha$  atoms in the different peptide chains are all in the range 0.15–0.36 Å for *hybrid*-

MDH and in the range 0.38–0.57 Å for *cv*-MDH. The dimeric and tetrameric organization is very similar, and the dimers and tetramers of the four MDH structures can be superimposed with C $\alpha$  differences of 0.6–1.1 Å.

#### Active site and coenzyme binding

The active-site architecture and active-site residues in the tetrameric bacterial MDHs are the same as for MDHs from other species. The His174–Asp147 couple and Arg150, which are essential for substrate binding and catalysis, are positioned in a manner similar to that of the corresponding residues in the dimeric MDHs. All



**Figure 4.** Final ( $2F_o - F_c$ ) electron density map. A close-up of the map around one of the four NAD cofactors in the *hybrid*-MDH showing the quality of the map (contoured at  $1.0\sigma$ ). Hydrogen bonds are drawn with broken lines. Residue Glu100 forms a hydrogen bond to the amino group in the adenine ring, while Asp32 forms two hydrogen bonds to two hydroxyl groups in the adenosine ribose. The two phosphate groups in NAD interact with main-chain nitrogen atoms in Asn10 and Val11 as well as with the side-chain amino nitrogen atom in Asn10. The NMN ribose is bonded to Asn119 while the carboxamide moiety of the nicotinamide ring is hydrogen bonded to His174 and the main-

chain carbonyl groups of residues Val117 and Met142. The two conformations of Glu100 and His174 illustrate the shifts in the *hybrid*-MDH relative to *cv*-MDH upon binding of cofactor.

structures except *cv*-MDH contain NAD. The NAD molecules have well-defined electron density and have a boomerang-shaped conformation with interactions similar to those in the case of dimeric MDHs (Figure 4). The adenine ring is sandwiched between Val/Ile at position 33 and Ala/Ser at position 77, with a hydrogen bond to Glu100; the adenosine ribose moiety is in the *2'-endo* conformation and forms two hydrogen bonds to Asp32, which clearly distinguish the specificity for NAD relative to NADP. The phosphate groups in NAD form hydrogen bonds with the main-chain nitrogen atoms of residues 10 and 11. The ribose moiety in the nicotinamide mononucleotide adopts a *2'-endo* conformation and is hydrogen bonded to Asn119. The carboxamide moiety of the nicotinamide ring is hydrogen bonded to His174 and to the main-chain carbonyl groups of residues 117 and 142.

There are some conformational differences in the protein around the coenzyme-binding site between the coenzyme-containing *hybrid*-MDH and the coenzyme-free *cv*-MDH. In the NAD-containing *hybrid*-MDH, both the N terminus of  $\alpha$ C, the loop between  $\beta$ 5 and  $\alpha$ D as well as the loop between

$\beta$ 7 and  $\beta$ 8 have moved towards the coenzyme-binding cleft. The latter two loops form contacts with helix  $\alpha$ I at the protein surface. As a consequence of the relocation of the loops, residues 297–302 in  $\alpha$ I in the *hybrid*-MDH are shifted 1.6–2.2 Å toward the NAD-binding site. Further, displacements in the range 1.5–3.0 Å are observed with regard to the side-chain atoms in residues Glu100 and His174, as illustrated in Figure 4.

### Comparison with other MDHs and LDHs

The monomeric structures of the present four MDHs are very similar to those of other known MDH and LDH structures. The rms differences in  $C^\alpha$  positions of superimposed subunits are in the range 1.1–2.2 Å. Generally, the subunits in the MDHs described are more similar to the subunits of LDHs than to the subunits of MDHs from other sources. For example, the subunit of LDH from *T. maritima*<sup>25</sup> has an rmsd of 1.4 Å when the  $C^\alpha$  atoms are superimposed on the *hybrid*-MDH. In contrast, the tetrameric MDH from *H. marismortui*,<sup>37,38</sup> which has a structure most similar to the present MDHs, has a corresponding rmsd

**Table 1.** Half-lives of thermal inactivation and apparent melting temperatures ( $t_m$ ) for MDH from *C. aurantiacus* (*ca*-MDH), *C. tepidum* (*ct*-MDH), *C. vibrioforme* (*cv*-MDH) and a hybrid between *ct*-MDH and *cv*-MDH (*hybrid*-MDH). n.a., not applicable

Enzyme	Half-life (min) at			$t_m$ (°C)	Optimal growth temperature of corresponding organism (°C)
	55 °C	60 °C	65 °C		
<i>ca</i> -MDH	> 60	n.d. <sup>a</sup>	25	67.8	55
<i>ct</i> -MDH	> 30	1.5	0	52.6	47
<i>hybrid</i> -MDH	11.5	1.0	0	48.7	n.a.
<i>cv</i> -MDH	0.5	0.5	0	44.5	32

<sup>a</sup> n.d., Not determined.

**Table 2.** Protein and cavity volume calculations for MDH from *C. aurantiacus* (*ca*-MDH), *C. tepidum* (*ct*-MDH) and *C. vibrioforme* (*cv*-MDH) as obtained using VOIDOO<sup>63</sup>

Enzyme	Protein volume (Å <sup>3</sup> )			Cavities per monomer	
	Monomer <i>a</i>	Tetramer <i>b</i>	$\Delta = b - 4a$	Total cavity volumes (Å <sup>3</sup> )	Number of cavities
<i>ca</i> -MDH	28,530	114,000	-120	128	2
<i>ct</i> -MDH	28,650	114,500	-100	172	3
<i>cv</i> -MDH	28,930	115,700	-20	208	5

of 1.5 Å. The monomer–monomer organizations in *H. marismortui* and the *hybrid*-MDH are somewhat different, so the rmsd for superimposing a dimer is slightly higher (1.9 Å).

### Thermostability and subunit structure

A major aim of the present investigation is to obtain structural information in order to understand the molecular basis for the observed differences in thermostability between the various tetrameric MDHs. Table 1 gives the half-lives of thermal inactivation and apparent melting temperatures ( $t_m$ ) for the four enzymes examined. As can be seen from the results obtained, the observed differences in thermostability are significant and are well in accordance with the optimum growth temperatures of the corresponding organisms. To our knowledge, no  $t_m$  value has been determined for dimeric MDHs.

The most thermostable of the four, *ca*-MDH, has a higher number of alanine and aromatic residues than the others, 39 versus 25/26 and 16 versus 12, respectively. Compared with the less thermostable MDHs, some of the aromatic amino acid residues in *ca*-MDH participate in additional aromatic interactions, like Phe10 and Tyr225, or fill small cavities and clefts, like Phe195 and Phe264. In addition, the positively charged Arg138 interacts with the  $\pi$ -system of Phe134, which also fills a small cleft on the protein surface. All these effects could very well stabilize the *ca*-MDH monomer relative to the other MDHs. The cation to  $\pi$ -system type of interaction was studied recently in relation to thermostability.<sup>45</sup> The backbone flexibility in the more thermostable *ca*-MDH is probably reduced due to a slightly higher proline content compared with other enzymes. There are 16 proline residues in *ca*-MDH, whereas there are only 11/13 in *ct*/*cv*-MDH.

Salt-bridges have been demonstrated to be an important factor for temperature stability.<sup>18–20</sup> Surprisingly, there is a higher number of charged residues in the less thermostable MDHs; on average, there are four less cationic and four less anionic residues in *ca*-MDH than in the other three. In *ca*-MDH there are only six salt-bridges on the monomer surface compared with ten in *ct*-MDH and seven in *hybrid*-MDH and *ct*-MDH.

It has been suggested that thermostable proteins are more tightly packed than less thermostable species. Table 2 gives the protein volumes of single

subunits for the four MDHs investigated. The calculations show that there is an increase in the packing density following a decrease in the protein volume, in each monomer in the order of increased thermal stability. This is consistent with cavity calculations, which show that the volume and number of cavities decrease in the order of higher thermostability (Table 2).

However, more important differences between the MDHs occur at the oligomeric level than at the monomeric level, as discussed below.

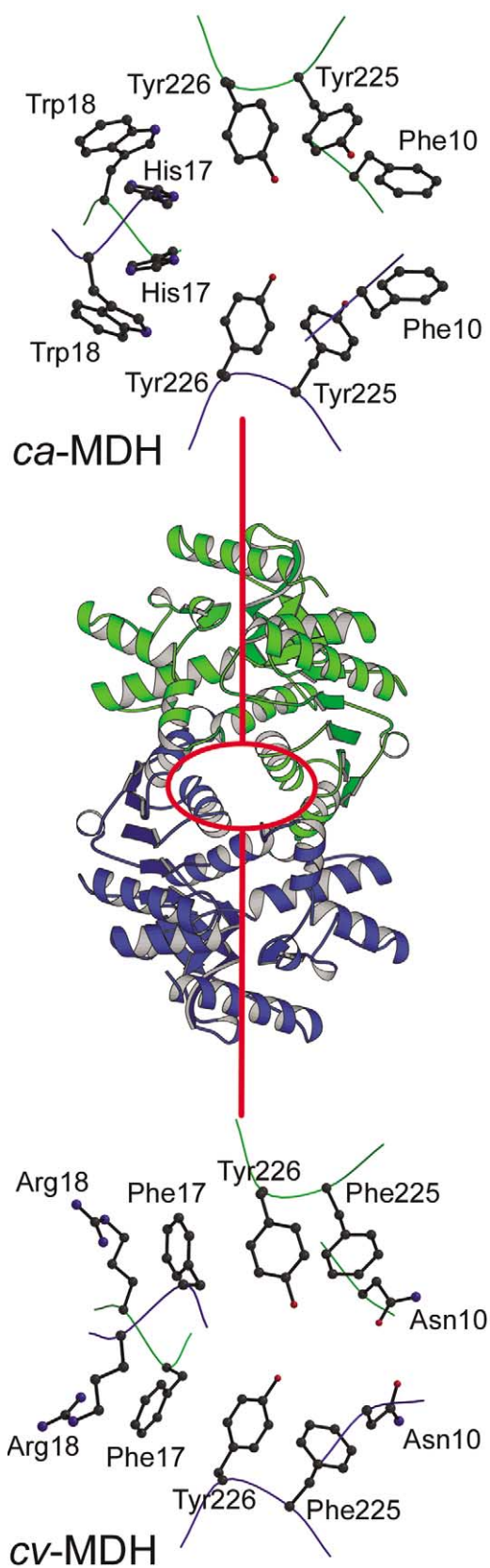
### Thermostability and quaternary structure

There is a clear correspondence between thermostability (Table 1) and molecular protein volume occupied by the full tetramers, as listed in Table 2. Actually, the difference between the tetrameric volume and the volume of the corresponding number of monomeric units,  $\Delta$ , is negative in all cases, demonstrating that the protein density increases slightly upon multimerization. Further, this contraction is larger in the more temperature-resistant enzymes, Table 2.

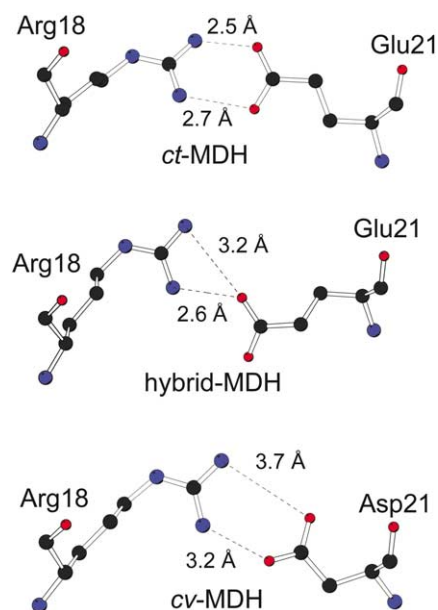
The central void in all four MDHs is connected to the solvent environment *via* a tunnel along the *R*-axis (Figure 3). Further, in *ct*, *hybrid* and *cv*-MDH this large void is connected to the surroundings through another tunnel along the *Q*-axis (i.e. along the 2-fold axis in each MDH-dimer). In *ca*-MDH this tunnel is closed at the His-Trp intercalation. Since the entrance to the interior along the *R*-axis is relatively large, no natural border can be used in volume calculations. Surprisingly, calculations of buried surfaces show no correlation with thermostability for the four enzymes.

### Intra-dimer interactions

The MDH dimers, comprised of subunits A + B or C + D, are formed by a high number of direct and water-mediated contacts across the *Q*-axis (Figure 3). The number and arrangement of hydrophobic residues that constitute the monomer–monomer interfaces in the dimer are, in general, the same in all four MDHs. However, the monomer–monomer interface in *ca*-MDH is improved through favorable interactions with amino acid residues from both monomers (Figure 5). His17 and Trp18 from each monomer intercalate to form a “zipper”. Furthermore, homologous



**Figure 5.** Ribbon diagram of one of the dimers in *ca*-MDH along the *Q*-axis (Figure 3). The two detailed pictures show selected interactions of this dimer interface in *ca*-MDH (upper panel) and *cv*-MDH (lower panel). For the sake of clarity, the detailed Figures are viewed with the *Q*-axis roughly horizontal. In *ca*-MDH, the His17-Trp18 forms a tight intercalation whereas the corresponding part of *cv*-MDH is open.



**Figure 6.** Arg18-Glu/Asp21 interactions. Ball-and-stick representation of the hydrogen bond interactions between Arg18 and Glu/Asp21 in *ct*-MDH, *hybrid*-MDH and *cv*-MDH, which is probably the major interaction that governs differences in thermostability between the closely related *hybrid* and *cv*-MDH. The hydrogen bond distances are mean distances for each interaction.

interactions of two Phe10 residues are part of an aromatic cluster, Phe10, Tyr225 and Tyr266 from each peptide chain. Recent studies show that additional aromatic clusters in thermophilic enzymes are usually small.<sup>45</sup>

Interestingly, in the series *ct*, *hybrid* and *cv*-MDH, the Arg18-Glu21 salt-bridge contains two strong hydrogen bonds in *ct*-MDH, one well-defined hydrogen bond in the *hybrid*-MDH while only a weak contact in *cv*-MDH, in accordance with the observed thermostability (Figure 6).

#### Inter-dimer interactions

When the tetrameric assembly is examined, the differences between the investigated MDHs are more pronounced. Formations of tetrameric MDHs occur through interactions between amino acid residues in loop regions  $\alpha$ A- $\beta$ 2,  $\alpha$ E- $\beta$ 7,  $\beta$ 8- $\alpha$ F and  $\alpha$ H- $\beta$ 9. Among the residues connecting the AB and CD dimers across the *PR*-plane, there are five ionic contacts in the most thermostable *ca*-MDH compared to only two in the other three enzymes (Table 3; Figure 3). Particularly, Glu23 in *ca*-MDH forms two salt-bridges, one to Lys241 and one to Lys243 in the neighboring peptide chain. This negatively charged glutamic acid residue is replaced by the neutral Gln23 in *ct*-MDH and *hybrid*-MDH, and with the positively charged Lys23 in the least stable *cv*-MDH. Hence, the ionic interaction gradually changes character from attractive to repulsive, in full agreement with the observed thermostability. Furthermore, the Arg

**Table 3.** Direct hydrogen bonds (per monomer) within dimers AB and CD as well as between dimers AB and CD in *ca*, *ct*, *hybrid* and *cv*-MDH

	<i>ca</i> -MDH	<i>ct</i> -MDH	<i>hybrid</i> -MDH	<i>cv</i> -MDH
<i>Hydrogen bonds within dimers AB and CD</i>				
	S13-Y226	R18-E21 <sup>a</sup> R18-E21	R18-E21	
	K22-E52			
	D44-225 <sub>N</sub>	D44-225 <sub>N</sub>	D44-225 <sub>N</sub>	D44-225 <sub>N</sub>
	D44-226 <sub>N</sub>	D44-226 <sub>N</sub>	D44-226 <sub>N</sub>	D44-226 <sub>N</sub>
	D44-227 <sub>N</sub>	D44-227 <sub>N</sub>	D44-227 <sub>N</sub>	D44-227 <sub>N</sub>
	E47-R150	44 <sub>O</sub> -S227 E47-R150 S48-S227 <sup>a</sup> S48-T230	E47-R150	E47-R150
	D55-S162	50 <sub>O</sub> -R152 D55-S162	50 <sub>O</sub> -R152 D55-S162	50 <sub>O</sub> -R152 D55-S162
	D55-163 <sub>N</sub>	D55-163 <sub>N</sub>	D55-163 <sub>N</sub>	D55-163 <sub>N</sub>
Sum	8	12	8	7
<i>Hydrogen bonds between dimers AB and CD</i>				
	R(-1)-241 <sub>O</sub>			
	R(-1)-242 <sub>O</sub>			
	E23-K241			
	E23-K243			
	D55-K243	D55-R243	D55-R243	D55-K243
		D55-R243	D55-R243	D55-K243
		159 <sub>O</sub> -K269	159 <sub>O</sub> -K269	159 <sub>O</sub> -K269
	162 <sub>N</sub> -244 <sub>O</sub>	162 <sub>N</sub> -244 <sub>O</sub>	162 <sub>N</sub> -244 <sub>O</sub>	162 <sub>N</sub> -244 <sub>O</sub>
	E164-K243	Q164-R243	Q164-R243	Q164-R243
		Q164-R243	Q164-R243	Q164-R243
	D165-R245	D165-R245	D165-R245	D165-R245
	D165-R245	D165-R245	D165-R245	D165-R245
	D165-246 <sub>N</sub>	D165-246 <sub>N</sub>	D165-246 <sub>N</sub>	D165-246 <sub>N</sub>
	165 <sub>O</sub> -Q167			
	Q167-T186			
	Q167-189 <sub>N</sub>			
	T186-T186			
	S188-246 <sub>O</sub>			
		D194-K282		
Sum	15	10	9	9
Total	23	22	17	16

Main chain atoms are noted with a subscript whereas side chain interactions are noted with the residue type in one letter code.

<sup>a</sup> Forms two hydrogen bonds, listed twice.

that is present at the N terminus in *ca*-MDH but not in the other MDHs, forms two well-defined hydrogen bonds with two main-chain carbonyl groups in helix  $\alpha$ H in the neighboring monomer.

Another feature of *ca*-MDH that may explain the increased thermostability through an improved dimer-dimer interface compared with the other three MDHs studied, is the presence of five hydrogen bonds between polar neutral residues unique to *ca*-MDH across the dimer-dimer interface (Table 3 and Figure 3).

#### Comparison of *cv*-MDH with *hybrid*-MDH and *ct*-MDH

Since *C. aurantiacus* is significantly more thermophilic than the *Chlorobium* species investigated, it is not surprising that *ca*-MDH differs from the other three MDHs with regard to stabilizing structural elements, particularly on the dimer-dimer interface. The stabilizing differences

between *ct*, *hybrid* and *cv*-MDH are, in general, less pronounced.

As can be seen from Figure 1, approximately 75% of *hybrid*-MDH is derived from that of the *cv*-MDH gene. It is not surprising that the amino acid sequences of *cv*-MDH and *hybrid*-MDH are dissimilar in only seven positions. For that reason, the observed difference in thermostability between the two MDHs (Table 1) is most likely due to structural changes introduced through one or more of these residues. The fact that *hybrid*-MDH has an alanine residue in position 42 whereas glycine is present in this position in *cv*-MDH probably stabilizes the helix  $\alpha$ B somewhat, since the residue in question is in the middle of the helix.<sup>46</sup> However, the amino acid residue most likely responsible for much of the increase in thermostability in the *hybrid*-MDH compared to that of *cv*-MDH, is Glu21 in the former. Glu21 makes one well-defined hydrogen bond to Arg18 across the monomer-monomer interface in each dimer. The corresponding

Asp21 in *cv*-MDH is less favorable, at a distance of about 3.6 Å (Figure 6).

*ct*-MDH and *cv*-MDH differ in their amino acid sequence at 43 positions, marked with boxed residues in Figure 1. In addition to the above-mentioned stabilizing factors in *hybrid*-MDH compared to *cv*-MDH, there are further structural elements distinguishing *ct*-MDH and *cv*-MDH. There are four more salt-bridges across each monomer surface in *ct*-MDH, of which several probably contribute towards a more rigid protein backbone, as the residues in question belong to different secondary structure elements, e.g. Glu158-Arg200, Asp242-Lys272. Interestingly, the number of Ser and Thr is somewhat higher in *ct*-MDH than in *hybrid/cv*-MDH with 37 versus 31/31 amino acid residues, respectively. Several of these residues participate in hydrogen bond stabilization of each subunit in *ct*-MDH. In total, there are eight more hydrogen bonds involving such hydroxyl groups in *ct*-MDH compared to *hybrid*-MDH and *cv*-MDH. There are only minor differences between *ct*-MDH and *cv*-MDH as far as interactions across subunit interfaces are concerned. A few serine residues contribute towards higher stability of the dimeric unit, and one additional salt-bridge, Asp194-Lys282, in the tetrameric state. The main stabilizing factors are thus probably located to each monomeric unit.

### Comparison of dimeric and tetrameric MDHs

Six tetrameric MDHs have been crystallized and determined by X-ray crystallography; the four MDHs presented here together with MDH from *H. marismorui*<sup>37,38</sup> and from *M. jannaschii*.<sup>27</sup> When comparing tetrameric and dimeric MDHs from prokaryotes, as in the structure-based multiple sequence alignment in Figure 1, possible determinants of tetrameric MDH formation are revealed. The loop in the region 51–55 is larger in all six tetrameric MDHs compared with the dimeric counterparts. The residues in this loop form several interactions with residues in helix  $\alpha$ A and  $\alpha$ H in the tetrameric state, e.g. the salt-bridge Lys22-Glu52 in *ca*-MDH. Furthermore, the aspartic acid residue in position 55 forms a well-defined salt-bridge with arginine/lysine 243 in all the present tetrameric MDHs. In the dimeric MDHs, the amino acid residues in the corresponding position are hydrophobic.

A C $\alpha$  superposition of two dimeric entities from each dimeric MDH onto the full tetrameric *ca*-MDH show that some of the largest differences occur in regions 190–196 and 241–245, which is two loops with several important monomer–monomer interactions in the tetrameric MDHs, e.g. the salt-bridges Asp55-Lys/Arg243 and Asp165-Arg245 in all our four structures. A few surface loops vary between the dimeric and tetrameric MDHs, e.g. around residues 24 and 133, but in general, most loops at the protein surface seems well conserved among all MDHs.

More specifically, the C $\alpha$  superposition reveals that the larger loops in MDH from *A. aquaticus* and *T. flavus* around position 24/25 get too close to their symmetry-equivalent partners to form tetramers. A few extra residues around position 245/246 have similar unfavorable contacts with loop 164–167 in another monomer. Furthermore, the molecular packing around loop 190–195 varies between the tetrameric and dimeric MDHs. Two loops in the *E. coli* MDH structure, around residues 165 and 190, have conformations that would prevent formation of a corresponding tetramer due to steric interference.

### Conclusions

Comparison of tetrameric malate dehydrogenases from two moderate thermophilic and one mesophilic phototropic bacterium, with temperatures for optimal growth in the limited range 32–55 °C, has highlighted a number of structural differences consistent with the current ideas for thermal stabilization. The study shows that thermal stability in these oligomeric MDHs arises from a combination of different mechanisms, such as increased number of salt-bridges, hydrogen bonds and aromatic interactions across the subunit interface, where the influence of each type of molecular interaction varies in the series. Furthermore, a slight increase in packing density as well as a reduction of the number and volumes of cavities within subunits is observed.

The more thermostable *ca*-MDH contains more proline and alanine residues than the others and this is in agreement with earlier studies of various proteins showing that proline and alanine appear to stabilize the protein backbone.<sup>47–50</sup> Compared to the less thermostable MDHs, *ca*-MDH has a larger number of ionic-pair and hydrogen bond interactions across the dimer–dimer interface that contribute to the stabilization of the quaternary structure. Similar stabilization by subunits interactions and higher oligomeric states were observed for other proteins.<sup>22,49,51</sup> In one study,<sup>52</sup> it was demonstrated by site-directed mutagenesis that only a few residues participating in the monomer–monomer interface could have a substantial influence on the temperature-stability of the oligomer. Two clusters at the monomer–monomer interface in the dimeric subunit, involving residues Phe10, His17 and Trp18 appear to contribute to the thermostability of *ca*-MDH. Protein volume calculations suggest that the multimerization process forms more compact tetramers for the most temperature-resistant species.

Comparison of the crystal structures suggests that the differences in thermostability between *cv/hybrid*-MDH on one hand and *ct*-MDH on the other hand, are mainly due to the presence of specific polar residues in *ct*-MDH that form additional hydrogen bonds that stabilize each monomeric subunit.

**Table 4.** Crystal data, data collection and refinement statistics

	<i>ca</i> -MDH	<i>ct</i> -MDH	<i>hybrid</i> -MDH	<i>cv</i> -MDH
<i>Crystal data</i>				
Space group	P3 <sub>1</sub> 21	C222 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub>
Cell parameters				
<i>a</i> (Å)	106.7	114.3	83.9	64.4
<i>b</i> (Å)	106.7	149.4	117.4	85.8
<i>c</i> (Å)	104.0	97.7	125.3	117.5
$\beta$ (deg.)				104.6
Content of the asymmetric unit	Dimer	Dimer	Tetramer	Tetramer
<i>Data collection</i>				
X-ray source	DESY	MAX-II	DESY	ESRF
Beamline	BW7B	I711	BW7A	ID14-1
Wavelength	0.956	0.9831	0.992	0.934
Temperature (K)	100	100	100	100
Resolution range (Å)	25–2.20	40.8–2.50	30.0–2.00	40–2.50
No. total reflections	80,712	117,908	502,001	388,140
No. unique reflections	35,062	28,459	82,230	37,824
No. reflections in working set	33,979	25,636	78,116	34,009
No. reflections in test set	1083	2823	4114	3815
Outer shell resolution range (Å)	2.33–2.20	2.64–2.50	2.07–2.00	2.59–2.50
$R_{\text{merge}}$ (%)	12.4 (28.4)	7.0 (25.4)	12.5 (29.7)	7.9 (25.1)
Completeness (%)	99.7 (87.1)	100 (100)	97.7 (85.6)	92.5 (84.1)
<i>Refinement statistics</i>				
Model content of the a.u.				
No. amino acid residues	596	596	1201	1192
No. protein non-H atoms	4418	4449	9022	8949
No. solvent water molecules	316	234	869	452
No. NAD-cofactor atoms	88	88	176	0
No. heterogen atoms (Cd)	5	–	–	–
$R$ -factor (%)	17.4	22.1	24.2	21.6
$R_{\text{free}}$	19.7	27.5	29.2	30.5
Average protein $B$ -factor (Å <sup>2</sup> )	19.9	31.4	28.4	26.6
<i>rmsd</i> from ideal values <sup>a</sup>				
Bond lengths (Å)	0.005	0.007	0.006	0.007
Bond angles (deg.)	1.13	1.15	1.19	1.26
Dihedrals (deg.)	21.3	21.9	21.8	22.6
Ramachandran plot <sup>b</sup>				
In most favored regions (%)	90.7	90.0	90.5	86.5
In other allowed regions (%)	8.9	9.6	8.8	13.1
In disallowed regions (%)	0.4	0.4	0.7	0.4
Mean coordinate error (Å)	0.21	0.33	0.31	0.33

<sup>a</sup> CNS 1.0.<sup>60</sup><sup>b</sup> PROCHECK.<sup>64</sup>

Comparison of the crystal structures of the present four MDHs with that of all other known prokaryotic MDH structures suggests that a few extra residues in each of three different loop regions have a major impact on the monomer multimerization process. In all tetrameric MDHs examined, analysis of various interactions in the tetramers demonstrates that residues at the subunit interfaces play an important role for the maintenance of the tetrameric state.

## Experimental

### Protein production, purification and characterization

The genes for MDH from *C. aurantiacus*, *C. tepidum*, *C. vibrioforme* and a *hybrid*-MDH was cloned, overexpressed in *E. coli* and purified as described.<sup>2,3</sup> After elution from a Blue Sepharose CL-6B column (Pharmacia, Uppsala, Sweden),

each protein was more than 99% pure, as judged by SDS-PAGE. The elution buffer contained 5 mM  $\text{NAD}^+$ , 10 mM malate, 1 mM dithiothreitol, 20 mM potassium phosphate at pH 7.5. The proteins were concentrated to 10–20 mg/ml, with a simultaneous change of phosphate buffer to Milli-Q water or 50 mM Tris-HCl buffer (pH 7.4) and stored frozen. The enzyme activity and thermostability measurements were performed as described,<sup>3</sup> at the temperatures indicated and after removal of substrate and cofactor by dialysis.

### Melting temperature measurements

Thermal unfolding of the enzymes were followed by circular dichroism (CD) spectroscopy using a Jasco J-810 spectropolarimeter calibrated with ammonium D-camphor-10-sulfonate. Measurements were done at 23 °C using a quartz cuvette with a path-length of 0.1 cm. All the measurements were done with a protein concentration of

0.10–0.15 mg/ml in 10 mM potassium phosphate buffer (pH 7.4). Samples were scanned five times at 50 nm/minute with a bandwidth of 1 nm and a response time of one second, over the wavelength range 190–240 nm. The data were averaged and the spectrum of a protein-free control sample was subtracted. All measurements were done at least twice. Thermal denaturation curves were determined by monitoring the change in the CD value at 222 nm with a temperature slope of 1 deg C/min. The enzymes unfold irreversibly in a cooperative process, and apparent melting temperatures ( $t_m$ ) were determined from the transition midpoint in each case. There is no significant difference in  $t_m$  values whether or not NAD is present.

### Crystallization

Crystals of all four MDHs were obtained by the hanging-drop, vapor-diffusion method. Crystal data are listed in Table 4. For *ca*-MDH, two crystal forms were obtained under different conditions. For the trigonal crystal form, the precipitant solution was 5–15% polyethylene glycol 400 (PEG400), 100 mM sodium acetate (pH 4.6), 40 mM cadmium acetate. The reservoir solution was 15% PEG400, 100 mM sodium acetate (pH 4.6). For the orthorhombic form, the reservoir and precipitant solution contained 10% PEG2000 in 2 M NaCl without additional buffer. The latter form has not been subjected to further investigation but was used in the initial multiple crystal averaging to improve the maps. The protein concentration of *ca*-MDH was 10 mg/ml.

For *ct*-MDH, crystallization conditions were screened with CRYSTOOL.<sup>53</sup> Plate-shaped crystals of the enzyme were grown at 15 °C. The protein solution (~20 mg/ml, 50 mM Tris-HCl, pH 7.4) was equilibrated against a reservoir containing 53% (v/v) methyl-2,4-pentenediol (MPD) and 100 mM Hepes (pH 7.5). For *cv*-MDH, small and thin plates were obtained by equilibrating the protein solution (~20 mg/ml, 50 mM Tris-HCl, pH 7.4) against a reservoir with 40% PEG-MME5000, 100 mM sodium succinate buffer (pH 6.0). For the hybrid form of MDH (10 mg/ml), crystals were obtained in drops with 0.15 M sodium acetate, 75 mM Tris-HCl (pH 8.5) with 22.5% PEG 4000 as a precipitant.

### Data collection

Both in-house X-ray sources and synchrotron radiation have been used in the current project. Final data for *ca*-MDH were collected at BW7B, EMBL, Hamburg. Data for *ct*-MDH were collected at beam-line 711 at the MAX-II Synchrotron in Lund, Sweden. For *cv*-MDH, the ID14-1 station at ESRF in Grenoble, France, was used for the diffraction experiment. Data for the *hybrid*-MDH were collected at BW7A, EMBL, Hamburg. Diffraction images for the various datasets were processed with either DENZO<sup>54</sup> and scaled and merged with

Scalepack<sup>55</sup> or using the MOSFLM package<sup>56</sup> with subsequent scaling and merging within the CCP4 software suite.<sup>57</sup> Further details on the data collection are summarized in Table 4.

### Structure solution

The structure of *ca*-MDH was determined with the molecular replacement method using the program AMoRe.<sup>58</sup> A homology model was used that was generated automatically by the SwissModel server, whereby the *ca*-MDH sequence was homology modeled onto the coordinates of *Bacillus stearothermophilus* LDH (PDBcode: 1ldb). The first residue, Lys4 of the *ca*-MDH homology model corresponds to Arg8 of LDH, and the last residue, Leu303 of the *ca*-MDH homology model corresponds to residue Leu311 of LDH. A solution with two positions for the monomer was found in the trigonal form. The structure of *hybrid*-MDH was determined with molecular replacement using AMoRe with a polyalanine model of *ca*-MDH as a search model. The asymmetric unit contains four crystallographically independent monomers (Table 4). The *ct*-MDH structure was solved with molecular replacement using the program EPMR<sup>59</sup> with one monomer from the *hybrid*-MDH as a search model. The model included side-chains, since less than 10% of the side-chains vary between the three molecules *ct*-MDH, *cv*-MDH and *hybrid*-MDH. The positions of both peptide chains in the *ct*-MDH dimer were found (correlation coefficient = 0.624,  $R = 0.396$  for both molecules). For *cv*-MDH, the full *hybrid*-MDH tetramer including side-chains was used in the molecular replacement module of CNS 1.0.<sup>60</sup>

### Refinement

Refinement of all four structures was performed with CNS 1.0<sup>60</sup> with the maximum likelihood function as the target function. In each case, the weight between the X-ray terms and the energy terms was optimized by the program. Initial rigid-body refinements of the various solutions from MR were followed by simulated annealing protocols. Except for the *cv*-MDH, all NAD cofactors could be placed in well-defined electron density calculated with amplitudes  $F_o - F_c$  after a simulated annealing refinement step. To begin with, tight non-crystallographic symmetry (NCS) restraints between the crystallographically independent monomeric protein-chains were applied. All NCS restraints were removed in the final rounds of model refinement.

The models were adjusted in the program O<sup>61</sup> using both  $F_o - F_c$ ,  $2F_o - F_c$  and, to a lesser extent, composite-omit maps (omitting 10%). Electron density maps were calculated within CNS and transformed to O format with MAPMAN.<sup>62</sup>

Solvent water molecules were identified with the water-picking procedure of CNS over several iterations. Only those water molecules with

well-defined electron density and reasonable hydrogen bond geometries have been retained in the final models. In our comparative analysis, we have not included water-mediated contacts across dimer interfaces since the *cv*-MDH and *ct*-MDH structures were determined to 2.5 Å resolution only.

For all four proteins, some amino acid side-chains on the surface of the tetramers have weak or no density at all, especially some Lys residues. The side-chains have, nevertheless, been retained in the final models. In general, all amino acid residues in the dimer-dimer interface relevant for the present discussion have well-defined density with no sign of alternative conformations. The only exception to this is Ser227 in *ct*-MDH, which has weak density.

As is observed in other MDH structures without bound substrate or inhibitor, the flexible loop around residues 80–90 is difficult to model.<sup>33</sup> In *cv*-MDH and the *hybrid*-MDH, this loop is recognizable, and included in the model for two of the four peptide-chains. In *ct*-MDH the loop is well defined for one monomer only. For all four enzymes, the last four or five residues at the C terminus have poor density and all models terminate at residue 305, except in *cv*-MDH, where two of the four crystallographically independent chains terminate at 299 and 302. Residue 120 is a *cis*-Pro in all structures.

### Volume calculations

Molecular protein volumes and cavity calculations were performed with the program VOIDOO.<sup>63</sup> The molecular volumes were calculated using a probe with radius 0 Å in order to get the protein volume *per se*. The cavity volumes were calculated using a probe with 1.4 Å radius, comparable to the size of a water molecule, and the volumes were defined as the probe-occupied regions, thus resembling Connolly-type of surfaces. Volumes were refined until shifts in volumes were less than 0.1 Å<sup>3</sup> or 0.1%. The flexible loop region 80–90, the flexible C termini 300–305/306, the NAD cofactor and solvent water molecules were all omitted from the calculations in order to get comparable values for all four MDHs. Accessible surfaces were calculated using SURFACE.<sup>57</sup>

### PDB Data Bank accession codes

The atomic coordinates and structure factors have been deposited in the PDB as entries 1guy (*ca*-MDH), 1guz (*hybrid*-MDH), 1gv0 (*ct*-MDH) and 1gv1 (*cv*-MDH).

### Acknowledgments

We thank the staff at I711 (Max-II, Lund, Sweden), ID14-1 (ESRF, Grenoble, France) and BW7A/B (DESY,

EMBL-Hamburg, Germany) for help with data-collection and Norwegian Research Council (to B.D, A.B. and R.S.) and Swedish Natural Science Council (to H.E.) for support. Thanks to Dimitrios Mantzilas (Department of Biochemistry, University of Oslo) for help with  $t_m$  measurements.

### References

1. Snaidr, J., Amann, R., Huber, I., Ludwig, W. & Schleifer, K. H. (1997). Phylogenetic analysis and *in situ* identification of bacteria in activated sludge. *Appl. Environ. Microbiol.* **63**, 2884–2896.
2. Synstad, B., Emmerhoff, O. & Sirevåg, R. (1996). Malate dehydrogenase from the green gliding bacterium *Chloroflexus aurantiacus* is phylogenetically related to lactic dehydrogenases. *Arch. Microbiol.* **165**, 346–353.
3. Naterstad, K., Lauvrak, V. & Sirevåg, R. (1996). Malate dehydrogenase from the mesophile *Chlorobium vibrioforme* and from the mild thermophile *Chlorobium tepidum*: molecular cloning, construction of a hybrid, and expression in *Escherichia coli*. *J. Bacteriol.* **178**, 7047–7052.
4. Sundaram, T. K., Wright, I. P. & Wilkinson, A. E. (1980). Malate dehydrogenase from thermophilic and mesophilic bacteria. Molecular size, subunit structure, amino acid composition, immunochemical homology, and catalytic activity. *Biochemistry*, **19**, 2017–2022.
5. Smith, K., Sundaram, T. K. & Kernick, M. (1984). Malate dehydrogenases from actinomycetes: structural comparison of *Thermoactinomyces* enzyme with other actinomycete and *Bacillus* enzymes. *J. Bacteriol.* **157**, 684–687.
6. Tayeh, M. A. & Madigan, M. T. (1987). Malate dehydrogenase in phototrophic purple bacteria: purification, molecular weight, and quaternary structure. *J. Bacteriol.* **169**, 4196–4202.
7. Rolstad, A. K., Howland, E. & Sirevåg, R. (1988). Malate dehydrogenase from the thermophilic green bacterium *Chloroflexus aurantiacus*: purification, molecular weight, amino acid composition, and partial amino acid sequence. *J. Bacteriol.* **170**, 2947–2953.
8. Breiter, D. R., Resnik, E. & Banaszak, L. J. (1994). Engineering the quaternary structure of an enzyme: construction and analysis of a monomeric form of malate dehydrogenase from *Escherichia coli*. *Protein Sci.* **3**, 2023–2032.
9. Vieille, C. & Zeikus, G. J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1–43.
10. Wynne, S. A., Nicholls, D. J., Scawen, M. D. & Sundaram, T. K. (1996). Tetrameric malate dehydrogenase from a thermophilic *Bacillus*: cloning, sequence and overexpression of the gene encoding the enzyme and isolation and characterization of the recombinant enzyme. *Biochem. J.* **317**, 235–245.
11. Langelandsvik, A. S., Steen, I. H., Birkeland, N. K. & Lien, T. (1997). Properties and primary structure of a thermostable L-malate dehydrogenase from *Archaeoglobus fulgidus*. *Arch. Microbiol.* **168**, 59–67.
12. Madern, D., Ebel, C., Mevarech, M., Richard, S. B., Pfister, C. & Zaccai, G. (2000). Insights into the molecular relationships between malate and lactate dehydrogenases: structural and biochemical properties of monomeric and dimeric intermediates of a

- mutant of tetrameric L-[LDH-like] malate dehydrogenase from the halophilic archaeon *Haloarcula marismortui*. *Biochemistry*, **39**, 1001–1010.
13. Jaenicke, R. (1996). Stability and folding of ultra-stable proteins: eye lens crystallins and enzymes from thermophiles. *FASEB J.* **10**, 84–92.
  14. Eriksson, A. E., Baase, W. A., Wozniak, J. A. & Matthews, B. W. (1992). A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature*, **355**, 371–373.
  15. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992). Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science*, **255**, 178–183.
  16. Takano, K., Ogasahara, K., Kaneda, H., Yamagata, Y., Fujii, S., Kanaya, E. *et al.* (1995). Contribution of hydrophobic residues to the stability of human lysozyme: calorimetric studies and X-ray structural analysis of the five isoleucine to valine mutants. *J. Mol. Biol.* **254**, 62–76.
  17. Serrano, L., Day, A. G. & Fersht, A. R. (1993). Step-wise mutation of barnase to binase. A procedure for engineering increased stability of proteins and an experimental analysis of the evolution of protein stability. *J. Mol. Biol.* **233**, 305–312.
  18. Vogt, G., Woell, S. & Argos, P. (1997). Protein thermal stability, hydrogen bonds, and ion pairs. *J. Mol. Biol.* **269**, 631–643.
  19. Pappenberger, G., Schurig, H. & Jaenicke, R. (1997). Disruption of an ionic network leads to accelerated thermal denaturation of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *J. Mol. Biol.* **274**, 676–683.
  20. Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T. *et al.* (1998). Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* **255**, 336–346.
  21. Chang, C., Park, B. C., Lee, D. S. & Suh, S. W. (1999). Crystal structures of thermostable xylose isomerases from *Thermus caldophilus* and *Thermus thermophilus*: possible structural determinants of thermostability. *J. Mol. Biol.* **288**, 623–634.
  22. Korkhin, Y., Kalb, A. J., Peretz, M., Bogin, O., Burstein, Y. & Frolow, F. (1999). Oligomeric integrity—the structural key to thermal stability in bacterial alcohol dehydrogenases. *Protein Sci.* **8**, 1241–1249.
  23. Britton, K. L., Yip, K. S., Sedelnikova, S. E., Stillman, T. J., Adams, M. W., Ma, K. *et al.* (1999). Structure determination of the glutamate dehydrogenase from the hyperthermophile *Thermococcus litoralis* and its comparison with that from *Pyrococcus furiosus*. *J. Mol. Biol.* **293**, 1121–1132.
  24. Tanner, J. J., Hecht, R. M. & Krause, K. L. (1996). Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5 angstrom resolution. *Biochemistry*, **35**, 2597–2609.
  25. Auerbach, G., Ostendorp, R., Prade, L., Korndorfer, I., Dams, T., Huber, R. & Jaenicke, R. (1998). Lactate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: the crystal structure at 2.1 Å resolution reveals strategies for intrinsic protein stabilization. *Structure*, **6**, 769–781.
  26. Wallon, G., Kryger, G., Lovett, S. T., Oshima, T., Ringe, D. & Petsko, G. A. (1997). Crystal structures of *Escherichia coli* and *Salmonella typhidurium* 3-isopropylmalate dehydrogenase and comparison with their thermophilic counterpart from *Thermus thermophilus*. *J. Mol. Biol.* **266**, 1016–1031.
  27. Lee, B. I., Chang, C., Cho, S.-J., Eom, S. H., Kim, K. K., Yu, Y. G. & Suh, S. W. (2001). Crystal structure of the MJ0490 gene product of the hyperthermophilic Archaeobacterium *Methanococcus jannaschii*, a novel member of the lactate/malate family of dehydrogenases. *J. Mol. Biol.* **307**, 1351–1362.
  28. Kelly, C. A., Nishiyama, M., Ohnishi, Y., Beppu, T. & Birktoft, J. J. (1993). Determinants of protein thermostability observed in the 1.9 Å crystal structure of malate dehydrogenase from the thermophilic bacterium *Thermus flavus*. *Biochemistry*, **32**, 3913–3922.
  29. Roderick, S. L. & Banaszak, L. J. (1986). The three-dimensional structure of porcine heart mitochondrial malate dehydrogenase at 3.0 Å resolution. *J. Biol. Chem.* **261**, 9461–9464.
  30. Gleason, W. B., Fu, Z., Birktoft, J. & Banaszak, L. (1994). Refined crystal structure of mitochondrial malate dehydrogenase from porcine heart and the consensus structure for dicarboxylic acid oxidoreductases. *Biochemistry*, **33**, 2078–2088.
  31. Birktoft, J. J., Bradshaw, R. A. & Banaszak, L. J. (1987). Structure of porcine heart cytoplasmic malate dehydrogenase: combining X-ray diffraction and chemical sequence data in structural studies. *Biochemistry*, **26**, 2722–2734.
  32. Chapman, A. D., Cortes, A., Dafforn, T. R., Clarke, A. R. & Brady, R. L. (1999). Structural basis of substrate specificity in malate dehydrogenases: crystal structure of a ternary complex of porcine cytoplasmic malate dehydrogenase, alpha-ketomalonate and tetrahydroNAD. *J. Mol. Biol.* **285**, 703–712.
  33. Hall, M. D., Levitt, D. G. & Banaszak, L. J. (1992). Crystal structure of *Escherichia coli* malate dehydrogenase. A complex of the apoenzyme and citrate at 1.87 Å resolution. *J. Mol. Biol.* **226**, 867–882.
  34. Kim, S. Y., Hwang, K. Y., Kim, S. H., Sung, H. C., Han, Y. S. & Cho, Y. (1999). Structural basis for cold adaptation. Sequence, biochemical properties, and crystal structure of malate dehydrogenase from a psychrophile *Aquaspirillum arcticum*. *J. Biol. Chem.* **274**, 11761–11767.
  35. Johansson, K., Ramaswamy, S., Saarinen, M., Lemaire-Chamley, M., Issakidis-Bourguet, E., Miginiac-Maslow, M. & Eklund, H. (1999). Structural basis for light activation of a chloroplast enzyme: the structure of sorghum NADP-malate dehydrogenase in its oxidized form. *Biochemistry*, **38**, 4319–4326.
  36. Carr, P. D., Verger, D., Ashton, A. R. & Ollis, D. L. (1999). Chloroplast NADP-malate dehydrogenase: structural basis of light-dependent regulation of activity by thiol oxidation and reduction. *Structure*, **7**, 461–475.
  37. Dym, O., Mevarech, M. & Sussman, J. L. (1995). Structural features that stabilize halophilic malate dehydrogenase from an archaeobacterium. *Science*, **267**, 1344–1346.
  38. Richard, S. B., Madern, D., Garcin, E. & Zaccari, G. (2000). Halophilic adaptation: novel solvent protein interactions observed in the 2.9 and 2.6 Å resolution structures of the wild type and a mutant of malate dehydrogenase from *Haloarcula marismortui*. *Biochemistry*, **39**, 992–1000.
  39. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to

- check the stereochemistry quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
40. Rossmann, M. G., Liljas, A., Brändén, C.-I. & Banaszak, L. J. (1975). Evolution and structural relationships among dehydrogenases. In *The Enzymes* (Boyer, P. D., ed.), vol. 11, pp. 61–102, Academic Press, New York.
  41. Rolstad, A., Howland, E. & Sirevåg, R. (1988). Malate dehydrogenase from the thermophilic green bacterium *Chloroflexus aurantiacus*: purification, molecular weight, amino acid composition, and partial amino acid sequence. *Bacteriology*, **170**, 2947–2953.
  42. Charnock, C., Refseth, U. H. & Sirevåg, R. (1992). Malate dehydrogenase from *Chlorobium vibrioforme*, *Chlorobium tepidum*, and *Heliobacterium gestii*: purification, characterization, and investigation of dinucleotide binding by dehydrogenases by use of empirical methods of protein sequence analysis. *J. Bacteriol.* **174**, 1307–1313.
  43. Wilks, H. M., Hart, K. W., Feeney, R., Dunn, C. R., Muirhead, H., Chia, W. N. *et al.* (1988). A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework. *Science*, **242**, 1541–1544.
  44. Holbrook, J. J., Liljas, A., Steindel, S. J. & Rossmann, M. G. (1975). Lactate dehydrogenase. In *The Enzymes* (Boyer, P. D., ed.), 3rd edit., vol. 11, pp. 191–292, Academic Press, New York.
  45. Kannan, N. & Vishveshwara, S. (2000). Aromatic clusters: a determinant of thermal stability of thermophilic proteins. *Protein Eng.* **13**, 753–761.
  46. Serrano, L., Neira, J. L., Sancho, J. & Fersht, A. R. (1992). Effect of alanine versus glycine in alpha-helices on protein stability. *Nature*, **356**, 453–455.
  47. Matthews, B. W., Nicholson, H. & Becktel, W. J. (1987). Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc. Natl Acad. Sci. USA*, **84**, 6663–6667.
  48. Nicholson, H., Tronrud, D. E., Becktel, W. J. & Matthews, B. W. (1992). Analysis of the effectiveness of proline substitutions and glycine replacements in increasing the stability of phage T4 lysozyme. *Biopolymers*, **32**, 1431–1441.
  49. Li, C., Heatwole, J., Soelaiman, S. & Shoham, M. (1999). Crystal structure of a thermophilic alcohol dehydrogenase substrate complex suggests determinants of substrate specificity and thermostability. *Proteins: Struct. Funct. Genet.* **37**, 619–627.
  50. Lo Leggio, L., Kalogiannis, S., Bhat, M. K. & Pickersgill, R. W. (1999). High resolution structure and sequence of *T. aurantiacus* xylanase I: implications for the evolution of thermostability in family 10 xylanases and enzymes with (beta)alpha-barrel architecture. *Proteins: Struct. Funct. Genet.* **36**, 295–306.
  51. Dams, T., Auerbach, G., Bader, G., Jacob, U., Ploom, T., Huber, R. & Jaenicke, R. (2000). The crystal structure of dihydrofolate reductase from *Thermotoga maritima*: molecular features of thermostability. *J. Mol. Biol.* **297**, 659–672.
  52. Kirino, H., Aoki, M., Aoshima, M., Hayashi, Y., Ohba, M., Yamagishi, A. *et al.* (1994). Hydrophobic interaction at the subunit interface contributes to the thermostability of 3-isopropylmalate dehydrogenase from an extreme thermophile, *Thermus thermophilus*. *Eur. J. Biochem.* **220**, 275–281.
  53. Segelke, B. (1995). Efficiency analysis of sampling protocols used in protein crystallization screening and crystal structure from two novel crystal forms of PLA2. PhD thesis, University of California, San Diego
  54. Otwinowski, Z. (1993). Data collection and processing. In *Proceedings of the CCP4 Study Weekend* (Sawyer, L., ed.), pp. 56–62, Daresbury Laboratory, Warrington, UK.
  55. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray data collected in oscillation mode. In *Methods in Enzymology, Macromolecular Crystallography, Part A* (Carter, C. W. Jr. & Sweet, R. M., eds), vol. 276, pp. 307–326, Academic Press, New York.
  56. Leslie, A. G. W. (1992). Mosfilm. *Joint CCP4 and ESF-EACMB Newsletter Protein Crystallography*, Daresbury Laboratory, Warrington, UK.
  57. Collaborative Computational Project, Number 4 (1994). The CCP4 Suite: programs for protein crystallography. *Acta Crystallog. Sect. D*, **50**, 760–763.
  58. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog.* **50**, 157–163.
  59. Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). Rapid automated molecular replacement by evolutionary search. *Acta Crystallog. Sect. D*, **55**, 484–491.
  60. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. Sect. D*, **54**, 905–921.
  61. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog.* **47**, 110–119.
  62. Kleywegt, G. J. & Jones, T. A. (1996). xdlMAPMAN and xdlDATAMAN. *Acta Crystallog. Sect. D*, **D52**, 826–828.
  63. Kleywegt, G. J. & Jones, T. A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallog. Sect. D*, **50**, 178–185.
  64. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemistry quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
  65. Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
  66. Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 132–134.
  67. Merritt, E. A. (1997). Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524.

Edited by R. Huber

(Received 4 October 2001; received in revised form 30 January 2002; accepted 31 January 2002)