J Mol Evol (2002) 54:825–840
DOI: 10.1007/s00239-001-0088-8

JOURNAL OF MOLECULAR EVOLUTION

Molecular Evolution Within the L-Malate and L-Lactate Dehydrogenase Super-Family

Dominique Madern

Laboratoire de Biophysique Moléculaire, Institut de Biologie Structurale CEA-CNRS-UJF, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France

Received: 2 October 2001 / Accepted: 17 December 2001

Abstract. The NAD(P)-dependent malate (L-MalDH) and NAD-dependent lactate (L-LDH) form a large super-family that has been characterized in organisms belonging to the three domains of life. In the first part of this study, the group of [LDH-like] L-MalDH, which are malate dehydrogenases resembling lactate dehydrogenase, were analyzed and clearly defined with respect to the other enzymes. In the second part, the phylogenetic relationships of the whole super-family were presented by taking into account the [LDH-like] L-MalDH. The inferred tree unambiguously shows that two ancestral genes duplications, and not one as generally thought, are needed to explain both the distribution into two enzymatic functions and the observation of three main groups within the super-family: L-LDH, [LDH-like] L-MalDH, and dimeric L-MalDH. In addition, various cases of functional changes within each group were observed and analyzed. The direction of evolution was found to always be polarized: from enzymes with a high stringency of substrate recognition to enzymes with a broad substrate specificity. A specific phyletic distribution of the L-LDH, [LDH-like] L-MalDH, and dimeric L-MalDH over the Archaeal, Bacterial, and Eukaryal domains was observed. This was analyzed in the light of biochemical, structural, and genomic data available for the L-LDH, [LDH-like] L-MalDH, and dimeric L-MalDH. This analysis led to the elaboration of a refined evolutionary scenario of the super-family, in which the selection of L-LDH and the fate of L-MalDH during mitochrondrial genesis are presented.

Key words: Phylogeny — Molecular evolution — Malate dehydrogenase — Lactate dehydrogenase — [LDH-like] L-MalDH

Introduction

In living organisms, the key role of central metabolism with its various pathways, in providing energy and biochemical precursors has been known for a long time. With a view to understanding the evolutionary relationships between the various pathways, it is now possible to establish whether or not a given transformation is acheived by homologous or analogous enzymes encoded in the genomes of a wide range of organisms (Galparin and Koonin 1999; Huynen et al. 1999). The goal of this paper is to refine the understanding, at the evolutionary level, of two enzymes acting in metabolism: the L-malate and L-lactate dehydrogenases. Its content is focused on the homologous NAD(P)H-dependent L-malate dehydrogenase (L-MalDH)[1] and NADH-dependent L-lactate

Email: dominique.madern@ibs.fr

[1] *Abbreviations:* L-LDH: L(+)-lactate: NAD^+ oxidoreductase. D-LDH: D(−)-lactate: NAD^+ oxidoreductase. L-MalDH: L(+)-malate: NAD^+ oxidoreductase. GAPDH: Glycerate 3-phosphate dehydrogenase. GAP-FdOR: Glycerate 3-phosphate ferredoxin oxidoreductase. PGK: 3-phosphoglycerate kinase. PGM: Phosphoglycerate mutase. Enolase: 2-phosphoglycerate dehydratase. PK: Pyruvate kinase. TCA: Tricarboxylic acid cycle; HacDH, hydroxyacid dehydrogenase

dehydrogenase (L-LDH). Non-homologous enzymes, such as NADH-dependent D-LDH, flavocytochrome *b2* LDH, malate:quinone oxidoreductase, etc., which act as analogs of L-LDH or L-MalDH, are not mentioned here.

NAD(P)H-dependent L-MalDH, NADH-dependent L-LDH, and alcohol dehydrogenase (ADH), are members of a large family of dehydrogenases (Banaszak and Bradshaw 1975; Birkoft et al. 1982). L-MalDH catalyse the NAD(P)H-dependent interconversion of malate and oxaloacetate. They have been purified, characterized, and sequenced from a wide variety of organisms representative of the Bacterial, Archaeal, and Eukaryal domains of life (for review see Goward and Nicholls 1994). L-MalDH are multimeric enzymes with subunit molecular masses of 30–40 KDa. Dimeric and tetrameric oligomeric states of L-MalDH have been identified in solution studies (Sundaram et al. 1980; Bonneté et al. 1993) and by X-ray cristallography (Chapman et al. 1999; Richard et al. 2000; Lee et al. 2001). L-MalDH enzymatic activity is associated with both oligomeric states (Madern et al. 2001 and references therein). L-LDH are also well-studied enzymes (for reviews see Garvie 1980; Jaenicke 1998). They are tetrameric enzymes, which recognize pyruvate and lactate as substrates. However, in contrast to L-MalDH, no L-LDH has been described to use NADPH as co-enzyme, and NADH-dependent L-LDH has been found and purified only from Bacteria and Eukarya. Work discussing LDH enzymatic activity in Archaea relates to D-lactate dehydrogenase (Hecht et al. 1990). Their sequences and mechanisms show that they are very different from the homologous L-LDH/L-MalDH super-family analyzed in this work (Taguschi and Ohta 1991). In a few bacterial species, L-LDH is activated by fructose 1,6-bisphosphate. Important active site residues involved in catalysis have been identified for both L-LDH and L-MalDH (Goward and Nicholls 1994). The substrate specificity of both L-LDH and L-MalDH has been altereted by mutagenesis of a single residue at position 102 (Wilks et al. 1988; Cendrin et al. 1993; Boernke et al. 1995). L-MalDH have a positively charged residue (Arg) that forms hydrogen bonds with the second carboxyl group of the oxaloacetate. In L-LDH, an uncharged residues (generally a Gln) was found at this position. Crystallographic studies have shown that a charge imbalance inside the catalytic vacuole is responsible for discrimination between their respective substrates (Wigley et al. 1992; Chapman et al. 1999).

Using structural and biochemical data, some authors have suggested that L-malate and L-lactate dehydrogenases derived from a common ancestor (Birkoft et al. 1982; McAlister-Henn 1988). Phylogenetic analyses have been previously performed on each of the L-LDH (Griffin et al. 1992; Tsuji et al. 1994) and L-MalDH families, individually. In the latter case, only the two clades (cytosolic and mitochondrial) of dimeric enzymes have been analyzed (Roger et al. 1999), whereas evidence for a third, lactate dehydrogenase-like [LDH-like] group of L-MalDH exists (Cendrin et al. 1993; Synstad et al. 1996; Naterstad et al. 1996; Wynne et al. 1996; Langelandsvik et al. 1997; Madern 2000). Combined molecular phylogenetic analyses of L-LDH and L-MalDH have been attempted (Iwabe et al. 1989; Golding and Dean 1998) but their conclusions remain controversial because only a limited set of sequences was used and/or because of the inclusion of incorrectly assigned sequences.

Before the present phylogenetic study, a series of experiments was performed in order to describe unambiguously the correct function, oligomeric state, and co-enzyme preference of some genomic ORF used in this analysis, and to analyze the folding pathway of [LDH-like] L-MalDH (Madern 2000; Madern et al. 2000; Madern et al. 2001).

A set of 107 NAD-dependent L-LDH, NAD(P)-dependent L-MalDH, and three homologous hydroxy-acid dehydrogenase sequences were used to perform a structure-based alignment, which was subjected to various phylogenetic analyses. The inferred phylogenetic trees, analysed in the light of recent biochemical and structural data, allowed a new and more accurate description of the evolution within the L-LDH and L-MalDH super-family.

Materials and Methods

Choice of Sequences

The choice of correctly assigned homologous L-MalDH or L-LDH sequences, which were extracted from SWISSPROT databank and ongoing genome sequence, was based on their sequence similarities with enzymes unambiguously characterized at both the biochemical and structural levels. The L-LDH from *Bacillus stearothermophilus* and two L-MalDHs from *Escherichia coli* and *Haloarcula marismortui* were taken as enzymes of reference. To extract sequence information by similarity searches, appropriate programs from BLAST Version 2.0 or WIT2, with the default settings were used. Data were obtained from the Institute for Genomic Research website (http://www.tigr.org) or at the GOLD database (http://geta.life.uiuc.edu) giving progress reports and access to the 250 genomes currently being sequenced, which is maintained by Dr N. Kyrpides. The genome site of the national center for biotechnology information (http//www.ncbi.nlm.nih.gov) was also used. The sequences and accession numbers are listed in Table 1.

Sequence Alignment

The selected sequences were aligned by using the CLUSTAL W software package (Thompson et al. 1994) and hand adjusted using structural information for L-LDH (Wigley et al 1992; Auerbach et al. 1997), tetrameric L-MalDH (Richard et al 2000; Lee et al. 2001), and dimeric L-MalDH (Carr et al. 1999; Sun-Yong et al. 1999; Hall et al. 1992). This information allowed to adjust more precisely the regions of low sequence similarity. Sequences of N and C terminal extensions were

Table 1. List of sequences and accession numbers

Assignment	Organisms	Accession numbers	Assignment	Organisms	Accession numbers[a]
LDH A	*H. vulgare*	P29988	MaIDH	*B. israeli*	CAA62129
LDH A	*X. laevis*	P42120	MaIDH	*R. prowaekii*	CAA14835
LDH B	*F. heteroclitus*	P20373	MaIDH	*G. sulfureducens*	ct1282
LDH B	*H. vulgare*	P22989	MaIDH	*P. putida*	ct0763*
LDH B	*X. laevis*	P42119	MaIDH	*D. ethanogenes*	ct6143*
LDH C	*F. heteroclitus*	Q06176	MaIDH	*R. leguminosarum*	CAA05717
LDH H	*A. platyrhychos*	P13743	MaIDH	*C. crecentus*	AAK25617
LDH H	*G. galus*	P00337	MaIDH	*A. pernix*	E72655
LDH H	*H. sapiens*	P07195	MaIDH	*A. aeolicus*	O67655
LDH H	*S. scrofa*	P00336	MaIDH	*A. aeolicus*	D70444
LDH H	*R. norvegicus*	P42123	MaIDH	*M. jananschii*	Q60176
LDH M	*B. taurus*	P19858	MaIDH	*A. fulgidus*	O08349
LDH M	*G. galus*	P00340	MaIDH	*H.* Sp. NRC1	AAG19885
LDH M	*H. sapiens*	P00338	MaIDH	*H. marismortui*	Q07841
LDH M	*S. scrofa*	P00339	MaIDH	*C. jejuni*	CAB75168
LDH M	*A. platyrhychos*	P13491	MaIDH	*H. sapiens*	P40925
LDH M	*R. norvegicus*	P04642	MaIDH	*S. scrofa*	A32472
LDH M	*S. acanthias*	P00341	MaIDH	*M. musculus*	P14152
LDH	*B. psychrosaccharolyticus*	P20619	MaIDH	*F. bidentis*	P46489
LDH X	*H. sapiens*	P07864	MaIDH	*P. sativum*	CAA52614
LDH	*B. megaterium*	P00345	MaIDH	*M. crystallinum*	CAA45572
LDH	*B. stearothermophilus*	P00344	MaIDH	*Z. mays*	P15719
LDH	*B. subtilus*	E69649	MaIDH	*S. vulgare*	CAA38270
LDH	*B. halodurans*	AP001520	MaIDH	*C. trachomatis*	AAF39479
LDH	*B. longum*	P19869	MaIDH	*D. radiodurans*	E75535
LDH	*T. aquaticus*	P13715	MaIDH	*T. vaginalis*	AAC46986
LDH	*L. casei*	P00343	MaIDH	*G. lamblia*	AAD44473
HacDH	*L. confusus*	P14295	MaIDH	*T. flavus*	P10584
LDH	*E. faecalis*	ORF1381	MaIDH	*A. arcticum*	AAD13225
LDH	*P. acidilactici*	CAA50278	MaIDH	*M. crystallinum*	CAA65384
LDH	*Z. mays*	P29038	MaIDH	*M. tuberculosis*	AAK45536
LDH	*L. esculentus*	CAA71611	MaIDH	*M. leprae*	CAC31472
LDH	*M. hyopneumoniae*	S33362	MaIDH	*Glycine max*	AAC37464
LDH	*M. genitalium*	NP073130	MaIDH	*C. lanatus*	P17783
LDH	*P. marinus*	P33571	MaIDH	*E. gunnii*	CAA55383
LDH	*S. mutans*	P26283	MaIDH	*M. musculus*	NP_032643
LDH	*T. caldophilus*	P06150	MaIDH	*S. scrofa*	AAA31071
LDH	*T. maritima*	P16115	MaIDH	*S. cerevisiae*	NP_012838
LDH	*D. radiodurans*	E75282	MaIDH	*E. coli*	P06994
LDH	*C. acetobutylicum*	ORF41	MaIDH	*H. influenzae*	P44427
LDH	*P falciparum*	Q27743	MaIDH	*P. profondum*	P37226
LDH	*T. gondii*	AAC47443	MaIDH	*S. typhymurium*	P25077
LDH	*C. parvum*	AAG17668	MaIDH	*Vibrio* sp. 5710	P48364
MaIDH	*M. thermoautotrophicum*	AAB84694	MaIDH	*P. multocida*	AAK02634
LDH	*S. solfataricus*	SSO2585	MaIDH	*S. putrefaciens*	Ct 833*
LDH	*T. vaginalis*	AAC72735	MaIDH	*Y. pestis*	AAG21998
MaIDH	*B. braunii*	AAB38970	MaIDH	*T. cruzi*	AAC05706
MaIDH	*R. capsulatus*	ORF2795*	MaIDH	*T. brucei*	AAC7934
MaIDH	*P. aerophilum*	Ct123*	MaIDH	*T. brucei*	AAC27101
MaIDH	*B. stearothermophilus*	Ct576*	MaIDH	*Phytomonas* sp.	AAG01146
MaIDH	*B. halodurans*	BAB06877	MaIDH	*L. Major*	CAB55506
MaIDH	*B. subitlis*	I40383	HacDH	*Phytomonas*	AAG01145
MaIDH	*T. acidophilum*	CAC12081	HacDH	*T. cruzi*	AAF36775
MaIDH	*T. volcanium*	BAB60263			
MaIDH	*C. vibrioforme*	P80038			
MaIDH	*C. tepidum*	P80039			
MaIDH	*C. aurantiacus*	CAA61436			
MaIDH	*Synechocystis* sp.	S75735			

[a]Sequences can be retrieved at http://www.ncbi.nlm.nih.gov.

*Sequences that can be retrieved at http://igweb.integratedgenomics.com/GOLD, http://genome.caltech.edu/pyrobaculum, http://www.tigr.org/, http:rhodoL.uchicago.edu/capsulapedia.

trimmed to yield a homogeneous block containing 110 sequences of 359 characters including gaps.

Numbering of Amino Acid Residues

The numbering of residues of Eventoff et al. (1977) and abbreviations for secondary structure elements were given according to those used for *Squalus acanthias* LDH. This system widely accepted was adopted at a time when the exact sequence of this LDH was not known. Consequently, some residue numbers are not used [21, 82, 104, 300] and in cases where the same number refers to more than one amino acid residue these are distinguished by letters. This numbering allows ready recognition of important active site residues, such as R109, D168, R171, and H195, which are strictly conserved in all L-MalDH and L-LDH and the discriminating residue at position 102 either a R with L-MalDH or a Q with L-LDH (Goward and Nicholls 1994; Eventoff et al. 1977) and [LDH like] L-MalDH (Richard et al. 2000).

Phylogenetic Analysis

Analyses were carried out by using various programs of the Phylogeny Inference Package (PHYLIP) Version 3.57C (Felsenstein 1993). The Dayhoff PAM option was used to generate distance matrices with PROTDIST. Neighbor-joining trees were generated with NEIGHBOR and DRAWGRAM. Branch lengths were obtained from the unbootstrapped distance matrix. Maximum parsimony analysis was done using PROTPARS. To estimate the confidence limits of nodes, 100 bootstrap resamplings were generated with SEQBOOT and the majority rule consensus tree generated by CONSENSE. Quartet maximum-likelihood (ML) analysis was done using TREE-PUZZLE 5.0 (Strimmer and von Hasseler 1997). Estimation of parameters were done using 1000 puzzling quartets, the Muller and Vingron matrix (Muller and Vingron 2000), and gamma distributed rates over eight catagories. For Fig. 2, the calculation of statistics values was done using the L-LDHs rooted with four [LDH-like] L-MalDH, and the [LDH-like] L-MalDHs rooted with four L-LDH. For Fig. 3, the calculation was done using the 40 sequences. The trees were drawn using the program NJ Plot (Perrière and Gouy 1996). In Fig. 1, the NJ-tree was infered from the 359 positions of the alignment, in Figs. 2 and 3, NJ-trees were infered from 241 homologous positions (gaps and some specific positions, analyzed in the main text, were not taken into account by using the weight option in Protdist).

Results

Evidence for Three Groups with Strong Correlations Between Primary Structure and Oligomeric State

A structure-based sequence alignment, available on request, of the 110 L-LDH, L-MalDH, and HacDH sequences was subjected to a first phylogenetic analysis. The inferred, unrooted neighbor-joining tree (Fig. 1) is similar to the one already observed for the super family (Golding and Dean 1998). The L-lactate dehydrogenases are mainly located on the left of the dashed line, whereas the L-malate dehydrogenase are on the right-hand side. The tree is therefore split into two main parts according to enzymatic function. The oligomeric state of the enzymes was then taken into account. The L-LDH are well known to be tetrameric (Jaenicke 1998). According to their oligomeric state, L-MalDH have been characterized as dimeric or tetrameric enzymes (Tayeh and Madigan 1987; Madern et al. 2001 and references therein). Enzymes which were characterized as tetrameric are located on the left side of the continuous line in Fig. 1, whereas the dimeric enzymes are on the right side. The structural information shows that the super family is in fact split into three groups. The first contains the well known, tetrameric L-LDH. The second identifies the dimeric L-MalDH. The intersection between the functional and structural information clearly defines the third group. It contains tetrameric [LDH-like] L-MalDH, L-MalDH sequences that present strong similarities with L-LDH, confirming individual observations (Cendrin et al. 1993; Synstad et al. 1996; Wynne et al. 1996; Langelandsvik et al. 1997; Madern 2000). In this group, the enzyme from *Archaeoglobus fulgidus* is the only dimeric [LDH-like] L-MalDH, an exception that can be explained by a large sequence deletion, not observable in the other dimeric [LDH-like] L-MalDH (Madern et al. 2001).

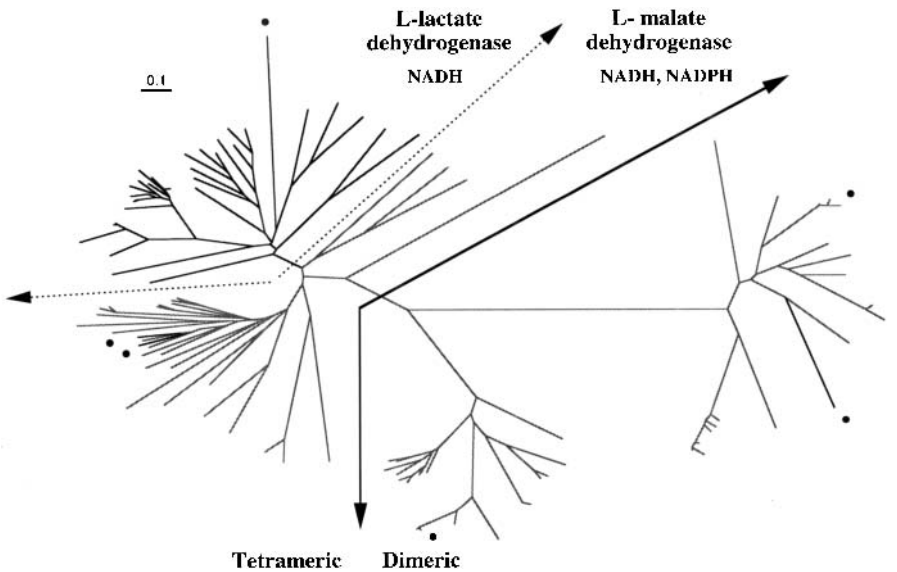

Fig. 1. Unrooted neighbor-joining tree of L-LDH, [LDH-like] L-MalDH, and L-MalDH. The dashed lines split the tree by enzymatic function, and the continuous black lines by oligomeric state. Lactate dehydrogenases are in **bold** and malate dehydrogenases in pale grey. Dots indicate some particular cases discussed in the text. The scale bar represents 0.1 amino acid substitution per site.

Some sequences (seven-cases indicated by squares, arrows, and dots in Fig. 1) present some specific variations with respect to the main characteristics described previously. This will be analyzed further in a specific paragraph: functional changes.

The analysis of the super-family can be refined when the coenzyme requirements for activity are considered. Two archaeal enzymes have been shown to use NADPH instead of NADH (Thompson et al. 1997; Madern 2000; Graupner and White 2000) demonstrating that enzymes which belong to [LDH-like] L-MalDH, can be either NADH or NADPH-dependent, as in the case of dimeric L-MalDH.

A clear discrimination between L-LDH and the [LDH-like] L-MalDH might be established by inspection of two specific positions in their sequences. The first is the Arg residue at conserved position [102] known to provide substrate specificity (Wilks et al. 1988; Cendrin et al. 1993; Boernke et al. 1995). The other residue is a Pro at position 250 conserved in all [LDH-like] L-MalDH, whereas it is most frequently Ile in L-LDH or Ser in dimeric L-MalDH. This position offers a new important target for mutagenesis studies. Up to the present, the signature sequences GXGXXG and GXXGXXG close to the N-terminus were thought to be specific for L-LDH and L-MalDH enzymes, respectively (Wierenga et al. 1986). However, in the 30 [LDH-like] L-MalDH sequences, four have the L-MalDH GXXGXXG sequence, and the 18 others contain the L-LDH GXGXXG sequence. The putative L-LDH from *Toxoplasma gondii,* and the well-characterized L-LDH from *P. falciparum* (Bzik et al. 1993), are strikingly clustered within the [LDH-like] L-MalDH (stars in Fig. 1). Their sequence analysis showed that they can be described as [LDH-like] L-MalDH and not as belonging to the L-LDH. In particular, they have a Lys and a Pro residue at position 102 and 250, respectively, instead of Gln and Ile for L-LDH. The Lys residue having the same charge than the Arg residue found in L-MalDH.

Phylogeny Based on the Common Scaffold

In L-LDH and in all groups of L-MalDH, the mobile catalytic loop (residues between conserved position 95 and 109) contains the substrate discriminating residues at position 102. In contrast to the whole sequence which give evidence for three groups, the analysis of the loop shows that only two kinds of mobile loop exists. They reflect the split by enzymatic function: L-LDH and L-MalDH. As observed with the whole sequence, the catalytic loop of LDH from *T. gondii* and *P. falciparum* are more similar to catalytic loops of L-MalDH, although they present an insertion of five residue at position 108. The region corresponding to the loop was not taken into account in the following phylogenetic analyses. In addition, the six residues corresponding to a structural loop which controls the association of dimers to form a tetramer in the [LDH-like] group of L-MalDH (Madern et al. 2001) were also removed. A rooted tree was infered from the 241 homologous sites that identify the common scaffold, by choosing the reciprocal rooting between the group of tetrameric (L-LDH and [LDH-like] L-MalDH) and the dimeric (cytosolic and mitochondrial L-MalDH) enzymes inside the super-family. This common scaffold corresponds to the secondary structure elements without their connecting loop and without the mobile catalytic loop. In this case, and in contrast to the first analysis and previous observations (Golding and Dean 1998), the [LDH-like] L-MalDH form a homogeneous group which can be rooted with L-LDH. For the sake of clarity, the tetrameric and dimeric clades are presented separately (Figs. 2 and 3).

Phyletic Distribution and Functional Changes Within the L-LDH and L-MalDH Super-Family

It is generally accepted that enzymes belonging to the L-LDH/L-MalDH super-family are present throughout the three domains of life: Archaea, Bacteria, and Eukarya. This is not the case, however, when the distribution is analyzed with respect to the three groups that make up the super-family.

The clade of tetrameric enzymes. All the L-LDH and [LDH-like] L-MalDH sequences are clustered to form a homogeneous clade containing mainly tetrameric enzymes (Fig. 2), which is different from the one formed by the dimeric L-MalDH (Fig. 3). The specific case of *A. fulgidus* [LDH-like] L-MalDH (a dimer) has been explain recently (Madern et al. 2001). The split by enzymatic function which depends on the discrimating residue at position 102 is well defined. All sequences with Gln 102 form the group of L-LDH and those with Arg 102 form the [LDH-like] group of L-MalDH. The position of the lowest branches in each the L-LDH and [LDH-like] L-MalDH groups gave slightly different branching pattern with the various methods used. Only three members of the Bacillus genus (*B. subtilis, B. stearothermophilus, B. halodurans*) have genes coding for both enzymes (stars). Attempts to find (by BLAST searches at http://igweb.integratedgenomics.com/GOLD), either an [LDH-like] L-MalDH ortholog or an L-LDH ortholog, in the 64 other published genome sequences, were unsuccessful.

NADH-Dependent L-LDH. The comparative organization within the L-LDH has been previously analyzed, especially with respect to tissue location of Eukaryal L-LDH. As it has already been shown, Eukaryal L-LDH form a homogeneous group, which is statistically robust (Griffin et al. 1992; Tsuji et al. 1994; Stock et al. 1997). Until the present study, the characterization of an archaeal NAD-dependent L-LDH had not been reported.

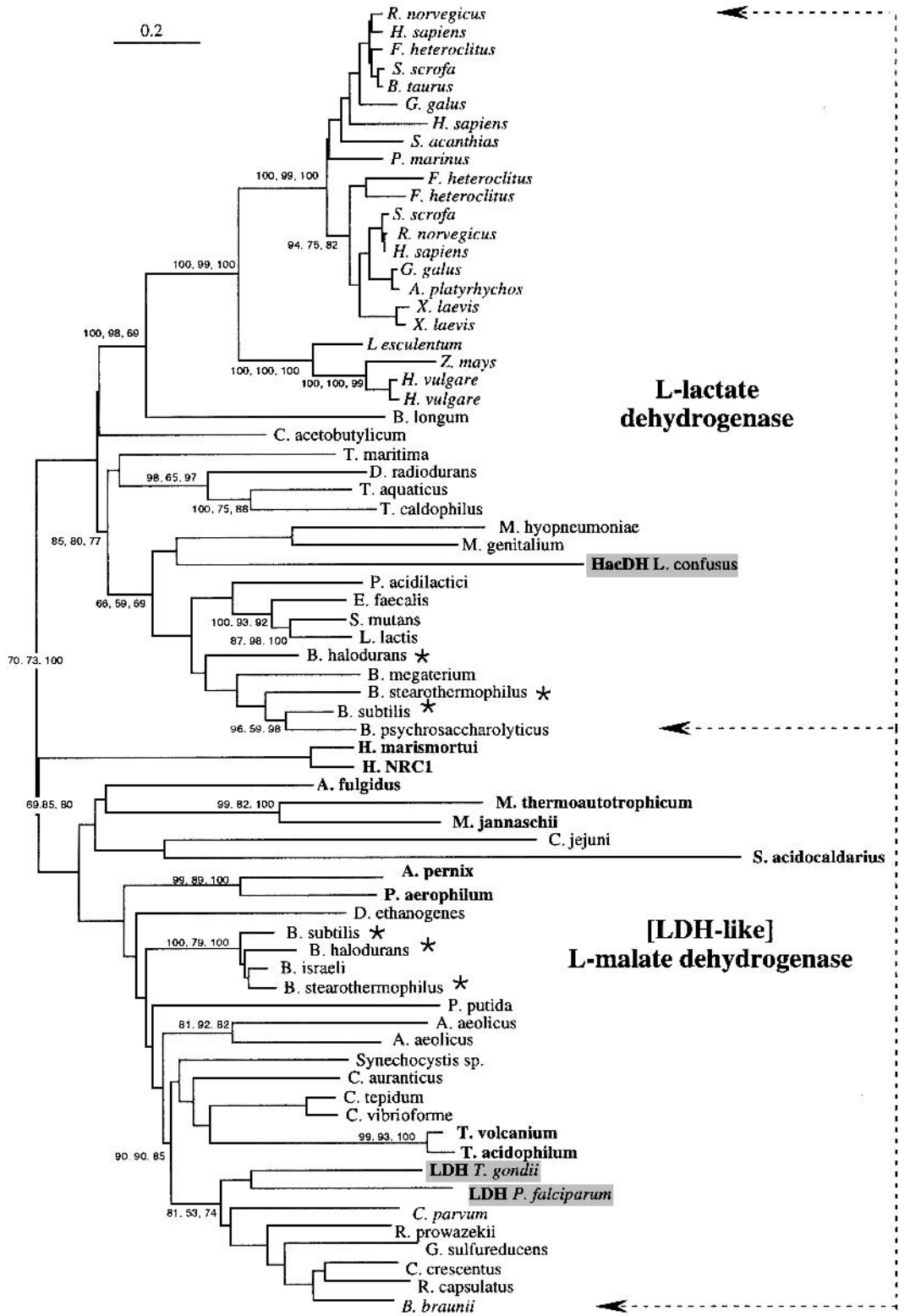

Fig. 2. Neighbor-joining tree of L-LDH, [LDH-like] L-MalDH. Eukaryotic and archaeal species are in italics and bold, respectively. The functional assignment (lactate dehydrogenase or [LDH-like] malate dehydrogenase) is indicated by shaded lines. The stars indicate species in which both enzymes were found. Functional changes are boxed in grey. The scale bar represents 0.2 amino acid substitution per site. The statistics values are provided from distance-based, maximum-likelihood, and parsimony analyses, respectively.

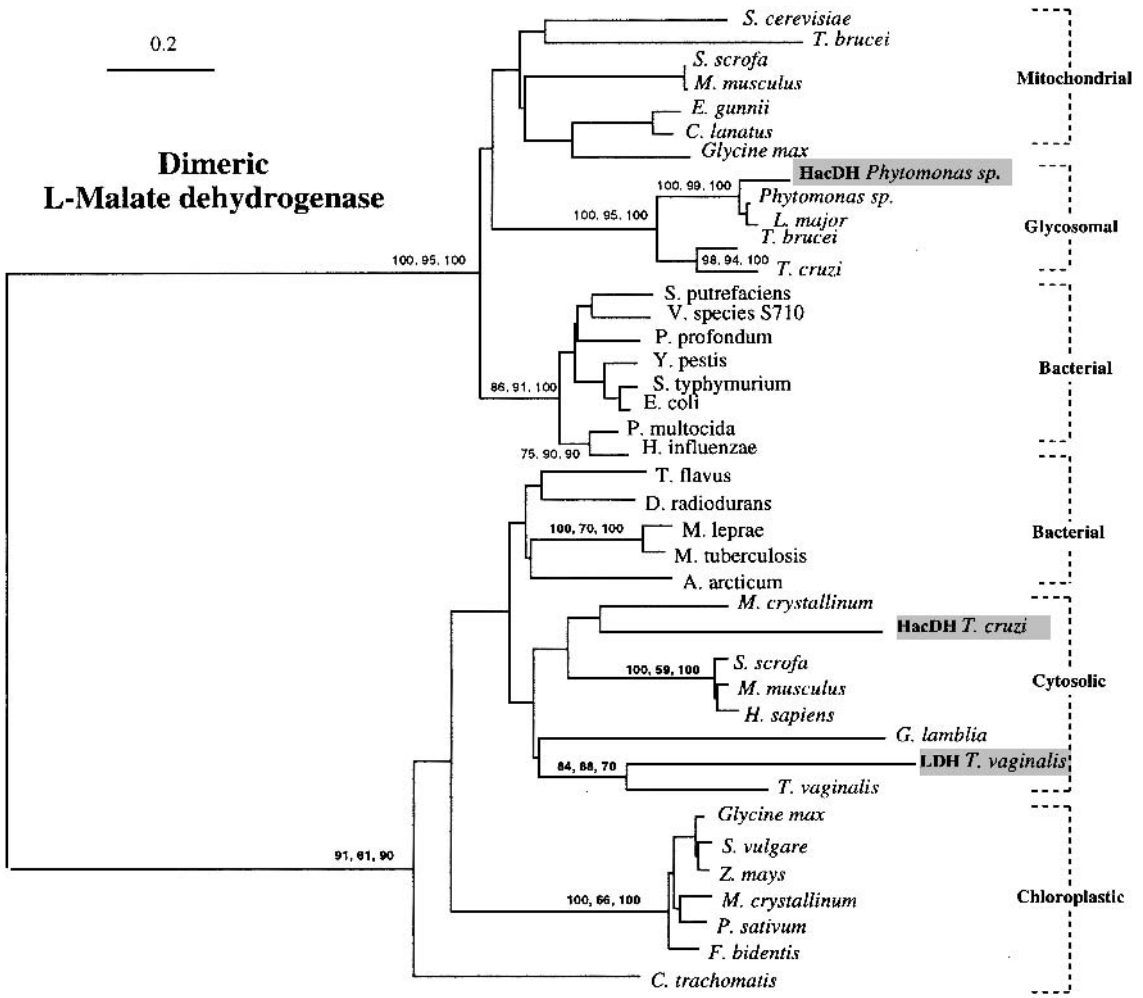

Fig. 3. Neighbor-joining tree of dimeric L-MalDH. Eukaryotic species are in italics. Functional changes are boxed in grey. The dashed lines indicate the various subgroups of enzymes. The scale bar represents 0.2 amino acid substitution per site. The statistics values are provided from distance-based, maximum-likelihood, and parsimony analyses, respectively.

Similarity searches have not detected genes or pseudogenes homologous to the L-LDH in the archaeal genome sequences presently available. These enzymes appear, therefore, to be found only in the Eukaryal and Bacterial domains.

In the case of the Bacterial domain, there is a specific distribution (an exclusion) of L-LDH. In particular, the characterization of a L-LDH from an α-proteobacterium has not been reported. Extensive similarity searches on the complete lactate dehydrogenase entries (>200) at the NCBI failed to identify L-LDH orthologs belonging to the α-proteobacterial species.

NAD(P)H-dependent [LDH-like] L-MalDH. The deepest branches of the group were formed by most of the Archaeal sequences. A DNA region encoding [LDH-like] L-MalDH can be detected by similarity search in the genome of *Sulfolobus acidocaldarius.* This sequence is incorrectly identified in the NCBI data bank as an L-LDH. However, the protein was purified and characterized as a tetrameric L-MalDH (Hartl et al. 1987) and the phenetic analysis has shown that it is not a L-LDH. The location of *Termoplasma acidophilum* and *Termoplasma volcanium* sequences with the bacterial [LDH like] L-MalDH probably reflects lateral gene transfer, an observation in agreement with the fact that the *T. acidophilum* genome contains a substantial proportion of bacterial genes (Ruepp et al. 2000). The *Methanobacterium thermoautotrophicum* and *Methanococcus jannaschii* enzymes have been shown to use NADPH (Thompson et al. 1997; Madern 2000; Graupner and White 2000). This is due to the lack of the well-conserved Asp residue in position 53 (observable in all the NADH-dependent enzyme), which is replaced by a glycine. It is known that phylogenetic methods are sensitive to evolutionary rates, and are therefore sensitive to the long branch attraction phenomenon (Philippe and Forterre 1999). The location of the Bacterial [LDH like] L-MalDH from *Campylobacter jejuni* within the Archaeal branches could be due to a highly divergent evolution rate in this species compared to the others. A tree (not shown) constructed with a version of the NJ algorithm less sensitive to high rate

of substitution (Gascuel 1997) locates the *C. jejuni* with the bacterial enzymes. In addition to this sequence, a second ORF with high sequence similarity could be detected in its genome indicating gene duplication. A careful inspection of this sequence showed that various nucleotide insertions and deletions produced frame shift mutations along the gene sequence. Moreover, some of the well-conserved active site residues were suppressed. This ORF, therefore, is neither [LDH-like] L-MalDH nor L-LDH, but a sequence that has been converted into a pseudogene. A unique case of gene duplication that gives two copies of [LDH like] L-MalDH is observable with *Aquifex aeolicus.*

The eukaryal sequences of a green alga sequence (*Botryococcus braunii*) and those of apiclomplexan parasites (*Cryptosporidium parvum, Toxoplasma gondii,* and *Plasmodium falciparum*) are clustered with α- and δ-proteobacterial sequences. The *T. gondii* and *P. falciparum* enzymes are discussed below.

The clade of dimeric L-MalDH. The phylogeny of dimeric L-MalDH has been studied for a long time. This family is split into cytosolic and mitochondrial clades, which contains various subgroups. The chloroplastic L-MalDH are located within the cytosolic clade and the bacterial enzymes are distributed in both the cytosolic and mitochondrial clades (Roger et al. 1999; Ocheretina et al. 2000). In order to refine the actual description of the dimeric L-MalDH, glycosomal L-MalDH (Hunter et al. 1999; Uttaro et al. 2000) were included in this study. The inferred phylogenetic tree is presented in Fig. 3. The previous description of various sub-groups is still valid, and, in addition, it is possible to show that the glycosomal L-MalDH form a new sub-group, highly supported by statistical scores, within the mitochondrial clade. Glycosomal L-MalDHs display the tripeptide-SKL at their C-terminus, which corresponds to the peroxisomal targeting signal (PTS-1) (Anderson et al. 1998). These data suggest that glycosomal MalDH were inherited from a mitochondrial gene duplication followed by the acquisition of a targeting signal.

Extensive sequence similarity searches using mitochondrial or cystosolic L-MalDH orthologs in the 10 completely sequenced archaeal genomes were unsuccessful. Databank searches showed that dimeric L-MalDH orthologs are also not present within the α-proteobacterial group within the Bacterial domain of life. In this case, the enzymatic function is sustained by [LDH-like] L-MalDH.

Functional Changes

In each tree, it is possible to identify sequences (grey shadow) that identify functional changes.

The biochemically well characterized L-LDH from *P. falciparum* (Bzik et al. 1993) and the putative L-LDH from *T. gondii,* are located in the [LDH-like] L-MalDH group and not in the L-LDH (Fig. 2), as was previously observed in the unrooted tree. As was discussed, there is a positively charged residue Lys at the discriminant [102] position (more related to the Arg of a L-MalDH catalytic site) which was found instead of the Gln present in all tetrameric L-LDH. Moreover, the specific Pro present in all [LDH-like] L-MalDH at position [250] was found in both sequences and the catalytic loop phylogeny classifies them as L-MalDH. Another functional change might be observed, in the group of L-LDH, with the L-2-hydroxyisocaproate dehydrogenase of *Lactobacillus confusus.* At the discriminant [102] position there is a Gln residue as always observed in L-LDH. For these two cases of functional changes ([LDH-like] L-MalDH toward L-LDH and L-LDH toward HacDH) there is a 5 and 4 residue insertion, respectively, within the mobile loop.

In Fig. 3, as it has already been established using the complete sequence (Wu et al. 1999), the L-LDH from the amitochondriate protist *Trichomonas vaginalis* is located within the cytosolic group of dimeric L-MalDH (grey shadow). At the discriminant [102] position in *T. vaginalis* L-LDH there is a Leu residue instead of an Arg as within L-MalDH. In the present analysis, its location is still the same when the sites corresponding to the common scaffold of the super family are taken into account. It corresponds therefore to a gene duplication of the cytosolic *T. vaginalis* L-MalDH that was associated, in one copy, with a functional change of the catalytic loop. This was indicated by the *T. vaginalis* loop analysis, which ressembles a L-LDH loop and not a L-MalDH loop. A hydroxy acid dehydrogenase from *Trypanosoma cruzi* which does not possess a malate dehydrogenase activity has been characterized (Cazzulo et al. 1999). At the discriminant [102] position in *T. cruzi* HacDH, there is an Ala residue instead of an Arg, as in L-MalDH. Its location in Fig. 4 shows, therefore that another functional change has occurred within the cytosolic subgroup of L-MalDH. The phylogenetic position of *T. cruzi* HacDH varies according to the various method used. The last functional change within the dimeric L-MalDH is observed within the sub group of glycosomal enzymes. A malate dehydrogenase isoform of *Phytomonas species* was characterized and shown to be a 2-hydroxyacid dehydrogenase (Uttaro et al. 2000). At the discriminant [102] position there is a Ile residue instead of Arg.

Two, and Not One, Ancestral Paralogous Gene Duplications Explain the Functional Distribution in the Super-Family

In contrast to the tree based on the whole sequence, the neighbor-joining tree based on the common scaffold (il-

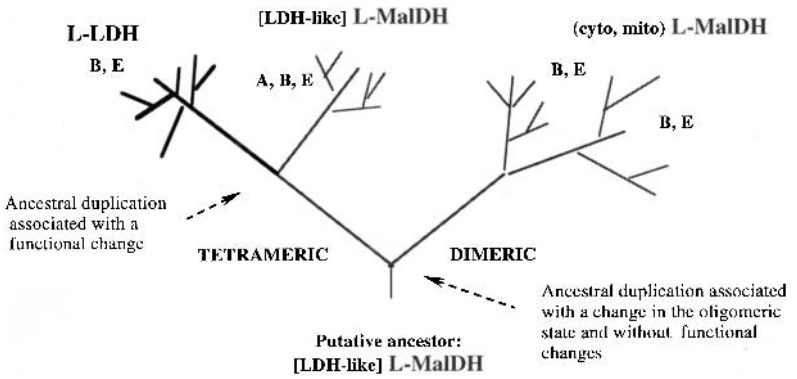

Fig. 4. Schematic rooted tree of the L-MalDH and L-LDH super-family. The L-MalDH are in grey and the L-LDH, in **bold**. The distribution over the three domain of life (Woese et al. 1990) is abbreviated by: A, Archaea; B, Bacteria; E, Eukarya. The arrows indicate the two ancestral gene duplications which afford for the enzymatic distribution, in three groups, by size and function.

lustrated in Fig. 4) identifies more clearly the relationships within the super-family.

A gene coding for a [LDH-like] L-MalDH is postulated to be the ancestor. After a first ancestral gene duplication, one copy has evolved to give an ancestral dimeric L-MalDH. This new kind of active L-MalDH was very likely selected from mutations that had disrupted the tetrameric assembly. These mutations were not deleterious with respect to function and had modified the biochemical properties (see discussion). Within the ancestral group dimeric of L-MalDH, a consecutive gene duplication has lead to the cytosolic and mitochondrial sub-groups. This overall tree topology of the dimeric L-MalDH is equivalent to those generaly obtained with these two clades (Roger et al. 1999). Only Eucaryal and Bacterial enzymes are present.

The second copy obtained from the duplication of the ancestral [LDH-like] L-MalDH has evolved by a gene duplication associated with a functional change and has led to the L-LDH on one hand and to the [LDH-like] L-MalDH on the other. The traces of this duplication are observable in some members of the genus *Bacillus.* Genes coding for a NADH-dependent L-LDH were not selected in the Archaea (see discussion) and therefore only Eukaryal and Bacterial enzymes are observable. In contrast, [LDH-like] L-MalDH are present in organisms of the three domains of life.

The generally accepted point of view is that L-LDH and L-MalDH are the consequence of a single ancestral gene duplication (Birkoft et al. 1982; Mc-Allister Henn 1988). This conclusion was made at a time when the [LDH-like] L-MalDH were unknown. Their use in present phylogenetic analysis, showed that two main paralogous duplications of an ancestral gene, and not one, explains the distribution of enzymatic function inside the L-MalDH and L-LDH super-family.

In addition to this, it is possible to observe some cases of functional change within each of the L-LDH, the [LDH-like] L-MalDH, and L-MalDH groups. These changes are polarized: from L-MalDH to L-LDH or HacDH, and from L-LDH to HacDH, and do not occur in the opposite direction (see discussion).

Discussion

Biochemical Properties Make Sense of the Three Groups

A folding and association pathway of dimeric L-MalDH, [LDH-like] L-MalDH, and L-LDH has been presented (Madern et al. 2000). Identical inactive monomeric subunits assemble to form dimeric species that may condense into active tetramers. Whereas the dimeric MalDH are always active over a broad pH range, the activity of equivalent [LDH-like] L-MalDH dimeric units has been

			k_{cat}/K_m expressed in (M^{-1} s^{-1})	
			Tetrameric	Dimeric
L-MalDH	Wild type (R102)	OAA	$2.0\ 10^5$	$2.6\ 10^7$
L-MalDH	Wild type (R102)	Pyruvate		
L-MalDH	Mutant (R102Q)	OAA	68	257
L-MalDH	Mutant (R102Q)	Pyruvate	5600	132
L-LDH	Wild type (Q102)	OAA	4000	
L-LDH	Wild type (Q102)	Pyruvate	$4.2\ 10^6$	
L-LDH	Mutant (Q102R)	OAA	$4.2\ 10^6$	
L-LDH	Mutant (Q102R)	Pyruvate	500	

Fig. 5. Catalytic efficiency of L-LDH, (LDH-like) L-MalDH, and dimeric L-MalDH. The three enzymes for which the influence of position 102 on the substrate specificity was probed are *B. stearothermophilus* L-LDH (Wilks et al. 1988), *H. marismortui* (LDH-like) L-MalDH (Cendrin et al. 1993), and *E. coli* L-MalDH (Nicholls et al. 1992).

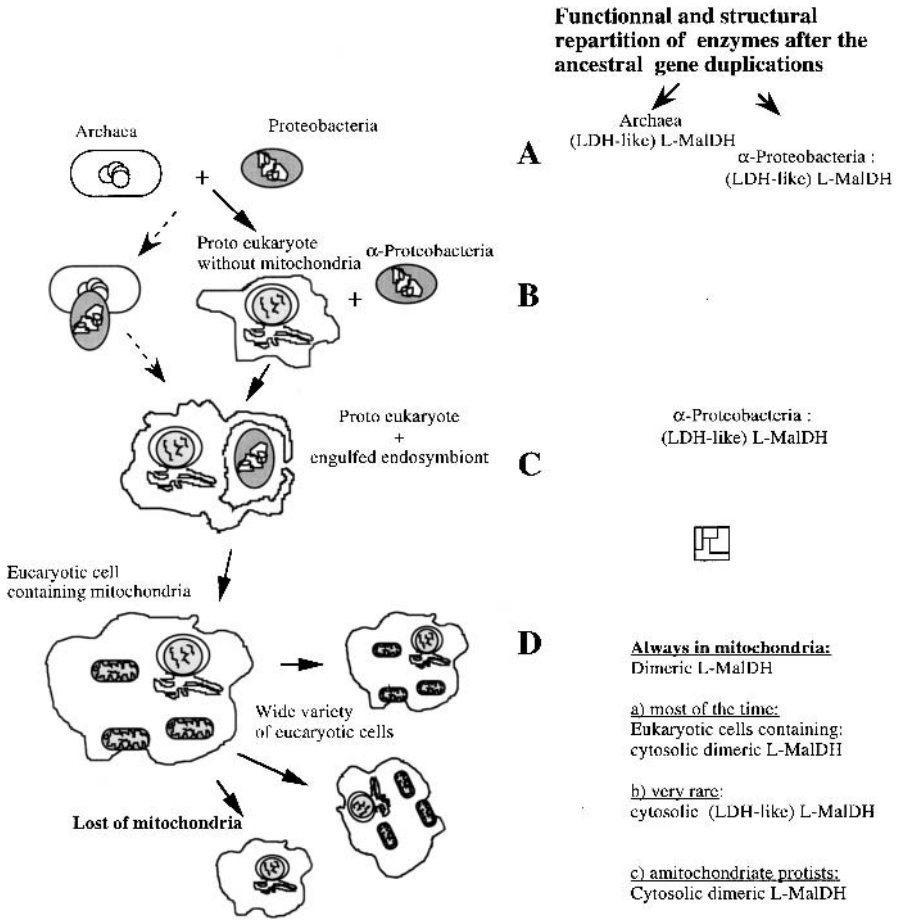

Fig. 6. Schematic description of hypotheses describing the genesis of mitochondria, redrawn from Gray et al. (1999) in light of the L-MalDH evolution. On the left side, the dashed lines illustrate the genesis of eukaryotic cells containing mitochondria in a single step, through the fusion between an ancestral methanogenic Archaea and an ancestral α-proteobacterium (gray). The continuous lines illustrate the endosymbiosis between an ancestral proto eucaryote and an ancestral α-proteobacterium. The main evolutionary steps are indicated by capitalized letters (**A, B, C, D**). Some important data on L-MalDH are located in parallell on the right side of the figure. The selective pressure that was acting on the ancestral α-proteobacterium (see Discussion) is symbolized by a puzzle.

shown to be regulated through a pH-dependent mechanism (Madern et al. 2000; Madern et al. 2001). The biochemical study of the pig heart enzyme, in the clade of mitochondrial L-MalDH, has shown that the protein dissociates into inactive monomers in a pH-dependent process (Wood et al. 1981; Steffan and McAllister-Henn 1991; Chen and Smith 2001). Compared to [LDH-like] L-MalDH, evolution has, therefore, eliminated the intermediate step that produces a pH-dependent (inactive) dimer upon dissociation and indeed, selected and favored the active one. In the case of bacterial L-LDH, the behavior of dimeric intermediates in the folding pathway is sometimes modified by the allosteric effector fructose 1,6-bis phosphate which promotes tetramerization and activation (Garvie 1980; Jaenicke 1998). A structural comparison of L-LDH (Jackson et al. 1992) and [LDH-like] L-MalDH (Richard et al. 2000; Madern et al. 2000) with dimeric L-MalDH has shown that dimeric MalDH cannot reach the tetrameric state because of steric hindrance by large loops. In the clade of chloroplast L-MalDH, N-, and C-terminal extensions in the dimeric L-MalDH from plants control the enzymatic activity (Carr et al. 1999). The properties of other bacterial enzymes have not been documented and the phylogenetic trees concerning the two cytosolic and mitochondrial dimeric L-MalDH clades (Roger et al. 1999; Wu et al. 1999) are not discussed here.

In the hypothesis of a [LDH-like] L-MalDH as common ancestor, why would evolution have selected a new dimeric species when L-MalDH enzymatic activity was already present on the tetrameric fold? Analysis of the catalytic properties of L-LDH, [LDH-like] L-MalDH, and dimeric L-MalDH correlated with their respective oligomeric state show that they make sense in the context of the emergence of a more efficient dimeric malate dehydrogenase. The k_{cat}/K_M values for oxaloacetate reduction by *E. coli* L-MalDH (dimer) show its higher catalytic efficiency compared with *H. marismortui* [LDH-like] L-MalDH tetramer (favored direction indicated by black arrows, Fig. 5). The natural evolution of a L-LDH from a [LDH-like] L-MalDH would have produced an efficient enzyme with the same overall fold. Due to a charge repulsion mechanism, L-MalDH does not accept pyruvate for catalysis (Chapman et al. 1999). In contrast to L-MalDH, the uncharged Gln[102] in L-LDH displays a preference for pyruvate over oxaloacetate. In the tetrameric L-LDH, the low level of oxaloacetate reduction should be considered, therefore, as a residual enzymatic activity of the original tetrameric fold of [LDH-like] L-MalDH. The L-LDH may also be seen as a less stringent enzyme, a

property that has been used for the design of enzymes with broader substrate specificity (Hogan et al. 1995).

Oxaloacetate reduction was strongly favored when the Gln to Arg 102 mutation was engineered into L-LDH (which therefore mimicked an ancestor [LDH-like] L-MalDH rather than a dimeric L-MalDH). The effect of Arg to Gln[102] mutation in a [LDH-like] L-MalDH or a dimeric L-MalDH are not similar (Fig. 5). With [LDH-like] L-MalDH, pyruvate reduction is favored (the enzyme becomes a not very efficient L-LDH), whereas oxalacetate reduction is always preferred in the recombinant dimeric Arg to Gln[102] L-MalDH. The sequence analysis of catalytic loops, shows that residues of both dimeric and [LDH-like] groups of L-MalDH are very similar. This may explain why the Arg to Gln [102] mutation is not sufficient to induce a high level of pyruvate reduction when introduced in a [LDH-like] L-MalDH. In contrast, the L-LDH loop which was shown to have great flexibility seems to be more tolerant to the Gln to Arg [102] mutation allowing a high rate of oxaloacetate reduction. An attempt to design an efficient dimeric L-LDH by four residue changes that mimic a L-LDH loop on the dimeric L-MalDH from *E. coli* was not successful (Boernke et al. 1995). The dimeric L-MalDH, therefore, would be seen as a more accomplished version of L-MalDH (compared to the ancestral [LDH-like] L-MalDH) in which the inactive intermediate was eliminated and from which the evolution toward L-LDH was less favored.

As shown above, the recent pathway of evolution has favored the enzymatic conversion from malate to lactate dehydrogenase, and so far, the reverse direction has never been observed. After the gene duplication of an ancestral [LDH-like] L-MalDH, one copy was kept coding for a tetrameric enzyme, whereas the other copy has evolved following various selective pressures, towards more active dimers (L-MalDH). In the lineage of tetrameric enzymes, an additional gene duplication of the ancestral [LDH-like] L-MalDH associated with a functional changes of the new copy has led to a dimeric species that could be assembled into tetramers with a different activity (L-LDH). When physical contacts between operating protein partners are loosened, their sequence evolution rate is increased (Dickerson 1971). The loosening of physical contacts within the ancestral enzyme would have favored intermediates of varying properties and potential, which were the precursors of these two new enzymes.

The Ancestral Birth of NADH-Dependent L-LDH

It has been proposed that, in some ecological niches, oxygen respiration developed early, even when the primitive earth's atmosphere was mainly anoxic (Towe 1990). In the framework of such a scenario, it is possible to suggest how the L-LDH have been selected in an ancestral cell. After the gene duplication of an ancestral [LDH-like] L-MalDH, a single base mutation has changed the sense of the codon at the discriminating position 102. This leads to a protein equivalent to the mutant Arg 102 Gln [LDH-like] L-MalDH (see Fig. 5) with a relaxed substrate specificity. If there is no selective pressure, copies containing deleterious mutations are generally eliminated. However, this mutation may have induced an advantage for an ancestral aerobic bacteria submitted to a temporary decrease in O_2, because it offers a way to regenerate the NAD^+ by the use of pyruvate produced by NADH-dependent phosphorylating GAPDH during glycolysis. The main advantage, therefore was not the use of an additional substrate but the new possibility to survive in a fluctuating environment. The evolution toward the true efficient L-LDH was achieved by a complete loop adaptation in the descendant.

Until now, no NADH-dependent L-LDH activity has been detected in Archaea. The following analysis provides a possible explanation for this. Biochemical and in silico analyses of the pathways for glucose degradation in Archaea have shown some modifications with respect to the classical Embden-Meyerhof-Parnas (EMP) and Entner-Doudoroff (ED) pathways characteristic of eukaryotes or bacteria (Danson 1993; Romano and Conway 1996; Selig et al. 1997; Koonin et al. 1998; Makarova et al. 1999). An analysis of the nature of MalDH involved in the TCA cycle has been published (Cordwell 1999; Huynen et al. 1999). This analysis was refined (not presented) to take into account the different MalDH (LDH-like, dimeric or non-homologous) as defined in this work. In particular, in both *Pyrococcus horikosii* and *Pyrococcus abysii* archaeal genomes the MalDH enzymatic function seems to be sustained by nonhomologous enzymes. It was shown that there is no evolution from the ancestral tetrameric L-MalDH towards a NADH-dependent L-LDH, when at least one enzyme (PGK, PGM, enolase, PK) is missing in the second part of the classical glycolytic pathway. Even if these activities have been detected in *P. furiosus* (Kengen et al. 1994) it is clear that the glyceraldehyde-3-phosphate oxidation is done via a GAP-FdOR (ferredoxin oxidoreductase) and not a GAPDH (Mukund and Adams 1995). The only cristallographic structure available for a GAPDH with strong preference for $NADP^+$ over NAD^+ is that of the enzyme from *Methanothermus fervidus* (Charron et al. 2000). The authors have shown that the *A. fulgidus, P. wosei, P. abyssi, M. thermoautotrophicum,* and *M. jannaschii* GAPDH are closer to *M. fervidus* GAPDH than to bacterial or eukaryal GAPDH, suggesting a preference for $NADP^+$ as coenzyme. In a similarity search, not presented here, which was done at the NCBI site (http//www.ncbi.nlm.nih.gov) using the *M. fervidus* GAPDH as query, it is possible to see that this holds also for GAPDH from *Halobacterium* NRC1, *T. acidophilum, S. solfataricus,* and *A. pernix.*

It can be suggested that, after the ancestral gene duplication, and in contrast to the bacterial lineages, the fixation and evolution of an NAD-dependent L-LDH from a NAD-dependent [LDH-like] L-MalDH did not occur in the archaeal lineages because the conditions required for their selection were never achieved:

- In most of the Archaea, the coenzyme preference was not common between GAPDH and [LDH-like] L-MalDH. The selection of a putative NAD-dependent L-LDH evolving from a [LDH-like] L-MalDH, in order to promote $NADP^+$ recycling was, therefore, not possible.
- Even though there is a putative common $NADP^+$ usage by both archaeal GAPDH and [LDH-like] L-MalDH, as in the case of the methanogenic archaea, there is no physiological reasons to select a L-LDH from a [LDH-like] L-MalDH in these organisms. As suggested in previous paragraphs, the NAD-dependent L-LDH selection was due to a decrease in the oxygen concentration, a condition never acheived with methanogenic Archaea which are strict anaerobes.
- In two *Pyrococcus* species (*P. abysii* and *P. horikoshii*), the malate dehydrogenase activity is due to a non-homologous enzyme (the gene coding for an [LDH-like] L-MalDH was not detected in their genome sequences), and therefore, it could not be possible to select in these species, a NADH-dependent L-LDH because the putative ancestor was absent.

Functional Changes

The analysis of the functional changes in the superfamily has shown some additional births of L-LDHs independently of the second ancestral divergence which gave the main group of orthologous L-LDH. The prediction of a generally unfavorable evolution of L-LDH from dimeric L-MalDH is reinforced by the report of a unique case of L-LDH evolution from a dimeric L-MalDH in the eucaryotic protist *Trichomonas vaginalis* (Wu et al. 1999). In the *T. vaginalis* L-LDH, a Leu instead a Gln was found at the discriminant position 102. This sequence is found in Fig. 3 in the clade of cytosolic-type dimeric L-MalDH. The authors measured an unusually high oxaloacetate activity with this L-LDH and concluded that this recent divergence did not achieve the same substrate selectivity as in tetrameric L-LDH. At position 102, the directed mutation Leu to Arg, produced an efficient L-MalDH (Wu et al. 1999).

Two other cases of functional changes, which gave rise to α-hydroxyacid dehydrogenases, are observed within the dimeric L-MalDH group (Fig. 3). These HacDH from *T. cruzi* and *Phytomonas* shared Arg to Ala and Arg to Ile mutations, respectively, at the discriminating position 102 with other mutations that modify the net charge and increase the volume of the active site are observed (Cazzulo Franke et al. 1999; Uttaro et al. 2000).

The present work has shown another appearance of L-LDH activities, which can be observed in the group of [LDH-like] L-MalDH (Fig. 2), with the putative L-LDH from *Toxoplasma gondii* and the L-LDH from *Plasmodium falciparum.* The evolution of these apicomplexan parasites has been recently described as involving an ancestral endosymbiotic event with an alga. Different analyses suggest a red or a green alga as endosymbiont (Köhler et al. 1997; Dzierszinski et al. 1999; Fast et al. 2001). The [LDH-like] L-MalDH tree topology suggests a link between apicomplexan parasites and the green algae (*Botryococcus braunii*). Until now, however no [LDH-like] L-MalDH has been reported in red algae. The presence of L-LDH in apicomplexan parasites is well described by a [LDH-like] L-MalDH gene duplication event associated with a functional change. The distinct biochemical and immunochemical properties of *P. falciparum* L-LDH compared to eukaryal L-LDH were assigned to novel features in its primary sequence (Bzik et al. 1993). The crystallographic structure of *P. falciparum* L-LDH was determined and shows a specific five residue insertion in the mobile catalytic loop (Cameron et al. 1996). This allowed these authors to propose an explanation for the catalysis in the absence of a Gln normally found at position [102] in L-LDH. The lysine in *P. falciparum* L-LDH projects away from the substrate site, because the longer catalytic loop distorts this region. Consequently, the Trp of the insertion mimics the role of a Gln in L-LDH, because its side chain fills the space at position [102] allowing, the recognition of pyruvate.

The functional change from L-LDH to an L-2-Hydroxyisocaproate dehydrogenase from *Lactobacillus confusus* (Fig. 2) is due to a topological modification of the substrate binding pocket structure which allowed the recognition of substrates with side-chains more bulky than pyruvate (Niefind et al. 1995). This modification is induced by a three residue insertion in the catalytic loop and by hydrophilic to hydrophobic mutations in the binding pocket. The *L. confusus* HacDH is clustered with the L-LDH sequences of Mollicutes. In the case of *Mycoplasma genitalium,* the gene coding for a L-MalDH is absent, whereas the activity is present. It has been proposed that a single glycine residue insertion in the L-LDH catalytic loop has induced a slight change in the *M. genitalium* L-LDH, which is sufficient to afford dual enzymatic function (Cordwell et al. 1997).

L-MalDH in the Framework of Mitochondrial Evolution

It is widely accepted that eukaryotes have inherited the oxidative respiratory system by the sequestration of a primitive endosymbiotic bacteria, to give, after various steps of evolution, mitochondria in their contemporary form and function (Fig. 6, step A to D) (Anderson and Kurland 1998; Martin and Müller 1998; Martin and Her-

man 1998; Gray et al. 1999; Kurland and Adersson 2000; Gray et al. 2001). When the engulfed endosymbiont became totally trapped (a protomitochondrion), a redundancy between some of its genes and those of the host genome existed (step C). It was followed by a mechanism of reductive evolution and gene loss, in which an important gene transfer from the endosymbiont genome to the host genome occured. The addition of a targeting signal to nuclear genes coding for proteins which have to be located inside the mitochondrion was one the later steps of this evolution. In such a scenario, an α-proteobacterium is proposed to be the closest species of the ancestral endosymbiont (Bonen et al. 1977; Scheritz-Pontén et al. 1998; Anderson et al. 1998; Gray et al. 1999; Karlberg, et al. 2000).

In the specific case of L-malate dehydrogenases, there is a striking discrepancy concerning such a relationship, which is illustrated on the right side of Fig. 6. The present study has shown that there is no close phylogenetic links between the mitochondrial malate dehydrogenases (Fig. 3) and those of α-proteobacterial origin that belong to the [LDH-like] group of L-MalDH (see in Fig. 2, *R. prowazekii, R. capsulatus,* and *C crescentus*). Moreover, the enzymes from *Bradyrhizobium meliloti, Rhodospirillum rubrum,* and *Rhodobacter spheroides* were shown to be tetrameric (Tayeh and Madigan 1987), strongly suggesting they are [LDH-like] L-MalDH.

It can be argued therefore, in favor of an hypothesis which explain the fate of malate dehydrogenase during eukaryogenesis:

Step from C to D_a: Biochemical and structural considerations explain the reason for which the [LDH-like] L-MalDH gene of the ancestral α-proteobacterial symbiotic species has not been selected in order to be functional inside the mitochondrion. *In vitro* and *in vivo* experiments strongly suggest that a quinary structure (the metabolon) of the soluble enzymatic components of the TCA cycle exists in mitochondria in order to promote substrate channeling during the TCA cycle (Ovadi and Srere 2000; Grandier-Vazeille et al. 2001 and references therein). The first mitochondrial enzyme involved in the oxidative TCA cycle is the dimeric citrate synthase. At the end of the cycle, the oxaloacetate is regenerated by the L-MalDH, to be one of the substrates of the citrate synthase (CS) in a further cycle. The electrostatic channeling of oxaloacetate, by a direct transfer between active sites, between the mitochondrial L-MalDH and mitochondrial CS is well documentated (Velot and Srere 2000; Shatalin et al. 1999; Morgunov and Srere 1998). A structural model for the mitochondrial metabolon has been proposed (Velot et al. 1997). In particular, there is an extensive topological surface complementarity, with thermodynamically favorable interactions, between MalDH and citrate synthase. The citrate synthases exist in two mains types. Eukaryotes, Gram positive bacteria, and Archaea have a small dimeric CS that is isosterically regulated by ATP, while a large hexameric enzyme regulated allosterically by NADH or 2-oxoglutarate is found in Gram negative bacteria (Danson, 1988). The tree of the CS (work not shown), using more than 50 taxon including six sequences of α-proteobacteria (*Rhodobacter capsulatus, Rhodospirilum rubrum, Rhizobium tropici, Sinorhizobium meliloti, Bradyrhizobium japonicum,* and *Rickettsia prowazekii*) showed that there is an large phylogenetic distance between the α-proteobacterial sequences and the clade of mitochondrial sequences. However, the *R. prowazekii* CS was characterized as a dimeric enzyme regulated by ATP, i.e the properties shared by mitochondrial CS (Phibbs et al. 1982; Wood et al. 1983). It can be suggested, therefore, that during mitochondrion genesis, a dimeric, ATP-controled citrate synthase of Rickettsial origin, rather than an hexameric CS generally found in α-proteobacteria, has been submitted to a rapid accumulation of mutations leading to the ancestral CS gene of mitochondrion. The physical association between such a CS and an efficient dimeric L-MalDH (already present in the host), instead of a less efficient and topologically unfavorable [LDH-like] L-MalDH (present in the symbiont), has promoted the successful formation of a protomitochondrial metabolon. This was a means to increase the TCA cycle turn-over in the first primitive eukaryotic cells, and obviously would have been a important advantage to compete with the rapidly increasing oxygen concentration of the earth's atmosphere.

Step from C to D_b: The DNA sequence analysis of two eukaryotic [LDH-like] L-MalDH from *Botryococcus braunii* and *Cryptosporidium parvum* shows that they encode enzymes without additional sequences at their N terminus. They display a close phylogenetic relationship (Fig. 2) to α-proteobacterial sequences from *Rickettsia prowazeckii, Caulobacter crescentus,* and *Rhodobacter capsulatus.* This might therefore, be seen as a relic of the intermediate step, in which the selective pressure acting on malate dehydrogenase (as presented above) was efficient. The ancestral [LDH-like] L-MalDH gene was transferred and fixed to the host nuclear genome to encode for a cytosolic enzyme, because it was not selected for mitochondrial targeting.

Step from C to D_c: The amitochondriate protist are eukaryotic cells which have very likely lost their mitochondria. This is suggested by studies of various protist genes which display a proteobacterial origin (Horner et al. 1996; Roger et al. 1998). The L-MalDH from *Giardia lamblia* and *Trichomonas vaginalis* are located in the cytosolic clade, as observed by Roger et al. (1999). These examples illustrate, therefore, the lost of the ancestral [LDH-like] L-MalDH of symbiotic origin.

Martin and Müller (1998) has proposed the birth of the first eukaryote containing mitochondria in a single step (Fig. 6, dashed lines). This hypothesis suggests a primitive fusion between an Archaea and an α-proteo-

bacterium. As shown in the present work, the nature of species involved (containing only [LDH-like] L-MalDH) cannot explain the presence of a dimeric L-MalDH in eukaryotes and their mitochondria.

To be present in mitochondria, L-MalDH orthologs have therefore been recruited from an unknown bacterial species (cryptic endosymbiosis, lateral gene transfer) or were already present in a protoeukaryotic ancestor.

The Fate of L-LDH

In agreement with the scenario developed in the previous chapter, various enzymes of the eukaryotic glycolytic pathway have been shown to be transferred from the ancestor of mitochondria to the nucleus (Henze et al. 1995; Brinkmann and Martin 1996; Henze et al. 1998; Figge et al. 1999; Liaud et al. 2000). In the present study, it was found that NADH-dependent L-LDH are not present in α-proteobacterial species. This suggests therefore, that ancestral NADH-dependent L-LDH gene was either present in the protoeukaryotic cell before the genesis of mitochondria or was acquired later by a lateral gene transfer from other bacterium.

Acknowledgments. I am indebted to Drs. G Zaccai, JC Willison, and P Forterre for stimulating discussions. A special acknowledgement is due to D Navizet of the Ecole Polytechnique, Palaiseau, for participation in this work during his student-training period in the laboratory.

References

Anderson SA, Carter V, Hagen CB, Parsons M (1998) Molecular cloning of the glycosomal malate dehydrogenase of *Trypanosoma brucei.* Mol Biochem Parasitol 96:185–189

Anderson SGE, Kurland CG (1998) Reductive evolution of resident genome. Trends Microbiol 6:263–268

Andersson SGE, Zomorodipour A, Andersson JO, Sicheitz-Pontén T, Alsmark UC, Podowski RM, Näslund K, Erikson AS, Winkler H, Kurland CG (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396:133–140

Auerbach G, Ostendorp R, Prade L, Kordorfer I, Dams T, Huber R, Jaenicke R (1998) Lactate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima:* the crystal structure at 2.1 Å resolution revels strategies for intrinsic protein stabilization. Structure 6:769–781

Banaszak LJ, Bradshaw RA (1975) Malate dehydrogenase. In: Boyer PDE (ed.) The enzymes, 3rd ed. Academic Press, New York, pp. 369–396

Birkoft JJ, Fernley RT, Bradshaw RA, Banaszak LJ (1982) Amino acid sequence homology among the 2-hydroxyacid dehydrogenase: mitochondrial and cytoplasmique malate dehydrogenase form a homologous system with lactate dehydrogenase. Proc Natl Acad Sci USA 79:6166–6170

Boernke WE, Sanville-Millard C, Wilkins-Stevens P, Kakar SN, Stevens FJ, Donnelly MI (1995) Stringency of substrate specificity of *Escherichia coli* malate dehydrogenase. Arch Biochem Biophys 332:43–52

Bonen L, Cunningham RS, Gray MW, Doolittle WF (1977) Wheat embryo mitochondrial 16S ribosomal RNA: evidence for its procaryotic nature. Nucleic Acids Res 4:663–671

Bonneté F, Ebel C, Zaccai G, Eisenberg H (1993) A biophysical study of halophilic malate dehydrogenase in solution: revised subunit structure and solvent interactions in native and recombinant enzyme. J Chem Soc Faraday Trans 89:2659–2666

Brinkmann H, Martin W (1996) Higher-plant chloroplast and cytosolic 3-phosphoglycerate kinases: a case of endosymbiotic gene replacement. Plant Mol Biol 30:65–75

Bzik DJ, Fox BA, Gonyer K (1993) Expression of *Plasmodium falciparum* lactate dehydrogenase in *Escherichia coli.* Mol Biochem Parasitol 59:155–166

Carr PD, Verger D, Ashton AR, Ollis DL (1999) Chloroplast NADP-malate dehydrogenase: structural basis of light-dependent regulation of activity by thiol oxidation and reduction. Structure 7:461–475

Cazzulo Franke MC, Vernal J, Cazzulo JJ, Nowicki C (1999) The NAD-linked aromatic alpha-hydroxy acid dehydrogenase from *Trypanosoma cruzi.* A new member of the cytosolic malate dehydrogenases group without malate dehydrogenase activity. Eur J Biochem 266:903–910

Cendrin F, Chroboczek J, Zaccaï G, Eisenberg H, Mevarech M (1993) Cloning, sequencing and expression in *Escherischia coli* of the gene coding for malate dehydrogenase of the extremely halophilic archaebacterium *Haloarcula marismortui.* Biochemistry 32:4308–4313

Chapman ADM, Cortés A, Dafforn TR, Clarke AR, Brady RL (1999) Structural basis of substrate specificity in malate dehydrogenases: crystal structure of a ternary complex of porcine cytoplasmic malate dehydrogenase, α-ketomalonate and tetrahydoNAD. J Mol Biol 285:703–712

Charron C, Talfournier F, Isupov MN, Littlechild JA, Branlant G, Vitoux B, Aubry A (2000) The crystal structure of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic archaeon *Methanothermus fervidus* in the presence of NADP(+) at 2.1 Å resolution. J Mol Biol 297:481–500

Chen J, Smith DL (2001) Amide hydrogen exchange shows that malate dehydrogenase is a folded monomer at pH 5. Protein Sci 10:1079–1083

Cordwell SJ (1999) Microbial genomes and "missing" enzymes: redefining biochemical pathways. Arch Microbiol 172:269–279

Cordwell SJ, Basseal DJ, Pollack JD, Humphery-Smith I (1997) Malate/lactate dehydrogenase in mollicutes: evidence for a multienzyme protein. Gene 195:113–120

Danson M (1993) Central metabolism. In: Kates M, Kushner D, Matheson AT (eds.) The biochemistry of archaea (archaebacteria). Elsvier Biomedical Press, Amesterdam, pp. 1–24

Danson MJ (1988) Archaebacteria: the comparative enzymology of their central metabolic pathways. Adv Microb Physiol 29:165–231

Dickerson RE (1971) The structure of cytochrome c and the rates of molecular evolution. J Mol Evol 1:26–45

Dunn CR, Banfield, MJ, Barker, JJ, Higham, CW, Moreton, KM, Turgut-Balik D, Brady L, Holbrook JJ (1996) The structure of lactate dehydrogenase from *Plasmodium falciparum* reveals a new target for anti-malarial design. Nat Struct Biol 11:912–915

Dzierszinski F, Popescu O, Toursel C, Slomianny C, Yahiaoui B, Tomavo S (1999) The protozoan parasite *Toxoplasma gondii* expresses two functional plant-like glycolytic enzymes. Implications for evolutionary origin of apicomplexans. J Biol Chem 274:24888–24895

Eventoff W, Rossmann MG, Taylor SS, Torff HJ, Meyer H, Keil W, Kiltz HS (1977) Structural adaptation of lactate dehydrogenase isozymes. Proc Natl Acad Sci USA 74:2677–2681

Fast NM, Kissinger JC, Roos DS, Keeling PJ (2001) Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. Mol Biol Evol 18:418–426

Felsenstein J (1989) PHYLIP-phylogeny inference package (version 3.2) Cladistics 5:164–166

Figge RM, Schubert M, Brinkmann H, Cerff R (1999) Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eu-

karyotes: evidence for intra- and inter-kingdom gene transfer. Mol Biol Evol 16:429–444

Frazer CM, Gocayne J, White O, Adams, MD, Clayton R, Fleischmann RD, Bult CJ, Galperin MY, Koonin EV (1999) Functional genomics and enzyme evolution. Homologous and analogous enzymes encoded in microbial genomes. Genetica 106:159–170

Girg R, Rudolph R, Jaenicke R (1983) The dimeric intermediate on the pathway of reconstitution of lactate dehydrogenase is enzymatically active. FEBS Lett 163:132–132

Garvie EL (1980) Bacterial lactate dehydrogenase. Microbiol Rev 44: 106–139

Gascuel O (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol 14:685–695

Golding GB, Dean AM (1998) The structural basis of molecular adaptation. Mol Biol Evol 15:355–369

Goward RC, Nicholls DJ (1994) Malate dehydrogenase: a model for structure, evolution and catalysis. Prot Sci 3:1883–1888

Grandier-Vazeille X, Bathany K, Chaignepain S, Camougrand N, Manon S, Schmitter JM (2001) Yeast mitochondrial dehydrogenases are associated in a supramolecular complex. Biochemistry 40:9758–9769

Graupner M, Xu H, White RH (2000) Identification of an archaeal 2-hydroxy acid dehydrogenase catalyzing reactions involved in coenzyme biosynthesis in methanoarchaea. J Bacteriol 182:3688–3692

Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. Science 283:1476–1481

Gray MW, Burger G, Lang BF (2001) The origin and early evolution of mitochondria. Genome Biol 2:10181–10185

Griffin HG, Swindell SR, Gasson MJ (1992) Cloning and sequence analysis of the gene encoding L-lactate dehydrogenase from *Lactococcus lactis:* evolutionary relationships between 21 different LDH enzymes. Gene 122:193–197

Hall MD, Levitt DG, Banaszak L (1992) Crystal structure of *Escherichia coli* malate dehydrogenase. J Mol Biol 226:867–882

Hartl T, Grossebuter W, Gorish H, Stezowski JJ (1987) Crystalline NAD/NADP-dependent malate dehydrogenase; the enzyme from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius.* Biol Chem Hoppe-Seyler 368:259–267

Hecht K, Langer T, Wrba A, Jaenicke R (1990) Lactate dehydrogenase from the extreme halophilic archaebacterium *Halobacterium marismortui.* Biol Chem Hoppe-Seyler 372:515–519

Henze K, Badr A, Wettern M, Cerff R, Martin W (1995) A nuclear gene of eubacterial origin in *Euglena gracilis* reflects cryptic endosymbioses during protist evolution. Proc Natl Acad Sci USA 92: 9122–9126

Henze K, Morrison HG, Sogin ML, Muller M (1998) Sequence and phylogenetic position of a class II aldolase gene in the amitochondriate protist, *Giardia lamblia.* Gene 222:163–168

Hogan JK, Pittol, CA, Jones JB, Gold M (1995) Improved specificity toward substrates with positively charged side chain by sie-directed mutagenesis of the L-lactate dehydrogenase of *Bacillus stearothermophilus.* Biochemistry 34:4225–4230

Horner DS, Hirt RP, Kilvington S, Lloyd D, Embley TM (1996) Molecular data suggest an early acquisition of the mitochondrion endosymbiont. Proc R Soc Lond B Biol Sci 1373:1053–1059

Hunter GR, Hellman U, Cazzulo JJ, Nowicki C (2000) Tetrameric and dimeric malate dehydrogenase isoenzymes in *Trypanosoma cruzi* epimastigotes. Mol Biochem Parasitol 105:203–214

Huynen MA, Dandekar T, Bork P (1999) Variation and evolution of the citric acid cycle: a genomique perspective. Trends Microbiol 7: 281–291

Iwabe N, Kuma KI, Hasegawa M, Osawa S, Miyata T (1989) Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc Natl Acad Sci USA 86:9355–9359

Jackson RM, Gelpi JL, Cortes A, Emery DC, Wilks H, Moreton KM, Halsall DJ, Sleigh RN, Behan-Martin M, Jones GR, Clarke AR, Holbrook JJ (1992) Construction of a stable dimer of *Bacillus stearothermophilus* lactate dehydrogenase. Biochemistry 31:8307–8314

Jaenicke R (1998) Oligomeric proteins. In: Fink, Goto (eds.) Molecular chaperones in the life cycle of proteins. Marcel Decker, New York, pp. 35–70

Karlberg O, Canbäck B, Kurland CG, Andersson SGE (2000) The dual origin of the yeast mitochondrial proteome. Yeast 17:170–187

Karlin S, Brocchieri L, Mrazek J, Campbell AM, Spormann AM (1999) A chimeric prokaryotic ancestry of mitochondria and primitive eukaryotes. Proc Natl Acad Sci USA 96:9190–9195

Kengen SWM, De Bok F, De Vos WM (1994) Evidence for the operation of a novel Embden-Meyerhof pathway that involves ADP-dependent kinases during sugar fermentation by *Pyrococcus furiosus.* J Biol Chem 269:17357–17541

Kerlavage AR, Sutton G, Kelley JM (1995) The minimal gene complement of *Mycoplasma genitalium.* Science 270:397–403

Köhler S, Delwiche CF, Denny PW, Tilney LG, Webster P, Wilson RJ, Palmer JD, Roos DS (1997) A plastid of probable green algal origin in Apicomplexan parasites. Science 275:1485–1489

Koonin EV, Mushegian AR, Bork P (1996) Non-orthologous gene displacement. Trends Gen 12:334–336

Kurland CG, Anderson SGE (2000) Origin and evolution of the mitochondrial proteome. Microbiol Mol Biol Rev 64:786–820

Langelandsvik AS, Steen IH, Birkeland NK, Lien T (1997) Properties and primary structure of a thermostable L-malate dehydrogenase from *Archaeoglobus fulgidus.* Arch Microbiol 168:59–67

Lee BI, Chang C, Cho SJ, Eom SH, Kim KK, Yu YG, Suh SW (2001) Crystal structure of the MJ0490 gene product of the hyperthermophilic archaebacterium *Methanococcus jannaschii,* a novel member of the lactate/malate family of dehydrogenases. J Mol Biol 307: 1351–1562

Liaud MF, Lichtlé C, Apt K, Martin W, Cerff R (2000) Compartment specific isoform of TPI and GAPDH are imported into diatom mitochondria as a fusion protein/evidence in favor of a mitochondrial origin of the eukaryotic glycolytic pathway. Mol Biol Evol 17:213–223

Madern D (2000) The putative L-lacatate dehydrogenase from *Methanococcus jannaschii* is an NADPH-dependent L-malate dehydrogenase. Mol Microbiol 37:1515–1520

Madern D, Ebel C, Dale HA, Lien T, Steen IH, Birkeland, NK, Zaccai G (2001) Differences in the oligomeric states of the [LDH-like] L-MalDH from the hyperthermophilic archaea *Methanococcus jannaschii* and *Archaeoglobus fulgidus.* Biochemistry 40:10310–10316

Madern D, Ebel C, Mevarech M, Richard SB, 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

Makarova KS, Aravind LA, Galperin MY, Grishin NV, Tatusov RL, Wolf YI, Koonin EV (1999) Comparative genomics of the archaea (euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. Genome Res 9:608–628

Martin W, Herrmann RG (1998) Gene transfert from organelles to the nucleus: how much, what happens, and why? Plant Physiol 118:9–17

Martin W, Müller M (1998) The hydrogen hypothesis for the first eukaryote. Nature 392:37–41

McAlister-Henn L (1988) Evolutionary relationships among the malate dehydrogenase. Trends Biochem Sci 13:178–181

Morgunov I, Srere PA (1998) Interaction between citrate synthase and malate dehydrogenase. Substrate channeling of oxaloacetate. J Biol Chem 273:29540–29545

Muller T, Vingron M (2000) Modeling amino acid replacement. J Comput Biol 7:761–776

Munkund S, Adams MWW (1995) Glyceraldehyde-3-phosphate ferredoxin oxidoreductase, a novel tungsten containing enzyme with a potential glycolytic role in the hyperthermophilic archaeon *Pyrococcus furiosus.* J Biol Chem 270:8389–8392

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 *Escherischia coli.* J Bact 178:7047–7052

Niefind K, Hecht HJ, Schomburg D (1995) Crystal structure of L-2-hydroxyisocaproate dehydrogenase from *Lactobacillus confusus* at 2.2 Å resolution. An example of strong asymmetry between subunits. J Mol Biol 251:256–281

Ocheretina O, Haferkamp I, Tellioglu H, Scheibe R (2000) Light-modulated NADP-malate dehydrogenases from mossfern and green algae: insights into evolution of the enzyme's regulation. Gene 258:147–154

Ovadi J, Srere PA (2000) Macromolecular compartmentation and channeling. Int Rev Cytol 192:255–280

Perrière G, Gouy M (1996): WWW-Query: an on-line retrieval system for biological sequence banks. Biochimie 78:364–369

Phibbs PV, Winkler HH (1982) Regulatory properties of citrate synthase from *Rickettsia prowazekii.* J Bact 149:718–725

Philippe H, Forterre P (1999) The rooting of the universal tree of life is not reliable. J Mol Evol 49:509–523

Richard SB, Madern D, Garcin E, Zaccai G (2000) Halophilic adaptation: novel solvent protein interactions observed in the 2.9 Å and 2.6 Å resolution structures of the wild type and a mutant of malate dehydrogenase from *Haloarcula marismortui.* Biochemistry 39: 992–1000

Roger AJ, Morrison HG, Sogin ML (1999) Primary structure and phylogenetic relationships of a malate dehydrogenase gene from *Giardia lamblia.* J Mol Evol 6:750–755

Roger AJ, Svard SG, et al. (1998) A mitochondrial-like chaperonin 60 gene in *Giardia lamblia:* evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. Proc Natl Acad Sci USA 95:229–234

Romano AH, Conway T (1996) Evolution of carbohydrate metabolic pathways. Res Microbiol 147:448–455

Ruepp A, Graml W, Santos-Martinez ML, Koretke KK, Volker C, Mewes HW, Frishman D, Stocker S, Lupas AN, Baumeister W (2000) The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum.* Nature 407:508–513

Scheritz-Pontén T, Kurland CG, Andersson SGE (1998) A phylogenetic analysis of the cytochrome b and cytochrome c oxidase I genes supports an origin of mitochondria from within the *Rickettsiaceae.* Biochem Biophys Acta 1365:545–551

Selig A, Xavier KB, Santos H, Schönheit P (1997) Comparative analysis of Embden-Meyerhof and Entner-Doudoroff glycolytic pathways of hyperthermophilic archaea and the bacterium *Thermotoga.* Arch Microbiol 167:217–232

Shatalin K, Lebreton S, Rault-Leonardon M, Velot C, Srere PA (1999) Electrostatic channeling of oxaloacetate in a fusion protein of porcine citrate synthase and porcine mitochondrial malate dehydrogenase. Biochemistry 38:881–889

Steffan JS, McAllister-Henn L (1991) Structural and functionnal effects of mutations altering the subunit interface of mitochondrial malate dehydrogenase. Arch Biochem Biophys 287:276–282

Stock DW, Quattro JM, Whitt GS, Powers DA (1997) Lactate dehydrogenase (LDH) gene duplication during chordate evolution: the cDNA sequence of the LDH of the tunicate *Styela plicata.* Mol Biol Evol 14:1273–1284

Strimmer K, von Haeseler A (1997) Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. Proc Natl Acad Sci USA 94:6815–6819

Sun-Yong K, Kwang YH, Sung-Hou K, Ha-Chin S, Ye Sun H, Ynje C (1999) Structural base for cold adaptation. J Biol Chem 272:11761–11767

Sundaram TK, Wright IP, Wilkinson AE (1980) Malate dehydrogenase from thermophilic and mesophilic bacteria. Molecular size, subunit structure, amino acid composition, immunochemical homology and catalytic activity. Biochemistry 19:2017–2022

Synstad B, Emmerhoff O, Sirevag R (1996) Malate dehydrogenase from the green gliding bacterium *Chloroflexus aurantiacus* is phylogenetically related to lactic dehydrogenases. Arch Microbiol 165: 346–353

Taguchi H, Ohta T (1991) D-lactate dehydrogenase is a member of the D-isomer-specific 2-hydroxyacid dehydrogenase family. Cloning, sequencing, and expression in *Escherichia coli* of the D-lactate dehydrogenase gene of *Lactobacillus plantarum.* J Biol Chem 266: 12588–12594

Tayeh MA, Madigan MT (1987) Malate dehydrogenase in phototrophic purple bacteria: purification, molecular weight, and quaternary structure. J Bact 169:4196–4202

Thompson H, Tersteegen A, Thauer RK, Hedderich R (1998) Two malate dehydrogenases in *Methanobacterium thermoautotrophicum.* Arch Microbiol 170:38–42

Towe KM (1990) Aerobic respiration in the archaean? Nature 348:54–56

Tsuji S, Qureshi MA, Hou EW, Fitch WM, Li SSL (1994) Evolutionary relationships of lactate dehydrogenases (LDHs) from mammals, birds, an amphibian, fish, barley, and bacteria: LDH cDNA sequences from *Xenopus,* pig, and rat. Proc Natl Acad Sci USA 91:9392–9396

Uttaro AD, Altabe SG, Rider MH, Michels PA, Opperdoes FR (2000) A family of highly conserved glycosomal 2-hydroxyacid dehydrogenases from Phytomonas sp. J Biol Chem 275:31833–31837

Vellot C, Mixon, MB, Teige M, Srere PA (1997) Model of a quinary structure between Krebs TCA cycle enzymes: a model for the metabolon. Biochemistry 36:14272–14276

Velot C, Srere PA (2000) Reversible transdominant inhibition of a metabolic pathway. In vivo evidence of interaction between two sequential enzyme tricarboxylic acid cycle enzymes in yeast. J Biol Chem 275:12926–12933

Wigley DB, Gamblin SJ, Turkenburg JP, Dobson, EJ, Piontek K, Muirhead HM, Holbrook JJ (1992) Structure of a ternary complex of an allosteric lactate dehydrogenase from *Bacillus stearothermophilus.* J Mol Biol 223:317–335

Wilks HM, Hart KM, Feeney R, Dunn CR, Muirhead H, Chia WN, Barstow DA, Atkinson T, Clarke AR, Holbrook JJ (1988) A specific, highly active malate dehydrogenase by redisgn of a lactate dehydrogenase framework. Science 242:1541–1544

Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains archaea, bacteria, and eukarya. Proc Natl Acad Sci USA 87:4576–4579

Wood DC, Jurgensen SR, Geesin, JC, Harrison JH (1981) Subunit interactions in mitochondrial malate dehydrogenase. J Biol Chem 256:2377–2382

Wood DG, Atkinson, WH, Sikorksi RS, Winkl K (1983) Expression of the *Rickettsia prowazekii* citrate synthase gene in *Escherichia coli.* J Bact 155:412–416

Wrba A, Jaenicke R, Huber R, Stetter, KO (1990) Lactate dehydrogenase from the extreme thermophile *Thermotoga maritima.* Eur J Biochem 188:195–201

Wu G, Fiser A, Ter Kuile B, Sali A, Muller M (1999) Convergent evolution of *Trichomonas vaginalis* lactate dehydrogenase from malate dehydrogenase. Proc Natl Acad Sci USA 96:6285–6290

Wynne SA, Nicholls, DJ, Scawen, MD, Sundaram TK (1996) Tetrameric malate dehydrogenase from a thermophilic *Bacillus:* cloning, sequence and overexpression of the gene encoding the enzyme and isolation and characterisation of the recombinant enzyme. Biochem J 317:235–245