Volume 247, number 2, 259–262 FEB 07036 April 1989

Archaebacterial malate dehydrogenase: the amino-terminal sequence of the enzyme from *Sulfolobus acidocaldarius* is homologous to the eubacterial and eukaryotic malate dehydrogenases

Helmut Görisch and Klaus-Dieter Jany*

Institut für Mikrobiologie der Universität Hohenheim, Garbenstr. 30, D-7000 Stuttgart 70 and **Institut für Biochemie der Technischen Hochschule Darmstadt, Petersenstr. 22, D-6000 Darmstadt, FRG*

Received 14 February 1989

42 residues of the N-terminal amino acid sequence of malate dehydrogenase from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius* have been determined as VKVAFIGVGRGVGQTIAYNTIVNGYADEVMLYDVV-PELPTKK. In eubacterial and eukaryotic enzymes this region is known to encompass residues involved in pyridine nucleotide binding. In the archaebacterial enzyme the residues Gly-7, Gly-11 and Asp-33 are also present. The data suggest that in the enzyme from *S. acidocaldarius* like in the other malate dehydrogenases the binding domain for NAD(H) is localized at the N-terminal part of the polypeptide chain. The archaebacterial enzyme is homologous to the other malate dehydrogenases, of which the amino acid sequences are known, however, it is only distantly related to the mitochondrial/*E. coli* group and the cytosolic/*Thermus flavus* group.

Malate dehydrogenase; Archaebacteria; N-terminal sequence; NAD-binding domain; (*Sulfolobus acidocaldarius*)

1. INTRODUCTION

Malate dehydrogenase catalyzes the pyridine nucleotide-dependent interconversion of malate and oxaloacetate. In eukaryotic cells a mitochondrial and a cytoplasmic form of the enzyme is found. The compartmentalized isozyme pairs have been isolated from mouse and pig and their amino acid sequences are known [1]. In addition amino acid sequence data are available for the mitochondrial malate dehydrogenases from rat [1] and yeast [2] and for the single prokaryotic enzymes from *E. coli* [3] and *Thermus flavus* [4]. With the recognition of the archaebacteria as a phylogenetically distinct lineage, it is important to extend studies of the molecular diversity of malate dehydrogenases to this group. Recently we purified and crystallized malate dehydrogenase from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius* [5]. Here we report the amino acid sequence of the N-terminus of the archaebacterial enzyme and compare it with the sequences of the mitochondrial and cytoplasmic forms from eukaryotes and with the known eubacterial enzymes.

Dedicated to Professor F. Radler on the occasion of his 60th birthday

Correspondence address: H. Görisch, Institut für Mikrobiologie, Universität Hohenheim, Garbenstr. 30, D-7000 Stuttgart 70, FRG

2. EXPERIMENTAL

Sulfolobus acidocaldarius (DSM 639) was grown aerobically at 70°C and pH 2.2 as described previously [6]. Malate dehydrogenase was purified as reported [5]. The purified enzyme was homogeneous as proved by SDS-polyacrylamide gel electrophoresis and dansylation. Prior to sequencing the protein was exhaustively dialyzed against 20 mM potassium phosphate, pH 7.0, and then lyophilized. 130 pmol protein was used for the sequence determination on the Applied Biosystems gas-phase

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

sequencer (model 470 A) equipped with an on-line HPLC system (Applied Biosystems model 120 A) for the identification of the PTH-amino acid derivatives. With slight modifications the sequencing cycles were run on the manufacturer's 03R PTH program.

3. RESULTS

The N-terminal amino acid sequence of malate dehydrogenase from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius* has been determined. The sequence of 42 residues is shown in table 1 and fig.1. A computer-aided comparison of this amino-terminal sequence with NAD- and FAD-dependent oxido-reductases using the algorithms of Lipman and Pearson [7] revealed a striking homology in sequences near the N-termini. This part is known to build up the $\beta\alpha\beta$ structural element that interacts with the pyrophosphate moiety of the coenzyme [8]. The alignment of the corresponding sequences of the eubacterial malate dehydrogenases and the mitochondrial and cytosolic forms from eukaryotes with that of the archaebacterial malate dehydrogenase is shown in table 1. Optimal alignment of all malate dehydrogenases shown in table 1 requires two large gaps in the enzymes from *S. acidocaldarius, E. coli* and mitochondria. When the cytosolic malate dehydrogenases and the enzyme from *Thermus flavus* are omitted only three gaps are required for alignment, each involving one residue. The archaebacterial malate dehydrogenase is more similar to the *E. coli* enzyme with identities at 26% of the residue positions known compared to the mitochondrial proteins with identities at 22% of the residue positions. An even lower similarity is observed with the cytosolic protein and the enzyme from *T. flavus*, with identities at 14% of the residue positions.

In table 2 a correlation matrix is given for all N-terminal sequences of the known malate dehydrogenases. The matrix shows homologies based on identical residue positions and also homologies calculated by counting the conservative exchanges Gly-Ala, Val-Leu-Ile, Asp-Glu, Asn-Gln and Arg-Lys as invariant residues. Both calculations show three distinct groups of malate dehydrogenases. The cytosolic malate dehydrogenase from mouse and the enzyme from *T. flavus* show a high degree of homology as is the case with the mitochondrial malate dehydrogenases and the enzyme from *E.*

Table 1

Comparison of the amino-terminal sequences of malate dehydrogenase from different sources

```
                                  7      11                                                      33
                                N N   N N N N                                                   N N N
                                       10                  20                              30                      40
S. acidocal.          V K V A F I G V G R G V G Q T I A Y N T I V N G Y A - - - - - - D E V M L Y D V V - - - P E L P T K K

                                       10                  20                              30                        40
mouse mit             A K V A V L G A S G G I G Q P L S L L L K - N S P L - - - - - V S R L T L Y D I A - - - - H T P G V A
yeast mit             Y K V T V L G A G G G I G Q P L S L L L K L N - H K - - - - - V T D L R L Y D L K - - - - G A K G V A
                                       10                  20                            30                        40
E. coli               M K V A V L G A A G G I G Q A L A L L L K T Q L P S - - - - - G S E L S L Y D I A - - - P V T P G V A

                               10                  20                    30                  40                  50
mouse cyt     M S C P I R V L V T G A A G Q I A Y S L L Y S I G - N G S V F G K D Q P I I L V L L D I T P M M G V L D G V L

T. flavus     M K A P V R V A V T G A A G Q I G Y S L L F R I A A - G E N L G K D Q P V I L Q L L E I P Q A M K A L E G V V
```

Dashes represent arbitrary gaps. Mitochondrial and cytosolic enzymes are indicated by mit and cyt. The N-terminal sequences of the mitochondrial enzymes from mouse, pig and rat [1] and the N-terminal sequences of the cytosolic enzymes from mouse and pig are almost identical [1,9,11,15]. The amino acid residues in boxes indicate chemically homologous but not identical residues with those in the *S. acidocaldarius* sequence. G/A, V/L/J, D/E, N/Q and K/R have been regarded as conservative exchanges. Specific residues thought to be involved in NAD(H) binding on the basis of X-ray crystallographic analysis of pig malate dehydrogenase [10,14] are indicated by N

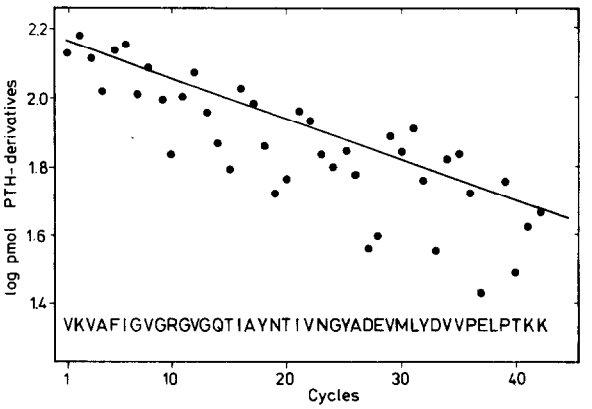

Fig.1. N-terminal gas-phase degradation of malate dehydrogenase (approx. 130 pmol) from *Sulfolobus acidocaldarius*. The semilogarithmic plot shows the yields of the PTH-derivatives in pmol per injection. The regression line corresponds to a mean repetitive yield of 96%.

coli, while malate dehydrogenase from *S. acidocaldarius* is characterized by significant lower sequence identities to both other groups of enzymes.

In table 1 residues are indicated, which are known to be involved in NAD(H) binding in mitochondrial malate dehydrogenase [9]. The amino acid residues Gly-7, Gly-11 and Asp-33 are thought to be critical [10]. These residues are also conserved in the archaebacterial enzyme. In addition the regions around these residues contain most of the identical residues and of the conservative exchanges within all of the archaebacterial, eubacterial and eukaryotic malate dehydrogenases, while the region between residues 22 and 29 shows a high degree of variability.

Table 2

Correlation matrix comparing the amino-terminal sequences of malate dehydrogenases from different sources

	Mouse c	*T. flavus*	Mouse m	Yeast m	*E. coli*	*S. acidoc.*
Mouse c	x	73	34	34	34	30
T. flavus	59	x	34	36	34	26
Mouse m	28	28	x	56	58	32
Yeast m	26	24	54	x	50	36
E. coli	30	30	58	46	x	38
S. acidoc.	14	14	22	22	26	x

The lower left side of the matrix shows the number of identical residues. Deletions have been treated as nonidentities. The upper right side shows the sum of the number of identical residues plus the number of conservative exchanges. Residues 5 to 54 based on counting of mouse cytosolic malate dehydrogenase have been compared

4. DISCUSSION

The striking similarity between the amino acid sequence at the N-terminal end of pig mitochondrial and *E. coli* malate dehydrogenase was first reported by Fernley et al. [11]. This similarity was regarded as an argument supporting the endosymbiotic theory for the origin of mitochondria [9,11]. Recently it was shown however, that malate dehydrogenase from another eubacterium, the thermophilic *T. flavus*, shows a striking similarity with the amino acid sequence of the cytosolic form of the pig enzyme [1]. Thus for each of the eukaryotic malate dehydrogenases, the mitochondrial and the cytosolic form, there exists a eubacterial enzyme, which is closely homologous to it.

We determined the sequence of a stretch of 42 amino acid residues at the N-terminus of malate dehydrogenase, isolated from the thermoacidophilic archaebacterium *S. acidocaldarius* [5]. The archaebacteria comprise extreme halophilic, methanogenic, thermophilic and sulfur-dependent thermophilic microorganisms [12] and represent a third form of life, in addition to the well-known eubacteria and eukaryotes [13].

The archaebacterial protein represents a third form of malate dehydrogenase, as is evident from the correlation matrix comparing the mitochondrial/*E. coli* enzymes (group 1) and the cytosolic/*T. flavus* enzymes (group 2) with the enzyme from the archaebacterium *S. acidocaldarius*. The archaebacterial malate dehydrogenase shows a lower degree of homology with both groups of enzymes, than the group 1 and group 2 malate dehydrogenases between each other (table 2).

The quantitative aspects shown in table 2 are derived from N-terminal segments which represent about 14% of the respective polypeptide chains and therefore will vary somewhat from those calculated from a comparison of the complete sequences. The main conclusion, however, that the archaebacterial enzyme represents a distinct form is unlikely to be altered. The homology data reported herein are basically the same as those

reported for malate dehydrogenase from *T. flavus* and mouse mitochondria and cytosol based on complete sequence data [1].

Malate dehydrogenase from the archaebacterium *S. acidocaldarius* is homologous to the eubacterial and eukaryotic enzymes. Residues Gly-7, Gly-11 and Asp-33, which are known to be critical for NAD(H) binding in the mitochondrial enzymes [10] are also present in the archaebacterial malate dehydrogenase (table 1). In addition most of the residues known to be localized in the hydrophobic pocket of the adenine-binding region for NAD(H) are conserved or occupied by residues which represent conservative exchanges. Thus in all malate dehydrogenases the regions from residues 2 to 21 and around residue 33 contain almost all of the identical residues and the conservative exchanges, while between residues 22 and 29 there is a more variable segment of the polypeptide chain. These data suggest that in the archaebacterial enzyme like in the other malate dehydrogenases the binding domain for NAD(H) is localized at the N-terminal part of the polypeptide chain. The archaebacterial enzyme shows sufficient relatedness with all other malate dehydrogenases to suggest a common ancestral precursor.

Acknowledgements: The authors wish to thank M. Hahn for skillful technical assistance. This work was supported by the Fonds der Chemischen Industrie.

REFERENCES

[1] Joh, T., Takeshima, H., Tsuzuki, T., Setoyama, C., Shimada, K., Tanase, S., Kuramitsu, S., Kagamiyama, H. and Morino, Y. (1987) J. Biol. Chem. 262, 15127–15131.
[2] Thompson, L.M., Sutherland, P., Steffan, J.S. and McAlister-Henn, L. (1988) Biochemistry 27, 8393–8400.
[3] McAlister-Henn, L., Blaber, M., Bradshaw, R.A. and Nisco, S.J. (1987) Nucleic Acids Res. 15, 4993.
[4] Nishiyama, M., Matsubara, N., Yamamoto, K., Iijima, S., Uozumi, T. and Beppu, T. (1986) J. Biol. Chem. 261, 14178–14183.
[5] Hartl, T., Großebüter, W., Görisch, H. and Stezowski, J.J. (1987) Biol. Chem. Hoppe-Seyler 368, 259–267.
[6] Großebüter, W. and Görisch, H. (1985) System. Appl. Microbiol. 6, 119–124.
[7] Lipman, P. and Pearson, W.R. (1985) Science 227, 1435–1441.
[8] Wierenga, R.K., De Maeyer, M.C.H. and Hol, W.G.I. (1985) Biochemistry 24, 1346–1357.
[9] McAlister-Henn, L. (1988) Trends Biochem. Sci. 13, 178–181.
[10] Birktoft, J.J. and Banaszak, L.J. (1984) Pept. Protein. Rev. 4, 1–46.
[11] Fernley, R.T., Lentz, S.R. and Bradshaw, R.A. (1981) Biosci. Rep. 1, 497–507.
[12] Woese, C.R., Gupta, R., Hahn, C.M., Zillig, W. and Tu, J. (1984) Syst. Appl. Microbiol. 5, 97–105.
[13] Woese, C.R. (1982) in: Archaebacteria, Proc. Ist. Internat. Workshop on Archaebacteria, München, 1981 (Kandler, O. ed.) pp.1–17, Fischer, Stuttgart, New York.
[14] Roderick, S.L. and Banaszak, L.J. (1986) J. Biol. Chem. 261, 9461–9464.
[15] Birktoft, J.J., Fernley, R.T., Bradshaw, R.A. and Banaszak, L. (1982) Proc. Natl. Acad. Sci. USA 79, 6166–6170.