the FEBS Journal

Organelle and translocatable forms of glyoxysomal malate dehydrogenase

The effect of the N-terminal presequence

Bryan Cox[1], Ma May Chit[2], Todd Weaver[3], Christine Gietl[4], Jaclyn Bailey[5], Ellis Bell[6] and Leonard Banaszak[1]

1 Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, MN, USA
2 Western University of Health Sciences, Pomona, CA, USA
3 Department of Chemistry, University of Wisconsin La Crosse, WI, USA
4 Institute of Botany, Technical University of Munich, Germany
5 Gustavus Adolphus College, St Peter, MN, USA
6 Department of Chemistry, University of Richmond, VA, USA

Keywords
glyoxysome organelle-precursor; malate dehydrogenase; protein translocation; X-ray diffraction

Correspondence
L. Banaszak, 6–155 Jackson Hall University of Minnesota 321 Church St. S.E. Minneapolis, MN 55455, USA
Fax: +1 612 6245121
Tel: +1 612 6266597
E-mail: banas001@umn.edu

(Received 24 June 2004, revised 2 November 2004, accepted 10 November 2004)

doi:10.1111/j.1742-4658.2004.04475.x

Many organelle enzymes coded for by nuclear genes have N-terminal sequences, which directs them into the organelle (precursor) and are removed upon import (mature). The experiments described below characterize the differences between the precursor and mature forms of watermelon glyoxysomal malate dehydrogenase. Using recombinant protein methods, the precursor (p-gMDH) and mature (gMDH) forms were purified to homogeneity using Ni^{2+}–NTA affinity chromatography. Gel filtration and dynamic light scattering have shown both gMDH and p-gMDH to be dimers in solution with p-gMDH having a correspondingly higher molecular weight. p-gMDH also exhibited a smaller translational diffusion coefficient (D_t) at temperatures between 4 and 32 °C resulting from the extra amino acids on the N-terminal. Differential scanning calorimetry described marked differences in the unfolding properties of the two proteins with p-gMDH showing additional temperature dependent transitions. In addition, some differences were found in the steady state kinetic constants and the pH dependence of the K_m for oxaloacetate. Both the organelle-precursor and the mature form of this glyoxysomal enzyme were crystallized under identical conditions. The crystal structure of p-gMDH, the first structure of a cleavable and translocatable protein, was solved to a resolution of 2.55 Å. GMDH is the first glyoxysomal MDH structure and was solved to a resolution of 2.50 Å. A comparison of the two structures shows that there are few visible tertiary or quaternary structural differences between corresponding elements of p-gMDH, gMDH and other MDHs. Maps from both the mature and translocatable proteins lack significant electron density prior to G44. While no portion of the translocation sequences from either monomer in the biological dimer was visible, all of the other solution properties indicated measurable effects of the additional residues at the N-terminal.

Abbreviations
D_t, translational diffusion coefficient; DSC, differential scanning calorimetry; DLS, dynamic light scattering; ER, endoplasmic reticulum; gMDH, MDH from watermelon glyoxysomes; MDH, malate dehydrogenase; p-gMDH, precursor of watermelon glyoxysomal MDH; PTS1 or PTS2, peroxisomal targeting signal 1 or 2; R_h, equivalent radius (sphere) of hydration.

Organelles are subcellular particles present in all eukaryotic cells, and are typically bounded by one or more protein-containing membranes. In eukaryotic cells, examples may include: nuclei, mitochondria, peroxisomes, chloroplasts and the endoplasmic reticulum (ER) system. With some exceptions, the internalized enzymes and membrane proteins are coded for by nuclear genes and translocated into the organelle from the cytosol [1]. The ER system is relatively unique as translocation appears to most often occur simultaneously with translation. This means that folding occurs in the lumen or membranes of the ER system.

Aside from the ER system, other organelles present in the cell rely on some partially defined recognition and translocation processes. The recognition element directing proteins to the appropriate organelle includes a variety of amino acid sequences, which may be located at the N-terminus, the C-terminus, or internally. Frequently, N-terminal fragments are removed during translocation and consequently the primary structure found in the organelle is different from the precursor form. A number of organelle systems fit this category, and there are translocation sequences on nuclear-coded mitochondrial, glyoxysomal and chloroplast proteins [2].

The assortment of proteins and metabolic enzymes present within these various types of microbodies characterizes their cellular functions. For example, in plants, the distinction is made between glyoxysomes, which are involved in mobilization of stored fat and leaf-type peroxisomes that are involved in photorespiration. The presence and amounts of organelle enzymes for different metabolic pathways may vary depending on species, growth conditions and upon the age of the cells [3].

Because there is overlap in catalytic requirements, there is also overlap in the class of enzymes present within different organelles. For example, malate dehydrogenases (MDHs) are found in mitochondria, glyoxysomes and chloroplasts, with each organelle having a different nuclear gene. It is presently unclear how much of the recognition and translocation processes are common for different organelle-types.

The study described below focuses on a malate dehydrogenase found within glyoxysomes but MDHs are needed wherever the citric acid and glyoxylate cycles are operating. Two peroxisomal targeting signals (PTS) are known to contribute to the import of matrix proteins into peroxisomes and glyoxysomes. PTS1 is a noncleavable, C-terminal tripeptide, generally having the amino acid sequence -S-K-L, while PTS2s are at the N-terminus and is generally removed upon import [4]. Like other extended N-terminal sequences, p-gMDH is cleaved upon import into peroxisomes and glyoxysomes in higher eukaryotes such as mammals and plants but not in lower eukaryotes such as yeast [5].

There are but a few crystal structures of translocatable proteins. Yeast thiolase, the last enzyme in the peroxisomal β-oxidation pathway contains a noncleavable PTS2 sequence on its N-terminus [6]. The crystal structure has been reported, but no visible electron density was found for the first 27 amino acids [6]. A crystallographic structure of aldolase from *Trypanosoma* glycosomes, a type of peroxisome, which contain the glycolytic enzymes, has also been obtained. Like yeast thiolase, the aldolase has a noncleavable peroxisomal targeting sequence [7]. Unlike yeast thiolase in the crystal structure of this tetrameric aldolase, the targeting segment forms an important part of a subunit interface.

In order to compare the properties of translocatable proteins and their processed forms, we have undertaken a study of the MDHs. As noted above, there exist unique isoforms of MDH for glyoxysomes, mitochondria and chloroplasts, and in many species there is also a cytosolic form [8]. Examples of the N-terminal presequences of organelle isoforms of MDH showing mainly their translocation segments are shown in Fig. 1. The alignment in Fig. 1 is based on the consensus sequence of the NAD^+ binding domains: –VLGAAGGIGQP–. This omnipresent amino acid sequence is a reliable common feature identifying the beginning of the coenzyme-binding domain. In general, the processed MDH subunits are very similar in terms of the start site of the N-termini. The precursor form, p-gMDH, has a 37-residue signal motif at the N-terminus and after

```
MGQPIPDVNQR IARISAHLHP PKSQMEESSA LRRANC***RA***KG GAPGFKVAIL GAAGGIGQ GLYOX.  WATMEL
          MKASILRSVR SAVSRSSSSN RLLSRS***FA***TE SVPERKVAVL GAAGGIGQ MITO.   WATMEL
                     MLS RVAKRAFSST VAN***PY***KVTVL GAGGGIGQ MITO.   YEAST
                                      MKVAVL GAAGGIGQ          E. Coli
                                   MSEPIRVLVT GAAGQIAY CYTO.   PIG
                  MLSALARPVG AALRRSFSTS AQN***NA***KVAVL GASGGIGQ MITO.   RAT
25-RPRRALL ATVR***CS***VDAA KCFGVFCTTY DLKAEDKTKS WKKLVNIAVS VAAVMISN CHLOR.  SORGHUM
```

Fig. 1. N-Terminal amino acid sequence of MDHs. The amino acid sequences at the N-termini of the MDHs from different organelles, and for a cytosolic and prokaryote enzyme are aligned using the NAD^+ binding consensus sequence -G-A-A-G-G-I-G-. The amino acids shown in italic/bold mark the position of the proteolytic cleavage site that commonly occurs upon organelle uptake. Glyox, enzyme derived from glyoxysomes; mito, MDHs found within mitochondria; chlor, an example of an MDH derived from chloroplasts; cyto, a form of cytosolic MDH. For the sake of brevity, 25 amino acids are missing from the N-terminal of the chloroplast enzyme derived from sorghum [26].

import, the proteolytically processed or organelle gMDH consists of 319 amino acids [9].

In the known structures of MDH, the N-termini of the mature form of the protein encompass a β-strand, followed by a turn and the beginning of an α-helix with this segment nestling close to the bound coenzyme in the holo-form. As it has no extra amino acids on the β-strand that starts the consensus structure (Fig. 1, *Escherichia coli*), the prokaryotic form of MDH may be regarded as the minimal fold required for the catalytic function. Note there is also a cytosolic form of MDH listed in Fig. 1. Like the prokaryotic form, it has essentially no N-terminal extension. The study described below evaluates the impact of the additional amino acids. We report the X-ray structure of the precursor and mature forms of the novel glyoxysomal MDH as well as the steady state parameters, the quaternary structure and the thermal stability of the two forms.

Results

The experiments described here focus on establishing the similarities and dissimilarities between gMDH and the organelle translocatable form, p-gMDH. Significant differences were found in the overall stability and the unfolding mechanism. On the other hand the catalytic parameters differed only slightly and no major conformational differences were observed in the X-ray crystal structures.

Purification and characterization of p-gMDH and gMDH

Approximately 2 mg of either gMDH or p-gMDH were obtainable from 1 g of *E. coli* cells (JM105) containing the appropriate plasmid. Three independent experiments were used to characterize the translocatable form of the protein. The results are summarized in Fig. 2. In the early stages of this study, many preparations showed evidence of proteolytic modification. A typical example is shown in the insert of Fig. 2. In this preparation represented by lanes 5–8, the presence of a proteolytic contaminant is visible especially in the lanes with higher concentrations of protein.

The MALDI-TOF mass spectrum also shown in Fig. 2 was obtained from a preparation that was free of any proteolytic cleavage product. A single species is present at an *m/z* ratio of 38 615 ($MW_{calculated}$ 38 517). N-Terminal sequencing of the protein prior to crystallization and after diffraction data collection, also verified the presence of the presequence.

To compare the quaternary structure of the two forms of the enzyme, dynamic light scattering was used

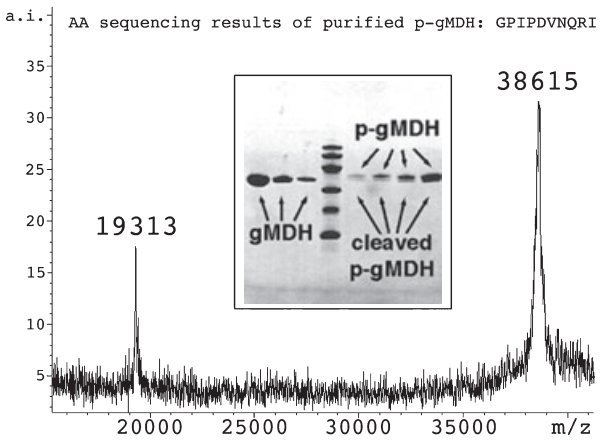

Fig. 2. Characterization of recombinant p-gMDH and gMDH. Three methods were used to demonstrate the presence of the N-terminal translocation segment in preparations of the recombinant p-gMDH. In this composite, data is shown for samples free of proteolytic changes and in the inset an example of an SDS/PAGE experiment on a purified sample that was not useful since some proteolysis had occurred during the purification. The graph represents a MALDI-TOF mass spectrum of purified p-gMDH free of any proteolytic modification. Theoretical mass of p-gMDH is 38 517 Da, gMDH is 34 686 Da and a frequently observed proteolytically modified form is 34 839 Da. The inset contains an SDS/PAGE stained with Coomassie Brilliant Blue with a sample showing some proteolytic modification. The eight lanes from left to right contained (1) 0.3 mg·mL^{-1} gMDH, (2) 0.15 mg·mL^{-1} gMDH, (3) 0.075 mg·mL^{-1} gMDH, (4) molecular mass markers, (5) 0.025 mg·mL^{-1} p-gMDH, (6) 0.05 mg·mL^{-1} p-gMDH, (7) 0.1 mg·mL^{-1} p-gMDH, (8) 0.2 mg·mL^{-1} p-gMDH. Different amounts of protein were applied to visualize any trace contamination. In lanes 5–8, a second band is visible indicating some proteolytic cleavage. The upper left corner shows the results of an N-terminal sequencing experiment, used in the early stages before mass spectroscopy became readily available. The results of the experiment indicate a preparation uncontaminated by proteolysis.

with the results shown in Fig. 3. Both the mature and translocatable forms of the protein were linear as expected in the 4–32 °C temperature range and the translational diffusion constants are similar enough to indicate the equivalent quaternary structure. As expected, p-gMDH has a somewhat smaller D_t than the mature enzyme, and this can only result from the presence of the translocation sequence. The D_t values may also be used to calculate the equivalent spherical radius of hydration (R_h), and the results are summarized in the insert of Fig. 3. The R_h calculations suggest an approximate 15% increase in size for the translocatable form, p-gMDH. To make more direct comparisons with the X-ray results, the R_h from the dynamic light scattering (DLS) studies were also compared to estimates obtained from the X-ray model and the results shown in the insert of Fig. 3.

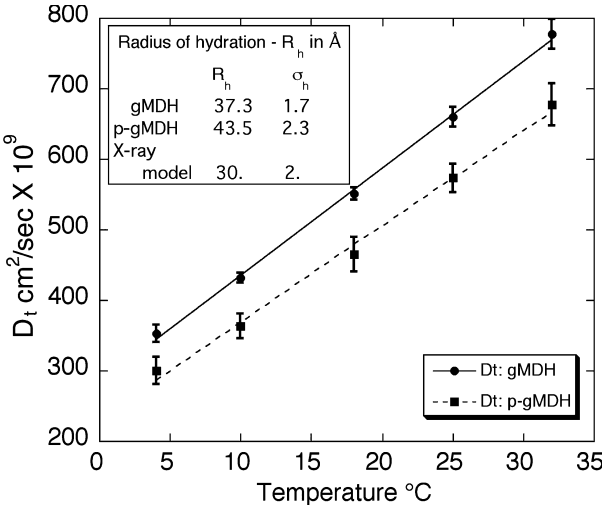

Fig. 3. Quaternary state of p-gMDH and gMDH using dynamic light scattering. The translational diffusion constants, D_t values, were determined by dynamic light scattering and the results are plotted as a function of temperature. Each data point is the average of three measurements with the standard deviation indicated by the vertical bars. The protein concentration for both samples was 0.94 mg·mL^{-1}, and both experiments were carried out in 20 mM $NaPO_4$ pH 7.4, 100 mM NaCl, 5 mM dithiothreitol, 1 mM EDTA. Inset: radius of an equivalent sphere, R_h. In addition, using the X-ray coordinates, the molecular structure was fitted to a prolate ellipsoid and R_h calculated using the relationship $R_h = (ab^2)^{1/3}$.

The DLS results demonstrate that the quaternary states of the two proteins are the same and both are therefore homodimeric. However, p-gMDH has a significantly different D_t and the difference must be attributed to the N-terminal translocation segment.

Catalytic properties of the precursor and mature forms

A comparison of the steady state parameters was made to test the effect of the N-terminal extension on the catalytic reaction. The v_0 values were obtained from the initial decrease in absorbance at 340 nm resulting from NADH oxidation. Assays were carried out using a final active site concentration of 4.25 nM, calculated using a subunit molecular weight of 39 000. The results are summarized in Fig. 4(A,B) and Table 1.

The steady state values for K_m and V_{max} were determined at two different concentrations of oxaloacetate (0.5 and 5.0 mM) or NADH (50 and 200 μM). The K_m (gMDH) for NADH was 146 μM at both oxaloacetate concentrations. The K_m values for oxaloacetate (gMDH) were 75.8 ± 6.7 μM and 147.6 ± 7.0 μM at low and high NADH concentrations, respectively. In the experiments where NADH was varied, the plots were linear over the entire concentration range as can be seen in Fig. 4A. However, a marked substrate inhibition at higher oxaloacetate concentrations was observed (Fig. 4B). The overall V_{max} for the organelle form of the enzyme was 9.2×10^2 min^{-1}.

However, with p-gMDH, significantly different values were found. A distinct upwards curvature of the Lineweaver-Burk plot was observed at both oxaloacetate concentrations, and extrapolation of the linear regions at concentrations of NADH gave K_m values of 75 and 26 μM at low and high oxaloacetate concentrations, respectively (Fig. 4A). With oxaloacetate concentration as the variable, again substrate inhibition was observed but somewhat higher K_m values were calculated: 172 and 1184 μM, respectively. The overall V_{max} for p-gMDH was found to be 2.3×10^3 min^{-1}. With defined differences in the catalytic properties, some conformational effects of the N-terminal extension on the active site must be present. One possibility is that any difference in the catalytic activity of the two forms of the enzyme is the result of electrostatic effects.

To test for charge effects, two additional experiments were carried out: (a) Attempts were made to determine the isoelectric points. (b) The pH dependence of V_{max} and K_m for oxaloacetate was determined using Tris buffers in the pH range 6.0–9.0. When samples of gMDH and p-gMDH were analyzed on an isoelectric focusing gel (pH 3–10), gMDH exhibited a p*I* of ~9.2 and p-gMDH did not enter the gel suggesting an even higher p*I*. In addition, the p*I* values were calculated to be 8.67 and 8.26 for p-gMDH and gMDH, respectively, using the PROTPARAM tool within http://www.expasy.ch.

Determination of the K_m for oxaloacetate (OAA) as a function of pH proved more informative. In these experiments, NADH concentrations were fixed at 100 μM and oxaloacetate concentrations were varied from 10 to 100 mM to avoid the substrate inhibition discussed above. Figure 4C shows the pH dependence of K_m^{OAA} for each form of the enzyme. In general, the curves are similar but there is a difference in the pH 7.0–8.0 region. Here the K_m for oxaloacetate is somewhat higher for the mature form, gMDH and it is possible that this is the consequence of the extra positive charges on p-gMDH. At elevated pH values (> 8.5) K_m is increasing rapidly and the experimental error becomes significant.

Stability of the precursor and mature forms

One possible function of the translocation segment is that it serves to destabilize the protein and thereby

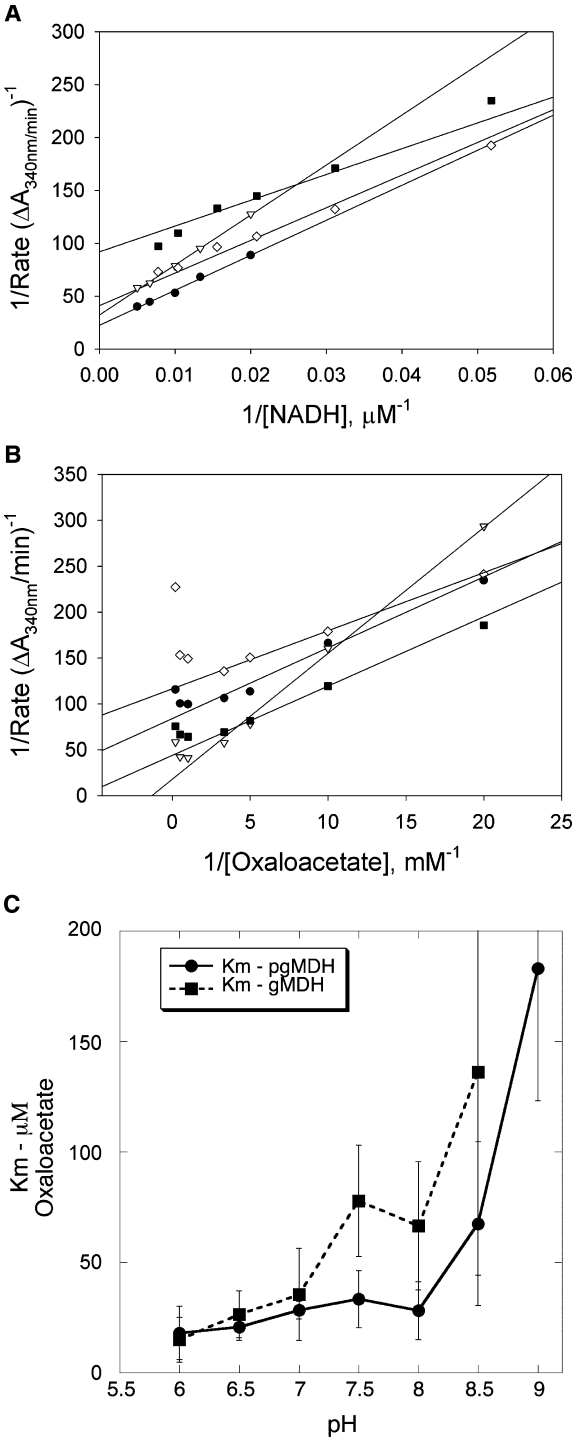

Fig. 4. Steady state kinetics of the MDH reaction for p-gMDH and gMDH. All steady state measurements were carried out at two different saturating concentrations of the second substrate: low (filled symbols) or high (open symbols). The data in (A) and (B) are shown as reciprocal plots. Reaction mixtures were pre-equilibrated to 25 °C and initiated by the addition of 0.17 μg·mL^{-1} of enzyme. p-gMDH is shown by circles and triangles while gMDH is shown by squares or diamonds. (A) Variations in the coenzyme concentration (NADH) are shown while using fixed concentrations of oxaloacetate of 0.5 mM and 5 mM. (B) Saturating concentrations of NADH of 50 μM and 200 μM were used as the oxaloacetate concentration was varied. In Fig. 4A data over comparable substrate concentrations are shown: the linear regression used to calculate kinetic parameters listed in Table 1 for gMDH are over a wider range of data as linearity was visible over a much wider range of substrate concentrations. The linear regressions shown in Fig. 4B omitted the three data points at highest oxaloacetate concentrations where there was a clear indication of substrate inhibition in all cases. Data in Fig. 4C resulted from three different preparations of the two enzymes. In nearly all cases, the v_0 values were measured four times for each pH, oxaloacetate concentration and the three enzyme preparations. The final data were averaged and the graph shows the mean and SD in the form of error bars. Each of the three preparations was checked for the correct form of the enzyme by mass spectrometry.

Table 1. Steady state activity constants for p-gMDH and gMDH.

Conditions[a]	Organelle: gMDH	Precursor: p-gMDH
K_m NADH	146 mM	75 mM
Low OAA	No cooperativity	Positive cooperativity
K_m NADH	146 mM	26 mM
High OAA	No cooperativity	Positive cooperativity
K_m Oxaloacetate	76 mM	172 mM
Low NADH	Substrate inhibition	Substrate inhibition
K_m Oxaloacetate	148 mM	1184 mM
High NADH	Substrate inhibition	Substrate inhibition
V_{max}	9.2×10^2 min^{-1}	2.3×10^3 min^{-1}

[a] Measurements were carried out at pH 8.0 in 0.02 M phosphate buffer.

facilitates unfolding prior to transport through the organelle membrane [10–12]. If so, the relative stability of the two forms of the enzyme comes into question. Here the stability of the translocatable and the mature forms of the MDHs were compared by two experimental methods.

In one set of experiments, the rate of heat inactivation of both gMDH and p-gMDH was determined by preincubating NaH_2PO_4 buffer at pH 8.0 and 45 °C, and determining the rate of loss of enzymatic activity. Rate constants for the inactivation of 0.20 ± 0.03 min^{-1} and 0.63 ± 0.12 min^{-1} were obtained for the p-gMDH and gMDH, respectively, with the amplitudes of either process showing 100% inactivation (data not shown). In these experiments, p-gMDH appeared to be the more stable of the two forms.

A second approach was to observe heat capacity, C_p, changes as a function of temperature. Such experiments provide information about the melting temperature(s), T_m, and estimates of the number of

transition(s) during the unfolding reaction [13]. If different transitions were apparent in the comparable unfolding processes for gMDH and p-gMDH, the results would suggest some unique structural components such as additional domains, dissimilar states of quaternary structure and/or different enthalpic properties of the folded components. Typical differential scanning calorimetry (DSC) results for both gMDH and p-gMDH are described in Fig. 5.

There is a definite difference in the observed temperature dependence of C_p for the p-gMDH compared to the mature form with the former experiencing additional transitions. Using ORIGIN 7.0 software and ignoring the endothermic transitions occurring after unfolding, a satisfactory interpretation suggested a minimum of two events for gMDH, X_1 and X_2 in Fig. 5A. The X_1 transition varied somewhat in magnitude but was clearly present in all experiments.

In the case of p-gMDH, the number of transitions involved in the unfolding process has increased and a satisfactory interpretation involved a minimum of 3 or 4 transitions: X_1 X_2, X_3 and X_4 as are visible in Fig. 5B. For both proteins, similar results were observed for comparable experiments carried out at 0.1 and 1.0 mg·mL^{-1}. In every DSC experiment and again for both forms of the enzyme, the protein precipitated at some temperature beyond the final transitions. The precipitation was accompanied by a decrease in C_p and prevented any attempt to study the reverse reaction.

For comparative purposes, the averaged transition temperatures with their standard deviation are given in Table 2. Multiple experiments, as described in Table 2, indicated that the changes in T_m for the different transitions were reproducible. Although subject to some variation because of the nature of curve fitting, the unfolding profiles of gMDH vs. p-gMDH have striking differences. The unfolding of the translocatable form, p-gMDH has additional transitions that occur at higher temperatures.

The marked differences in the DSC profiles must be related directly to the translocation segment on the N-termini of p-gMDH. The added conformational transformations, visible as the third and fourth peak in Fig. 5B, have midpoints of about 59.1 and 63.0 °C. As will be described below, the crystal structures of the two enzymes are essentially identical. The significance of this additional transitions will be argued in the Discussion section, but in general, multiple transformations in the temperature dependent ΔC_p values are thought to be related to the unfolding of some unique composite secondary structure or tertiary domains [14].

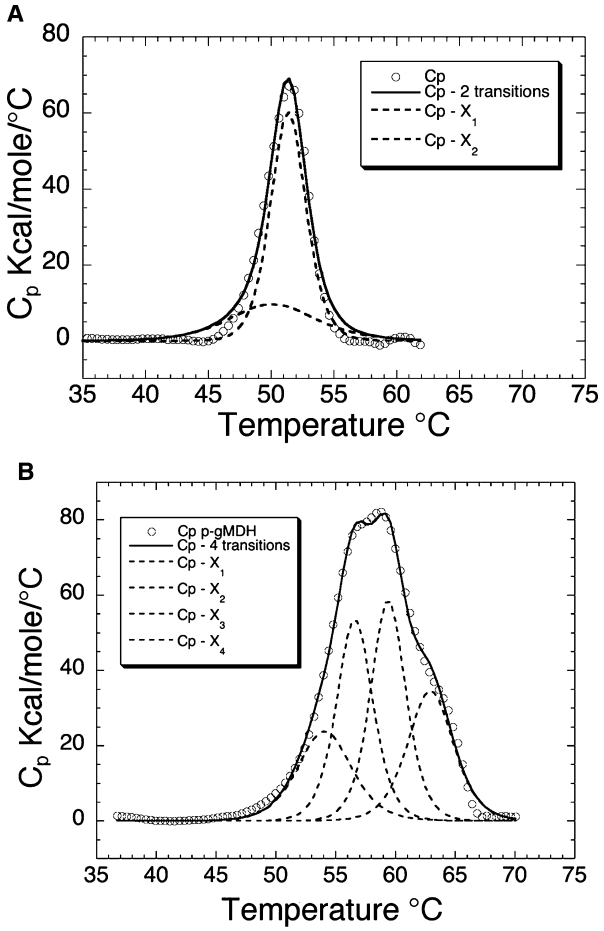

Fig. 5. Thermodynamics of unfolding gMDH and p-gMDH. (A) The heat capacity, C_p, of a solution of gMDH is shown as a function of temperature. The overall melting profile was obtained as described in the Experimental procedures. (B) The same results for p-gMDH under essentially identical conditions are presented. The fitted transitions (X1, etc) are the dashed curves read from low to higher temperature. The curve analyses were carried out using ORIGIN as described in the Experimental procedures.

Table 2. Mean unfolding transitions temperatures for p-gMDH and gMDH. ND, not determined.

MDH	T_{mx1} °C	σ[a]	T_{mx2} °C	σ	T_{mx3} °C	σ	T_{mx4} °C	σ
gMDH	52.4	2.4	55.0	2.2	–	–	–	–
p-gMDH	53.5	2.1	57.4	1.8	59.1	2.3	63.0	ND

[a] Standard deviation of the results in the conventional sense.

The crystal structures of gMDH and p-gMDH

The translocatable form under identical chemical conditions, crystallized in a different space group from the organelle form of the enzyme (Table 3). The X-ray

Table 3. X-ray data collection statistics.

Data set	p-gMDH	gMDH
Source of data	19-ID, SBC, ANL	19-ID
Space group	$P4_12_12$	$P2_1$, pseudo $C222_1$
Cell dimensions		
a (Å)	96.8	137.43
b (Å)	96.8	88.05
c (Å)	213.32	138.82
β (°)	90	91.53
Resolution (Å)	2.55	2.50
% Complete	99.8	88.54
% Highest shell	99.9	88.3
Number of reflections	34 759	100 256

studies are summarized as follows. p-gMDH crystals diffracted to 2.55 Å and have a Matthews' coefficient (V_m) of 3.3 $Å^3 \cdot Da^{-1}$, with 1 dimer in the crystallographic asymmetric unit. The structure of p-gMDH was solved by molecular replacement using the polyalanine equivalent of the dimer coordinates from porcine heart mitochondrial MDH (1mld, PDB) [15].

GMDH crystals diffracted to a resolution of 2.50 Å under cryogenic conditions, and gave a V_m of 2.45 $Å^3 \cdot Da^{-1}$ with four dimers in the asymmetric unit. The molecular packing in the crystal lattice produces a pseudo space group, $C222_1$. Many hours were spent trying to process the raw data using the centered orthorhombic system but the refinement and predictions were always short of satisfactory, and the X-ray structure was solved in the monoclinic space group using the coordinates of p-gMDH and the method of molecular replacement. The refinement results for both crystal structures and their PDB accession codes are given in Table 4.

Discussion of the quaternary and tertiary structure of the gMDHs seems unnecessary *in lieu* of the many crystal structures in the PDB. Instead using the available coordinates for the enzymes shown in Fig. 1 and the method of least squares, studies to compare the main chain conformations were carried out. The main chain conformation of gMDH is shown by the red tubular representation in Fig. 6. In both forms of the enzyme, electron density was first visible beginning at G44 (precursor numbering scheme), seven amino acids into the N-termini of the mature form. At least in these two crystal structures, the dynamically disordered translocation segment begins at an amino acid present in the N-terminus of the cleaved enzyme.

The different crystal packing between gMDH and p-gMDH under identical chemical conditions corroborates but does not prove the presences of the presequence in the one lattice. In the crystal structure of p-gMDH, the first visible N-terminal residue points directly into a solvent channel of the lattice suggesting that there is adequate space for the substantial but disordered translocation segment.

Table 4. X-ray and coordinate refinement statistics for crystalline p-gMDH and gMDH. RMSD, root mean square deviation.

	p-gMDH	gMDH
Resolution (Å)	20–2.55	20–2.5
σ Cut-off	0.0	0.0
Number of reflections	33 921	98,022
R	0.207	0.205
R_{free}	0.258	0.255
Number of protein atoms	4620	18480
Number water molecules	205	398
Average B-factor ($Å^2$)	36.8	41.7
RMSD bond lengths (Å)	0.018	0.006
RMSD bond angles (°)	2.05	1.27
RMSD dihedral (°)	23.0	22.0
RMSD improper (°)	1.33	0.88
Ramachandran geometry		
Most favored (%)	90.6	88.6
Allowed (%)	7.8	10.4
Generously allowed (%)	0.8	0.8
Disallowed (%)	0.8	0.1
PDB accession code	1SEV	1SMK

Because of the large asymmetric unit for gMDH (four molecules, eight subunits) coupled with the two subunits determined for p-gMDH, the coordinates of a single subunit of this glyoxysomal enzyme are tenfold redundant. The reliability of the coordinates and the site of any significant differences could be estimated by cross-comparisons (45 in total). To do this, crystal coordinates of each independent subunit were overlaid pairwise by the method of least squares and the resulting root mean square deviations (Å) compared. The outcome is summarized in Fig. S1 (supplementary data). Overall, the coordinates for the two forms of the protein are in close agreement. Among the 10 copies of the gMDH monomer in the crystal structures, the root mean square deviation between C_α atoms is between 0.2 and 0.5 Å, near to the anticipated experimental error. However there are two regions that appear to contain conformational variability.

The first region is around residues 123–132, the so-called active site loop. The electron density in this region was generally low in both of the crystal structures and most of the subunits suggesting that this loop is not always in the same conformation. This disordered segment has the amino acid sequence: -P-R-K-P-G-M-T-R-D-D- (123–132), and overall these residues have the highest B-factors of the entire structure. R124 near the tip of a loop is believed to be

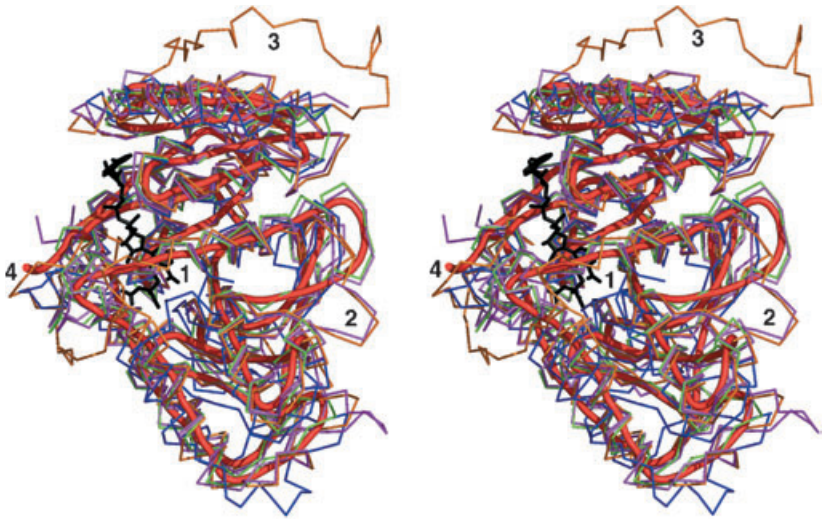

Fig. 6. GMDH and other organelle and prokaryotic MDHs. The stereo image shows an overlay of five MDH structures. After least squares fitting to the *E. coli* form of MDH, the root mean square deviations ranged from 0.7 to 1.7 Å. GMDH is represented by the red-maroon cartoon structure. The other conformations are represented by Cα stick models using the following color coding: *E. coli* MDH, light green (1EMD) [27]; pig mitochondrial MDH, violet (1MLD) [15]; sorghum chloroplast MDH, orange (7MDH) [26]; *Archea*, blue (1HYE) [28]; and cytosolic MDH, purple (5MDH) [29]. The numbered positions were highlighted as follows: 1, the active site, for eMDH both the coenzyme and a substrate molecule are indicated in black in the active site; 2, marks the location of a significant conformational difference found in the mitochondrial and chloroplast enzymes from other family members; 3, a major conformational insertion occurs in the chloroplast MDH; 4, marks the loop region noted in the text where weak electron density was found in both p-gMDH and gMDH.

involved in interaction with the dicarboxylic acid substrates during the formation of catalytic intermediates. The binding of substrate is believed to draw the loop further into the active site and the conformational change may be necessary for catalysis [16].

The second conformational variation occurs in the polypeptide chains at residues 256–268. In the dimeric MDH family, the interface is formed by the packing of three α-helices from each subunit. In the amino acid sequence, the last of the three helices begins at residue 268. To examine the dimer interface for each of the two enzymes, the X-ray structures were used to estimate the change in the solvent accessible surface area as a result of dimer formation. For gMDH the loss was 3180 ± 63 Å^2. For p-gMDH crystallographic coordinates indicated a loss of 3160 ± 41 Å^2, and using this criterion, the two interfaces are identical.

Discussion

Recognition and translocation of nuclear-encoded organelle proteins appear to have many aspects in common even though the current dogma explaining these processes differs for organelle classes [17]. Common elements include a variety of protein factors, both cytosolic and membranous, that appear to be specific for a given organelle type. For the organelles mentioned in the Introduction and proteins destined for translocation, segments of extra amino acids mostly on the N-termini are required for the import process. Questions arise as to whether or not there are discernible differences in the properties of these proteins with and without these so-called translocation segments.

In the experiments described above, one enzyme was compared in both its translocatable and organelle form. In the crystallographic portion of the experiments, two conformational variations were identified between the two forms, but no electron density was visible for the translocation segment. Lack of electron density in a protein crystal structure is explainable by dynamic disorder; that is, the structural element is in multiply interchangeable conformations in the crystal lattice.

The comparison of the physical and catalytic properties of the two nearly identical enzymes has nonetheless produced some surprising differences. For example, one possible consequence of the presence of the translocation segment would be to alter the quaternary structure. GMDH and most but not all MDHs are dimers. The DLS results as described above indicated that p-gMDH is homodimeric like the organelle form. However, the precursor had a significantly smaller D_t, and using the assumption of a spherical

molecule in analyzing translational diffusion, it is approximately 15% larger. Although the crystallographic studies show that the dimers are not spherical, the additional N-terminal segment has measurable effects on the overall hydrodynamic properties of the enzyme.

Enzyme catalysis is generally a function of a finely tuned atomic array at the active site and the translocation segment appears to have no dramatic effect on this constellation. However, comparison of the steady state properties demonstrated that the active sites are nearly the same but small differences are identifiable. At high concentrations of substrates, the turnover number for p-gMDH was improved by nearly a factor of two over the organelle form. How could this happen? In most of the NAD^+-dependent dehydrogenase reactions, the catalytic step of hydride transfer, or a linked conformational change, is probably rate limiting. During substrate inhibition, there must be either formation of abortive complexes or substrate hindrance of normally rate limiting conformational changes.

One explanation for changes in catalytic parameters is the fact that the precursor with the translocation segment adds seven extra positive charges (if two histidines are included), but only three acidic residues. Depending on the pK values of the histidines, the net charge change ranges from + 2 to + 4. V_{max} for both forms of the enzyme showed similar pH dependence with a broad optimum between pH 7.5 and 8.0. In this relatively narrow pH region, the K_m for oxaloacetate is smaller for the precursor form of the enzyme. Assuming the same mechanism occurs in both forms of the enzyme, it is possible that the additional basic residues present in the translocation segment are increasing the affinity for the anionic substrate, oxaloacetate.

The results of superimposition of the available crystallographic coordinates of MDHs presented in Fig. 6 above was carried out to analyze both conformation and properties that might distinguish organelle enzymes from other forms. In terms of molecular structure, only two sites on a MDH monomer were visibly different. At one site, the mitochondrial and chloroplast enzymes have a small insertion labeled as '2' in Fig. 6. The chloroplast MDH also contains a large insertion at '3' that, in fact, is adjacent to the N-terminus. However overall, the conformations including the prokaryotic forms of MDH are very similar. The conclusion is that if translocation operates with folded proteins [17,18], the presence of the translocation segment may be related solely to the recognition process.

The removal of an N-terminal segment from many of the nuclear-coded proteins upon reaching the matrix of an organelle is another interesting step of the translocation phenomena. Because it is proteolytically cleaved during movement into the organelle, a region of the translocation segment must be accessible to the cellular protease whether or not the protein is already folded. The crystallographic results on p-gMDH and gMDH indicate that conformational variation near the N-terminal appears to begin several residues into the chain of the mature protein. This would make the upstream cleavage site more available to a protease.

The proposed mitochondrial import mechanism is purported to involve threading the random coil form through a preformed pore [1]. On the other hand, recent evidence seems to show that glyoxysomal/peroxisomal import occurs with fully folded proteins [18]. The studies described above cannot attest to either proposal. However, the gMDH vs. p-gMDH comparison has shown that unfolding of the precursor form is more energy requiring than the unfolding of the organelle form of the protein lacking the translocation segment.

To summarize, the results described above illustrate methods for preparing quantities of translocatable protein by recombinant methods. For glyoxysomal MDH, the precursor form folds normally but the N-terminal segment is extremely sensitive to proteolysis. Hydrodynamic and enzymatic characterization of p-gMDH compared to gMDH showed that the two proteins are similar but not identical. The most surprising result stems from the DSC studies. The results of the heat capacity changes as a function of temperature point to additional structure in p-gMDH. Preliminary results (not shown) suggest that NADP-dependent isocitrate dehydrogenase belonging to yeast mitochondria produce similar DSC results when compared to the identical enzyme without the N-terminal translocation segment. If these observations can be demonstrated in other organelle enzymes, unfolding prior to translocation does not appear to be a viable hypothesis.

Experimental procedures

Reagents

The plasmid pQE 60 containing the cDNA sequence of either p-gMDH or gMDH and a His_6 tag at the C-terminus was generated as previously described [19]. Ni^{2+}–nitrilotriacetate (Ni^{2+}–NTA) resin was purchased from Qiagen (Valencia, CA, USA). Pefabloc was purchased from Roche Biochemicals (Indianapolis, IN, USA). All other reagents were obtained from Sigma Chemical Co. (St. Louis, MO, USA).

Purification of p-gMDH and gMDH

P-gMDH and gMDH were expressed in the *E. coli* strain JM105 from a derivative of pQE 60 that carries either p-gMDH or gMDH cDNA with an inserted codon for glycine (GGA) between the start codon and the second codon of both cDNAs [19]. Cells were grown at 37 °C in LB medium in six 1 L flasks supplemented with ampicillin (100 $\mu g \cdot mL^{-1}$). When the cells reached $D_{600} \approx 0.6$ the temperature was lowered to 32 °C and grown for 5 h after induction with 1.0 mM isopropyl thio-β-D-galactoside. Cells were harvested by centrifugation at 4000 ***g*** for 30 min and stored at −80 °C.

To isolate the MDHs, cell pastes were thawed and re-suspended to 0.5 $g \cdot mL^{-1}$ in Buffer A (50 mM Tris pH 7.2, 300 mM NaCl). All steps were conducted at 4 °C unless otherwise noted. A cocktail of protease inhibitors was added to Buffer A to a final concentration of 4 mM Pefabloc, 100 $mg \cdot mL^{-1}$ phenylmethylsulfonyl fluoride, 1.0 $mg \cdot mL^{-1}$ pepstatin, 2.0 mM leupeptin, 1.0 mM benzamidine, and 1 mL of Sigma protease inhibitor cocktail P8849 per 20 g cell paste. The suspension was subjected to five, 30 s cycles of sonication in a dry ice/isopropanol bath followed by the addition of 0.35% polyethylenimine to precipitate the DNA. Cell debris and DNA were pelleted at 17 500 ***g*** for 15 min.

The supernatant was mixed for 5 min with Ni^{2+}–NTA resin (1 mL resin per 5 g cells) equilibrated with the Buffer A. The column was washed three times with three column volumes of Buffer B (50 mM Tris pH 7.2, 300 mM NaCl, 60 mM imidazole). p-gMDH or gMDH was eluted with two column volumes of Buffer C (50 mM Tris pH 7.2, 300 mM NaCl, 500 mM imidazole). Protease inhibitors were not included beyond the cell lyses step. Both proteins were then dialyzed against 4 L of buffer containing 50 mM NaH_2PO_4 pH 7.4, 300 mM NaCl, then dialyzed into 50 mM NaH_2PO_4 pH 7.4, 300 mM NaCl, 1 mM dithiothreitol, and 1 mM EDTA, using four, one liter buffer changes. One way of insuring success in the purification of p-gMDH was to complete the entire procedure, cells to pure protein, in about 1 h.

Steady state measurements

The catalytic activities of p-gMDH and gMDH were determined by following the disappearance of NADH by monitoring the absorbance at 340 nm in a reaction solution containing 50 mM phosphate at pH 8.0, 100 μM NADH and 1 mM oxaloacetate. To obtain the steady state parameters as a function of pH, 50 mM Tris HCl buffers were used. Reactions were initiated by the addition of 10–60 ng of protein to 3 mL of assay solution. All measurements were minimally carried out in triplicate and values of K_m and V_{max} are averages of three to five independent determinations of v_0.

Characterization of p-gMDH and gMDH

The purity of the protein was verified by SDS/PAGE with Coomassie brilliant blue staining. The concentrations of purified p-gMDH and gMDH were determined spectrophotometrically using an $\varepsilon_{280nm} = 0.475$ $mL \cdot mg^{-1} \cdot cm^{-1}$. The translational diffusion coefficient (D_t) was determined for both p-gMDH and gMDH at different temperatures by dynamic light scattering at a concentration of 0.94 $mg \cdot mL^{-1}$ using a Protein Solutions DynaPro Micro Sampler. All protein samples used for light scattering experiments were prefiltered with 0.02 μm Whatman Anotop Plus filter discs.

Verification of the presequence

Because of the sensitivity of p-gMDH to proteases, the purified fractions were always examined either by N-terminal sequencing analysis (University of Minnesota Micro-Chemical Facility) or mass spectrometry (Mass Spectrometry Consortium for the Life Sciences, University of Minnesota) and SDS/PAGE. Protein samples were sequenced from solutions containing 50 mM NaH_2PO_4 pH 7.4, and 300 mM NaCl. In one experiment, upon completion of X-ray data collection, a p-gMDH crystal was re-dissolved into 50 mM NaH_2PO_4 pH 7.4, 300 mM NaCl and the N-terminal sequence determined (data not shown).

The instrument used for the collection of MALDI-TOF data was a Bruker Biflex III, equipped with a N_2 laser (337 nm, 3 nanosecond pulse length) and a microchannel plate (MCP) detector. The data was collected in the linear mode, positive polarity, with an accelerating potential of 19 kV. Calibration was performed using bovine serum albumin ($MH^+/Z = 66431$, $MH^{2+}/Z = 33216$).

Thermal transitions

Heat capacity changes as a function of temperature were obtained with a VP-DSC differential scanning calorimeter (MicroCal, Northampton, MA, USA) in 50 mM NaH_2PO_4 pH 7.4 and 100 mM NaCl, with a heating rate of 2 $°C \cdot min^{-1}$. The resulting curve was analyzed with the ORIGIN 7.0 (http://www.micocal.com) software using the minimal number of transitions to fit the observations. For both proteins, some precipitate was present at the completion of the experiment.

Crystallization and X-ray data collection

Crystallization was carried out by first dialyzing the purified enzyme against a buffer containing 10 mM NaH_2PO_4 pH 7.4, 100 mM NaCl, 5 mM dithiothreitol, 1 mM EDTA. The dialyzed protein was concentrated to 10 $mg \cdot mL^{-1}$ using an Amicon concentrator (PM10, 43 mm). Crystals of p-gMDH and gMDH were grown at 20 °C by hanging drop – vapor diffusion from mother liquor containing

100 mM sodium citrate pH 4.5, 1 mM dithiothreitol, and PEG 8000. Needle-shaped crystals of both p-gMDH and gMDH appeared in a few days in the range of 6.5–11.0% and 9.0–13.5% (w/v) of PEG 8000, respectively.

Crystals of p-gMDH and gMDH were cryo-protected by adding a 50% glycerol solution to the crystallization drop resulting in a final concentration of 33% (v/v) glycerol. Data sets for p-gMDH and gMDH were collected at the SBC Beamline 19-ID (ANL-APS) with the 3×3 CCD area detector set at $2\theta = 0°$. Data were indexed and integrated with DENZO and scaled with SCALEPACK [20].

Phases were found for p-gMDH by molecular replacement using the dimeric structure of porcine heart mitochondrial malate dehydrogenase substituted with poly-alanine as a search model using X-PLOR [21]. One dimer was present in the asymmetric unit. The phases of gMDH were also found by molecular replacement using the partially refined model of p-gMDH and the program EPMR [22]. Four dimers were present in the asymmetric unit of a $P2_1$ unit cell. During the refinement, the space group was found to be pseudo C-centered. A further explanation of the space group ambiguity is given in the Results section.

Both X-ray structures were refined using CNS v.0.3 [23]. The coordinates were refined by rigid body, then simulated annealing, positional refinement, and B-factor refinement. Five percent of the data were omitted from refinement and used for R_{free} calculations [24]. Non-crystallographic symmetry was used to restrain the four dimers of gMDH during the early rounds of positional and B-factor refinement, and then released during subsequent macro-cycles. The program O was used for model rebuilding after each round of crystallographic refinement [25].

Acknowledgements

The authors recognize the Mass Spectrometry Consortium for the Life Sciences at the University of Minnesota and various supporting agencies, including the National Science Foundation for Major Research Instrumentation grant 9871237, used to purchase the instruments described in this study. We thank the Microchemical Facility (http://mcf.ahc.umn.edu) at the University of Minnesota for help in characterizing the proteins. We remain continually grateful to Ed Hoeffner for the maintenance of hardware and software associated with the X-ray diffraction, DSC and DLS instrumentation. The SBC station at the Argonne National Laboratory was used for collection of some X-ray data. We are also grateful to the Minnesota Supercomputer Institute for continuing use of their facilities necessary for X-ray crystallography especially the use of the Biomedical Computing Science Laboratory and software maintenance by Dr Patton Fast. This work was partially supported by a grant from the National Institutes of Health (GM 13925). Bryan Cox is grateful for support from the NIH Biophysical Training Grant (GM08277) and initially by a research grant from the NSF.

References

1 Schatz G (1998) The doors to organelles. *Nature* **395**, 439–440.
2 Steer C & Hanover J (1991) Intracellular Trafficking of Proteins. Cambridge University Press, Cambridge, UK.
3 Veenhuis M, Kiel JAKW & Van Der Klei IJ (2003) Peroxisome assembly in yeast. *Microsc Res Technique* **61**, 139–150.
4 Faber KN, Haima P, Gietl C, Harder W, Geert AB & Veenhuis M (1994) The methylotrophic yeast *Hansenula polymorpha* contains an inducible import pathway for peroxisomal matrix proteins with an N-terminal targeting signal (PTS2 proteins), *Proc Natl Acad Sci USA* **91**, 12985–12989.
5 Fransen M, Terlecky SR & Subramani S (1998) Identification of a human Pts1 receptor docking protein directly required for peroxisomal protein import. *Proc Natl Acad Sci USA* **95**, 8087–8092.
6 Mathieu M, Zeelen J, Pauptit RA, Erdmann R, Kunau WH & Wierenga RK (1994) The 2.8 Å crystal structure of peroxisomal 3-ketoacyl-CoA thiolase of *Saccharomyces cerevisiae*: a five-layered ABABA structure constructed from two core domains of identical topology. *Structure* **2**, 797–808.
7 Chudzik DM, Michels PA, de Walque S & Hol WG (2000) Structures of type 2 peroxisomal targeting signal in two trypanosomatid aldolases. *J Mol Biol* **300**, 697–707.
8 Gietl C (1992) Malate dehydrogenase isoenzymes: cellular locations and role in the flow of metabolites between the cytoplasm and cell organelles. *Biochim Biophys Acta* **1100**, 217–234.
9 Gietl C (1990) Glyoxysomal malate dehydrogenase from watermelon is synthesized with an amino-terminal peptide. *Proc Natl Acad Sci USA* **87**, 5773–5777.
10 Huang S, Ratliff K & Matouschek A (2002) Protein unfolding by the mitochondrial membrane potential. *Nat Struct Biol* **9**, 301–307.
11 Huang S, Ratiff KS, Schwartz M, Spenner J & Matouschek A (1999) Mitochondria unfold precursor proteins by unraveling them from their N-termini. *Nat Struct Biol* **6**, 1132–1138.
12 Pfanner N, Rassow J, Guiard B, Sollner T, Hartl F & Neupert W (1990) Energy requirements for unfolding and membrane translocation of precursor proteins during import into mitochondria. *J Biol Chem* **265**, 16324–16329.

13 Makhatadze G & Privalov P (1995) Energetics of protein structure. *Adv Protein Chem* **47**, 307–425.
14 Privalov P (1989) Thermodynamic problems of protein structure. *Ann Rev Biophys Biophys Chem* **18**, 47–69.
15 Gleason WB, Fu ZJ, Birktoft J & Banaszak L (1994) Refined crystal structure of mitochondrial malate dehydrogenase from porcine heart and the consensus structure for dicarboxylic acid oxidoreductases. *Biochemistry* **33**, 2078–2088.
16 Bell J, Yennawar H, Wright SK, Thompson J, Viola R & Banaszak L (2001) Structural analyses of a malate dehydrogenase with a variable active site. *J Biol Chem* **276**, 31156–31162.
17 Smith MD & Schnell DJ (2001) Peroxisomal protein import: the paradigm shifts. *Cell* **105**, 293–296.
18 Dammai V & Subramani S (2001) The human peroxisomal targeting signal receptor, Pex5p, is translocated into the peroxisomal matrix and recycled to the cytosol. *Cell* **105**, 187–196.
19 Gietl C, Seidel C & Svendsen I (1996) Plant glyoxysomal but not mitochondrial malate dehydrogenase can fold without chaperone assistance. *Biochim Biophys Acta* **1274**, 48–58.
20 Otwinowski Z & Minor W (1997) Processing of X-ray data collected in oscillation mode. *Methods Enzymol* **276**, 307–326.
21 Brunger AT, Adams PD, Clore GM, Delano WL, Gros P, Grossekunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T & Warren GL (1998) Crystallography and NMR system – a new software suite for macromolecular structure determination. *Acta Crystallogr Sect D Biol Crystallogr* **54**, 905–921.
22 Kissinger CR (2001) Molecular replacement by evolutionary search. *Acta Crystallogr Sect D Biol Crystallogr* **57**, 1474–1479.
23 Brunger AT, Adams PD & Rice LM (1998) Recent developments in the efficient crystallographic refinement of macromolecular structures. *Curr Opin Struct Biol* **8**, 606–611.
24 Kleywegt GJ & Brunger AT (1996) Checking your imagination: applications of the free *R* value. *Structure* **4**, 897–904.
25 Jones T, Zou J, Cowan S & Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr Sect A Found Crystallogr* **A47**, 110–119.
26 Johansson K, Ramaswamy S, Saarinen M, Lemaire-Chamley M, Issakidis-Bourguet E, Miginiac-Maslow M & Eklund H (1999) Structural basis for light activation of a chloroplast enzyme: the structure of sorghum NADP-malate dehydrogenase in its oxidized form. *Biochemistry* **38**, 4319–4326.
27 Hall MD & Banaszak LJ (1993) Crystal structure of a ternary complex of *Escherichia coli* malate dehydrogenase and NAD at 1.9 Å resolution. *J Mol Biol* **232**, 213–222.
28 Panisko E & Mcalister-Henn L (2001) Subunit interactions of yeast NAD + -specific isocitrate dehydrogenase. *J Biol Chem* **276**, 1204–1210.
29 Chapman ADM, Cortes 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, alpha-ketomalonate and tetrahydonad. *J Mol Biol* **285**, 703–712.

Supplementary Material

The following material is available from http://www.blackwellpublishing.com/products/journals/suppmat/EJB/EJB4475/EJB4475sm.htm

Fig. S1. Pairwise comparison of the X-ray crystal structures of gMDH and p-gMDH. The X-ray studies described in the accompanying text resulted in eight sets of atomic coordinates for the gMDH subunit. After superimposition by the method of least squares over the entire monomer, pairwise comparisons of the coordinates were calculated at each Cα position. These were averaged and the standard deviation calculated. The monomers of p-gMDH were also superimposed and a pairwise comparison was done. The results are shown in the plot. The solid line is σ, the standard deviation, at each Cα position and the mean RMSD positional difference is represented by the dotted line. To avoid confusion, the baseline of the p-gMDH data was raised from 0 to 2.