Structure, Vol. 10, 105–113, January, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S0969-2126(01)00696-7

Insights into Enzyme Evolution Revealed by the Structure of Methylaspartate Ammonia Lyase

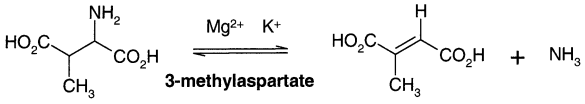
C.W. Levy,[1] P.A. Buckley,[1] S. Sedelnikova,[1] Y. Kato,[2] Y. Asano,[2] D.W. Rice,[1] and P.J. Baker[1,3]
[1]Krebs Institute for Biomolecular Research
Department of Molecular Biology
and Biotechnology
The University of Sheffield
Sheffield S10 2TN
United Kingdom
[2]Biotechnology Research Centre
Toyama Prefectural University
5180 Kurokawa, Kosugi
Toyama 939-0398
Japan

Summary

Methylaspartate ammonia lyase (MAL) catalyzes the magnesium-dependent reversible α,β-elimination of ammonia from L-*threo*-(2S,3S)-3-methylaspartic acid to mesaconic acid. The 1.3 Å MAD crystal structure of the dimeric *Citrobacter amalonaticus* MAL shows that each subunit comprises two domains, one of which adopts the classical TIM barrel fold, with the active site at the C-terminal end of the barrel. Despite very low sequence similarity, the structure of MAL is closely related to those of representative members of the enolase superfamily, indicating that the mechanism of MAL involves the initial abstraction of a proton α to the 3-carboxyl of (2S,3S)-3-methylasparic acid to yield an enolic intermediate. This analysis resolves the conflict that had linked MAL to the histidine and phenylalanine ammonia lyase family of enzymes.

Introduction

Methylaspartate ammonia lyase (MAL) catalyzes the reversible α,β-elimination of ammonia from L-*threo*-(2S,3S)-3-methylaspartic acid to give mesaconic acid.



MAL was first isolated from *Clostridium tetanomorphum* and was subsequently found to be narrowly distributed throughout both obligate and facultative anaerobes [1], where it forms a component of glutamate catabolism. MALs isolated from *Citrobacter amalonaticus* and *C. tetanomorphum* have been shown to be dimeric in solution, with a subunit molecular mass of approximately 45.5 kDa and requiring both divalent and monovalent cations, such as Mg^{2+} and K^{+}, for activity [2, 3]. Crystals of the *C. tetanomorphum* MAL have recently been reported [4]. Sequence comparisons (Figure 1) of MALs across a range of species show that the enzymes share considerable sequence identity (37% over five species), indicating that the structures of MAL from these species are closely related.

Currently, two possible reaction mechanisms for MAL have been proposed. The first places MAL in the family of amino acid lyases, which include histidine ammonia lyase (HAL) and phenylalanine ammonia lyase (PAL). The reaction mechanism of PAL and HAL is thought to proceed via the formation of a covalent intermediate between the substrate and a posttranslationally modified residue in the active site. In the case of PAL, biochemical studies indicate that a serine is modified to a catalytically essential dehydroalanine prosthetic group [5]. In HAL, a similar modification was proposed [6, 7]. However, the structure revealed the presence of 4-methylidene-imidazole-5-one at the active site, a novel posttranslational polypeptide modification derived from an active site serine [8]. Biochemical studies of MAL by Gani and coworkers have suggested that MAL, like PAL, utilizes a dehydroalanine prosthetic group for catalysis, following modification of an active site serine (Ser-173 in *C. amalonaticus* MAL [9]). Furthermore, a considerable primary deuterium isotope effect for (2S,3S)-3-methylaspartic acid, the natural substrate, has been cited as indicative of the presence of a covalently bound intermediate [9]. These findings are in contrast to the earlier reports of Bright et al. [10], who found no primary deuterium isotope effect for the deamination reaction of *C. tetanomorphum* MAL.

The alternative reaction mechanism for MAL is based on the recognition of very low-level sequence similarities (less than ten identities in total) between MAL and members of the enolase superfamily, particularly focusing on a number of important catalytic residues [11]. These studies suggest that the fold of MAL is based on a TIM barrel architecture, in keeping with the known structures for other members of the superfamily and different from HAL and PAL, which are composed of a core of 20 nearly parallel α helices reminiscent of fumarase C and related proteins [12]. In this enzyme mechanism, the MAL chemistry would be initiated by the abstraction of the 3-proton of (2S,3S)-3-methylaspartic acid, which is α to the 3-carboxyl of the substrate, to form an enolic intermediate, which would then collapse to eliminate ammonia [11].

In order to resolve the conflict in the mechanism for MAL, we initiated a structure determination program on the *C. amalonaticus* enzyme. MAL has also been shown to catalyze the deamination of a variety of 3-substituted aspartic acids, including (2S,3S)-3-ethylaspartic acid. A further aim of the study was to define the determinants of substrate specificity of the enzyme, which might ultimately lead to the engineering of a modified enzyme for the chiral synthesis of novel 3-substituted (S)-aspartic

[3]Correspondence: p.baker@sheffield.ac.uk

Key words: methylaspartate ammonia lyase; enolase superfamily; evolution; X-ray crystallography; MAD phasing; selenomethionine

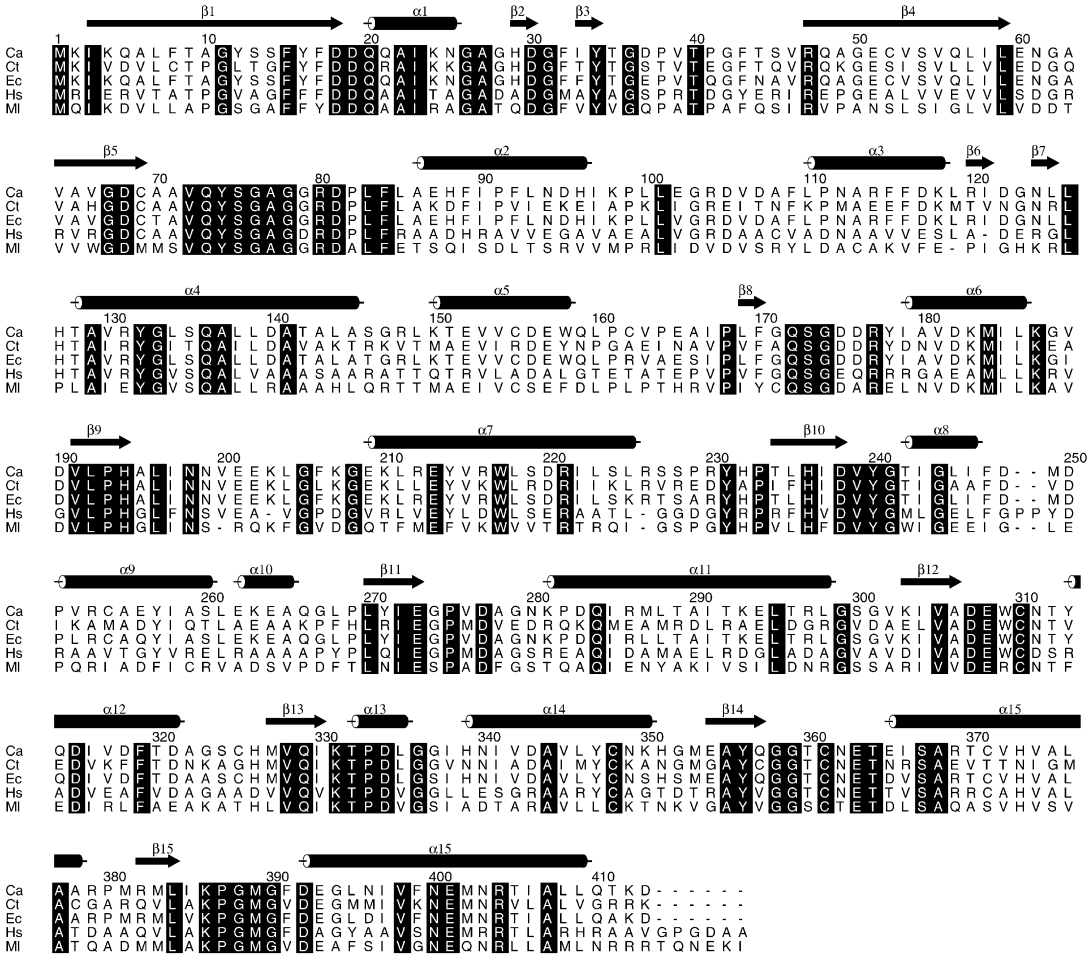


Figure 1. The Aligned Sequences of the *Citrobacter amalonaticus* (Ca), *Clostridium tetanomorphum* (Ct), *E. coli* (Ec), *Halobacterium species* (Hs), and *Mesorhizobium loti* (Ml) MALs

Invariant residues are highlighted in black, and the secondary structure elements are shown above the sequences.

acids from their corresponding fumaric acid derivatives [2], an endeavor with considerable biotechnological potential. In this paper, we report the 1.3 Å resolution structure determination of the *C. amalonaticus* MAL and an analysis of the binary complex with (2S,3S)-3-methylaspartic acid and discuss the implications of the structure for the enzyme mechanism and substrate specificity.

Results and Discussion

Structure Determination

MAL from *C. amalonaticus* was overexpressed in *E. coli* in the presence of selenomethionine and crystallized as previously described [13] to give two distinct crystal forms, form A ($P4_122$, with a monomer in the asymmetric unit) and form B (C222, with a dimer in the asymmetric unit). The structure was determined by multiwavelength anomalous diffraction (MAD) techniques using data collected from a single form A SeMet-labeled crystal to 2.16 Å resolution (data set MALP4). High-resolution data to 1.3 Å were collected from a single form B crystal (again selenomethionine derivatized [data set MALC2]). The form B structure was solved by molecular replacement using the coordinates of the form A structure as the starting model. The A and B crystal forms gave rise to virtually identical molecular models and, in this paper, only the structure of the form B crystals, which corresponds to a dimer in the asymmetric unit, is described.

Overall Fold, Secondary and Quaternary Structure

The final model for the free enzyme consists of 822 of the expected 826 residues (no density was present for the two C-terminal residues within each subunit), 1564 water molecules, and one sulfate ion. The *C. amalonaticus* MAL subunit consists of a single polypeptide chain of 411 of the expected 413 residues and folds into two domains with approximate overall dimensions of length 70 Å and maximum width of 60 Å. It is composed of 15 α helices and 15 β strands, together with a small number of 3_{10} helices, helical turns, and a number of loops (Figure 2). These elements of regular secondary structure comprise 61% of the polypeptide chain. The N-terminal domain (residues 1–160) has an architecture composed of a three-stranded antiparallel β sheet and an antiparallel, four α helix bundle. This domain is connected by a single

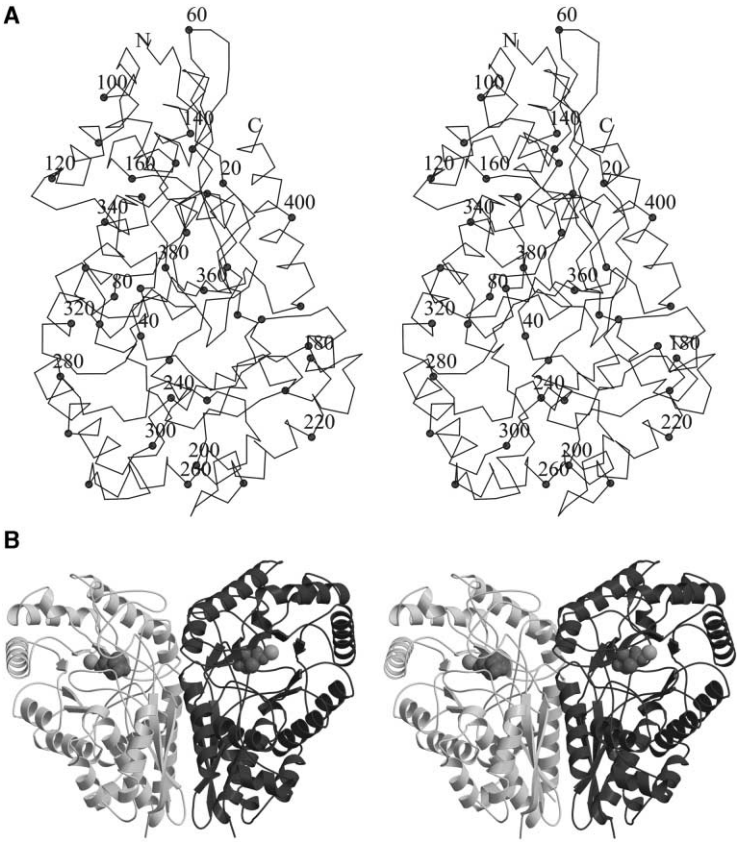


Figure 2. Stereo Diagrams

(A) The Cα backbone trace of a single subunit of MAL, with every 10th residue marked by a black sphere, and every 20th numbered.

(B) The two MAL subunits (black and white) in the dimer, showing the position of the active site Mg^{2+} (gray sphere) and (2S,3S)-3-methyl aspartic acid (dark gray spheres).

strand of extended β-type structure to the first β strand of the second domain (residues 170–411), which folds into an eight-stranded TIM barrel structure [13] (Figure 2A).

Gel filtration studies of *C. amalonaticus* MAL have suggested that the enzyme is a dimer in solution [1]. The solvent-accessible surface area of an isolated monomer of MAL is 16,200 Å^2 (figures quoted to nearest 50 Å^2), and, in the crystal, 13% of this surface (2100Å^2) is buried between two molecules related by a noncrystallographic 2-fold axis. In the dimer, a substantial interface is made between the C-terminal helix of the TIM barrel domain from one monomer and the three-stranded antiparallel β sheet of the N-terminal domain of the second monomer (Figure 2B). The interface between the two monomers involves both a complex network of hydrogen bonding and ionic interactions as well as a number of hydrophobic contacts.

Active Site

To identify the active site of MAL, a form B crystal (C222) was soaked in cryoprotectant containing 10 mM (2S,3S)-3-methylaspartic acid plus 5 mm $MgCl_2$ (data set β-Me-Asp; 2.1 Å). Analysis of the $2F_o - F_c$ and difference electron density maps revealed a clear electron density feature consistent with a bound metal and the adjacent (2S,3S)-3-methylaspartic acid (Figure 3). The divalent cation is octahedrally coordinated by interactions to one carboxyl oxygen of three side chains (Asp-238, Glu-273, and Asp-307), two water molecules, which are both linked to the carboxyl of Glu-308, and one of the carboxyl oxygens of the substrate (Figure 4A). Given the reaction chemistry, the unambiguous interpretation of the electron density is somewhat surprising, as the enzyme could potentially turn over under these conditions. However, it may be that the reaction is inhibited in the crystal by crystal lattice contacts or reduced by either the pH of the solution (pH 7.0), at which the relative activity of *Citrobacter amalonaticus* MAL for the deamination reaction is only 10% of the maximum [2], or by the presence of high concentrations of cryoprotectant and other crystallization components.

On the basis of the fit to the electron density alone, there are two ways in which the substrate can be oriented, which arise as a consequence of the pseudosymmetry of the electron density for the molecule, given the similar electron density of the 2-amino and 3-methyl moieties (Figure 3). The two orientations correspond to an exchange in position of the 2-amino and 3-methyl groups on the enzyme. However, a number of pieces of independent evidence suggests that the substrate adopts the orientation highlighted in Figure 4A, with the 3-methyl group closest to the bound metal atom and with the 3-carboxyl coordinating the metal center. These include the following:

First, in one orientation, the 3-methyl of the substrate occupies a hydrophobic depression in the protein surface, which could also accommodate the 3-substituted C_2H_5 group of (2S,3S)-3-ethylaspartic acid, which is also a substrate of MAL [2]. In this orientation, the 2-amino group makes favorable interactions to the enzyme and

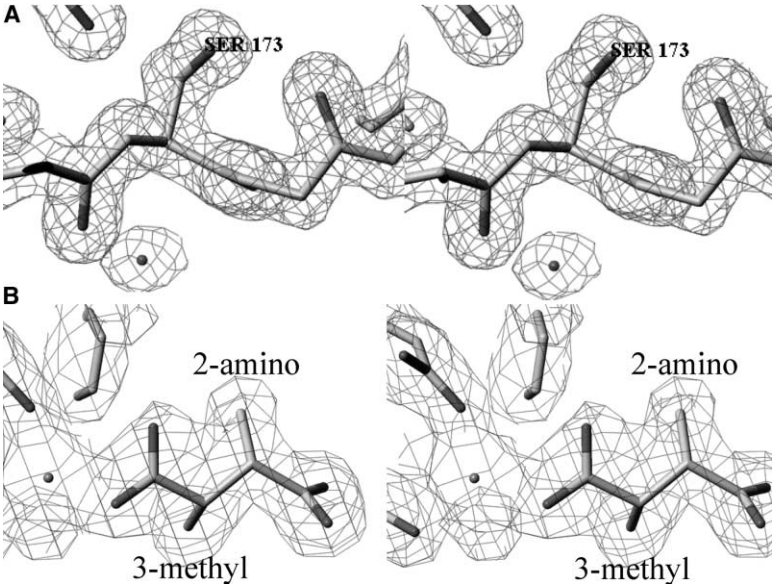


Figure 3. Stereo Diagrams of the Final $|2F_O - F_C|$ Electron Density Maps Contoured at 2.0 σ

(A) A representative portion of the 1.3 Å *C. amalonaticus* MAL structure.

(B) The density for the Mg^{2+} and substrate in the 2.1 Å *C. amalonaticus* MAL/(2S,3S)-3-methyl aspartic acid binary complex.

is also partially exposed to the solvent. In the other orientation, however, the 2-amino moiety would occupy the hydrophobic pocket, with no residues available to stabilize this group, and the 3-methyl group would be exposed to solvent.

Second, the reaction mechanism of *C. tetanomorphum* MAL is thought to be ordered with the release of ammonia occurring prior to the release of the product mesaconic acid [9]. When orientated as shown in Figure 4A, the 2-amino group occupies a region of the active site pocket, which is solvent exposed and therefore consistent with the proposed mechanism. Whereas, in the other orientation, the 2-amino group would be buried on the enzyme surface and thus unable to be released prior to the product.

We thus believe that the substrate adopts the orientation shown in Figure 4A, with the (2S,3S)-3-methylaspartic acid moiety lying at the C-terminal end of the β strands of the TIM barrel domain adjacent to β14. The surface of the protein that packs against the substrate is provided by residues Gln-73, Gln-172, His-194, Gln-329, Lys-331, Gly-359, Thr-360, Cys-361, and Met-389 (Figure 4B). Specific hydrogen bonds are made between Nϵ2 of Gln-329 and Nϵ2 of His-194 to the 3-carboxyl oxygens of the substrate, between the 2-amino group and Oϵ1 of Gln-172 and, via a water molecule, to Nϵ2 of Gln-73, and, finally, between the 1-carboxylate and the main chain NH of Cys-361, the Oγ of Thr-360 and Nϵ2 of Gln-172. Further stabilization in the active site may be provided by favorable van der Waals packing interactions between the 3-methyl group and the side chain of Leu-384 and Gln-172. In addition, one of the six ligands to the Mg^{2+} is provided by one of the 3-carboxyl oxygen atoms of the substrate.

Substrate Specificity

C. amalonaticus MAL is known to catalyze the α,β-elimination of ammonia from a range of 3-substituted (S)-aspartic acids, such as (2S,3S)-3-methylaspartic and (2S,3S)-3-ethylaspartic, but not (S)-aspartic acid [2]. It is apparent from the structure that there is a small cavity adjacent to the carbon at the 3 position that could accommodate the increase in side chain bulk from a methyl to an ethyl as encountered in going from (2S,3S)-3-methylaspartic to (2S,3S)-3-ethylaspartic acid. The enzyme surface lining the 3-methyl binding site is composed almost entirely of side chains (including those of Tyr-356, Phe-170, and Leu-384), with little or no interaction between the 3-methyl group and main chain atoms. There is thus considerable potential for the engineering of this site to alter the substrate specificity of MAL, providing a route for the enzymic production of novel 3-substituted aspartic acids (Figure 4B).

MAL Is a Member of the Enolase Superfamily

Having identified the active site of MAL, we immediately recognize that MAL is not structurally related to either HAL or PAL. Furthermore, there is no evidence in the MAL electron density map for any posttranslational modification of any active site residue, including Ser-173, previously proposed to be posttranslationally modified to dehydroalanine (Figure 3A). Ser-173 lies on strand β8 of the TIM barrel, with its side chain remote to the active site, packing against the side chains of Arg-177, Ala-195, and Pro-193, clearly playing no part whatsoever in contacts to the substrate.

The TIM barrel architecture of MAL is very similar to that seen throughout the enolase superfamily, confirming the suggestions that MAL might adopt this fold [11]. To examine the relationship between MAL and members of this family, MAL was superimposed onto the known structures of *Pseudomonas putida* mandelate racemase (MR, Protein Data Bank code 1MDL, rmsd 1.79 Å for 194 overlapped Cα [15]), *Pseudomonas putida* muconate lactonising enzyme I (MLEI, Protein Data Bank code 1F9C, rmsd 1.84 Å for 196 overlapped Cα [16]), and *Saccharomyces cerevisiae* enolase (Protein Data Bank code 1EBG, rmsd 1.87 Å for 329 overlapped Cα [17]) (Figure 4C). The relationship between the two subunits in

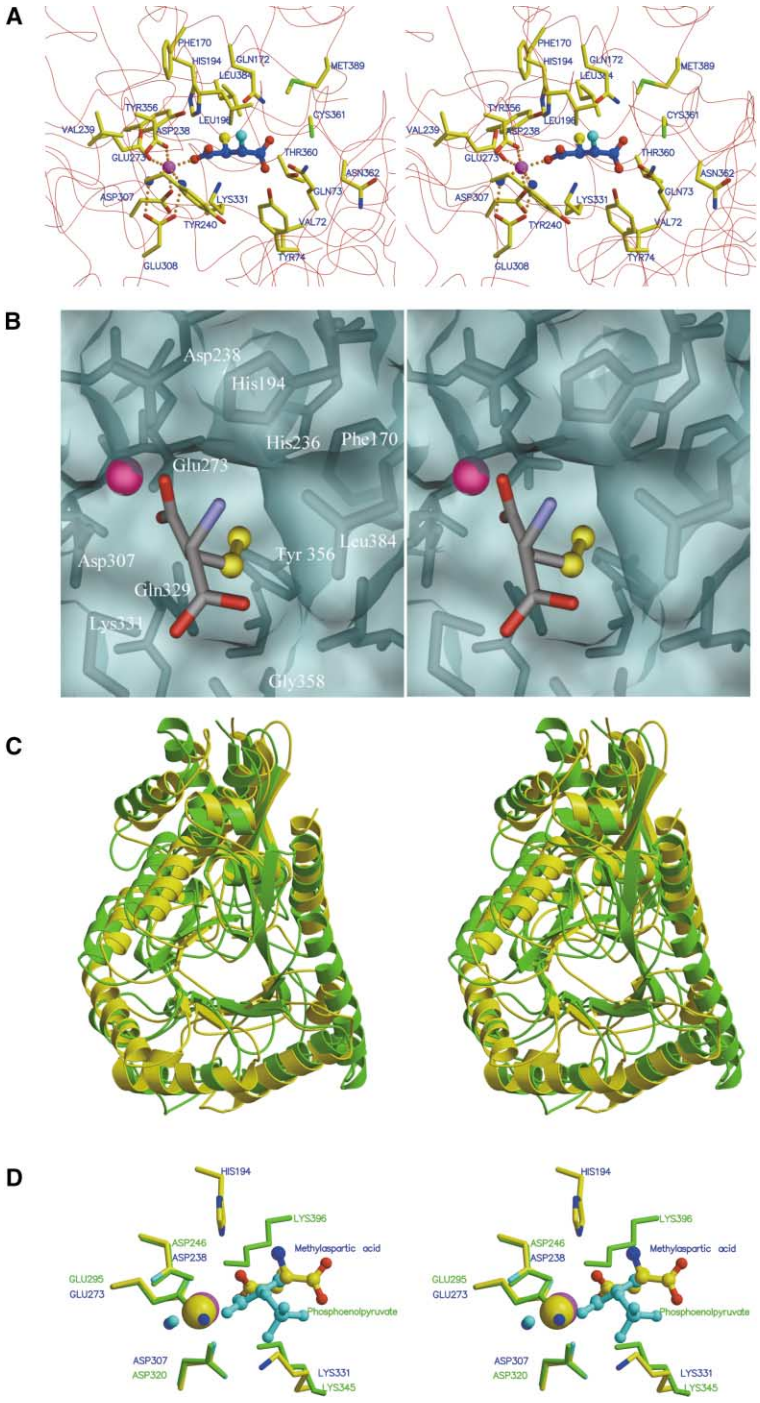


Figure 4. Stereo Diagrams

(A) The active site of *C. amalonaticus* MAL showing the Mg^{2+} (magenta), waters coordinating the metal (blue), the Cα backbone (red), the substrate (blue), and the side chains of all residues within 7.0 Å of the substrate (atom colors). Of these 25 residues, 22 are identical in all the MAL sequences, and the remaining three are conservative substitutions.

(B) A surface representation of the active site of MAL with the Mg^{2+} (magenta) and the substrate, (2S,3S)-3-methyl aspartic acid (atom colors). A clear surface depression can be seen adjacent to the 3-methyl of the substrate, which is able to accommodate the ethyl group of the alternative substrate, (2S,3S)-3-ethyl aspartic acid (shown modeled, yellow).

(C) An overlap between a single subunit of MAL (yellow) and enolase (green, Protein Data Bank code 1ONE), showing the close similarity in structure between these two enzymes.

(D) A close up of the superimposed metal binding sites in MAL (atom colors, cyan Mg^{2+}, and dark blue water) and enolase (green, yellow Mg^{2+}, and light blue water), together with their respective substrates, (2S,3S)-3-methyl aspartic acid (atom colors) and the enolase substrate analog, phosphenolpyruvate (cyan), showing the close similarity in the active sites. The three carboxyl residues and the water molecules that provide ligands to the metal are shown, together with the general base Lys-345 in enolase and the proposed, functionally analogous residue Lys-331 in MAL.

the MAL dimer is also identical to that between subunits related by one of the 2-fold axes in the enolase tetramer.

Members of the enolase superfamily have been shown to exhibit a number of conserved features. First, like those of MAL, their substrates bind in related positions at the C-terminal end of the TIM barrel. Second, they each contain a cluster of strongly conserved carboxyls that act to form the binding site for the divalent cation essential for catalysis, which, while differing somewhat in detail, contain one acidic residue whose spatial position is conserved (Asp-198 in MLEI, Asp-195 in MR, and Asp-246 in enolase [11]). Comparisons with MAL show that the carboxyls that ligate the metal (Asp-238, Glu-273, and Asp-307) superimpose to within 1 Å (for all atoms) with the metal ligating residues of these other enzymes (Figure 4D). Third, for each of the enzymes, the initial step of the reaction involves the abstraction of a proton α to the carboxylic acid of the substrate by a base to generate the enolic intermediate. Some enzymes of the superfamily contain two such residues, as dictated by their chemistry (for example, the racemization reaction catalyzed by MR). The (2S,3S)-3-methyl-

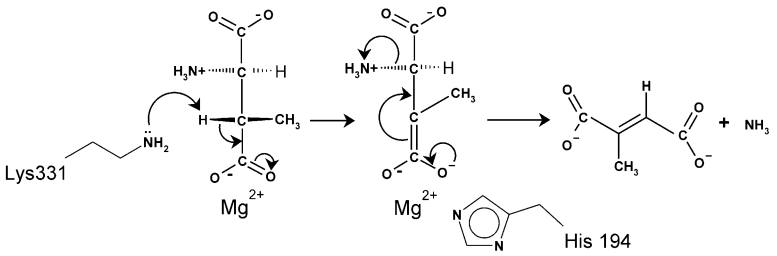


Figure 5. A Schematic of the Proposed Reaction Mechanism of MAL

The 3-proton of (2S,3S)-3-methyl aspartic acid is abstracted by Lys-331 acting as a base to give the enolic intermediate shown in the middle panel. The negative charge on the *aci*-carboxylate is stabilized by the metal ion and possibly by His-194 acting as an electrophile. The enolic intermediate collapses with the elimination of ammonia to yield mesaconic acid (right hand panel).

aspartic acid-bound complex of MAL reveals clearly that Lys-331 is ideally positioned to act as the base to remove the proton and thus carry out the first step of catalysis, as had been previously predicted [11]. Comparison with the other family members reveals that Lys-331 in MAL superimposes to within 1 Å (based on all atom superposition of the general bases and divalent cation ligating residues) of the general bases previously identified in enolase and MR. Fourth, an electrophilic residue (K164 in MR and K167 in MLEI) whose role is to neutralize the charge on the *aci*-carboxylate formed following abstraction of the 3-proton has been identified in some, but not all, members of the family. Analysis of the structure of MAL indicates that His-194 may perform this role in MAL.

On the basis of this analysis, the likely mechanism for MAL therefore involves abstraction of the 3-proton of the substrate by Lys-331, the stabilization of the enolic intermediate by the metal ion and possibly His-194, and its subsequent collapse and elimination of ammonia to give mesaconic acid (Figure 5). These data, together with the lack of any evidence for the posttranslational modification of any residues in the structure of MAL, confirms that the chemistry of MAL is completely unlike that of the other amino acid lyases, PAL and HAL.

Having shown that MAL is indeed a member of the enolase superfamily, we find it interesting to speculate why MAL is not structurally similar to the other amino acid lyases, HAL and PAL. The reactions catalyzed by the latter two enzymes cannot be accommodated by a mechanism related to the enolase superfamily, as neither of their substrates (histidine and phenylalanine, respectively) contains a carboxylate group α to the abstracted proton, and, thus, an enolic intermediate cannot be formed in these enzymes.

The α,β-elimination of ammonia from aspartate is catalyzed by aspartate amino acid lyase (AAL) [18], the structure of which is similar to that of HAL [8]. This would suggest that AAL has a chemistry like that of HAL and PAL, although there is currently no evidence for a posttranslationally modified residue at the active site of AAL. Our analysis on the chemistry of MAL suggests that the α,β-elmination of ammonia from aspartate could, in principle, proceed via an enolic intermediate, since, like 3-methyl aspartate, aspartic acid contains a carboxylate group α to the abstracted proton. However, since it is unclear whether a metal is involved in catalysis by AAL, a mechanism like HAL and PAL is perhaps more probable. Further studies on AAL will be valuable in establishing which of the two mechanisms are used, but, in principle at least, evolution could have led to convergence of a MAL-like mechanism using a structure derived by divergent evolution from a HAL-like precursor enzyme.

Whatever the mechanism of AAL, the analysis presented here clearly indicates that nature has evolved two different strategies for carrying out the related chemistries of the ammonia lyases, each of which exploits the inherently different potential of the substrates in dictating the chosen chemistry. One intriguing question to ask is why nature has chosen to use the enolase superfamily chemistry for MAL. Is this entirely serendipitous or is it a choice dictated by other factors, currently unknown?

Biological Implications

Methylaspartate ammonia lyase (MAL) catalyzes the highly stereo selective magnesium-dependent reversible addition of ammonia to mesaconic acid to give L-*threo*-(2S,3S)-3-methylaspartic acid, a reaction with considerable biotechnological potential for the chiral synthesis of novel amino acids. The crystallographic

Table 1. Data Collection Statistics

Data Set	MALP4			MALC2	β-Me-Asp
Space group	$P4_122$	$P4_122$	$P4_122$	C222	C222
Wavelength	λ_1(0.97963)	λ_2(0.97946)	λ_3(0.96863)	λ(0.979)	λ(1.54179)
Resolution (Å)	99–2.16	99–2.16	99–2.16	20–1.33	18–2.10
Highest resolution (Å)	2.21–2.16	2.21–2.16	2.21–2.16	1.36–1.33	2.15–2.10
Number of unique reflections	30925	30926	30913	245233	60480
Completeness (%)	97.3 (99.6)	97.3 (99.6)	99.2 (99.6)	95.8 (97.2)	93.4 (91.6)
Redundancy	3.8	3.8	3.6	3.1	3.0
R_{merge}[1] (%)	3.2 (5.0)	3.3 (5.2)	3.3 (5.4)	6.0 (33.9)	5.6 (21.8)
I/σI	36.5 (25.20)	36.3 (24.9)	34.4 (23.0)	18.9 (3.5)	18.2 (5.1)

Where appropriate, the values in parentheses are for data in the highest resolution shell.
[1] $R_{merge} = \Sigma_{hkl}|I_i - I_m|/\Sigma_{hkl}I_m$, where I_m is the mean intensity of the reflection.

Table 2. Refinement Statistics

Data Set		MALC2	MALP4	β-Me-Asp
Resolution (Å)		1.3	2.16	2.1
Number of reflections used		222972	30184	53530
Number of atoms		7811	3177	6866
Completeness (%)		95.7	100	93.2
R_{cryst}[1] (%)/R_{free}[2] (%)		17.4/19.8	22.5	16.7/22.7
Ramachandran plot[3]		90.6/8.4/0.4/0.6	88.2/10.8/0.3/0.6	90.5/8.3/0.6/0.6
RMS distance deviations[4] (Å)		0.011	0.037	0.045
RMS angles deviations (°)		1.4	2.6	2.0
Mean B values (Å^2)	Protein mainchain	15.3	11.5	27.9
	Protein sidechain	16.8	12.5	29.8
	Water	33.5	39.5	39.1
	SO_4	22.7	—	—
	Mg^2	—	—	26.6
	(2S-3S)-3-methylaspartic acid	—	—	26.0

[1] $R_{cryst} = \Sigma_{hkl}(|F_{obs}| - |F_{calc}|)/\Sigma_{hkl}|F_{obs}|$.
[2] R_{free} was calculated on 5% of the data omitted randomly.
[3] Ramachandran plot according to PROCHECK [32] statistics, percent of residues in most favored/additional allowed/disallowed regions, respectively.
[4] Root-mean-square (rms) bond lengths and angles are the rms deviations from ideal bond lengths and angles, respectively.

study of *Citrobacter amalonaticus* MAL presented here shows that the structure of this enzyme is based on a TIM barrel. Analysis of its binary complex with (2S,3S)-3-methylaspartic acid has led to the identification of residues lining the substrate 3-methyl binding site to guide the engineering of this site to alter the substrate specificity.

Structure comparisons show that MAL is closely related to members of the enolase superfamily, which suggests that the first step in the mechanism is the removal of the proton α to the 3-carboxyl of the substrate to generate an enolic intermediate. In contrast, the reactions catalyzed by other amino acid lyases, such as histidine ammonia lyase and phenylalanine ammonia lyase, are thought to proceed via the formation of a covalent intermediate between the substrate and a post-translationally modified residue in the active site. These latter two enzymes have a completely different three-dimensional structure to MAL, and it is clear that nature has evolved two different strategies for carrying out the related chemistries of these three enzymes, each of which exploits the inherently different potential of their substrates in dictating the chosen chemistry.

Experimental Procedures

Preparation of SeMet-MAL

E. coli strain DL41 (F-, lambda-, met-A28; CGSC #7177) was a gift from Yale University *E. coli* Genetic Stock Center [19, 20] and was transformed with the plasmid pMAL [1, 2, 21, 22] encoding the gene for *C. amalonaticus* MAL.

The transformed cells were grown in media supplemented with selenomethionine and harvested, and the protein was purified as previously described [13].

Protein Crystallization

Selenomethionine-labeled MAL from *Citrobacter amalonaticus* was purified and crystallized as described previously for the unlabeled protein [13]. Crystals were obtained from hanging drops after approximately 72 hr using a well solution consisting of 0.1 M potassium phosphate buffer (pH 8.0) and 32% ammonium sulfate. The crystals grew in two different forms under the same conditions; tetragonal form A crystals (a = 66.0 Å, b = 66.0 Å, and c = 233.1 Å, with a monomer in the asymmetric unit) and C-centered orthorhombic form B crystals (a = 129.5 Å, b = 238.9 Å, and c = 66.3 Å, with a dimer in the asymmetric unit). Prior to data collection, the crystals were cryoprotected in 25% glycerol and 45% ammonium sulfate in 0.1 M potassium phosphate buffer (pH 8.0) before flash freezing in a stream of nitrogen at 100 K.

Data Collection and Processing

Data were collected to 2.1 Å resolution at three wavelengths (0.979628 Å, 0.979463 Å, and 0.968634 Å) from a single cryofrozen SeMet form A crystal at station BM 30 of the ESRF at Grenoble, using the inverse beam rotation camera method and a MAR345 image plate (Table 1). The three wavelengths corresponded to the Se K-edge f′ minimum, f′′ maximum, and high-energy remote positions, as determined on the basis of an X-ray fluorescence spectrum collected directly from the crystal, which was analyzed by the program Chooch [23]. A second data set was collected to a resolution of 1.3 Å from a single cryofrozen SeMet form B crystal at station PX 9.6 at the SRS synchrotron (Daresbury), using the rotation camera method and an ADSC Quantum 4 CCD detector (Table 1). Data to 2.1 Å were also collected on a single form B crystal soaked in cryoprotectant containing 10 mM (2S,3S)3-methylaspartic acid plus 5 mM $MgCl_2$, using CuKα X-rays generated from a Rigaku RU200 generator, MSC/YALE focusing mirrors, and a MAR345 image plate (Table 1). All data sets were scaled and integrated using the HKL suite [24], and, unless otherwise stated, handled subsequently using the CCP4 suite [25].

Form A Crystal Structure Solution

MAD Phasing

The form A diffraction data were analyzed using the program SOLVE [26], and the anomalous signal was found to extend to the resolution limit of these data (2.1Å). The SeMet substructure was determined using SOLVE and all data from the form A crystal. A solution for all ten of the selenium atoms in a single polypeptide chain was found only in space group $P4_122$, confirming this to be the space group of the form A crystals. Electron density maps using the experimental phases were calculated to 2.1 Å, in which the protein conformation could be clearly discerned.

Automatic Model Building and Refinement

To produce an atomic model of MAL, automatic model building was performed with the program ARP/wARP 5.0 [27] using the phases output from SOLVE at 2.1 Å. The warpNtrace procedure within ARP/wARP 5.0 produced a partial model containing 380 of the 413 residues with a mixture of fully and partially built side chains. The remaining parts of the main chain structure were built manually using O [28]. Side chains were not rebuilt at this point. This model was then subjected to the side chain identification and sequence-docking

procedure associated with warpNtrace [27]. The sequence was docked with 100% certainty, and the resulting model was subjected to several rounds of rebuilding and refinement with REFMAC [29] using data in the resolution range 20–2.1 Å. Water molecules were added with the program ARP [30], which was run in conjunction with REFMAC. The final model of the form A crystals contained 4156 atoms, including 992 solvent, with R = 0.225 (Table 2).

Form B Crystal Structure Solution

Molecular Replacement

The final model of the form A monomer was used as the starting point for the molecular replacement solution of the high-resolution form B data. Initially, a dimer was generated by rotation about one of the 2-fold axes present in the $P4_122$ crystal form. This model was then subjected to a cross-rotation search using the program CNS [31]. Two clear peaks, related by the molecular 2-fold axes, gave correlation coefficients of 0.0990 and 0.0961, respectively, and one was taken forward into the CNS translation search. A clear solution was only obtained in space group C222, confirming this to be the space group of the form B crystals, at $\theta_1 = 269.93°$, $\theta_2 = 91.38°$, $\theta_3 = 193.25°$, X = −83.75 Å, Y = −107.59 Å and Z = 0.79 Å, with a correlation coefficient of 0.5. The resulting model was subjected to simulated annealing and density modification, again using the CNS procedure. This produced a map of excellent quality, clearly describing the expected structure. Following several rounds of rebuilding, solvent addition, and refinement using REFMAC5, the final dimeric model of the form B crystals contained 7811 atoms, including 1564 solvent, with R = 0.174 and R_{free} = 0.198 (Table 2). No discernible electron density could be seen for the two C-terminal residues (Lys-412 and Asp-413) in either subunit, and these residues have been omitted from the final model. The structure of the MAL (2S,3S)-3-methylaspartic acid binary complex was initially determined by subjecting the high resolution form B model to rigid body refinement, using the binary complex data. Rounds of refinement and rebuilding resulted in a final model of 6866 atoms, including one Mg^{2+} and one (2S,3S)-3-methylaspartic acid per subunit and 707 solvent molecules, with R = 0.167 and R_{free} = 0.227 (Table 2). Analysis of the stereochemical quality of the models was accomplished using the program PROCHECK [32]. The refinement statistics are summarized in Table 2. The program LSQKAB [33] was used for the superposition of different structures, and sequence alignments was undertaken with ALSCRIPT [34]. All figures have been produced using TURBO FRODO [35] and MOLSCRIPT [36].

Acknowledgments

We thank Michel Roth of the ESRF, Grenoble, France, for station alignment and help in data collection. This work was supported by BBSRC, the New Energy and Industrial Development Organization, (NEDO), the British Council/ The Royal Society Anglo-Japanese Scientific Exchange Scheme, the Wolfson Foundation, and the Wellcome Trust. C.W.L. is supported by a BBSRC industrial case award with SYNGENTA. The Krebs Institute is a BBSRC-funded molecular recognition center and a member of the North of England Structural Biology Consortium.

Received: August 21, 2001
Revised: November 27, 2001
Accepted: November 28, 2001

References

1. Asano, Y., and Kato, Y. (1994). Crystalline 3-methylaspartate from a facultative anaerobe, *E. coli* strain YG-1002. FEMS Microbiol. Lett. *118*, 255–258.
2. Kato, Y., and Asano, Y. (1995). Purification and properties of crystalline 3-methylaspartase from two facultative anaerobes, *Citrobacter sp.* strain YG-0504 and *Morganella morganii* strain YG-0601. Biosci. Biotechnol. Biochem. *59*, 93–99.
3. Goda, S.K., Minton, N.P., Botting, N.P., and Gani, D. (1992). Cloning, sequencing, and in *Escherichia coli* of the *Clostridium tetanomorphum* gene encoding β-Methylaspartase and characterization of the recombinant protein. Biochemistry *31*, 10747–10756.
4. Asuncion, M., Barlow, J.N., Pollard, J., Staines, A.G., McMahon, S.A., Blankenfeldt, W., Gani, D., and Naismith, J. H. (2001). Overexpression, purification, crystallization and data collection of 3-methylaspartase from *Clostridium tetanomorphum*. Acta Crystallogr. D Biol. Crystallogr. *57*, 731–733.
5. Schuster, B., and Rétey, J. (1995). The mechanism of action of phenylalanine ammonia-lyase: the role of prosthetic dehydroalanine. Biochemistry *92*, 8433–8437.
6. Langer, M., Lieber, A., and Retey, J. (1994). Histidine ammonia-lyase mutant S143C is posttranslationally converted into fully active wild type enzyme, evidence for serine-143 to be the precursor of active site dehydroalanine. Biochemistry *33*, 14034–14038.
7. Langer, M., Reck, G., Reed, J., and Retey, J. (1994). Identification of serine-143 as the most likely precursor of dehydroalanine in the active site of histidine ammonia lyase-A study of the overexpressed enzyme by site-directed mutagenesis. Biochemistry *33*, 6462–6467.
8. Schwede, T.F., Rétey, J., and Schulz, G.E. (1999). Crystal structure of histidine ammonia-lyase revealing a novel polypeptide modification as the catalytic electrophile. Biochemistry *38*, 5355–5361.
9. Pollard, J.R., Richardson, S., Akhtar, M., Lasry, P., Neal, T., Botting, N.P., and Gani, D. (1999). Mechanism of 3-methylaspartate probed using deuterium and solvent isotope effects and active site directed reagents: identification of an essential cysteine residue. Bioorg. Med. Chem. *7*, 949–975.
10. Bright, H.J. (1964). The mechanism of the β-methylaspartase reaction. J. Biol. Chem. *239*, 2307–2315.
11. Babbitt, P.C., Hasson, M.S., Wedekind, J.E., Palmer, D.R.J., Barrett, W.C., Reed, G.H., Rayment, I., Ringe, D., Kenyon, G.L., and Gerlt, J.A. (1996). The enolase superfamily: A general strategy for enzyme-catalyzed abstraction of the α-protons of carboxylic acids. Biochemistry *35*, 16489–16501.
12. Weaver, T., Levitt, D., Banaszak, L., Donnelly, M., and Wilkens-stevens, P. (1995). The crystal structure of fumarase-C from *E. coli.* FASEB J. *9*, A1466–A1466.
13. Levy, C.W., Buckley, P.A., Baker, P.J., Sedelnikova, S., Rodgers, F., Li, Y.F., Kato, Y., Asano, Y., and Rice, D.W. (2001). Crystallization and preliminary X-ray analysis of *Citrobacter amalonaticus* methylaspartate ammonia lyase. Acta Crystallogr. D Biol. Crystallogr. *57*, 1922–1924.
14. Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E., et al. (1975). Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. Nature *255*, 609–614.
15. Kallarakal, A.T., Mitra, B., Kozarich, J.W., Gerlt, J.A., Clifton, J.G., Petsko, G.A., and Kenyon, G.L. (1995). Mechanism of the reaction catalysed by mandelate racemase: structure and mechanistic properties of the K166R mutant. Biochemistry *34*, 2788–2797.
16. Kajander, T., Kahn, P.C., Passila, S.H., Cohen, D.C., Lehtio, L., Adolfsen, W., Warwicker, J., Schell, U., and Goldman, A. (2000). Buried charge surface in proteins. Structure *8*, 1203–1214.
17. Wedekind, J.E., Poyner, R.R., Reed, G.H., and Rayment, I. (1994). Chelation of serine 39 to Mg^{2+} latches a gate at the active site of enolase: structure of the bis(Mg^{2+}) complex of yeast enolase and the intermediate analog phosphonoacetohydroxamate at 2.1 Å resolution. Biochemistry *33*, 9333–9342.
18. Shi, W., Dunbar, J., Jayasekera, M.K., Viola, R.E., and Farber, G.K. (1997). The structure of L-aspartate ammonia lyase from *Escherichia coli.* Biochemistry *36*, 9136–9144.
19. Hendrickson, W., Stoner, C., and Schleif, R. (1990). Characterization of the *E. coli* ARAFGH and ARAJ promoters. J. Mol. Biol. *215*, 497–510.
20. LeMaster, D.M., and Richards, F.M. (1985). H-1-N15 heteronuclear NMR-studies *E. coli* thioredoxin in samples isotopically labelled by residue type. Biochemistry *24*, 7263–7268.
21. Kato, Y., and Asano, Y. (1995). 3-methylaspartate ammonia-

lyase from a facultative anaerobe, strain YG-1002. Appl. Microbiol. Biotechnol. *43*, 901–907.
22. Kato, Y., and Asano, Y. (1997). 3-methylaspartate ammonia-lyase as a marker enzyme of the mesaconate pathway for (*S*)-glutamate fermentation in Enterobacteriaceae. Arch. Microbiol. *168*, 457–463.
23. Evans, G. (1994). Chooch (http://lagrange.mrc-lmb.cam.ac.uk/doc/gwyndaf/Chooch.html).
24. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.
25. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.
26. Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. *55*, 849–861.
27. Perrakis, A., Morris, R., and Lamzin, A. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. *6*, 458–463.
28. Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.
29. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.
30. Lamzin, V.S., and Wilson, K.S. (1997). Automated building of solvent structure combined with standard restrained refinement. Methods Enzymol. *277*, 269–305.
31. Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system. A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.
32. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
33. Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. Acta Crystallogr. A *32*, 922–923.
34. Barton, G. (1993). ALSCRIPT: a tool to format multiple sequence alignments. Protein Eng. *6*, 37–40.
35. Roussel, A., and Cambillau, C. (1991). Silicon Graphics Geometry Partners (Mountain View, CA: Silicon Graphics).
36. Kraulis, P.J. (1991). MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures. J. Appl. Crystallogr. *24*, 946–950.

Accession Numbers

The atomic coordinates for MAL and the MAL (2S,3S)-3-methylaspartic acid binary complex have been deposited in the Protein Data Bank (accession numbers 1KKO and 1KKR).