The Structure of 3-Methylaspartase from Clostridium tetanomorphum Functions via the Common Enolase Chemical Step*

Received for publication, November 21, 2001, and in revised form, December 12, 2001
Published, JBC Papers in Press, December 17, 2001, DOI 10.1074/jbc.M111180200

Miryam Asuncion‡§, Wulf Blankenfeldt‡§, John N. Barlow¶, David Gani¶, and James H. Naismith¶†

From the ‡The Centre for Biomolecular Sciences, The University, St. Andrews, Scotland, United Kingdom KY16 9ST and the §School of Chemical Sciences, The University of Birmingham, Edgbaston, Birmingham, United Kingdom B15 2TT

Methylaspartate ammonia-lyase (3-methylaspartase, MAL; EC 4.3.1.2) catalyzes the reversible anti elimination of ammonia from l-threo-(2S,3S)-3-methylaspartic acid to give mesaconic acid. This reaction lies on the main catabolic pathway for glutamate in Clostridium tetanomorphum. MAL requires monovalent and divalent cation cofactors for full catalytic activity. The enzyme has attracted interest because of its potential use as a biocatalyst. The structure of C. tetanomorphum MAL has been solved to 1.9-Å resolution by the single-wavelength anomalous diffraction method. A divalent metal ion complex of the protein has also been determined. MAL is a homodimer with each monomer consisting of two domains. One is an αβ-barrel, and the other smaller domain is mainly β-strands. The smaller domain partially occludes the C terminus of the barrel and forms a large cleft. The structure identifies MAL as belonging to the enolase superfamily of enzymes. The metal ion site is located in a large cleft between the domains. Potential active site residues have been identified based on a combination of their proximity to a metal ion site, molecular modeling, and sequence homology. In common with all members of the enolase superfamily, the carboxylic acid of the substrate is co-ordinated by the metal ions, and a proton adjacent to a carboxylic acid group of the substrate is abstracted by a base. In MAL, it appears that Lys831 removes the α-proton of methylaspartic acid. This motif is the defining mechanistic characteristic of the enolase superfamily of which all have a common fold. The degree of structural conservation is remarkable given only four residues are absolutely conserved.

The facile synthesis of optically pure chiral compounds remains one of the holy grails of organic chemistry. One approach is to use stereospecific or highly stereoselective enzyme catalysts to control, for example, the stereochemical course of addition reactions (1, 2) or the selection of specific enantiomers of the substrate in kinetic resolutions. However, the exquisite substrate selectivity of enzymes is also one of their principle limitations, their catalytic activity often being restricted to a few substrates. Rational redesign of protein catalysts has had several high profile successes but has so far failed to open up large areas of organic chemistry to biocatalysis. An alternative approach that shows considerable promise is to couple rational design with directed evolution methods to select enzyme activity (e.g. Altamirano et al. (3)). In such examples enzyme residues are chosen for random mutagenesis on the basis of function (e.g. substrate binding or catalysis), and the required enzyme activity is selected using a suitable screen. Of course, the ability to assign function to enzyme amino acid residues is greatly aided by a three-dimensional structural model.

One enzyme that shows considerable potential for use in organic synthesis is 3-methylaspartic acid ammonia-lyase (MAL). MAL is a 45.5-kDa enzyme found on the main catabolic pathway for glutamate in Clostridium tetanomorphum (4) and a number of other anaerobic microorganisms. MAL catalyzes the reversible anti elimination of ammonia from l-threo-(2S,3S)-3-methylaspartic acid to give mesaconic acid (Fig. 1). MAL also catalyzes the stereo- and regioselective addition of ammonia to several derivatives of mesaconic acid to form a limited number of homochiral substituted aspartic acids (1, 5).

Access to these synthetically useful compounds by conventional synthesis is extremely difficult and not well developed. The accessible range of synthetic homochiral substituted aspartic acids could potentially be extended using engineered MAL. However, engineering of MAL is greatly hindered by the lack of structural information and the absence of homologues in the Protein Data Bank.

The mechanism of MAL has attracted some interest. It was postulated that MAL belonged to the enolase superfamily of enzymes on the basis of distant sequence homology (6) and the requirement of MAL for two metal ion cofactors (Mg2+ and K+) (7). The enolase superfamily of enzymes catalyzes a wide variety of transformations, including racemization, β-elimination of water, and cycloisomerization, all of which are initiated by a common metal-assisted, general base-catalyzed abstraction of the α-proton of a carboxylic acid to generate a stabilized enolate anion intermediate (8). This enolate intermediate can be partitioned into different products by suitable modification of active site residues as demonstrated with mandelate racemase (9).

Alternatively to this mechanism and on the basis of chemical modification studies, it had been proposed that SeO32− of MAL is dehydrated post-translationally to dehydroalanine (9, 10). The unusual dehydroalanine prosthetic group, which is absent

© 2002 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.
in other members of the enolase superfamily, could function as a Michael acceptor in the deamination reaction (11, 12). Strong support for a similar post-translational modification in the ENO-like enzyme histidine ammonia-lyase (histidase) has been provided by x-ray crystallography (13). In histidase, an internal Ala-Ser-Gly tripeptide (residues 142–144) undergoes intrachain cyclization accompanied by dehydration of Ser to form a 4-methylidene-imidazol-5-one moiety. Interestingly, Ser of 3-methylaspartase is also present in an Ala-Ser-Gly tripeptide.

A fuller understanding of the enzyme specificity and mechanism clearly requires a three-dimensional structure. Here we report the structure of MAL from C. tetanomorphum. The structure of the protein from a different organism has been determined and is to be published shortly (14). The structure establishes that MAL belongs to the enolase superfamily, having the same α/β-barrel topology characteristic of the family, and suggests that Lys is the general base that abstracts the C-3 proton of 3-methylaspartate. It is evident from the crystal structure of MAL that Ser is unmodified. The active site of MAL is similar in architecture to enolase. Both use two metal ions and a catalytic base during catalysis. A sequence alignment based on structural similarity emphasizes the tolerance of the enolase-type reaction for sequence substitution. This strongly suggests that directed evolution experiments with this superfamily, where the first key common chemical step is preserved, should be successful.

MATERIALS AND METHODS

Protein Expression, Crystallization, and Structure Determination—MAL from C. tetanomorphum was purified and crystallized as described previously (15). Crystals were grown by sitting drop vapor diffusion against a well solution of 20–25% (w/v) polyethylene glycol 6000, 100 mM sodium acetate, pH 7.0, 25 mM Tris-HCl, pH 7.0, as precipitating solution and using 16–22% ethylene glycol as additive.

Se-methionine-labeled protein was expressed in the methionine auxotroph Escherichia coli BB84 following the protocol of Doublie (16). Diffraction data were measured from a shock-cooled crystal that was grown by sitting drop vapor diffusion against a well solution of 20% (v/v) 25% (w/v) polyethylene glycol 6000, 100 mM sodium acetate, and using 16% (v/v) 22% ethylene glycol as additive.

The peptide chain comes back and forms a helix (Leu85–Gly103), then a very highly conserved Pro99. The bend induced by Glu163 wraps around the outside of the main domain (Fig. 2). It starts with a three-stranded anti-parallel α-zigzag. At the end of the first strand, a compact structural unit (Asp19–Gln45) sits high above the entrance to the C-terminal α/β-barrel domain. Continuing our analogy of cylinder and arm, this region represents the hand. The last strand of the meander continues with a long loop (residues Val72–Leu84) extending toward the entrance to the active site. This loop is highly conserved among the different MAL sequences and contains the conserved Arg. The peptide chain comes back and forms a helix (Leu84–Gly87), which is packed against the zigzag. The helix is broken in the middle by a very highly conserved Pro. The bend induced by the Pro allows the helix to fit closely to the first two strands of the zigzag.

RESULTS AND DISCUSSION

Overall Structure—MAL is a dimeric enzyme. Both monomers are found in the asymmetric unit and are related by a 2-fold axis. They are effectively identical with a root mean square deviation between the 412 Ca atoms of each monomer of 0.35 Å. The principal differences between the monomers occur in a loop at residue 28 and a rigid body movement of residues Ala–Ala. The differences in these regions are consistent with crystal contacts. The monomer can be further decomposed into two subunits: a smaller N-terminal domain and a C-terminal α/β-barrel (Fig. 2). The domains are linked by a short helix (Met–Tyr) flanked by two stands. The C-terminal end of the barrel is partially obscured by two loops from the N-terminal domain. The result is the formation of a large open cleft (−20 Å long, 16 Å wide, and 19 Å deep) with the loops from the N-terminal domain on one side and the C-terminal end of the barrel on the other side (Fig. 2). The overall shape of the molecule resembles a squat cylinder with a long arm attached to the cylinder rim. The arm is bent so that it reaches over the middle of the cylinder.

N-terminal Domain—The smaller N-terminal domain, Met–Glu wraps around the outside of the main domain (Fig. 2). It starts with a three-stranded anti-parallel β-zigzag. At the end of the first strand, a compact structural unit (Asp–Gln) sits high above the entrance to the C-terminal α/β-barrel domain. Continuing our analogy of cylinder and arm, this region represents the hand. The last strand of the meander continues with a long loop (residues Val–Leu) extending toward the entrance to the active site. This loop is highly conserved among the different MAL sequences and contains the conserved Arg. The peptide chain comes back and forms a helix (Leu–Gly), which is packed against the zigzag. The helix is broken in the middle by a very highly conserved Pro. The bend induced by the Pro allows the helix to fit closely to the first two strands of the zigzag.

α/β-Barrel Domain—The main C-terminal domain Ile–Gly is an 8-fold α/β-barrel. The topology of the barrel is β0α (βα)8. The inner β-barrel is not entirely parallel with the second strand being antiparallel to the other strands, and the direction of the first helix is reversed with respect to the other helices. There are additional secondary structural elements on the outside of the barrel that do not belong to the core barrel fold. Ser, the residue postulated to be modified to dehydroalanine, is located on the first β-strand of the barrel. The side chain points toward the helix against which the β-strand stacks. The electron density is entirely consistent with a Ser side chain, and the main chain shows no unusual structural features, which might be expected if there was unsaturation present. Mass spectroscopy of the protein confirms that this residue in solution is Ser. The conserved Cys is oxidized in all structures (Se-Met, native, and metal ion) despite the presence of antioxidants in all protein solutions. The F′ − F″ electron density maps shows three clear peaks distributed around the sulfur atom. At the resolution of our study it is impossible to absolutely distinguish between single oxidation with static disorder or multiple oxidation with no disorder. However, a single oxygen at three positions, each with a third occupancy, refines with a substantially lower B-factor than the connecting sulfur (2.5 versus 13 Å2). Three oxygen atoms, each with full occupancy, refine to approximately the same B-factor as the
residues are in these conformations. His194 is part of the metal
tron density clearly indicates clearly in both cases that the
Val391, adopt disallowed

sulfur (17 versus 15 Å²). This suggests complete oxidation to
sulfenic acid. Mass spectroscopic data (not shown) shows a +16
m/z peak for a trypic digest of protein in solution immediately
after purification. This suggests that this residue is unusually
sensitive to oxidation. Two residues in the barrel, His194 and
Val201, adopt disallowed ψ/φ backbone conformations. The
electron density clearly indicates clearly in both cases that the
residues are in these conformations. His194 is part of the metal
binding site and is discussed below. Val201 is part of a loop at
the dimer interface.

Dimer Structure—The two molecules in the asymmetric unit
make extensive contacts, burying 4700 Å² of surface area, 62%
of which is apolar (values from www.biochem.ucl.ac.uk/bsm/
PP/server/). This extensive contact area is consistent with so-
lution scattering results (not shown), which indicate that the
molecule exists as a dimer in solution. The dimer has dimen-
sions of 76 × 60 × 60 Å. Particularly striking is that the clefts
seen in each monomer connect to form a 60-Å-long channel in
the dimer. There are a number of contacts between the monom-
ers. These contacts occur between the outer surfaces of the
last, first, and second helix of the barrel and the

| Data collection | MAD (ESRF BM14)* |
|-----------------|------------------|
| Wavelength (Å)  | 0.978997 (peak)  | 0.979173 (inflection) | 0.946439 (remote) |
| Resolution (highest shell, Å) | 48.795–2.32 | (2.38–2.32) | 42.258–2.24 |
| (2.38–2.32) | 48.795–2.31 | (2.37–2.30) | |
| Space group     | P212121         | P212121         | P212121 |
| Cell constants (Å; °) | a = 67.27, b = 109.26, c = 108.99; |
|                 | a = β = γ = 90 |
| Vm              | 2.19            |
| Total measurements | 134,146        | 237,234          | 148,995 |
| Unique reflections | 30,325         | 34,228           | 33,640 |
| Average redundancy | 4.3 (2.4)     | 6.8 (2.6)       | 4.3 (2.4) |
| I/σ             | 11.0 (5.1)     | 11.6 (3.4)      | 9.3 (4.7) |
| Completeness (%) | 95.2 (54.8)    | 98.3 (78.3)     | 95.1 (53.1) |
| Anom. completeness (%)^b | 90.4 (31.8) | 87.4 (43.1) | 90.2 (31.0) |
| Rmerge^c         | 4.5 (12.8)     | 5.6 (20.5)      | 4.9 (15.6) |
| f^c/F^c          | −9.30/5.86     | −9.91/2.23      | −4.82/3.32 |

a MAD, multiwavelength anomalous diffraction; ESRF, European Synchrotron Radiation Facility.
^b Anomalous completeness corresponds to the fraction of possible acentric reflections for which an anomalous difference has been measured.
^c Rmerge = ΣΣI(hj) − (I(h)/ΣΣI(hj)), where I(hj) is the measured diffraction intensity and the summation includes all observations.
Mg$^{2+}$. In the metal-free structure the water molecules fill the metal coordination site, and the side chains adopt slightly different conformations. In the dimer the divalent metal ion binding site from each monomer sits at almost opposite ends of the large channel. His194 hydrogen bonds to Glu238, presumably to assist in the stabilization of the enolate.

**MAL Is a Member of the Enolase Superfamily**—Structure comparisons identify MAL as a member of the enolase superfamily. This family carries out a diverse series of chemical reactions with one step in common, the abstraction of a proton to a carboxylic acid function. MAL is structurally most similar to muconate-lactonizing enzyme (29) and mandelate racemase (9) (Z score above 20) but less similar to enolase (28), glucarate dehydratase (30), and o-succinylbenzoate synthase (31) (Z score below 10). A sequence alignment based on structural superposition is shown in Fig. 3. Despite the structural similarity, only four residues are conserved, and the sequence identity of MAL to the others members of the family is below 20%. This strikingly illustrates the tolerance of this fold for amino acid substitution. Two of the conserved residues, Asp238 and Glu273, are ligands to the divalent metal ion. The other two are Pro275 and Glu352. These residues fulfill structural roles in the protein and are chemically inert. In the structural superposition of MAL with enolase, the divalent metal ion binding site is effectively identical. This supports and extends the point made by Babbit et al. (6) and developed by Hasson et al. (29) that the chemistry carried out by the enzyme depends only on the metal ion to which the carboxylic acid of the substrate binds. In addition to the recognition of the carboxylic acid, the metal ion stabilizes the enolate intermediate that develops after abstraction of the proton in effect lowering the pK$_a$ of the α-proton. This stabilization may be assisted by His194, which hydrogen bonds to Glu238, and His236, which hydrogen bonds to Glu273. His194 occupies the same position as Lys164 in mandelate racemase (9) and Lys213 in glucarate dehydratase (30), residues which are proposed to perform the same function. The enzyme can then use an appropriately positioned base to remove the α-proton. The enzyme can change the position and nature of the other catalytic residues to achieve catalysis (9, 29).

**The Active Site Location in MAL**—An alignment of various MAL sequences shows that conserved residues cluster around the metal binding site and the open cleft. Comparison with the other members of the superfamily, particularly the structure of enolase with its substrate (28), confirms that these residues constitute the active site. The muconate-lactonizing enzyme study (29) confirmed the predictions based on the enolase structure (28) that the principal catalytic base, which abstracts the proton α to the carboxylic acid, is located close to the metal binding site in the same spatial orientation in all enolase-type enzymes.

![Structure of Methylaspartase](http://www.jbc.org/content/8309/1/MalB.png)
enzymes. Hasson et al. (29) further showed by structural superposition that the nature of the principal catalytic base varies (Asp, which acts on nearby His in mandelate racemase (9), and Lys in the others). In MAL Lys331 occupies the same spatial relationship relative to the metal ion site seen in the other enzymes, and therefore we assign it as the principal catalytic base. This assignment supports the inference drawn from the fold of the enzyme that MAL proceeds by an enolase-type mechanism.

**Mechanistic Implications of This Study**—We have been unable to obtain a convincing electron density for a substrate complex, and thus we have had to model in the substrate to the enzyme active site. This is relatively straightforward given the enolase substrate complex structure. It is clear that on substrate binding the protein cleft closes up and that in our structures we have an open conformation. The cleft is closed as the two domains move toward each other hinging on the loop connecting them (Asn159–Ala166). A "closed" model can be constructed by superimposing each domain of MAL separately onto the closed enolase structure. The enolase substrate 2-phospho-D-glycerate and L-methylaspartate have the carboxylic acid and the proton to that function in common. There is a further similarity as enolase catalyzes the elimination of water and MAL catalyzes the elimination of ammonia (Fig. 4). In our model the carboxylic acid of the side chain of L-methylaspartic acid is bound in a bidentate manner to the Mg$^{2+}$ ion (Fig. 4). Lys$^{331}$ is positioned to abstract the proton α to this carboxylate (the 3S proton). The β-methyl group of the MAL substrate (same location as the hydroxyl in the enolase substrate) is bound in an open pocket formed by residues Gln$^{172}$, Phe$^{170}$, Leu$^{384}$, and Tyr$^{356}$. This binding pocket is quite large, consistent with observation that the enzyme can tolerate re-
placement of the methyl group with propyl groups (5, 9–11). Recognition of the “main chain” carboxylic acid (C1) of β-methylaspartate is accomplished by the conserved residue Arg<sup>80</sup>. We have not been able to identify the location of the second metal ion binding site, and its role in catalysis remains obscure. Although enolase requires two divalent metal ions (32), the second ion in MAL is monovalent. Interestingly MAL has been shown to turn over the other diastereomer of the substrate, the 2S,3R-epimer L-erythro-3-methylaspartic acid (33–36). In this molecule the α-proton is presented to the opposite face of the protein than for the usual substrate. This means that Lys<sup>331</sup> is on the wrong face for such catalysis, and there must be an alternative base. Interestingly, although the main chain of His<sup>194</sup> superimposes on Lys<sup>164</sup> of mandelate racemase (Lys<sup>211</sup> of glucarate dehydratase), the side chain occupies a similar position in space to the N<sub>C</sub> atom of Lys<sup>166</sup> (Lys<sup>213</sup> of glucarate dehydratase). Lys<sup>166</sup> in mandelate racemase is the second base (9), and Lys<sup>213</sup> of glucarate dehydratase is the principal base. It remains to be tested whether His<sup>194</sup> has this function in addition to its normal one of hydrogen bonding to the Asp<sup>238</sup> (metal ligand).

In summary, the data confirm the enolase superfamily has a new member, 3-methylaspartate lyase. This enzyme, contrary to some predictions, uses a metal ion to promote proton abstraction from the carbon α to the carboxylic acid function. The structure identifies the key catalytic residues and shows the very minimal degree of sequence conservation in the enolase superfamily. Only four residues are absolutely conserved of which only two are involved at the active site. However, the fold of the family is quite clear, and the key chemical step, abstraction of the proton α to a carboxylic acid, is apparently conserved. This implies that this class of enzyme should be particularly attractive for directed evolution experiments.

REFERENCES
1. Akhtar, M., Botting, N. P., Cohen, M. A. & Gani, D. (1987) Tetrahedron 43, 5899–5908
2. Akhtar, M., Cohen, M. A. & Gani, D. (1987) Tetrahedron Lett. 28, 2413–2416
3. Altamirano, M. M., Blackburn, J. M., Aguayo, C. & Fersht, A. R. (2000) J. Mol. Biol. 294, 1061–1067
4. Barker, H. A., Smyth, R. D., Wilson, R. M. & Weissbach, H. (1958) Proc. Natl. Acad. Sci. U. S. A. 45, 1159–1161
5. Babbit, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Knoy, G. L. & Gerlt, J. A. (1995) Science 267, 1159–1161
6. Goda, S. K., Minton, N. P., Botting, N. P. & Gani, D. (1992) Biochemistry 31, 10747–10756
7. Pollard, J., Richardson, S., Laery, P., Neal, T., Akhtar, M., Botting, N. P. & Gani, D. (1999) Bioorg. Med. Chem. 7, 949–957
8. Gulzar, M. S., Morris, K. & Gani, D. (1995) J. Chem. Soc. Perkin Trans. 2, 8017–8026
9. Babbit, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R., Barrett, W. C., Reed, G. H., Rainey, I., Ringe, D., Kenyon, G. L. & Gerlt, J. A. (1996) Biochemistry 35, 16489–16501
10. Gerlt, J. A. & Babbitt, P. C. (1998) Curr. Opin. Chem. Biol. 2, 607–612
11. Poyner, R. R., Cleland, W. W. & Reed, G. H. (1998) Biochemistry 37, 10649–10657
12. Archer, C. H. & Gani, D. J. (1993) Chem. Soc. Chem. Commun. 140–142
13. Archer, C. H., Johnson, R. B., Petsko, G. A. & Ringe, D. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 10396–10401
14. Gulick, A. M., Hubbard, B. K., Gerlt, J. A. & Rainey, I. (2001) Biochemistry 40, 10654–10662
The Structure of 3-Methylaspartase from *Clostridium tetanomorphum* Functions via the Common Enolase Chemical Step

Miryam Asuncion, Wulf Blankenfeldt, John N. Barlow, David Gani and James H. Naismith

*J. Biol. Chem.* 2002, 277:8306-8311.
doi: 10.1074/jbc.M111180200 originally published online December 17, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M111180200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 33 references, 2 of which can be accessed free at http://www.jbc.org/content/277/10/8306.full.html#ref-list-1