New Reactions in the Crotonase Superfamily: Structure of Methylmalonyl CoA Decarboxylase from *Escherichia coli*[†,‡]

Matthew M. Benning,[§] Toomas Haller,[||] John A. Gerlt,*[,||] and Hazel M. Holden*[,§]

Department of Biochemistry, University of Illinois, Urbana, Illinois 61801, and Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53705

Received December 16, 1999; Revised Manuscript Received February 8, 2000

ABSTRACT: The molecular structure of methylmalonyl CoA decarboxylase (MMCD), a newly defined member of the crotonase superfamily encoded by the *Escherichia coli* genome, has been solved by X-ray crystallographic analyses to a resolution of 1.85 Å for the unliganded form and to a resolution of 2.7 Å for a complex with an inert thioether analogue of methylmalonyl CoA. Like two other structurally characterized members of the crotonase superfamily (crotonase and dienoyl CoA isomerase), MMCD is a hexamer (dimer of trimers) with each polypeptide chain composed of two structural motifs. The larger N-terminal domain contains the active site while the smaller C-terminal motif is α-helical and involved primarily in trimerization. Unlike the other members of the crotonase superfamily, however, the C-terminal motif is folded back onto the N-terminal domain such that each active site is wholly contained within a single subunit. The carboxylate group of the thioether analogue of methylmalonyl CoA is hydrogen bonded to the peptidic NH group of Gly 110 and the imidazole ring of His 66. From modeling studies, it appears that Tyr 140 is positioned within the active site to participate in the decarboxylation reaction by orienting the carboxylate group of methylmalonyl CoA so that it is orthogonal to the plane of the thioester carbonyl group. Surprisingly, while the active site of MMCD contains Glu 113, which is homologous to the general acid/base Glu 144 in the active site of crotonase, its carboxylate side chain is hydrogen bonded to Arg 86, suggesting that it is not directly involved in catalysis. The new constellation of putative functional groups observed in the active site of MMCD underscores the diversity of function in this superfamily.

Members of the crotonase (enoyl CoA hydratase) superfamily catalyze a diverse range of chemical reactions, each involving the formation of an oxyanion intermediate derived from an acylated coenzyme A. Enolate anion intermediates, or transition states with significant localization of negative charge on the thioester carbonyl group, are thought to occur in the reactions catalyzed by crotonase (*1*), $\Delta^{3,2}$-enoyl and $\Delta^{3,5}$,$\Delta^{2,4}$-dienoyl CoA isomerases (*2, 3*), carnitinyl CoA racemase (*4*), feruloyl CoA hydratase/lyase (*5*), 1,4-dihydroxynaphthoyl CoA synthase (*6, 7*), 2-ketocyclohexylcarboxyl CoA hydrolase (*8*), and 4-chlorobenzoyl CoA dehalogenase (*9, 10, 11*). 3-Hydroxyisobutyryl CoA hydrolase (*12, 13*) catalyzes the hydrolysis of a thioester bond: although a mechanism involving an anionic tetrahedral intermediate is likely, the initial formation of an enolate anion intermediate is possible. A recent review of the crotonase superfamily can be found in ref *14*.

High-resolution X-ray structures are available for three members of this superfamily: rat mitochondrial crotonase (*15, 16*), 4-chlorobenzoyl CoA dehalogenase (*17*), and $\Delta^{3,5}$,$\Delta^{2,4}$-dienoyl CoA isomerase (*18*). In the case of the crotonase and the isomerase, the six identical subunits are arranged as a dimer of trimers. The dehalogenase packs in the crystalline lattice as a trimer. In all three of these enzymes, the subunit architecture can be described in terms of two distinct motifs: the larger N-terminal domain that is characterized by 10 strands of β-pleated sheet and the smaller C-terminal domain that is composed of three α-helices. The N-terminal domain contains the binding site for the acylthioester moiety of the substrate. Two peptidic NH groups located in this domain form an oxyanion hole in which the thioester carbonyl group of the substrate is bound, and the anionic intermediate/transition state is stabilized. This conserved structural feature allows the hypothesis that the evolution of this superfamily has been dominated by the need to stabilize enolate anion intermediates that otherwise would be too unstable to be kinetically competent. The acid/base functional groups that catalyze the diverse range of reactions characteristic of this superfamily are also located within the N-terminal domain. The smaller C-terminal domain provides various hydrophobic and hydrophilic interactions that stabilize the quaternary structure of the enzyme. Additionally, some of the functional groups in the C-terminal motif interact with the pyrophosphate moiety of a substrate molecule in a neighboring molecule. As such, the active sites in these enzymes are formed by two separate subunits comprising the trimer or hexamer.

Both three-dimensional structural comparisons as well as amino acid sequence alignments reveal that no active site

[†] This research was supported in part by NIH Grants GM-40570 and GM-52594 (to J.A.G.) and GM-55513 (to H.M.H.).

[‡] X-ray coordinates have been deposited with the Protein Data Bank.

* Address correspondence to these authors. (H.M.H.) E-mail: Hazel_Holden@biochem.wisc.edu. Fax: 608-262-1319. Phone: 608-262-4988. (J.A.G.) E-mail: j-gerlt@uiuc.edu. Fax: 217-265-0385. Phone: 217-244-7414.

[§] University of Wisconsin.

[||] University of Illinois.

10.1021/bi9928896 CCC: $19.00
Published on Web 03/29/2000

catalytic groups are strictly conserved within this superfamily. In some members, but not all, a glutamate residue is thought to be involved in proton removal from or delivery to carbon-2 of the thioester substrate, e.g., Glu 164 in the rat mitochondrial crotonase catalyzed reaction. In those reactions that involve the addition of water to an enoyl CoA, a conserved glutamate residue sometimes activates the attacking water molecule for addition to the 3-carbon, e.g., Glu 144 in the crotonase catalyzed reaction. In the reactions catalyzed by $\Delta^{3,5},\Delta^{2,4}$-dienoyl CoA isomerase and 4-chlorobenzoyl CoA dehalogenase, a conserved Asp residue serves either as an acid/base catalyst (Asp 204 in the isomerase) or as a nucleophile (Asp 145 in the dehalogenase). However, for many members of the superfamily, the correlation between structure and function is still too ill-defined to allow either the prediction of mechanism or the identities of the acid/base catalysts. The fact that the same secondary structural elements within the N-terminal domain deliver both the groups necessary for catalysis and for substrate specificity adds to the ambiguities.

Our laboratories have initiated a study of the structural/functional relationships in additional members of the crotonase superfamily in order to explore the manner in which similar active sites have evolved to catalyze different overall reactions. These investigations complement and contrast ongoing studies of members of the enolase superfamily (*19*). In the enolase superfamily, the active sites are found at the interfaces between two domains: the catalytic groups are located in conserved positions at the ends of the β-strands in an $(\alpha/\beta)_8$-barrel motif, while the specificity determinants are positioned in flexible loops in an N-terminal domain. This active site design, in some instances, has allowed the prediction of function for unknown members of the superfamily identified in genome sequence projects (*20*).[1]

For the investigation described here, a "hypothetical enzyme" designated YgfG in the *Escherichia coli* genome sequencing project (*21*) was selected as the initial target for detailed characterization of the crotonase superfamily. The gene encoding YgfG is located in an operon that was expected to be amenable to functional discovery. Alignment of the amino acid sequence of YgfG with those of other members of the superfamily suggested the presence of a conserved homologue of Glu 144 (Glu 113) but not of Glu 164 in the rat mitochondrial crotonase. On the basis of amino acid sequence alignments, it was clear that the configuration of the functional groups in YgfG was unique in the crotonase superfamily, however, and suggested that YgfG catalyzed a new chemical reaction. As reported in the companion paper, YgfG (hereafter referred to as MMCD[2]) is indeed a novel, biotin-independent methylmalonyl CoA decarboxylase. The operon also encodes methylmalonyl CoA mutase (Sbm) and a propionyl CoA:succinate CoA transferase (YgfH), thereby providing a previously unknown pathway for the decarboxylation of succinate in *E. coli*.

Here we describe the three-dimensional structure of MMCD, both in the apo- and liganded state. In sharp contrast to the previously solved structures of members of this superfamily, the active site of MMCD is contained within a single polypeptide. The active site contains at least one functional group that likely is involved in the decarboxylation of methylmalonyl CoA, namely, Tyr 140. Strikingly, the side chain of the conserved Glu 113 is hydrogen bonded to Arg 86, thereby suggesting that this carboxylate group has no direct role in catalysis.

MATERIALS AND METHODS

Preparation of 2S-Carboxypropyl CoA. 2*S*-Carboxypropyl CoA, an inert thioether analogue of methylmalonyl CoA, was synthesized as described for the 2*R*-diastereomer (*22*) except that methyl 3-bromo-2*S*-methylpropionate instead of its enantiomer was used to alkylate coenzyme A. The analogue was purified by reverse-phase HPLC using a semipreparative C_{18} column (10 × 250 mm; 218TP from Vydac) at a flow rate of 4 mL/min. After an injection of the sample and an initial wash with 1% acetonitrile/99% 100 mM ammonium acetate, pH 6.5, for 5 min, the analogue was eluted with a linear gradient of acetonitrile and 100 mM ammonium acetate, pH 6.5, in which the acetonitrile concentration was increased from 5% to 15% over 50 min. The retention time of the analogue was 10.5 min.

Crystallization and Preparation of Heavy-Atom Derivatives. MMCD for crystallization trials was purified via an N-terminal His-tag with a Ni^{2+}-chelating column followed by cleavage of the His-tag with thrombin and a final anion exchange chromatographic step (*23*). Large single crystals of apo-MMCD were grown at room temperature by batch methods using macroseeding techniques from 3% poly(ethylene glycol) 8000, 100 mM ammonium sulfate, 100 mM CHES (pH 9.0), and 5 mM sodium azide. Crystal growth was generally complete within 2 weeks, with some achieving dimensions of 0.2 mm × 0.4 mm × 0.6 mm. These crystals belonged to the space group *C*2 with unit cell dimensions of $a = 130.3$ Å, $b = 115.4$ Å, $c = 80.3$ Å, and $\beta = 127.8°$. The asymmetric unit contained three subunits.

Single crystals of the liganded enzyme were obtained at room temperature by cocrystallizing MMCD with 2*S*-carboxypropyl CoA. The crystals were grown via hanging drop methods from 5% poly(ethylene glycol) 8000, 50 mM magnesium chloride, 100 mM MES (pH 6.0), and 5 mM sodium azide. The concentration of the inhibitor was 2.5 mM. These crystals belonged to the space group $P2_1$ with unit cell dimensions of $a = 85.7$ Å, $b = 142.4$ Å, $c = 89.7$ Å, and $\beta = 109.2°$. The asymmetric unit contained six subunits.

For preparation of heavy-atom derivatives, apo-MMCD crystals were transferred to a synthetic mother liquor containing 8% poly(ethylene glycol) 8000, 200 mM ammonium sulfate, 100 mM CHES (pH 9.0), and 5 mM sodium azide. Two isomorphous heavy-atom derivatives were prepared with 5 mM potassium dicyanoaurate (I) and 5 mM ammonium tetrathiotungstate (VI).

X-ray Data Collection and Processing. All X-ray data sets for the apo-form of MMCD, the heavy-atom derivatives, and the thioether complex were collected at −5 °C with a Bruker HiStar area detector system. The X-ray source was nickel-filtered Cu Kα radiation from a Rigaku RU200 X-ray generator operated at 50 kV and 90 mA and equipped with a 300-μm focal cup. Friedel pairs were measured for the

[1] B. K. Hubbard, D. Z. Schmidt, J. Delli, and J. A. Gerlt, unpublished observations.

[2] Abbreviations: CHES, 2-(cyclohexylamino)ethanesulfonic acid; MES, 2-(*N*-morpholino)ethanesulfonic acid; MMCD, methylmalonyl CoA decarboxylase; rms, root mean square.

Table 1: X-ray Data Collection Statistics

	resolution (Å)	independent reflections	completeness (%)	av I/ av σ (I)	R_{sym} (%)[a]	R_{iso} (%)[b]	phasing power[c]
apo-MMCD	30−2.70	24 197	91	19.4	3.13		
apo-MMCD (−172 °C)	30−1.85	74 572	89	23.3	2.25		
ammonium tetrathiotungstate	30−2.70	24 656	93	19.0	3.10	20.4	1.31
potassium dicyanoaurate	30−2.70	24 429	92	22.6	2.80	18.5	1.25
MMCD/thioether complex	30−2.70	49 603	87	15.3	4.30		

[a] $R_{sym} = (\sum|I - \bar{I}|/\sum I) \times 100$. [b] $R_{iso} = (\sum|F_N - F_H|/\sum|F_N|) \times 100$. $|F_N|$ is the native structure factor amplitude, and $|F_H|$ is the derivative structure factor amplitude. [c] Phasing power is the ratio of the rms heavy-atom scattering factor amplitude to the rms lack of closure error.

potassium dicyanoaurate(I) derivative. All the X-ray data sets were initially collected to 2.7 Å resolution and were processed with the software package XDS (*24*, *25*). These X-ray data sets were internally scaled with XCALIBRE.[3] Relevant X-ray data collection statistics are presented in Table 1. Only one crystal was required per X-ray data set.

Each heavy-atom derivative data set was placed on the same scale as the native data set by a "local" scaling procedure developed by Wesenberg et al. The *R* factors between the native and the heavy-atom derivative data sets are listed in Table 1. An additional X-ray data set for apo-MMCD was collected at −172 °C. Prior to flash-cooling, an apo-MMCD crystal was transferred to a synthetic mother liquor containing 12% poly(ethylene glycol) 8000, 300 mM ammonium sulfate, 23% ethylene glycol, 100 mM CHES (pH 9.0), and 5 mM sodium azide. The apo-MMCD X-ray data set extended to a nominal resolution of 1.85 Å. It was not possible to extend the resolution of the X-ray data set for the MMCD/thioether complex beyond 2.7 Å resolution however. The $P2_1$ crystal form inherently diffracts X-rays less well.

Structural Determination and Least-Squares Refinement of Apo-MMCD. The positions of the heavy-atom binding sites were determined by inspection of difference Patterson maps calculated with X-ray data from 30 to 5.0 Å. The ammonium tetrathiotungstate(VI) and the potassium dicyanoaurate(I) derivatives each contained three heavy-atom binding sites. The heavy-atom derivatives were placed on a common origin by difference Fourier maps, and the positions and occupancies for each heavy-atom binding site were refined by the origin-removed Patterson-function correlation method to 2.7 Å resolution (*26*). The heavy-atom derivatives had similar but not identical binding sites. Anomalous difference Fourier maps calculated from 30 to 5.0 Å were employed for determining the correct hand of the heavy-atom constellation. Protein phases were calculated with the program HEAVY (*26*). Relevant phase calculation statistics can be found in Table 1.

An electron density map calculated to 2.7 Å resolution clearly revealed the overall course of the polypeptide chain. The map was subsequently improved by the techniques of molecular averaging around the local 3-fold rotational axis and solvent flattening according to the method of Bricogne (*27*). From the density-modified map, it was possible to trace the entire polypeptide chain for one subunit. The MMCD subunit contains 261 amino acid residues (*21*). There were only three breaks in the electron density, which corresponded to the loops delineated by Leu 71−Asp 77 in subunit I, Ile 68−Asp 77 in subunit II, and Asp 67−Leu 79 in subunit III. Using the X-ray data set collected at low temperature, the model was subsequently refined by least-squares analysis to 1.85 Å resolution with the software package TNT (*28*). Relevant refinement statistics are given in Table 2. The final *R* factor was 18.0% for all measured X-ray data from 30 to 1.85 Å resolution. The only significant outlier on the Ramachandran plot of all non-glycinyl residues for subunit I is Asn 12, which adopts dihedral angles of $\phi = 63°$ and $\psi = -112°$. The election density is unambiguous in this region, and this asparagine adopts the same conformation in subunits II and III as well. For the refinement of the apo-MMCD model, 10% of the X-ray data were excluded for the required calculation of R_{free} as listed in Table 2. In that all X-ray data are important for the Fourier synthesis, however, these data were ultimately included in the final stages of the refinement and model building. A representative portion of the electron density map is shown in Figure 1a. The average temperature factors for the waters and the polypeptide chain backbone atoms were 28.5 and 18.4 Å^2, respectively.

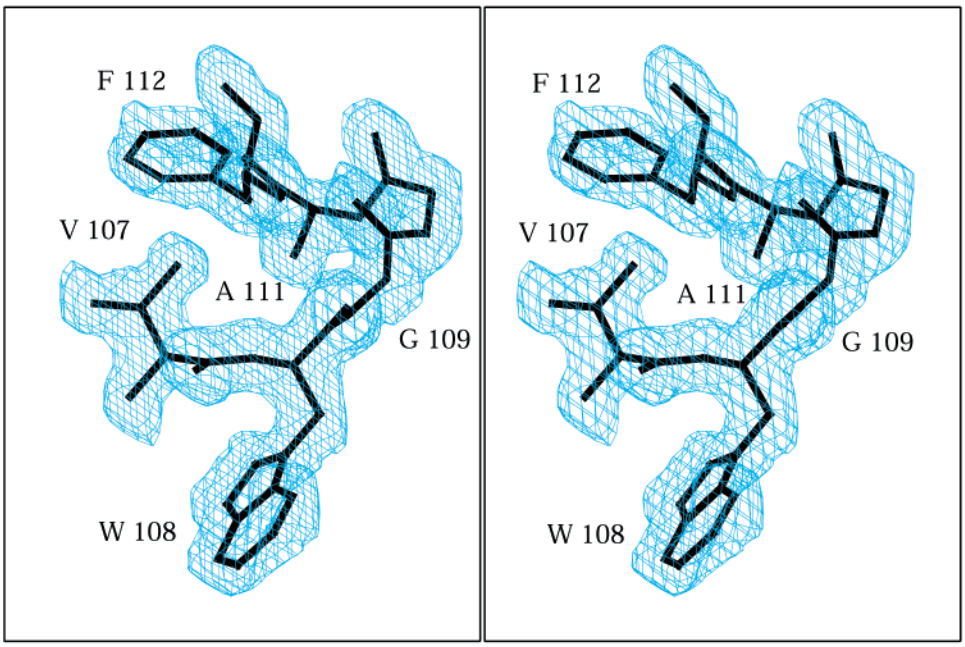
Table 2: Least-Squares Refinement Statistics

	apo-MMCD	MMCD/thioether complex
resolution limits (Å)	30.0−1.85	30.0−2.70
R factor (%)[a]	18.0	19.0
R-free (%)	23.7	
no. of reflections used	72 476	49 630
no. of protein atoms	5975	12 624
no. of solvent atoms[b]	604	0
Weighted rms Deviations from Ideality		
bond length (Å)	0.012	0.012
bond angle (deg)	2.03	2.27
planarity (trigonal) (Å)	0.007	0.007
planarity (other planes) (Å)	0.007	0.009
torsional angle (deg)[c]	16.3	20.2

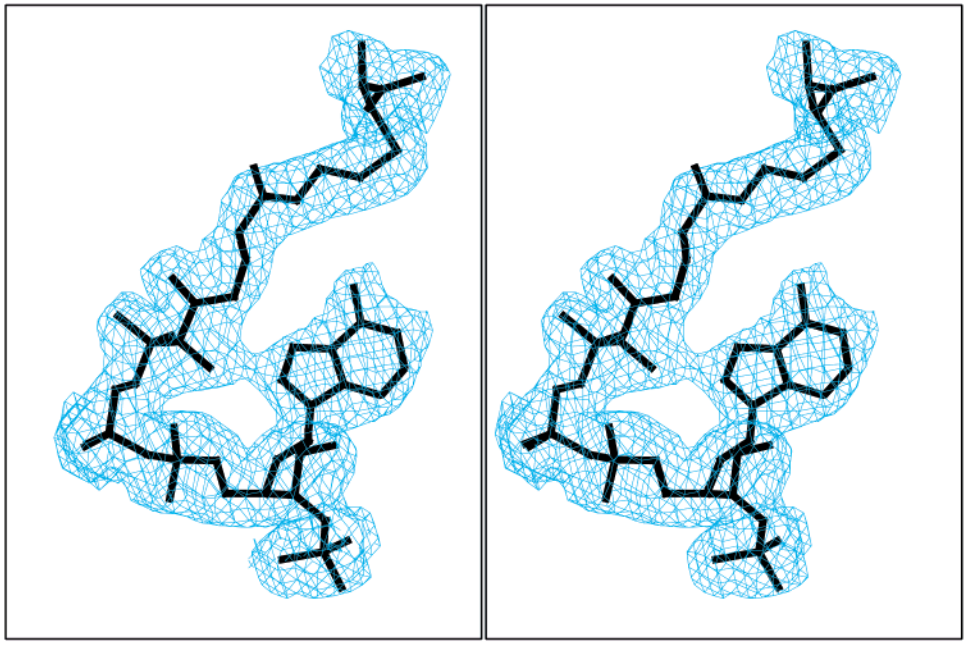
[a] R factor $= (\sum|F_o - F_c|/\sum|F_o|) \times 100$ where F_o is the observed structure factor amplitude and F_c is the calculated structure factor amplitude. [b] In addition to 603 water molecules, there was one nickel ion included in the refinement. [c] The torsional angles were not restrained during the refinement.

Structural Determination and Least-Squares Refinement of the MMCD/Thioether Complex. The structure of the MMCD/thioether complex was solved by molecular replacement with the software package AMORE (*29*, *30*) and employing the apo-MMCD trimer as a search model. A search sphere radius of 15 Å was utilized. Both the cross-rotational functions and translational searches were conducted with X-ray data from 10 to 4.0 Å. Two solutions were obtained corresponding to the two trimers in the asymmetric unit. Following rigid-body refinement with X-ray data from 20 to 4.0 Å, the correlation coefficient was 68.7% and the *R* factor was 37.6%. To expedite the model-building process and to remove bias due to the molecular replacement procedure, the electron densities (calculated with coefficients of the form $[2F_o - F_c]$, where F_o was the native structure

[3] G. Wesenberg and I. Rayment, unpublished results.



(a)



(b)

FIGURE 1: Representative electron densities. A portion of the electron density map corresponding to apo-MMCD is shown in panel a. The map was contoured at 1σ and calculated with coefficients of the form ($2F_o - F_c$), where F_o was the native structure factor amplitude and F_c was the calculated structure factor amplitude. With the exception of the surface loop delineated by Leu 71 to Asp 77 in subunit I (and the corresponding regions in subunits II and III), the electron density was of this quality throughout the map. Shown in panel b is the electron density corresponding to the bound thioether analogue of methylmalonyl CoA. This electron density is based on an averaged map calculated to 2.7 Å resolution as described in the Materials and Methods section.

factor amplitude and F_c was the calculated structure factor amplitude) corresponding to the six subunits in the asymmetric unit were averaged according to the algorithm of Bricogne (*27*). From this modified electron density map, one subunit of the MMCD/thioether complex was built. This averaged model was placed back into the unit cell and subjected to least-squares refinement at 2.7 Å resolution. Following refinement, the electron densities corresponding to the six monomers in the asymmetric unit were averaged again. The model for one subunit was adjusted on the basis of this averaged map. From this map it was possible to build those amino acid residues lying between Leu 71 and Asp 77. A final cycle of least-squares refinement reduced the *R* factor to 19.0% for all measured X-ray data from 30 to 2.7 Å. Due to the limited resolution of the X-ray data, no solvent molecules were included in the refinement, and the model

was never manually adjusted within the unit cell. As such, the following discussion concerning the MMCD/thioether complex will refer only to the averaged model. Electron density corresponding to the bound ligand is displayed in Figure 1b. Relevant refinement statistics are given in Table 2.

RESULTS AND DISCUSSION

Overall Three-Dimensional Structure of Apo-MMCD. Crystals of the apo-form of MMCD contain three subunits in the asymmetric unit that are arranged about a local 3-fold rotational axis. These subunits are virtually identical and can be superimposed upon one another with an average rms deviation of 0.24 Å for all backbone atoms. Upon further examination of the packing arrangement within the crystalline lattice, it is obvious that one such trimer packs against a symmetry-related one to form a hexameric species. MMCD is a hexamer in solution as judged by its elution volume in gel filtration chromatography.[4] A ribbon representation of the MMCD hexamer is depicted in Figure 2a. The orientation of Figure 2a is such that the crystallographic dyad running parallel to the *b*-axis lies perpendicular to the plane of the page. As can be seen in Figure 2a, the hexamer, with overall dimensions of approximately 89 Å × 85 Å × 101 Å, can be envisioned as a dimer of trimers. The solvent-accessible surface area lost upon dimerization of the trimers is approximately 3000 Å^2 as calculated according to the method of Lee and Richards (*31*) with a probe sphere radius of 1.4 Å. The trimer−trimer interface is formed primarily by a β-strand−α-helix motif delineated by Glu 218−Gln 231. As can be seen in Figure 2a, two β-strands, one contributed by each subunit, form a short stretch of antiparallel β-sheet. Specifically subunits I and VI, II and V, and III and IV interact in such a manner as to produce three regions of antiparallel β-sheet within the trimer−trimer interface. There are quite striking stacking interactions that occur between Phe 227 (subunit I) and Phe 227 (subunit VI) and between Phe 227 (subunit III) and Phe 227 (subunit IV). Because of the static disorder in the crystalline lattice, however, the phenylalanines at positions 227 in subunits II and V could only be modeled as alanine residues.

In the initial electron density maps, it was immediately obvious that there was a metal bound to the trimer contained within the asymmetric unit. The identity of this cation was assumed to be nickel because the protein employed for crystallization trials was purified via an N-terminal His-tag with a Ni^{2+}-chelating column. The metal is positioned along the local 3-fold rotational axis, and as indicated in Figure 2b, the histidines positioned at 220 in subunits I−III serve as ligands to the nickel. As described later, this region of polypeptide chain deviates significantly from that observed in other members of the crotonase superfamily. To ensure that this observed change in the three-dimensional structure in MMCD was not an artifact of metal-binding, a site-directed mutant protein was prepared whereby His 220 was changed to a glutamine residue. Subsequent structural analysis of the H220Q protein (possessing >50% the activity of wild-type MMCD) demonstrated that the observed changes in the three-dimensional structure of MMCD were not a function of metal binding. Indeed, the structures of the apo-MMCD and H220Q proteins were virtually identical with the exception that the H220Q form no longer contained bound metal ions.[5] Like that observed for the trimer−trimer interface in the hexamer, the subunit−subunit interface within the individual trimers is quite extensive with a buried surface area of ∼3400 Å^2 for each monomer. The molecular interfaces of the trimer are formed primarily by Met 116−Asp 119, Arg 150−Thr 163, Ile 177−His 180, His 197−Glu 212, and Met 222−Arg 229 in each subunit.

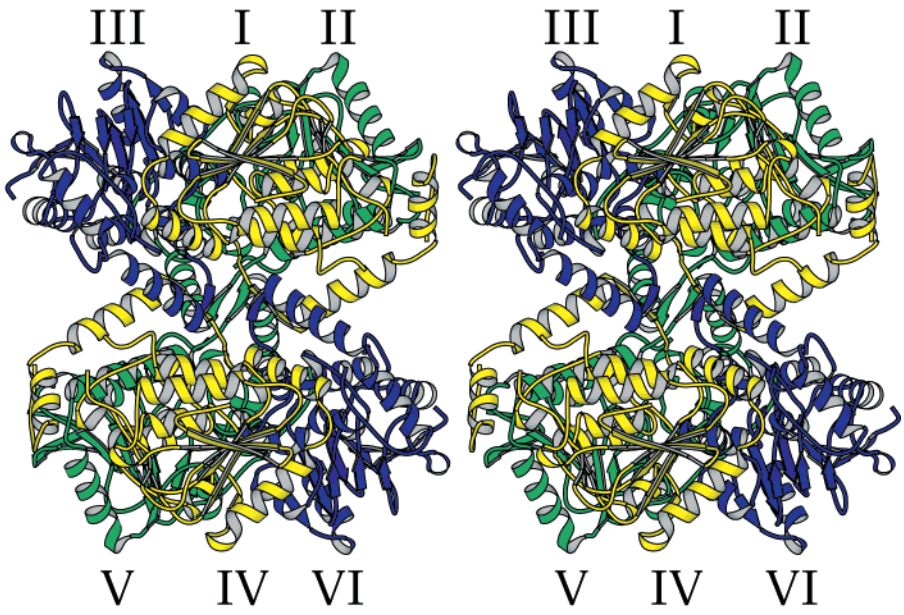
Table 3: List of Secondary Structural Elements

amino acid	secondary structural element
Tyr 3−Val 6	type I turn
Asn 7−Thr 10	β-strand (A)
Ile 11−Val 14	type II′ turn
Ala 15−Phe 19	β-strand (B)
Gly 22−Leu 25	type I turn
Lys 24−Ala 27	type III′ turn
Lys 30−Asp 42	α-helix (A)
Arg 45−Ile 48	type I turn
Cys 50−Ala 55	β-strand (C)
Pro 56−Ser 59	type II turn
Val 61−Ser 63	β-sheet (D)
Asp 67−Glu 70	type III turn
Ile 68−Leu 71	type III turn
Asp 77−Ser 80	∼type I turn
Ser 80−Asp 83	type I turn
Pro 84−Lys 94	α-helix (B)
Ile 99−Trp 108	β-strand (E)
Gly 110−Ser 117	α-helix (C)
Leu 120−Ala 123	β-strand (F)
Ala 124−Ser 127	type I turn
Thr 128−Met 131	β-strand (G)
Pro 133−Asn 135	α-helix (D)
Leu 142−Ile 145	α-helix (E)
His 146−Thr 149	type I turn
Phe 154−Thr 163	α-helix (F)
Ser 165−Thr 168	β-strand (H)
Ala 169−Ala 172	α-helix (G)
Leu 173−Gly 176	type I turn
His 180−Glu 183	β-strand (I)
Val 184−Ile 198	α-helix (H)
Pro 203−Glu 218	α-helix (I)
His 220−Asn 223	β-strand (J)
Ser 224−Leu 251	α-helix (J)

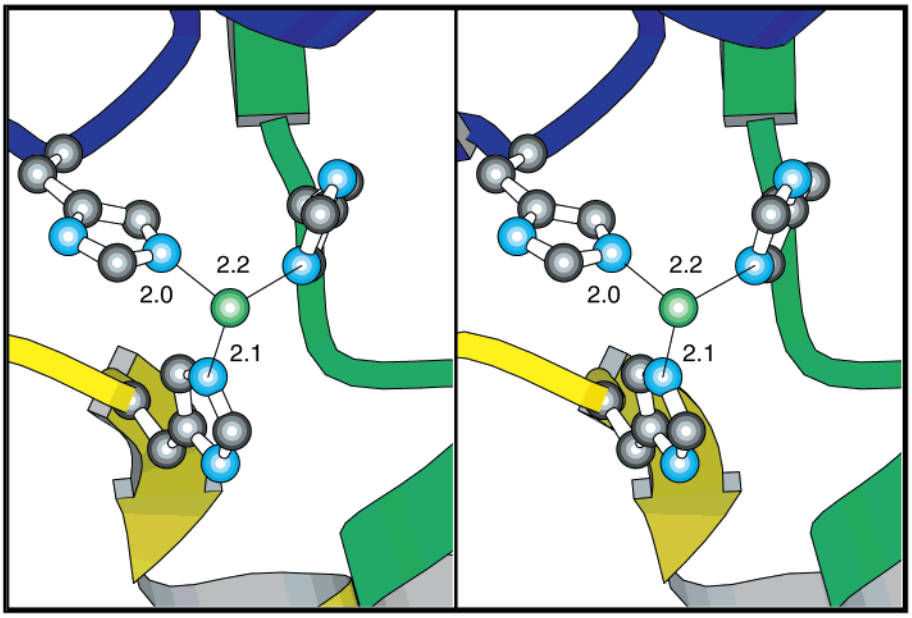
A ribbon representation of an MMCD subunit with the bound thioether analogue of methylmalonyl CoA is shown in Figure 3. The MMCD subunit is roughly spherical with overall dimensions of approximately 52 Å × 52 Å × 55 Å. A list of the secondary structural elements describing the overall fold of the molecule is given in Table 3. There are 10 strands of β-sheet that range in length from 3 to 10 amino acid residues. The first nine β-strands form two layers of mixed β-sheet, composed of six and four β-strands, respectively. β-Strand E is shared between the two layers of sheet. Both β-strands D and E are interrupted by β-bulges as a result of the dihedral angles adopted by Phe 62 ($\phi = -69°$ and $\psi = -49°$) and Gly 105 ($\phi = 102°$ and $\psi = -179°$), respectively. As described above, the last β-strand of the monomer, namely, that defined by His 220−Asn 223, plays a key structural role in the trimer−trimer interface of the hexameric species. The 10 β-strands of the subunit are connected together by 10 α-helices that range in length from

[4] T. Haller and J. A. Gerlt, unpublished observations.

[5] M. M. Benning, T. Haller, J. A. Gerlt, and H. M. Holden, unpublished results.



(a)

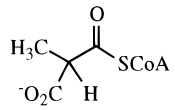


(b)

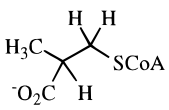
FIGURE 2: Ribbon representation of apo-MMCD. (a) The MMCD hexamer can be envisioned as a dimer of trimers. Subunits I−III form the trimeric species observed in the asymmetric unit and are related to subunits IV−VI by crystallographic dyad symmetry. As displayed here, the orientation of the hexamer is such that the crystallographic 2-fold running parallel to the *b*-axis lies perpendicular to the plane of the page. (b) Within the individual trimeric species, there is a single nickel-binding site located along the local 3-fold rotational axis. The imidazole ligands are provided by His 220 from all three subunits. Ligand−bond distances are given in ångströms.

3 to 28 amino acid residues. The last α-helix, helix J, displays a decided kink as a result of Ser 240, which adopts torsional angles of approximately $\phi = -78°$ and $\psi = 164°$. The loop defined by Leu 71−Asp 77, which was disordered in the apo-form of MMCD, folds into a series of type I and type III turns as listed in Table 3. Note that the polypeptide chain backbones for apo-MMCD and the MMCD/thioether complex superimpose with an rms value of 0.46 Å.

The Active Site Region. MMCD was cocrystallized with 2*S*-carboxypropyl CoA, an inert thioether analogue of methylmalonyl CoA. This analogue inhibits the MMCD-catalyzed reaction with a dissociation constant comparable to the K_m for methylmalonyl CoA (12 μM).



Methylmalonyl CoA



Thioether Analog

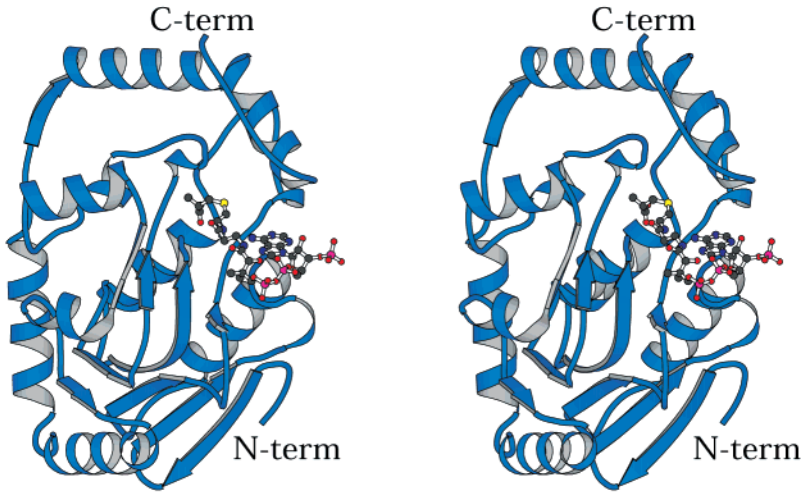


FIGURE 3: Ribbon representation of one subunit of the MMCD/thioether complex. The substrate analogue of methylmalonyl CoA used in this investigation is shown in a ball-and-stick representation.

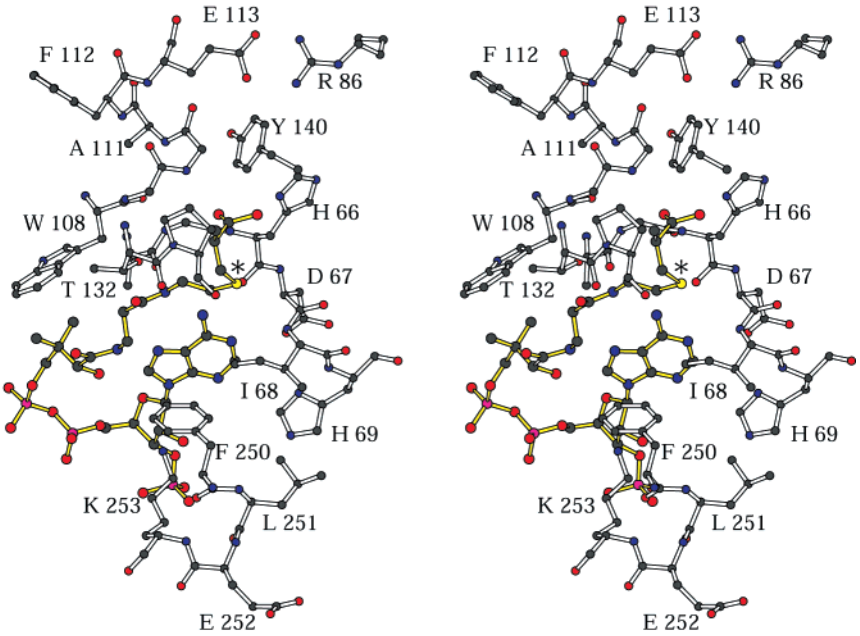


FIGURE 4: Close-up view of the MMCD active site. Those amino acid residues that lie within approximately 4.0 Å of atoms of the thioether analogue are shown. The position of the methylene bridging group characterizing the MMCD inhibitor is indicated by an asterisk.

A close-up view of the binding pocket for the thioether analogue of methylmalonyl CoA analogue is shown in Figure 4. Note that in this substrate analogue, the carbonyl moiety of the substrate has been replaced with a methylene group as indicated by the asterisk. Clearly, while this inhibitor allowed identification of the active site region of MMCD, it must be kept in mind that it may not necessarily mimic true substrate binding as will be described further below. What is absolutely clear, however, is that the carboxylate group of Glu 113 forms a salt bridge with the guanidinium group of Arg 86 and is not in a position to serve as an active site base. This same salt bridge occurs in apo-MMCD as well. In the case of 2-enoyl-CoA hydratase, the homologue to Glu 113, namely, Glu 144, is proposed to activate a water molecule that subsequently adds across the double bond of the substrate. The fact that this glutamate residue is conserved among members of the crotonase superfamily yet plays different biochemical roles in MMCD and 2-enoyl-CoA hydratase emphasizes the fact that amino acid sequence alignments alone are not sufficient to fully understand a superfamily as is being proposed by some proponents of structural genomics. The structure of the MMCD/ligand complex was solved only to 2.7 Å resolution, and as such, definition of the solvent structure within the active site is not possible (Figure 1b). Even from this limited resolution, however, it is clear that the adenine ring of the inhibitor interacts with the protein via the loop defined by Ala 64−Ile 68 and that the 2′-hydroxyl group of the ribose lies within hydrogen bonding distance to N^{ζ} of Lys 253. Both the 3′-phosphate group of the ribose and the pyrophosphate moiety project into the solvent while the terminal carboxylate of the inhibitor is juxtaposed to His 66 and the peptidic NH group of Gly 110. Unlike that observed in 4-chlorobenzoyl CoA dehalogenase, where the active site regions lie between two monomers, the ligand binding pocket of MMCD is contained within one subunit. Approximately 60% of the surface area of the ligand is buried upon complexation with the enzyme.

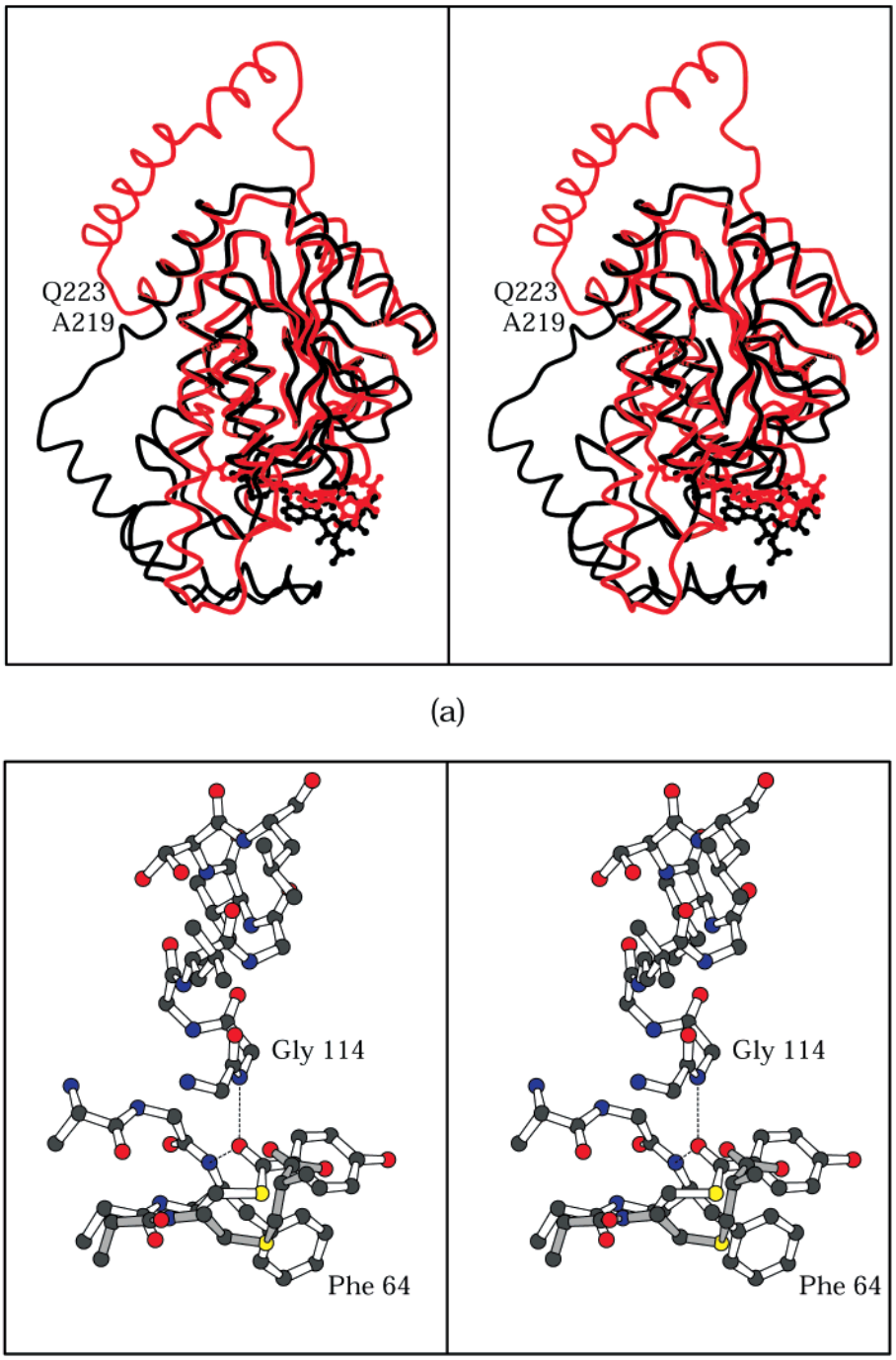


FIGURE 5: Comparison of MMCD with 4-chlorobenzoyl CoA dehalogenase. (a) The polypeptide chains for MMCD and the dehalogenase are displayed in black and red, respectively. A close-up view of the difference in ligand binding to MMCD and the dehalogenase is given in panel b. For the sake of clarity, only the reactive portions of the ligands are shown. The dehalogenase complex is highlighted in open bonds. Hydrogen bonds thought to be important for the catalytic mechanism of the dehalogenase are indicated by the dashed lines. Both the backbone amide groups of Phe 64 and Gly 114 are in the proper position for such electrostatic interactions. As can be seen, the thioether analogue of MMCD (shown in gray bonds) binds in a different orientation. Note the differences in the positions of the sulfurs between the two ligands. This different binding mode for the thioether analogue most likely results from the replacement of the thioester carbonyl of the substrate with a bridging methylene group.

Comparison with 4-Chlorobenzoyl CoA Dehalogenase. The first structures to be solved for members of the crotonase superfamily were that of 4-chlorobenzoyl CoA dehalogenase (*17*) followed immediately by that of 2-enoyl CoA dehydratase (*15*). These two structures were strikingly similar in their overall topological arrangements. It was anticipated, on the basis of primary structural alignments, that MMCD would likewise be quite similar with only the active site architecture differing to any significant extent. Indeed, the three-dimensional structures of MMCD and the dehalogenase are quite similar up to residue Val 215 of MMCD. Specifically, the rms deviation between α-carbons for these two proteins is approximately 1.4 Å for 178 structurally equivalent atoms. After residue Val 215, however, the polypeptide

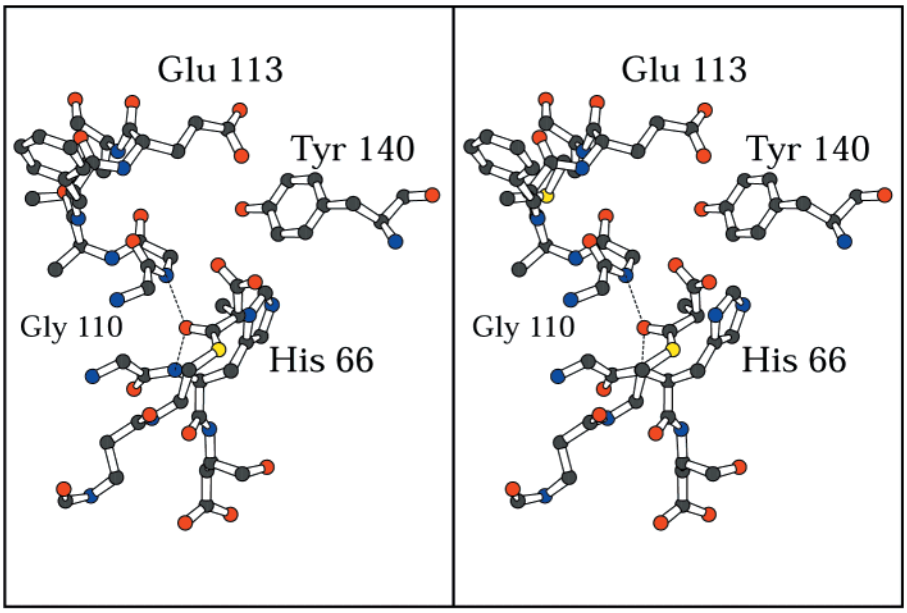


FIGURE 6: Possible binding mode of methylmalonyl CoA within the active site of MMCD. On the basis of the observed binding orientation of 4-hydroxybenzoyl CoA to the dehalogenase, a model of methylmalonyl CoA was positioned into the active site of MMCD as shown here. Possible key interactions between the protein and the substrate include the hydrogen bonds between the backbone amide groups of His 66 and Gly 110 and the ligand thioester carbonyl and a hydrogen bond between O^{η} of Tyr 140 and the leaving carboxylate group.

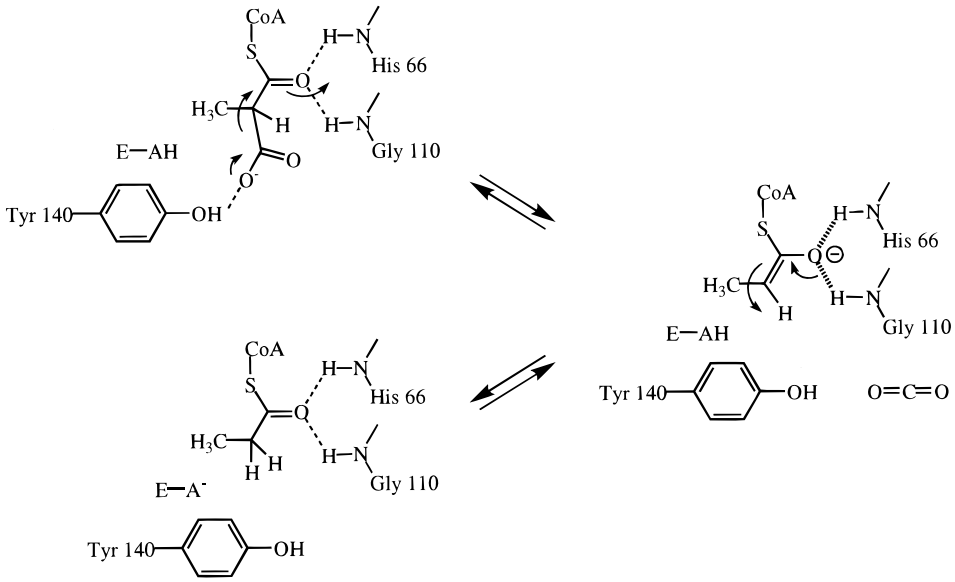


FIGURE 7: Proposed mechanism for the MMCD-catalyzed reaction.

chains diverge significantly as indicated in Figure 5a. In MMCD, Ala 219 adopts dihedral angles of $\phi = -64°$ and $\psi = 40°$, whereas in the dehalogenase the structurally equivalent residue, Gln 223, assumes torsional angles of $\phi = -113°$ and $\psi = 154°$. While the last C-terminal residues in both enzymes adopt α-helical conformations, the dispositions of these helices with respect to the N-terminal domains of the molecules are completely different. Clearly, this type of molecular rearrangement was not obvious from amino acid sequence alignments and yet has profound consequences with respect to the active site pockets.

In 4-chlorobenzoyl CoA dehalogenase, the thioester carbonyl of the 4-hydroxybenzoyl CoA product lies within hydrogen bonding distance to the backbone amide hydrogen atoms of Phe 64 and Gly 114 as indicated in Figure 5b. In addition, Gly 114 lies at the N-terminal end of an α-helix, and it has been speculated that both hydrogen bonding interactions and the helix dipole moment play critical roles for polarizing the thioester carbonyl of the substrate (*17*). This polarization serves both to activate the benzoyl ring C(4) toward nucleophilic attack by Asp 145 and to stabilize the Meisenheimer intermediate. Strikingly, a similar α-helix is found in MMCD with the N-terminal residue being Gly 110. As can be seen in Figure 5b, however, the thioether analogue of methylmalonyl CoA employed in this investigation of MMCD binds in a different orientation within the active site. It can thus be speculated that the substitution of a carbonyl functionality with a bridging methylene group results in this thioether analogue binding in a manner that does not mimic substrate positioning.

Proposed Reaction Mechanism. For decarboxylation of methylmalonyl CoA to occur, two key structural elements

are required: (i) polarization of the thioester carbonyl and/or stabilization of a thioester enolate anion intermediate and (ii) displacement of the leaving carboxylate group from the plane of atoms defined by the thioester linkage. On the basis of the observed binding mode of 4-hydroxybenzoyl CoA to dehalogenase, it is possible to speculate about the manner in which the substrate, methylmalonyl CoA, binds to MMCD. Shown in Figure 6 is a model of methylmalonyl CoA bound within the active site pocket of MMCD. This model was constructed according to the assumption that the orientation of methylmalonyl CoA in MMCD, and in particular the thioester carbonyl group, should be similar to that observed for 4-hydroxybenzoyl CoA in the dehalogenase. As can be seen, the backbone amide groups of His 66 and Gly 110 are in positions to form hydrogen bonds with the carbonyl oxygen of the substrate thereby producing the required polarization of this bond. In addition, O^{η} of Tyr 140 is within hydrogen bonding distance to the carboxylate group of the substrate and most likely plays a key role in orienting this moiety orthogonal to the plane of the thioester carbonyl group. Other than Tyr 140, the active site immediately surrounding the carboxylate group of methylmalonyl CoA is decidedly hydrophobic.

This nonpolar environment would destabilize the negatively charged carboxylate group and enhance the decarboxylation process, leading to formation of a thioester enolate anion intermediate that is stabilized by hydrogen bonding interactions with the backbone amide groups of His 66 and Gly 110 (Figure 7). Note that, in 4-chlorobenzoyl CoA dehalogenase, the structurally equivalent amino acid residues are Phe 64 and Gly 114, respectively. It is not presently clear which residue serves as the general acid to deliver a solvent-derived proton to the α-carbon to generate the propionyl CoA product; as modeled, the imidazole functional group of His 66 appears to be improperly positioned to function as the acidic catalyst. While this mechanism for the MMCD-catalyzed reaction is speculative, it can be tested and expanded by stereochemical studies, construction and mechanistic characterization of site-directed mutants, and crystallographic studies using additional substrate analogues. This work is in progress.

CONCLUSIONS

A new member of the crotonase superfamily, MMCD from *E. coli*, has been identified and its molecular structure solved by X-ray crystallographic analysis. While approximately the first 200 residues of MMCD fold into the canonical "crotonase" topology, significant molecular changes occur near His 220 such that in the case of MMCD the active site is contained wholly within one subunit of the hexameric species. In addition, the "conserved" Glu 113 apparently does not participate in catalysis as an acid/base catalyst but forms a salt bridge with the guanidinium group of Arg 86. These unexpected changes in structure and function could not have been predicted on the basis of amino acid sequence alignments alone. Accordingly, these results emphasize the caution that must be applied to the growing field of structural genomics: knowledge of the three-dimensional architecture for only one, or even a few, members of a superfamily may not be sufficient to fully define either the structures or functions of other members.

ACKNOWLEDGMENT

We thank Dr. W. W. Cleland for critical discussions throughout the course of this investigation.

REFERENCES

1. Stern, J. R. (1955) *Methods Enzymol. 1*, 559−566.
2. Stoffel, W., and Grol, M. (1978) *Hoppe-Seyler's Z. Physiol. Chem. 359*, 1777−1782.
3. Chen, L.-S., Jin, S.-J., and Tserng, K. Y. (1994) *Biochemistry 33*, 10527−10534.
4. Kleber, H.-P. (1997*) FEMS Microbiol. Lett. 147*, 1−9.
5. Gasson, M. J., Kitamura, Y., McLauchlan, W. R., Narbad, A., Parr, A. J., Parsons, E. L. H., Payne, J., Rhodes, M. J. C., and Walton, N. J. (1998) *J. Biol. Chem. 273*, 4163−4170.
6. Sharma, V., Suvarna, K., Meganathan, R., and Hudspeth, M. E. (1992) *J. Bacteriol. 174*, 5057−5062.
7. Igbavboa, U., and Leistner, E. (1990) *Eur. J. Biochem. 192*, 441−449.
8. Pelletier, D. A., and Harwood, C. S. (1998) *J. Bacteriol. 180*, 2330−2336.
9. Crooks, G. P. and Copley, S. D. (1994) *Biochemistry 33*, 11645−11649.
10. Loffler, F., Lingens, F., and Muller, R. (1995) *Biodegradation 6*, 203−212.
11. Chang, K.-H., Liang, P.-H., Beck, W., Scholten, J. D., and Dunaway-Mariano, D. (1992) *Biochemistry 31*, 5605−5610.
12. Shimomura, Y., Murakami, T., Fujitsuka, N., Nakai, N., Sato, Y., Sugiyama, S., Shimomura, N., Irwin, J., Hawes, J. W., and Harris, R. A. (1994*) J. Biol. Chem. 269*, 14248−14253.
13. Hawes, J. W., Jaskiewicz, J., Shimomura, Y., Huang, B., Bunting, J., Harper, E. T., and Harris, R. A. (1996) *J. Biol. Chem. 271*, 26430−26434.
14. Xiang, H., Luo, L., Taylor, K. L., and Dunaway-Mariano, D. (1999) *Biochemistry 38*, 7638−7652.
15. Engel, C. K., Mathieu, M., Zeelen, J. P., Hiltunen, J. K., and Wierenga, R. K. (1996) *EMBO J. 15*, 5135−5145.
16. Engel, C. K., Kiema, T. R., Hiltunen, J. K., and Wierenga, R. K. (1998) *J. Mol. Biol. 275*, 847−859.
17. Benning, M. M., Taylor, K. L., Liu, R.-Q., Yang, G., Xiang, H., Wesenberg, G., Dunaway-Mariano, D., and Holden, H. M. (1996) *Biochemistry 35*, 8103−8109.
18. Modis, Y., Filppula, S. A., Novikov, D. K., Norledge, B., Hiltunen, J. K., and Wierenga, R. K. (1998) *Structure 6*, 957−970.
19. 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) *Biochemistry 35*, 16489−16501.
20. Babbitt, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Kenyon, G. L., and Gerlt, J. A. (1995) *Science 267*, 1159−1161.
21. Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science 277*, 1453−1474.
22. Mancia, F., Smith, G. A., and Evans, P. R. (1999) *Biochemistry 38*, 7999−8005.
23. Haller, T., Buckel, T., Rétey, J, and Gerlt, J. A. (2000) *Biochemistry 39*, 4622−4629.
24. Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 67−71.
25. Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 916−924.
26. Terwilliger, T. C., and Eisenberg, D. (1983) *Acta Crystallogr. A39*, 813−817.
27. Bricogne, G. (1976) *Acta Crystallogr. A32*, 832−837.
28. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. Sect. A 43*, 489−501.
29. Rossmann, M. G. (1972) *The Molecular Replacement Method*, Gordon and Breach, New York.
30. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
31. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.

BI9928896