Structural Basis for Substrate Specificity in Adenosylcobalamin-dependent Isobutyryl-CoA Mutase and Related Acyl-CoA Mutases

Marco Jost, David A. Born, Valentin Cracan, Ruma Banerjee, and Catherine L. Drennan

Background: Acyl-CoA mutases catalyze radical-based carbon skeleton rearrangements.

Results: Crystal structures of isobutyryl-CoA mutase in complex with four different substrates reveal active site architecture and determinants of substrate specificity.

Conclusion: Identification of specificity-determining residues allows for prediction of new acyl-CoA mutase activities.

Significance: Improved understanding of acyl-CoA mutase substrate specificity is critical for biotechnological and engineering applications.

Acyl-CoA mutases are a growing class of adenosylcobalamin-dependent radical enzymes that perform challenging carbon skeleton rearrangements in primary and secondary metabolism. Members of this class of enzymes must precisely control substrate positioning to prevent oxidative interception of radical intermediates during catalysis. Our understanding of substrate specificity and catalysis in acyl-CoA mutases, however, is incomplete. Here, we present crystal structures of IcmF, a natural fusion protein variant of isobutyryl-CoA mutase, in complex with the adenosylcobalamin cofactor and four different acyl-CoA substrates. These structures demonstrate how the active site is designed to accommodate the aliphatic acyl chains of each substrate. The structures suggest that a conformational change of the 5'-deoxyadenosyl group from C2- to C3- could contribute to initiation of catalysis. Furthermore, detailed bioinformatic analyses guided by our structural findings identify critical determinants of acyl-CoA mutase substrate specificity and predict new acyl-CoA mutase-catalyzed reactions. These results expand our understanding of the substrate specificity and the catalytic scope of acyl-CoA mutases and could benefit engineering efforts for biotechnological applications ranging from production of biofuels and commercial products to hydrocarbon remediation.

Adenosylcobalamin (AdoCbl, coenzyme B12) is an organometallic enzyme cofactor for radical chemistry. Its reactivity is based on a unique covalent cobalt-carbon (Co–C) bond that is sufficiently weak to allow for reversible homolytic cleavage in enzyme active sites, generating a 5'-deoxyadenosyl radical in the presence of an appropriate substrate. The radical can then abstract a substrate hydrogen atom and initiate difficult chemical transformations such as the carbon-skeleton rearrangements of acyl-CoA thioesters catalyzed by a major group of AdoCbl-dependent enzymes as part of primary and secondary metabolic pathways. These so-called acyl-CoA mutases allow for the degradation and biosynthesis of branched-chain compounds by altering the level of branching and therefore have biotechnological potential ranging from remediation of hydrocarbon pollution to synthesis of commercial products such as solvents and fragrances. To realize this potential and to engineer new activities, a better understanding of the factors governing substrate specificity in acyl-CoA mutases is required.

The best characterized acyl-CoA mutase is methylmalonyl-CoA mutase (MCM), which is found in species ranging from bacteria to humans. MCM interconverts (R)-methylmalonyl-CoA and succinyl-CoA in the degradation of odd-chain fatty acids, cholesterol, and branched amino acids or to supply methylmalonyl-CoA units for polyketide biosynthesis.
Several related acyl-CoA mutases catalyze the interconversion of other branched-chain CoA thioesters (Fig. 1, (8)). Several related acyl-CoA mutases catalyze the interconversion of other branched-chain CoA thioesters (Fig. 1, (8)). Several related acyl-CoA mutases catalyze the interconversion of 2-(2'-aminophenyl)succinyl-CoA and (2'-aminobenzyl)malonyl-CoA and related compounds (Fig. 1f) in the anaerobic degradation of aromatic compounds such as indoleacetate (18, 19) as well as the interconversion of 2-(1'-methylalkyl)succinyl-CoA and (2'-methylalkyl)malonyl-CoA (Fig. 1g) in the anaerobic degradation of alkanes (20–23). Although these activities have not been biochemically verified, these studies suggest that the catalytic scope of acyl-CoA mutases might be much larger than originally anticipated.

Of these acyl-CoA mutases, only MCM (24–27), IcmF (28), and HCM1 (29) have been structurally characterized, and only MCM and HCM1 have been visualized with substrates bound. All three enzymes require two domains for catalytic activity as follows: a Rossmann-fold cobalamin (Cbl)-binding domain, which binds the AdoCbl cofactor in the “base-off/His-on” mode (24, 27, 30), and an (α/β)₈ triose-phosphate isomerase (TIM) barrel, which binds the substrate. These two domains can be encoded on a single polypeptide (MCM, IcmF) or on separate polypeptides (HCM1). The Cbl-binding domain positions the AdoCbl into the TIM barrel, forming a buried active site cavity in which the free radical intermediates of catalysis are protected from oxidative quenching. Intriguingly, the TIM barrels of MCM and IcmF can undergo a dramatic conformational change, from a catalytically active closed state that resembles a typical TIM barrel to an unusual open state, in which the TIM barrel is split into two halves of four β-strands each with a cavity in the center of the barrel (25). For the substrate-free structure of homodimeric IcmF, one chain has AdoCbl positioned into a closed barrel, and the other has an open barrel with the Cbl cofactor displaced out of the active site (28). In contrast, for MCM, the open barrel is associated with the substrate-free form of the enzyme (25), whereas the substrate-bound structure is in the closed configuration with the substrate threaded through the barrel (24–26). It has been proposed that this conformational change may afford at least part of the substantial 10¹² rate acceleration of Co–C bond homolysis that occurs upon substrate binding in MCM (25, 31, 32).

The large substrate-induced conformational rearrangement in MCM lies in stark contrast to the more subtle conformational changes that occur upon substrate binding in a different AdoCbl-dependent enzyme, glutamate mutase. Here, the ribose of the AdoCbl 5'-deoxyadenosyl group (5'-dAdo) undergoes pseudorotation from the C2'-endo conformation to the C3'-endo conformation when substrate binds, breaking the Co–C bond and repositioning the resulting 5'-deoxyadenosyl radical for hydrogen atom abstraction from substrate (33). No structural data exist as to whether the same change in ribose conformation is involved in reactivity of MCM or other acyl-CoA mutases. Given the limited number of mutases that have been structurally characterized, we still have much to learn.
about how these enzymes generate and control the highly reactive 5′-deoxyadenosyl radical upon substrate binding.

Furthermore, to engineer mutases for biotechnological applications, an understanding of the substrate binding determinants will also be required. Substrate-bound structures of MCM (25, 26) and HCM1 (29) as well as bioinformatic analyses (4, 6, 11, 13) have suggested that substrate specificity in acyl-CoA mutases is determined by the identity of a few key amino acids. Indeed, HCM1 carrying a single active site mutation has both considerable PCM and HCM2 activity and reduced HCM1 activity (29). All other attempts to rationally alter the substrate specificity of acyl-CoA mutases by mutagenesis, however, have failed (13, 34), indicating that our understanding of substrate specificity in acyl-CoA mutases is incomplete.

We recently reported crystal structures of IcmF from Cupriavidus metallidurans, which contains a G-protein domain in addition to the mutase domains, with AdoCbl in the ICM active site and GDP-Mg<sup>2+</sup> in the G-protein active site (holo-IcmF-GDP) but without acyl-CoA substrates (28). Here, we report crystal structures of IcmF bound to AdoCbl, GDP-Mg<sup>2+</sup>, and all four known acyl-CoA substrates (pivalyl-CoA, isovaleryl-CoA, isobutyryl-CoA, and n-butyryl-CoA), revealing the mode of substrate binding and the determinants of substrate specificity in IcmF. Guided by the structural insight and bioinformatic analyses, we identify two classes of acyl-CoA mutases that likely catalyze novel AdoCbl-dependent reactions.

**Experimental Procedures**

**Materials**—Isobutyryl-CoA, n-butyryl-CoA, and isovaleryl-CoA were obtained from Sigma. Pivalyl-CoA was synthesized in a one-step procedure from pivalic anhydride (Sigma) and CoA (Sigma) (Fig. 2). Briefly, to a solution of 0.20 mmol of pivalic anhydride (40.6 µl) in 2 ml of anhydrous dimethylformamide were added 0.040 mmol of solid CoA hydrate (32 mg), 0.12 mmol of triethylamine (16.8 µl), and a catalytic amount of dimethylaminopyridine. The reaction was stirred for 30 min at 25 °C. Reaction progress was followed by thin layer chromatography in 1:1:1.1 butanol/acetatic acid/ethyl acetate/water. The reaction was stopped by addition of 0.10 mmol of HCl followed by dilution with water, and the water/dimethylformamide mixture was removed by lyophilization. The solid product was dissolved in 400 µl of 95:5 water/acetonitrile and purified by HPLC on a 250 × 10-mm Targa C<sub>18</sub> (5-µm pore size) reversed-phase column (Higgins Analytical). The final yield of pivalyl-CoA was 0.018 mmol (45%). ESI-MS (m/z): [M−2H]<sup>2+</sup>− calculated for C<sub>26</sub>H<sub>44</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>S, 424.58; found, 424.58.

**Protein Expression, Purification, and Crystallization**—N-terminally His-tagged IcmF from *C. metallidurans* was expressed and purified as described previously (15, 16, 28). Purified IcmF was supplemented with AdoCbl (Sigma), GDP (Sigma), and MgCl<sub>2</sub> to generate holo-IcmF-GDP and crystallized at 25 °C using the hanging drop vapor diffusion technique. 1 µl of a protein solution (11.7 mg/ml IcmF in 100 mM NaCl, 50 mM HEPES, pH 7.5, 1 mM GDP, 3 mM MgCl<sub>2</sub>, 300 µM AdoCbl) was mixed with 1 µl of a precipitant solution (0.7–0.75 M potassium sodium tartrate, 0.2 M ammonium acetate, 0.1 M imidazole, pH 7.0–7.7, 3% (v/v) ethylene glycol) on a glass coverslip. The cover slip was sealed with grease over a reservoir containing 500 µl of the precipitant solution without ethylene glycol. Triangular crystals appeared within 3 weeks and grew to full size within 6 weeks. To generate crystals of holo-IcmF-GDP bound to n-butyryl-CoA, isobutyryl-CoA, isovaleryl-CoA, or pivalyl-CoA, pre-formed holo-IcmF-GDP crystals were transferred to 2 µl of a soak solution containing the precipitant, 2 mM GDP, 3 mM MgCl<sub>2</sub>, 5 mM substrate, and 20% (v/v) glycerol, incubated in that solution for 15 s, and then flash-frozen in liquid nitrogen. All crystallization and soaking procedures were carried out in a dark room under red light to prevent cleavage of the AdoCbl-Co−C bond before ligand binding.

**Data Collection and Processing**—All IcmF crystals belong to space group R32 (denoted as H32 by the PDB). All data were collected at the Advanced Photon Source (Argonne, IL) at beamline 24ID-C at a temperature of 100 K and a wavelength of 0.9795 Å (12,658 eV). Data for holo-IcmF-GDP bound to isobutyryl-CoA, n-butyryl-CoA, and isovaleryl-CoA were collected using a Quantum 315 detector in 0.5° (isobutyryl-CoA) or 1° oscillation steps (n-butyryl-CoA and isovaleryl-CoA). Data for holo-IcmF-GDP bound to pivalyl-CoA were collected using a Pilatus 6M detector in wedges of 20° in 0.5° oscillation steps. The crystal was displaced along its major macroscopic axis after each wedge.

All data were integrated in XDS and scaled in XSCALE (35). The same reflections as in the previously determined holo-IcmF-GDP data set (28) were marked for the free set of reflections in all data sets, corresponding to 5% of total reflections. All data collection statistics are summarized in Table 1.

**Structure Building and Refinement**—All structures were determined to resolutions ranging from 3.40 to 3.50 Å resolution (Table 1) by molecular replacement. First, the structure of holo-IcmF-GDP bound to isobutyryl-CoA was determined by molecular replacement with the structure of substrate-free holo-IcmF-GDP (PDB code 4XC6) (28) using rigid body refinement in PHENIX (36). To minimize existing model bias, 10 cycles of simulated annealing refinement were carried out in PHENIX. There was clear electron density for isobutyryl-CoA in one of the two protomers in the asymmetric unit. After insertion of the substrate, the model was adjusted to account for any changes in the protein environment by iterative cycles of manual model building in COOT (37, 38) and refinement in PHENIX. The structures of holo-IcmF-GDP bound to n-butyryl-CoA, isovaleryl-CoA, and pivalyl-CoA were determined by molecular replacement with the structure of isobutyryl-CoA-bound holo-IcmF-GDP using rigid body refinement in PHENIX. For each structure, there was clear electron density for the corresponding substrate in one of the two protomers in the asymmetric unit. We only modeled the given substrate in our structures instead of a substrate/product mixture because the crystal/substrate incubation times were short relative to IcmF’s turnover rate (15, 16). The models were adjusted to account for any changes in the protein environment by iterative
cycles of manual model building in COOT and refinement in PHENIX. Side chains with limited electron density were truncated at the last atom with visible electron density. Initial stages of refinement included B-factor refinement for individual atoms. Final stages of refinement included TLS parameterization using one TLS group per chain (39). Strict noncrystallographic symmetry restraints were applied in early cycles of refinement. In advanced stages of refinement, noncrystallographic symmetry restraints were loosened for residues involved in crystal contacts as well as selected residues that were in substantially different environments due to conformational differences between the two chains of IcmF in the asymmetric unit.

Parameter files for Cbl were generously provided by Oliver Smart at Global Phasing (Cambridge, UK). Refinement restraints for 5′-dAdo, isobutyryl-CoA, n-butyryl-CoA, isovaleryl-CoA, and pivalyl-CoA were generated using the Grade Web Server (40). Refinement restraints for GDP were generated using the electronic Ligand Builder and Optimization Workbench (eBOW) (41) implemented in PHENIX.

Crystallographic refinement of the four structures of holo-IcmF-GDP bound to substrates yielded models that possess low free R-factors, excellent stereochemistry, and small root mean square deviations (r.m.s.d.) from ideal values for bond lengths and angles. The final models of holo-IcmF-GDP bound to isobutyryl-CoA, n-butyryl-CoA, isovaleryl-CoA, or pivalyl-CoA include residues 21–1093 (of 1093) for chain A and residues 22–1093 (of 1093) for chain B, lacking the hexahistidine tag and residues at the N terminus. The model of holo-IcmF-GDP bound to isobutyryl-CoA additionally lacks residues 285, 530–537, and 1013–1014 in chain A and residues 592, 904–906, and 1011–1018 in chain B. The model of holo-IcmF-GDP bound to n-butyryl-CoA lacks residues 285, 530–536, and 1014 in chain A and residues 592–593, 905–906, and 1011–1018 in chain B. The model of holo-IcmF-GDP bound to isovaleryl-CoA lacks residues 284–285, 530–536, and 1013–1014 in chain A and residues 592–593, 904–906, and 1011–1018 in chain B. The model of holo-IcmF-GDP bound to pivalyl-CoA lacks residues 530–536 and 1012–1014 in chain A and residues 592–593 and 1011–1018 in chain B. For all models, each chain contains bound cobalamin and GDP•Mg²⁺ in the GDP-binding site; chain A contains bound substrate and 5′-dAdo, and chain B contains the nucleotide portion of the substrate. All refinement statistics are summarized in Table 1. The models were validated using simulated annealing composite omit maps calculated in CNS (42, 43). Model geometry was analyzed using MolProbity (44) and ProCheck (45). Figures were generated using PyMOL (46). Crystallography software packages were compiled by SBGrid (47).

Phylogenetic and Bioinformatic Analyses—For calculation of a phylogenetic tree, sequences of different acyl-CoA mutase substrate-binding domains were retrieved from the genomic BLAST interface (February 5, 2015) at the National Center for Biotechnology Information (NCBI) or from the Integrated Microbial Genomes and Metagenomes database of the Joint Genome Institute of the United States Department of Energy (img.jgi.doe.gov). A representative set of 200 sequences was manually selected for alignment. MCM sequences were chosen to cover eukaryotes, archaea, and all major bacterial phyla that contain MCM and had >5 genome sequences reported (Proteobacteria (including α-, β-, γ-, and δ-proteobacteria), Actinobacteria, Bacteroidetes, Chlorobi, Chloroflexi, Cyanobacteria, Deferribacteres, Deinococci, Firmicutes, Planctomycetes, Spirochaeta, Synergistia, and Verrucomicrobia). Additional sequences were chosen to cover characterized HCMs, ECMs, ICMs, IcmFs, and PCMs as well as putative uncharacterized mutases. A phylogenetic tree was then calculated using the following workflow implemented in the Phylogeny.fr web server (48). Sequences were aligned using MUSCLE (49), and the alignment was trimmed to homologous regions using Gblocks (50). A phylogenetic tree was constructed using the bootstrap method (100 bootstraps) and the LG substitution model (four substitution rate categories; γ-distribution parameter and proportion of invariant sites were estimated by the program) in PhyML (51) and visualized using TreeDyn (52). For validation, a complete phylogenetic tree was calculated using all sequences annotated as mutase substrate-binding domains (InterPro group IPR006098, accessed May 31, 2015, grouped together by >85% sequence identity) and using the same workflow. The two phylogenetic trees exhibited the same overall structure, confirming the observed grouping of sequences. The tree calculated from manually selected sequences was used for visualization.

Homology Modeling—Homology models of the uncharacterized mutase from Aromatoleum aromaticum (NCBI accession code WP_011236985.1) and of uncharacterized mutase 1 from Desulfatibacillum alkenivorans (NCBI accession code WP_012610856.1) were generated using the SWISS-MODEL Automated Comparative Protein Modeling Server (53) with the structure of substrate-bound MCM from Propionibacterium freudenreichii subsp. shermanii (PDB code 4REQ (26)) as template. The uncharacterized mutases from A. aromaticum and D. alkenivorans share 35 and 28% sequence identity with this template.

Results

Acyl-CoA Substrates Bind to the Catalytically Active Chain of IcmF—To visualize how IcmF binds its substrates, we sought to determine crystal structures of IcmF bound to its AdoCbl cofactor and the four different acyl-CoA molecules that all serve as substrates (15, 16). Whereas isobutyryl-CoA, n-butyryl-CoA, and isovaleryl-CoA are commercially available, the fourth substrate, pivalyl-CoA (2,2-dimethylpropionyl-CoA), was synthesized from pivalic anhydride via a one-step synthetic procedure (Fig. 2, see “Experimental Procedures”). Pre-formed holo-IcmF-GDP crystals were incubated with the different substrates, and structures were determined to resolutions ranging from 3.40 to 3.50 Å (Fig. 3 and Table 1). The resulting structures depict IcmF in complex with Cbl, the 5′-dAdo, and substrate in the mutase active site as well as GDP•Mg²⁺ in the G-protein.
Substrate-bound Structures of Isobutyryl-CoA Mutase

The substrate binding mode, we focused on chain A in our structures. To further evaluate the substrate binding mode, we focused on chain A in our structures. Acyl-CoA Substrates Are Threaded through the TIM Barrel Substrate-biding Domain—The four substrate-bound IcmF structures are nearly identical, with Cα r.m.s.d. values smaller than 0.2 Å between all structures. The acyl-CoA substrates are bound in the same overall fashion: the nucleotide portion of the substrate is positioned on the surface of the N-terminal face of the TIM barrel, the phosphopantetheine moiety is threaded through the center of the IcmF TIM barrel substrate-binding domain, and the acyl group is positioned adjacent to the Cbl cofactor and the 5′-deoxyadenosyl group in the active site cavity (Figs. 4, a–d and 5a). This mode of substrate binding is very similar to that observed in the related acyl-CoA mutases MCM (25, 26) and HCM1 (Fig. 4, e and f) (29), and many of the specific interactions are conserved. In IcmF, the thioester carbonyl is stabilized by hydrogen bonds from Gln-732 and His-780, securing the acyl group in the active site (Fig. 4, a–d). The interaction with His-780 in particular is likely important for catalysis, as mutation of the homologous His (His-244) to Gln or Ala in MCM reduced kcat by 102 to 103 and drastically increased the rate of oxidative inactivation (54). The phosphopantetheine arm and the nucleotide portion of the substrate are stabilized by additional interactions such as electrostatic interactions between the phosphophate groups and the positively charged residues Arg-589, Arg-622, Arg-728, Arg-856, and Lys-861 (Figs. 4, a–d, and 5a). Furthermore, Tyr-772 hydrogen bonds with a phosphate oxygen, Phe-585 engages in π–π stacking interactions with the adenine base, and Ser-821 forms a hydrogen bond to the phosphopantetheine hydroxyl group (Figs. 4, a–d, and 5a). Together, the interactions to the phosphopantetheine and the nucleotide plug the access tunnel to the active site, thereby protecting the reactive catalytic intermediates.

IcmF Substrate-bound Structures Suggest Two Conformations of the 5′-Deoxyadenosyl Group—To gain insight into the effect of substrate binding on the AdoCbl cofactor, the electron density for AdoCbl in the substrate-bound IcmF structures was analyzed. The electron density is best fit by a mixture of species: cleaved AdoCbl with 5′-dAdo in the C3′-endo conformation and a Co–C distance of 3.2–3.5 Å and uncleaved AdoCbl with 5′-dAdo in the C2′-endo conformation and a Co–C distance of 2.2 Å (Fig. 6a). Although the resolutions of our structures are moderate, this assignment is supported by the previous observation of 5′-dAdo in both the C3′-endo and C2′-endo conformations in the substrate-bound structure of the related AdoCbl-dependent enzyme glutamate mutase (Fig. 6, b–d) (33). In the C3′-endo conformation, the 5′-dAdo C5′ atom is close to the hydrogen atom abstraction site on the substrate and could initiate catalysis (Fig. 6c, see below). Notably, the 5′-dAdo is stabilized by different interactions in the C2′-endo and C3′-endo conformations. The C2′-endo conformation is stabilized by hydrogen bonds from Tyr-779 to the ribose O2′ and from Glu-905 to the ribose O3′ (Fig. 6d). In the C3′-endo conformation, the interaction with Tyr-779 is disrupted; instead, now the ribose O3′ also forms a hydrogen bond to Glu-905, and the ribose O4′ forms a hydrogen bond to Gln-865 (Fig. 6c). Similar changes in the hydrogen bonding patterns of the two 5′-dAdo conformers were observed for glutamate mutase (Fig. 6, c and d) (33). Glu-905 is also conserved in glutamate mutase as Glu-330, and mutation of this Glu-330 to Asp, Gln, or Ala leads to a drastic reduction in activity, supporting its importance for catalysis (55).
acetyl-CoA substrates are bound in the active site in a similar orientation (Fig. 7, a and b), with a β-carbon of the acyl group positioned within 3.6 Å of the 5′-dAdo C5′ (in the C3′-endo conformation, see above), which is in agreement with other distances reported for hydrogen atom transfer (56). The binding site for the acyl groups is lined by Gln-732 and His-780 (see above), which is in agreement with other stereochemical investigations on ICM (57, 58). Our observed modes of substrate binding also provide an explanation for the observed partial breakdown of stereospecificity in ICM, for which a small amount of hydrogen atom abstraction occurs from the pro-S methyl group, which is located within 3.5 Å of the 5′-dAdo C5′, whereas the pro-R methyl group is farther away at a distance of 3.9 Å (Fig. 8, left box). For n-butyryl-CoA, modeling of hydrogen atoms with ideal geometry positions the pro-S hydrogen on C3 toward the 5′-dAdo C5′, an ideal position for hydrogen atom abstraction (Fig. 8, right box). Both of these observations match previous stereochemical investigations on ICM (57, 58). Our observed modes of substrate binding also provide an explanation for the observed partial breakdown of stereospecificity in ICM, for which a small amount of hydrogen atom abstraction occurs from the pro-R methyl group of isovaleryl-CoA (57).

Finally, comparing the positioning of the isobutyryl-CoA and n-butyryl-CoA substrates allows us to probe the stereospecificity of the ICM reaction, which has been studied in stand-alone ICM (as opposed to the IcmF fusion protein) from Streptomyces cinnamoneus. IcmF-bound isobutyryl-CoA is positioned for hydrogen atom abstraction from the pro-S methyl group, which is located within 3.5 Å of the 5′-dAdo C5′, whereas the pro-R methyl group is farther away at a distance of 3.9 Å (Fig. 8, left box). For n-butyryl-CoA, modeling of hydrogen atoms with ideal geometry positions the pro-S hydrogen on C3 toward the 5′-dAdo C5′, an ideal position for hydrogen atom abstraction (Fig. 8, right box). Both of these observations match previous stereochemical investigations on ICM (57, 58). Our observed modes of substrate binding also provide an explanation for the observed partial breakdown of stereospecificity in ICM, for which a small amount of hydrogen atom abstraction occurs from the pro-R methyl group of isobutyryl-CoA (57).

The isobutyryl group likely has rotational flexibility in the active site, with the preferred mode of binding as observed in our structure and an alternative mode of binding with the pro-S methyl group pointing toward Phe-598, as observed for pivalyl-CoA (Fig. 7a). The two possible modes of binding would lead to the observed breakdown of stereospecificity, as originally hypothesized (57).

### TABLE 1

**Crystallographic data collection and refinement statistics**

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Holo-IcmF-GDP with isobutyryl-CoA</th>
<th>Holo-IcmF-GDP with n-butyryl-CoA</th>
<th>Holo-IcmF-GDP with isovaleryl-CoA</th>
<th>Holo-IcmF-GDP with pivalyl-CoA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>R32h</td>
<td>R32h</td>
<td>R32h</td>
<td>R32h</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>318.0, 318.0, 343.4</td>
<td>316.8, 316.8, 342.7</td>
<td>317.6, 317.6, 343.5</td>
<td>317.5, 317.5, 343.0</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 120</td>
<td>90, 90, 120</td>
<td>90, 90, 120</td>
<td>90, 90, 120</td>
</tr>
<tr>
<td>Resolution (Å)**</td>
<td>35.0-3.40</td>
<td>35.0-3.50</td>
<td>35.0-3.45</td>
<td>35.0-3.40</td>
</tr>
<tr>
<td>No. of reflections**</td>
<td>87,533 (6578)</td>
<td>81,568 (6052)</td>
<td>86,230 (6411)</td>
<td>90,203 (6660)</td>
</tr>
<tr>
<td>Rcryst (%)**</td>
<td>12.6 (67.1)</td>
<td>14.8 (95.0)</td>
<td>13.7 (101.5)</td>
<td>13.7 (101.5)</td>
</tr>
<tr>
<td>Rfree (%)**</td>
<td>14.3 (76.2)</td>
<td>16.6 (106.0)</td>
<td>15.7 (112.7)</td>
<td>14.5 (112.7)</td>
</tr>
<tr>
<td>CC1/2 (%)</td>
<td>99.3 (66.0)</td>
<td>99.4 (67.5)</td>
<td>99.2 (65.2)</td>
<td>99.8 (69.4)</td>
</tr>
<tr>
<td>Ramachandran statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>0.59</td>
<td>0.59</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>Rotamer outliers (%)</td>
<td>8 (0.5%)</td>
<td>9 (0.6%)</td>
<td>7 (0.4%)</td>
<td>8 (0.5%)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)**</td>
<td>35.0-3.40</td>
<td>35.0-3.50</td>
<td>35.0-3.45</td>
<td>35.0-3.40</td>
</tr>
<tr>
<td>Rcryst/Rfree (%)</td>
<td>0.189/0.209</td>
<td>0.185/0.209</td>
<td>0.192/0.214</td>
<td>0.183/0.201</td>
</tr>
<tr>
<td>No. of atoms</td>
<td>16,176</td>
<td>16,169</td>
<td>16,160</td>
<td>16,143</td>
</tr>
<tr>
<td>Protein</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>AdoCbl</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>GDP</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>Magnesium</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Substrate</td>
<td>80</td>
<td>80</td>
<td>81</td>
<td>81</td>
</tr>
<tr>
<td>Average B-factors (Å²)</td>
<td>89.9</td>
<td>98.7</td>
<td>89.5</td>
<td>108.7</td>
</tr>
<tr>
<td>Protein</td>
<td>86.0</td>
<td>98.5</td>
<td>90.6</td>
<td>108.5</td>
</tr>
<tr>
<td>Cbl</td>
<td>102.2</td>
<td>114.3</td>
<td>113.0</td>
<td>174.6</td>
</tr>
<tr>
<td>GDP</td>
<td>78.6</td>
<td>87.6</td>
<td>78.0</td>
<td>98.4</td>
</tr>
<tr>
<td>Magnesium</td>
<td>58.1</td>
<td>67.2</td>
<td>56.9</td>
<td>81.7</td>
</tr>
<tr>
<td>Substrate</td>
<td>120.1</td>
<td>120.2</td>
<td>128.2</td>
<td>133.7</td>
</tr>
<tr>
<td>r.m.s.d.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>0.59</td>
<td>0.59</td>
<td>0.60</td>
<td>0.59</td>
</tr>
<tr>
<td>Rotamer outliers (%)</td>
<td>8 (0.5%)</td>
<td>9 (0.6%)</td>
<td>7 (0.4%)</td>
<td>8 (0.5%)</td>
</tr>
<tr>
<td>Ramachandran statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favored</td>
<td>2031 (96.4%)</td>
<td>2028 (96.2%)</td>
<td>2023 (96.1%)</td>
<td>2025 (96.0%)</td>
</tr>
<tr>
<td>Allowed</td>
<td>72 (3.5%)</td>
<td>76 (3.6%)</td>
<td>78 (3.7%)</td>
<td>83 (3.9%)</td>
</tr>
<tr>
<td>Disallowed</td>
<td>3 (0.1%)</td>
<td>4 (0.2%)</td>
<td>5 (0.2%)</td>
<td>1 (0.1%)</td>
</tr>
<tr>
<td>Average estimated coordinate error (Å)</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
</tbody>
</table>

*Values in parentheses indicate highest resolution bin.
Specific Active Site Substitutions Allow for Binding of Aliphatic CoA Thioesters—Comparing IcmF to MCM and HCM1 reveals that the active site architectures and the substrate positions of these acyl-CoA mutases are nearly identical (Fig. 7, b–d). The identity of a few amino acid side chains, however, is distinct to account for the different substrates. In particular,
Phe-598 and Gln-742 in IcmF are replaced by a tyrosine (Tyr-89) and an arginine (Arg-207) in MCM (numbering as in MCM crystal structures from *P. freudenreichii* subsp. *shermanii*), which form specific contacts to the carboxylate groups of the MCM substrates (Fig. 7, b and c) (15, 25). In IcmF, the smaller Gln and Phe side chains increase the size and hydrophobicity of the active site, thereby allowing for accommodation of the hydrophobic substrates (Fig. 7, a and b). In HCM1, the glutamine is conserved (Gln-208), accommodating the substrate methyl group, but IcmF Phe-598 is replaced by an isoleucine (Ile-90). This replacement creates space for the substrate hydroxyl groups and an additional hydrogen-bonding asparagine (Asp-117, Fig. 7b). Superposition of IcmF and MCM also shows that the 5’-dAdo group is slightly shifted (Fig. 7b), but the significance of this shift, if any, is unclear.

Overall, the high structural similarity between MCM, HCM1, and IcmF as well as the sequence similarity (≥20–30% identity) between substrate-binding domains of acyl-CoA mutases suggest that other acyl-CoA mutases have similar structures and active site architectures. Within this architecture, the substrate binding specificity of these acyl-CoA mutases is likely governed by the identity of a few residues in the substrate-binding domain, as suggested by a number of studies.

For example, a Y89F/R207Q MCM double mutant was shown to bind isobutyryl-CoA and n-butyryl-CoA, whereas wild-type MCM does not appear to bind these thioesters (34). We note that in this case the mutant enzyme undergoes suicide inactivation during turnover, likely because second-sphere interactions are important to control the radical intermediates, but mutase binding specificity was clearly altered. Recently, it was also demonstrated that a single point mutation in HCM1, D117V, confers pivalyl-CoA mutase activity (29), further suggesting that these mutants can be tailored toward different substrates.

**Bioinformatic Analyses Reveal New Classes of Acyl-CoA Mutases**—To classify acyl-CoA mutases and predict their substrate specificities, we performed phylogenetic and bioinformatic analyses on the substrate-binding domains. The currently available sequences with homology to substrate-binding domains of characterized MCMs, ECMs, ICMs, IcmFs, or HCMs cluster into distinct groups in a phylogenetic tree according to the reaction catalyzed (Figs. 9 and 10 and supplemental material). Each cluster with characterized members contains a series of signature sequences indicating the substrate specificity (Figs. 9 and 11). Two major determinants are at the positions of MCM Tyr-89/IcmF Phe-598.

**FIGURE 6. 5’-dAdo conformational changes in IcmF and glutamate mutase.** a, 2Fo – Fc, omit electron density (orange mesh) contoured at 1.0 σ around Cbl and 5’-dAdo of n-butyryl-CoA bound IcmF. 5’-dAdo can be modeled in the C3’-endo conformation (cyan carbons) and in the C2’-endo conformation (light blue carbons). In the C2’-endo conformation, the C5’ is close to the Cbl cobalt, whereas in the C3’-endo conformation, the C5’ is pointed toward the substrate (orange carbons, dashed red line). Cbl is shown with carbons in pink and cobalt in purple. b, glutamate mutase active site (PDB code 119C) (33), revealing the presence of two 5’-dAdo conformers, C2’-endo (pink carbons) and C3’-endo (purple carbons), in the presence of glutamate (gray carbons). As in IcmF, the 5’-dAdo C5’ is close to the Cbl cobalt in the C2’-endo conformation (dashed red line) and pointed toward the location of hydrogen abstraction on the substrate in the C3’-endo conformation (dashed red line). Cbl is shown with carbons in light pink and Co in purple. c, comparison of the 5’-dAdo C3’-endo conformations in IcmF (cyan carbons) and glutamate mutase (purple carbons). Dashed red lines connect the 5’-dAdo C5’ and the corresponding substrate. Cbl is shown as in b. In both proteins, 5’-dAdo is stabilized by interactions (dashed black lines) to amino acid side chains (IcmF in green and glutamate mutase in pink). IcmF Gln-865 contributes to 5’-dAdo binding, but the corresponding Arg-66 in glutamate mutase does not. d, comparison of the 5’-dAdo C2’-endo conformations in IcmF (light blue carbons) and glutamate mutase (pink carbons). Protein side chains and Cbl colored as in c. Again, 5’-dAdo is stabilized by specific interactions (dashed black lines) to amino acid side chains. IcmF Tyr-779 contributes to 5’-dAdo binding, but the corresponding Pro-218 in glutamate mutase (hidden for clarity) does not. IcmF Asn-901 corresponds to glutamate mutase Lys-326 but does not contribute to 5’-dAdo binding.
FIGURE 7. IcmF active site and comparison with MCM and HCM1 in wall-eyed stereo view. a, overlay of four substrate-bound IcmF structures, revealing similar substrate orientation. Shown are isobutyryl-CoA (yellow carbons), n-butyryl-CoA (orange carbons), pivalyl-CoA (maroon carbons), and isovaleryl-CoA (light pink carbons). The locations of hydrogen abstraction are shown as spheres, located within 3.6 Å of the 5′-dAdo (C3′-endo conformation, cyan carbons) C5′ atom (cyan sphere), as indicated by the dashed red lines. Residues in the substrate-binding site are shown with dark green carbons. The third methyl group of pivalyl-CoA clashes with Phe-598 (yellow dashed line), leading to a small rotation of the side chain (maroon carbons) in this structure. Hydrogen bonds from Gln-732 and His-780 to the thioester carbonyl are indicated as dashed black lines. Cbl is shown with carbons in pink and cobalt in purple. b, overlay of substrate-bound IcmF and MCM (26) (PDB code 4REQ). Isobutyryl-CoA, n-butyryl-CoA, IcmF-bound 5′-dAdo, IcmF-bound Cbl, and IcmF residues are shown as in a. The locations of hydrogen atom abstraction in both methylmalonyl-CoA (gray carbons) and succinyl-CoA (purple carbons) overlay with those of IcmF substrates (spheres). MCM-bound 5′-dAdo, MCM-bound Cbl, and MCM substrate-binding residues are shown with gray carbons. Gln-197 and His-244 are conserved in MCM and IcmF, whereas IcmF Phe-598 is replaced by MCM Tyr-89 and IcmF Gln-742 is replaced by Arg-209, putatively accounting for the switch in substrate binding specificity. Hydrogen bonds and ionic interactions are shown as dashed black lines. c, active site of substrate-bound MCM, shown as in b, highlighting interactions to substrate carboxylate groups. Only interactions to methylmalonyl-CoA are shown for clarity. d, active site of substrate-bound HCM1 (PDB code 4R3U) (29), highlighting interactions to substrate hydroxyl groups. Hydroxyisobutyryl-CoA is shown in lilac and (S)-3-hydroxybutyryl-CoA in blue. Only interactions to hydroxyisobutyryl-CoA are shown for clarity. Locations of hydrogen atom abstraction are shown as spheres.
The chemical mechanism shown at the bottom was established based on stereochemical studies (57). Following Co–C bond homolysis (step not shown), the 5′-dAdo radical abstracts a hydrogen atom (red) from the pro-S methyl group of isobutyryl-CoA (blue). The isobutyryl-CoA radical rearranges to the n-butryl-CoA radical, which then re-abstracts the hydrogen atom from 5′-deoxyadenosine. The hydrogen atom ends up in the pro-S position. In the reverse reaction, the 5′-dAdo radical abstracts the pro-S hydrogen from n-butryl-CoA. The structures of IcmF bound to isobutyryl-CoA (left) and n-butryl-CoA (right) support the proposed stereochemistry. Isobutyryl-CoA (yellow carbons) positions its pro-S methyl group next to the 5′-dAdo group (cyan carbons), whereas n-butryl-CoA positions its pro-S hydrogen (white sticks) toward the 5′-dAdo group. The red dashed line connects the 5′-dAdo C5′ to the closest hydrogen atom. Hydrogens are modeled based on ideal geometry. Cobalamin is shown with pink carbons and cobalt as a purple sphere.

FIGURE 8. Stereochemical course of isobutyryl-CoA mutase reaction. The chemical mechanism shown at the bottom was established based on stereochemical studies (57). Following Co–C bond homolysis (step not shown), the 5′-dAdo radical abstracts a hydrogen atom (red) from the pro-S methyl group of isobutyryl-CoA (blue). The isobutyryl-CoA radical rearranges to the n-butryl-CoA radical, which then re-abstracts the hydrogen atom from 5′-deoxyadenosine. The hydrogen atom ends up in the pro-S position. In the reverse reaction, the 5′-dAdo radical abstracts the pro-S hydrogen from n-butryl-CoA. The structures of IcmF bound to isobutyryl-CoA (left) and n-butryl-CoA (right) support the proposed stereochemistry. Isobutyryl-CoA (yellow carbons) positions its pro-S methyl group next to the 5′-dAdo group (cyan carbons), whereas n-butryl-CoA positions its pro-S hydrogen (white sticks) toward the 5′-dAdo group. The red dashed line connects the 5′-dAdo C5′ to the closest hydrogen atom. Hydrogens are modeled based on ideal geometry. Cobalamin is shown with pink carbons and cobalt as a purple sphere.

(determinant 1) and MCM Arg-207/IcmF Gln-742 (determinant 2); MCMs and ECMs have Tyr and Arg at these positions, allowing them to bind charged carboxylate-bearing substrates (see above); ICMs and IcmFs have Phe and Gln; HCM1s have Ile or Val with an additional Asp (determinant 1) and Gln (determinant 2); and HCM2s have Ile and Ser. The most recently characterized PCM from *Xanthobacter autotrophicus* clusters with six other mutases. All of these mutases have Leu and Asn in the specificity positions, which provides space to bind the bulky pivalyl group (Figs. 9 and 11). MCMs and ECMs are further distinguished by two additional substitutions near the active site: a His and an Asn in MCMs are replaced by a Gly and a Pro in ECMs (Fig. 11), as described previously (11). Beyond these differences, catalytically important residues, such as the His and Glu contacting the substrate thioester carbonyl (Fig. 7) and the Glu contacting the 5′-dAdo ribose hydroxyl groups (Fig. 6d), are conserved in all sequences (Fig. 11), indicating that all sequences represent functional mutases.

This bioinformatic analysis allows for a number of interesting observations. First, archaeal MCMs cluster with ICMs rather than with bacterial and eukaryotic MCMs (Fig. 9) as noted previously (5, 11, 59). Archaeal MCMs encode the substrate-binding and Cbl-binding domains on a separate polypeptide, in notable contrast to most bacterial (see below) and all eukaryotic MCMs, which encode both domains on a single polypeptide. Most sequences annotated as archaeal MCMs have the characteristic features of MCMs, and MCM from the archaeon *Pyrococcus horikoshii* was recently shown to indeed have MCM activity (60). Notably, archaea appear to contain additional acyl-CoA mutases, including ICMs and several yet-uncharacterized mutases (Fig. 9), indicating that archaea use a variety of different AdoCbl-dependent reactions. The metabolic roles of these reactions remain to be determined.

Second, we observe two groups of MCMs, one containing MCMs from Firmicutes such as several *Clostridium* species and one containing MCMs from Thermotogae, that do not cluster with other bacterial MCMs (Figs. 9 and 10). Closer inspection reveals that these mutases contain the active site determinants of MCMs (Fig. 9) but are encoded on two separate polypeptides, in contrast to other known bacterial MCMs. Thus, both archaeal MCMs and a subgroup of bacterial MCMs resemble ICMs more closely than other MCMs, highlighting the complex evolutionary history of acyl-CoA mutases.

Third, the bioinformatic analysis reveals two phylogenetically distinct sequence clusters that do not contain characterized members (uncharacterized mutase clusters 1 and 2, see Fig. 9) and thus could represent new AdoCbl-dependent mutases. Analysis of the genomic context reveals that these mutase substrate-binding domains are encoded in larger operons that also encode a corresponding Cbl-binding domain, suggesting that they are active mutases. Uncharacterized mutase cluster 1 contains six mutases from different archaea and bacteria, including *A. aromaticum* (formerly *Azotococcus* strain EbN1). Uncharacterized mutase cluster 2 currently contains four sequences from different Deltaproteobacteria, including *D. alkenivorans*, a metabolically versatile bacterium (23, 61). The mutases in these clusters contain Tyr and Arg in the determinant positions, likely allowing them to bind carboxylate-bearing substrates (Fig. 11).

To further examine these uncharacterized mutases, we analyzed the sequences and generated homology models of these mutases. Mutases in uncharacterized cluster 2 have relatively low sequence similarity to MCM and IcmF, limiting the reliability of homology models. Current homology models reveal several structural changes in the active site, but do not provide conclusive evidence on the active site architecture. Notably,
mutases in uncharacterized cluster 1 contain a stretch of six residues in the active site with the sequence AGGGGG (Fig. 11), replacing several residues, including an otherwise strictly conserved Gln that contacts the 5’-dAdo group (Gln-330 in MCM or Gln-865 in IcmF, Fig. 6d) by small Ala and Gly residues (Fig. 11). In addition, an otherwise conserved Phe in the active site (Phe-287 in MCM and Phe-823 in IcmF) is replaced by Asn. A homology model generated using the structure of substrate-bound Structures of Isobutyryl-CoA Mutase
bound MCM (see under “Experimental Procedures”) (25) reveals that the Gln to Ala substitution enlarges the active site cavity, which now appears ideally suited to bind substrates carrying larger substituents, and that the Phe to Asn substitution positions an additional hydrogen bonding partner in the active site (Fig. 12). Thus, these mutases may catalyze novel AdoCbl-dependent interconversions.

Discussion

Acyl-CoA mutases are a growing family of AdoCbl-dependent enzymes that perform challenging carbon skeleton rearrangements. Despite extensive studies, our understanding of catalysis and substrate specificity of acyl-CoA mutases remains incomplete, currently limiting their utility for biotechnological applications.
applications. Here, we report crystal structures of IcmF, an ICM variant, bound to four acyl-CoA substrates, revealing how this acyl-CoA mutase positions its substrates for catalysis. Together with bioinformatic analyses, these structures expand our understanding of catalysis and substrate specificity in acyl-CoA mutases.

As observed previously in MCM (25, 26) and HCM1 (29), the IcmF acyl-CoA substrates are threaded through the 8-stranded β-barrel of the substrate-binding domain. Notably, TIM barrels typically feature a tightly packed hydrophobic core with the active site formed by loop regions at the periphery of the barrel and do not use the barrel core for substrate binding. The unique use of TIM barrels by acyl-CoA mutases and other AdoCbl-dependent enzymes is highlighted by the IcmF structures.
pendent enzymes is likely an adaptation to protect the radical-based intermediates that form during catalysis. The TIM barrels of acyl-CoA mutases are further distinguished by their ability to undergo a dramatic conformational change from a closed to an open conformation, splitting the barrel into two halves of four strands each (Fig. 5c). Initially observed in MCM (24, 25), we subsequently also captured IcmF in both these conformations (28), suggesting that TIM barrel flexibility is a general feature of acyl-CoA mutases. HCM1 has so far only been observed in the closed state (29), but only a single structure has been reported. In MCM, the barrel was captured in the open and closed conformations in the absence and presence of substrate, respectively, leading to the suggestion that the conformational change is substrate-induced (24, 25). In IcmF, however, we captured both open and closed conformations in the same structure in the absence of substrate (28), and here we again capture both conformations in the presence of substrate, with substrate binding to the already closed TIM barrel. The TIM barrel open conformation instead correlates with loss of the Cbl 5’-dAdo group and displacement of the Cbl out of the active site into a catalytically inactive conformation, possibly mediated by the cognate G-protein chaperone, which is absent in structures of MCM. Thus, we now have a series of snapshots depicting the TIM barrel in both open and closed conformations, indicating that these two conformations are in equilibrium, affected by the presence of substrates as well as by other factors such as presence of the G-protein chaperone and the cofactor state. Although TIM barrel opening could help product release and substrate binding and may play a role in triggering Co–C bond homolysis upon substrate binding (25), our IcmF structures indicate that the barrel does not absolutely need to open and close for every catalytic cycle. It appears that these barrel dynamics are inherent to acyl-CoA mutases, but further studies will be required to determine their role during catalysis and cofactor recycling.

In the active site, substrate binding and Co–C bond homolysis need to be tightly coupled to ensure a high catalytic rate while preventing generation of the 5’-deoxyadenosyl radical without substrate. Our structures suggest that in IcmF, the 5’-dAdo undergoes a conformational change from C2’-endo in intact AdoCbl to C3’-endo upon Co–C bond homolysis that propels the C5’ radical from its position above the Cbl cobalt toward the substrate for hydrogen atom abstraction. This pseudorotation of the 5’-dAdo ribose group appears ideally suited to bridge the 5.5–6.5 Å distance between the substrates and the Cbl in IcmF. Other AdoCbl-dependent mutases similarly position their substrates at the same distance from the Cbl, as determined from crystal structures (24, 29, 33) and by electron paramagnetic resonance spectroscopic studies of glutamate mutase (62) and MCM (63) under catalytic conditions. Pseudorotation of the 5’-dAdo ribose has also been suggested from structural and biochemical studies of glutamate mutase (33, 55) and from computational studies on MCM (64). Thus, given the combined structural, biochemical, and computational evidence, it appears that this mechanism of moving the active C5’ radical toward substrate is conserved in AdoCbl-dependent mutases such as acyl-CoA mutases. Notably, another group of AdoCbl-dependent enzymes, the eliminases, appear to employ a different conformational change; here, movement is proposed to occur by rotation about the 5’-dAdo N-glycosidic bond to bridge the larger distance of 11 Å between the substrate and the Cbl (65–67).

To accelerate Co–C bond homolysis, substrate binding likely modulates the interactions between the protein and the 5’-dAdo, for example by inducing large scale conformational changes such as the TIM barrel motions or by altering active site electrostatics or dynamics to destabilize the C2’-endo form or to stabilize the C3’-endo form (55, 68, 69). It is unclear how many molecular mechanisms AdoCbl enzymes use to afford the substantial 1012 enhancement in Co–C bond homolysis that accompanies substrate binding (1, 3, 70). The IcmF structures reported here suggest that the C2’-endo to C3’-endo transition that is promoted by substrate binding may be a more common mechanism for increasing Co–C homolysis rates than previously thought. These structures also cast doubt on the relevance of the TIM barrel motions to homolysis rates, if barrel opening and closing need not accompany every turnover. Although more studies are needed to understand the relationship between substrate binding and Co–C bond homolysis, it is clear that substrate positioning with respect to the AdoCbl is universally important. Both substrate radical generation by AdoCbl and AdoCbl regeneration following turnover require precise positioning of the substrate in the active site. Our structures of the acyl-CoA mutase IcmF reveal that all four substrates are positioned similarly, with a β-carbon pointed toward the 5’-dAdo for hydrogen atom abstraction. It appears that the active site has some flexibility, in particular at the position of Phe-598, allowing it to accommodate both sets of acyl-CoA substrates. Nevertheless, it appears that Phe-598 is a critical determinant for substrate specificity; smaller residues at this position allow for more facile binding of substrates with tertiary α-carbons, as observed in HCM and PCM.

With three different acyl-CoA mutases now known to have the same overall structure and mode of substrate binding, we can more reliably identify determinants of substrate specificity for other members of this class. Our bioinformatic analyses identify critical sequence determinants, similar to previous analyses (4, 11), for substrate specificity; the presence of a charged Arg and an additional Tyr is required for binding of substrates with carboxylate groups, whereas a panel of smaller groups allows for binding of different aliphatic substrates such as isobutyryl-CoA and hydroxysobutyryl-CoA. These analyses allow us to look more closely at two groups of thus far uncharacterized mutases (Fig. 9) and to predict their activities.

The first cluster of uncharacterized mutases is found within larger operons that encode a putative hydantoinase, a thiolaose, a CoA transferase, a tungsten-dependent (in archaea) or molybdenum-dependent oxidoreductase (in bacteria), and other enzymes. Although strains carrying these mutases are not well characterized, recent studies suggest that a strain related to A. aromaticum, Azotobacter evansii, employs this operon for anaerobic degradation of indoleacetate (18). The authors proposed that the degradation pathway involves the carbon skeleton rearrangement of 2-(2-aminophenyl)succinyl-CoA to (2-amino-phenyl)malonyl-CoA (Fig. 1f) by a novel acyl-CoA mutase and used similar bioinformatic analyses to identify the same
cluster of uncharacterized mutases (18). Our understanding of substrate binding in acyl-CoA mutases now allows for a re-examination of this proposal. These mutases likely bind carboxylate-bearing substrates, as indicated by the presence of Tyr and Arg in the determinant positions (Fig. 11). Additional replacements around the active site, conserved within this cluster but not in other mutases, lead to a substantially enlarged active site cavity, which may be able to bind the aminophenyl group of the proposed substrate (Fig. 12, a and b). Similarly, these or related mutases could be involved in anaerobic degradation of ethylbenzene, which likely requires a carbon skeleton isomerization after it gets metabolized to (1-phenylethyl)succinyl-CoA (6, 19). Together, our bioinformatic and modeling studies suggest that these mutases represent a new class of acyl-CoA mutases and accept substrates with aromatic groups.

Similarly, the second cluster of uncharacterized mutases is encoded in operons responsible for anaerobic oxidation of long-chain alkanes. These operons contain an alkysuccinate synthase of the glycyl radical enzyme family that is proposed to convert alkanes and fumarate to 2-(1'-methylalkyl)succinyl-CoA, which could then be isomerized by an acyl-CoA mutase to (2'-methylalkyl)malonyl-CoA (Fig. 1g) and further processed by β-oxidation (20–23). Notably, D. alkenivorans is indeed known to degrade long-chain alkanes under anaerobic conditions using such a pathway (22, 23). Thus, these uncharacterized mutases likely represent novel (2'-methylalkyl)malonyl-CoA mutases. Indeed, they contain Tyr and Arg in the determinant positions, which would allow them to accept carboxylate-containing substrates (Fig. 11). Unfortunately, our attempts to further model the active sites of these mutases failed due to the low sequence similarity to MCM and IcmF. Further biochemical characterization of these proteins will be required to establish their role in anaerobic alkane degradation.

Notably, A. aromaticum and D. alkenivorans as well as other strains bearing these novel mutases encode several additional acyl-CoA mutases in their genomes. A. aromaticum, for example, encodes MCM and IcmF (Figs. 9, 11), whereas D. alkenivorans encodes MCM as well as another uncharacterized mutase (Fig. 9). All of these strains are known to be metabolically flexible, and it is tempting to speculate that this ability in part stems from a diverse array of acyl-CoA mutases. These analyses highlight the complex evolutionary history of acyl-CoA mutases, which likely underwent specialization, frequent horizontal gene transfer, and different domain fusion events.

Altogether, our studies reveal important design principles of acyl-CoA mutases. The improved understanding of different acyl-CoA mutase classes and their metabolic versatility could help facilitate the rational and directed engineering of these acyl-CoA mutases for applications ranging from generation of branched-chain biofuels to hydrocarbon remediation.

Acknowledgment—We thank Marcus Gibson for help with generation of the phylogenetic tree and Professor Barbara Imperiali for advice and help regarding the synthesis of pivalyl-CoA.

References

20. Jarling, R., Sadeghi, M., Drozdowska, M., Lahme, S., Buckel, W., Rabus, R.,


So, C. M., and Young, L. Y. (1999) Isolation and characterization of a


