Structures of 2-Hydroxyisobutyric Acid-CoA Ligase Reveal Determinants of Substrate Specificity and Describe a Multi-Conformational Catalytic Cycle

Michael Zahn[1,†], Nadya Kurteva-Yaneva[2,†], Judith Schuster[2], Ulrike Krug[3], Tina Georgi[2], Roland H. Müller[2], Thore Rohwerder[2] and Norbert Sträter[1,]

1 - Institute of Bioanalytical Chemistry, *Center for Biotechnology and Biomedicine, Leipzig University, Deutscher Platz 5, 04103 Leipzig, Germany*
2 - Department of Environmental Microbiology, *Helmholtz Centre for Environmental Research—UFZ, Permoserstraße 15, 04318 Leipzig, Germany*
3 - Institute for Medical Physics and Biophysics, *Faculty of Medicine, Leipzig University, Härtelstraße 16-18, 04107 Leipzig, Germany*

Correspondence to Thore Rohwerder and Norbert Sträter: *Permoserstraße 15, 04318 Leipzig, Germany. Deutscher Platz 5, 04103 Leipzig, Germany. thore.rohwerder@ufz.de, strater@bbz.uni-leipzig.de*
https://doi.org/10.1016/j.jmb.2019.05.027
Edited by Dan Tawfik

Abstract

2-Hydroxyisobutyric acid (2-HIBA) is a biomarker of adiposity and associated metabolic diseases such as diabetes mellitus. It is also formed in the bacterial degradation pathway of the fuel oxygenate methyl *tert*-butyl ether (MTBE), requiring thioesterification with CoA prior to isomerization to 3-hydroxybutyryl-CoA by B_{12}-dependent acyl-CoA mutases. Here, we identify the adenylating enzymes superfamily member 2-HIBA-CoA ligase (HCL) in the MTBE-degrading bacterium *Aquincola tertiaricarbonis* L108 by knockout experiments. To characterize this central enzyme of 2-HIBA metabolism, ligase activity kinetics of purified HCL and its X-ray crystal structures were studied. We analyzed the enzyme in three states, which differ in the orientation of the two enzyme domains. A 154° rotation of the C-terminal domain accompanies the switch from the adenylate- into the thioester-forming state. Furthermore, a third conformation was obtained, which differs by 50° and 130° from the adenylation and thioesterification states, respectively. Phylogenetic and structural analysis reveals that HCL defines a new subgroup within phenylacetate-CoA ligases (PCLs) thus far described to exclusively accept aromatic acyl substrates. In contrast, kinetic characterization clearly demonstrated that HCL catalyzes CoA activation of several aliphatic short-chain carboxylic acids, preferentially 2-HIBA. Compared to the classical PCL representatives PaaK1 and PaaK2 of *Burkholderia cenocepacia* J2315, the acyl binding pocket of HCL is significantly smaller and more polar, due to unique active-site residues Y164 and S239 forming H-bonds with the OH-group of the acyl substrate moiety. Furthermore, HCL and PaaK topologies illustrate the evolutionary steps leading from a homodimeric to the fused monomeric core fold found in other ligases.

Introduction

CoA thioesters play an important role in primary and secondary metabolism. Acetyl-CoA, succinyl-CoA and other acyl-CoA thioesters are intermediates in the tricarboxylic acid cycle, carbohydrate catabolism and fatty acid synthesis as well as degradation via beta oxidation. In addition, carboxylic and amino acids as building blocks in polyketide [1] and non-ribosomal peptide [2,3] synthesis are activated by CoA thioester formation prior to incorporation into these secondary metabolites. Another example for a CoA thioester-dependent reaction is the carbon skeleton rearrangement of carboxylic acids catalyzed by B_{12}-dependent acyl-CoA mutases through a radical mechanism [4]. Here, the relatively large CoA group guarantees a tight binding of the acyl-CoA substrate, even if the acyl moiety itself lacks functional groups for a strong interaction with the enzyme. An activation of carboxylic acids to the corresponding CoA thioesters is also required in biotin precursor synthesis [5] and dissimilatory pathways of specific substrates,

such as phenylacetic and benzoic acid [6]. In total, CoA thioesters may be involved in about 4% of all enzymatic reactions [7] and the use of these building blocks seems to be a very ancient invention in biology [8].

Many carboxylic acids are activated by AMP-forming acyl-CoA synthetases, also called acid-CoA ligases, using the free energy of ATP hydrolysis for thioester bond formation. Most of these enzymes belong to the large superfamily of adenylating or acyl-AMP/thioester-forming enzymes (also designated as ANL superfamily), including also luciferases and the adenylation domains of non-ribosomal peptide synthetases [9]. In addition, two other classes of adenylate-forming enzymes are the aminoacyl-tRNA synthetases and specific subunits of non-ribosomal peptide synthetase-independent siderophore synthetases [8]. Recently, pimelate-CoA ligase BioW has been identified as a fourth class of adenylate-forming enzymes [5]. Interestingly, all these enzyme classes are structurally not related and, consequently, the strategy of carboxylate activation via an acyl-AMP intermediate has been adopted several times in convergent evolution.

On principle, the energetically difficult acyl-AMP formation from the free carboxylic acid through attack of the ATP α-phosphate requires a proper positioning of all reactive groups [8,9]. This challenge has been met in different ways. In all ANL enzymes, albeit overall sequence similarity among the diverse members of this superfamily only ranges between 20% and 30% [9,10], catalysis is carried out through a conserved two-step mechanism (Fig. 1). After formation of an enzyme-bound acyl-AMP intermediate (adenylation), a large-scale rotation of the C-terminal domain of about 140° occurs and the carboxylic acid residue is transferred to CoA or another thiol-bearing carrier molecule (thioesterification) [9]. This domain movement results in the presentation of different faces of the C-terminal domain to the active-site part formed by the N-terminal domain, thus enabling successive catalysis of the two part reactions. The loop bearing the catalytic Lys residue required for the adenylating half-reaction, for example, moves about 25 Å away from the active site. In contrast, ANL-unrelated ligases generally achieve the two-step catalysis by much smaller conformational changes, for example, the pimeloyl-CoA synthetase BioW by <3 Å movements of few active-site residues [5]. It has been speculated that this discrepancy results from the evolution of today's ANL enzymes from ligases which catalyzed the activation of simple carboxylic acid substrates having only “one handle” (functional polar group), that is, the carboxylate group, for specific interactions [9]. Proper positioning for the α-phosphate attack is achieved by positioning a large aromatic residue (His, Phe or Trp) close to the carboxylate [9,11]. Because of this arrangement, the large domain orientation is required to move the aromatic residue out of the active site and position the nucleophile for the thioesterification step. An advantage of the large domain rotation in the ANL enzymes might be that evolutionary adaptations to new substrates are more easily achieved compared to the other enzymes that have more restrictions concerning the active-site geometry for catalyzing the two chemical steps. Indeed, the ANL-unrelated enzyme classes present much less divergent groups of adenylate-forming enzymes. ANL superfamily enzymes are often promiscuous, but the substrate spectrum can also be quite narrow and several subgroups, such as enzymes with specificity to aromatic compounds or aliphatic carboxylic acids with short-, medium- or long-chain acyl residues, can be distinguished [10].

The ANL active site including the acyl substrate binding pocket is located at the interface of the N- and C-terminal domains involved in acyl-AMP and CoA thioester formation [9]. The size and structure of this acyl binding pocket mainly determine substrate specificity. The thus-far characterized representatives of the phenylacetate-CoA ligase (PCL) group, for example, initiate bacterial degradation of phenylacetic acid derived from phenylalanine, lignin-related phenylpropane units, and other natural or xenobiotic aromatic compounds (EC 6.2.1.30) [6]. Crystal structures of two paralogous PCLs from *Burkholderia cenocepacia* J2315, PaaK1 and PaaK2, have been analyzed [12]. In PaaK1, 13 amino acid residues have been identified defining the aryl substrate binding pocket, while only 11 residues are involved in PaaK2.

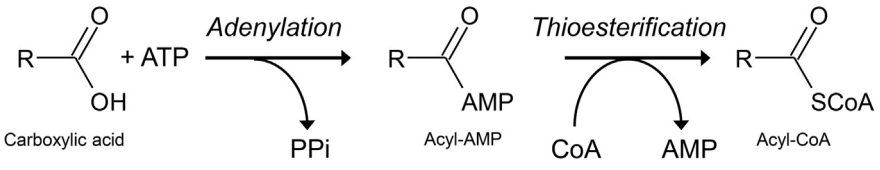

Fig. 1. Two-step mechanism of carboxylic acid-CoA ligase catalysis. During adenylation, an enzyme-bound acyl-AMP mixed anhydride intermediate and pyrophosphate (PPi) are formed from ATP and carboxylic acid catalyzed by the adenylate-forming conformation. Then, in ANL enzymes a large-scale rotation of the C-terminal domain to the thioester-forming conformer occurs and the acyl residue is transferred to CoA in the thioesterification step resulting finally in the release of acyl-CoA and AMP. *In vivo*, coupling of this energetically unfavorable reaction with the highly exergonic hydrolysis of PPi by pyrophosphatase assures efficient thioester synthesis.

This divergent extension of the aryl substrate pocket in PaaK1 resulted in a broader substrate spectrum allowing enhanced CoA activation of phenylacetic acid derivatives with polar ring substitutions. In addition, a closely related PCL enzyme from *Bacteroides thetaiotaomicron* VPI-5482 (Bt_0428), which accepts preferentially only phenylacetic acid [13], has been structurally characterized (PDB ID: 4RVN).

Another specific acid-CoA ligase has been postulated for the activation of 2-hydroxyisobutyric acid (2-HIBA), which is an intermediate in bacterial degradation of the fuel oxygenate methyl *tert*-butyl ether (MTBE). In this pathway (Fig. 2), the CoA-activated branched carboxylic acid is isomerized to the common metabolite 3-hydroxybutyryl-CoA by a B_{12}-dependent mutase reaction [15]. In the bacterial strain *Aquincola tertiaricarbonis* L108 and also in other MTBE-degrading bacteria such as *Methylibium petroleiphilum* PM1, the two-component enzyme HcmAB is responsible for the rearrangement of 2-hydroxyisobutyryl-CoA [14,16]. In addition, the recently discovered B_{12}-dependent mutase RcmAB catalyzes the isomerization of 2-hydroxyisobutyryl-CoA in the bacterial strains *Kyrpidia tusciae* DSM 2912 [17] and *Bacillus massiliosenegalensis* JC6 [18], indicating that 2-HIBA degradation pathways evolved several times. In all these strains, however, the mutase genes are associated in an operonic organization together with genes encoding the putative G protein chaperone MeaH and a putative 2-HIBA-CoA ligase (HCL) [17]. The latter enzyme would be the first acid-CoA ligase showing specificity to a short-chain carboxylic acid bearing a tertiary OH-group.

Phylogenetic analysis reveals a significant distance of HCL to other bacterial short- and medium-chain acid-CoA ANL ligases characterized thus far. For sequences of propionate-CoA ligase PrpE from *Salmonella enterica* [19] and short-chain acid-CoA ligase FadK from *Escherichia coli* [20], for example, a BLAST analysis [21] shows only very low similarity to HCL, with less than 20% coverage of homologous sequence segments. However, HCL is more closely related to PCLs, for example, sharing 34% and 33% identical residues and 91% sequence coverage with PaaK1 and PaaK2 from *B. cenocepacia* J2315, respectively.

In order to identify the role of HCL in MTBE and 2-HIBA degradation by *A. tertiaricarbonis* L108, we constructed an insertion knockout mutant strain L108(*Δhcl*)K4. This strain was unable to grow on the MTBE metabolite *tert*-butyl alcohol (TBA) but accumulated 2-HIBA on this substrate stoichiometrically. The purified ligase showed high specificity for 2-HIBA. To characterize the enzyme's molecular determinants of substrate specificity, its adenylate-forming complexes with 2-HIBA and other short-chain carboxylic acids were analyzed. In addition, crystal structures of the thioesterification conformation and an open conformation were obtained. These structures show the high flexibility of the domain orientation necessary for the whole reaction cascade in agreement with previous studies.

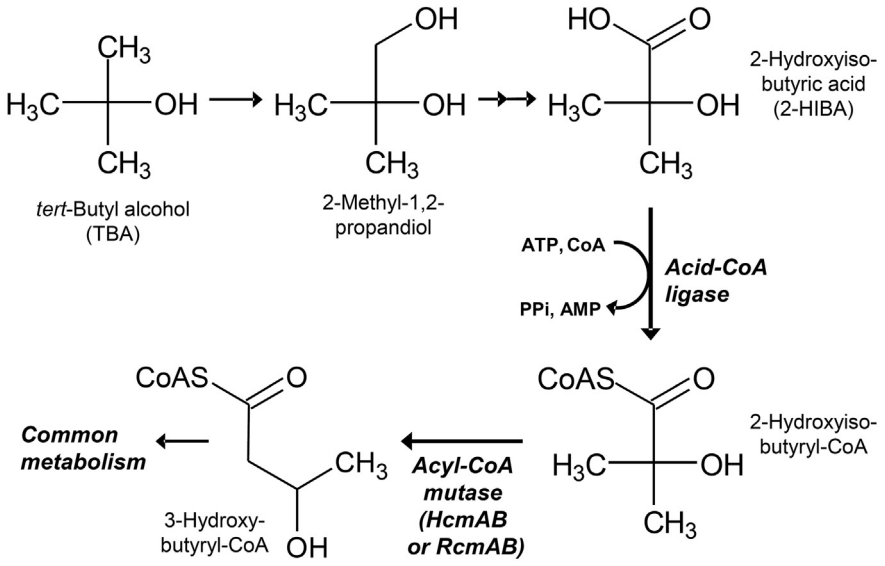

Fig. 2. Bacterial degradation of the MTBE metabolite TBA. After hydroxylation and two dehydrogenase steps, 2-HIBA is formed, which is converted to the corresponding CoA thioester by a carboxylic acid-CoA ligase. Then, 2-hydroxyisobutyryl-CoA is isomerized by a B_{12}-dependent acyl-CoA mutase to the common metabolite 3-hydroxybutyryl-CoA [14].

Results

Insertional inactivation of *hcl* and characterization of mutant strain

Transposon-mediated mutagenesis resulted in the *hcl* mutant *A. tertiaricarbonis* L108*(Δhcl)*K4 bearing an insertion of an active kanamycin resistance gene in the *hcl* gene (Fig. S1). In resting-cell experiments, mutant strain K4 stoichiometrically accumulated 2-HIBA from TBA (Fig. S2), establishing the postulated role of HCL in TBA degradation via 2-HIBA [14,15]. In line with this, growth of the mutant strain on TBA as well as 2-HIBA was insignificant (data not shown), while the C5 tertiary alcohol homolog *tert*-amyl alcohol still supported growth, as the latter is not degraded via the acyl-CoA mutase pathway requiring specific activation by a 2-hydroxyacid-CoA ligase [22]. These results strongly support the role of HCL as specific ligase for 2-HIBA activation in *A. tertiaricarbonis* L108, although we cannot completely rule out any polar effects of the insertion on neighboring genes, such as the 2-hydroxyisobutyryl-CoA mutase genes *hcmA* and *hcmB*.

Structure of the adenylate-forming complex

We analyzed the crystal structure of HCL in the adenylate-forming state in complex with 2-HIBA to 2.3 Å resolution (Fig. 3, Table 1). HCL consists of two protein domains, an N-terminal domain of residues 1 to 361 and a C-terminal domain from residues 362 to 467. In agreement with the close phylogenetic relationship to PCLs, the most similar previously determined structures are those of PaaK1 (PDB ID: 2Y27, rmsd for C_α-coordinates 1.9 Å; Figs. S3 and 4c) and PaaK2 (PDB ID: 2Y4O, rmsd 2.1 Å) [12] and of the PCL member from *B. thetaiotaomicron* VPI-5482 (PDB ID: 4RVN, rmsd 2.0 Å). All three enzymes form a symmetrical dimer and a characteristic 7-stranded β-sheet between the two subunits, such that one subunit contributes two and the other subunit five strands to the β-sheet (sheet a, Figs. 4a–c and S4). Interestingly, this β-sheet is present in all members of the ANL superfamily, but it is formed within one subunit in the other members, as noted previously [12]. This is due to the internal symmetry present in the other ANL subunit structures [23], which is most likely the result of a gene duplication and fusion event in the evolution of the ANL enzymes (Figs. S4 and S5 and discussion). Two additional β-sheets (denoted sheets b and c) are present in each subunit further away from the central two-fold dimer axis. These sheets are strongly twisted and bent and they contain β-strands of different length, some of which form part of both sheets. Thus, there is no unique solution to represent these two sheets in a topology diagram such as Fig. S4.

Substrate specificity and active-site architecture of the adenylate-forming complex

At optimal pH and temperature conditions (Fig. S6), that is, pH 7.8 and 35 °C, recombinant HCL is able to convert 2-HIBA to the corresponding CoA thioester at high rates showing V_{max} and k_{cat} values of 20.3 μmol min^{-1} mg^{-1} and 18.9 s^{-1}, respectively (Table 2). Only a few other short-chain carboxylic acids can be CoA-activated by the enzyme (Fig. S7), although at significantly lower rates than 2-HIBA (Table 2, Fig. S8). Still, between 2.6% and 6.5% of the 2-HIBA V_{max} is obtained with 2-HBA (2-hydroxybutyric acid), 3-HIBA (3-hydroxyisobutyryic acid) and pivalic acid. Less than 1% of the 2-HIBA V_{max} is obtained with isovaleric acid. In the case of 3-HBA (3-hydroxybutyryic acid), only catalysis of CoA thioester formation from the (*S*)-enantiomer is observed, showing a V_{max} value of about 0.15% of the rate obtained with the main substrate 2-HIBA. Other carboxylic acids tested, phenylacetic, isobutyric, butyric and acetic acid, cannot be CoA-activated by HCL. With a K_m value of about 200 μM, HCL also shows the highest affinity to 2-HIBA, resulting in a catalytic efficiency of about 60- to 4000-fold higher than with all other carboxylic acids tested. This clearly establishes HCL as an acid-CoA ligase highly specific for 2-HIBA. Hence, HCL forms a new subgroup within the PCL family of aromatic acid-CoA ligases, however, showing high specificity toward the short-chain aliphatic carboxylic acid 2-HIBA.

To analyze the structural basis of HCL substrate specificity, we determined complex crystal structures of the adenylate-forming state of HCL after cocrystallization with ATP and the substrates 2-HIBA, *R*3-HIBA [(*R*)-3-hydroxyisobutyric acid] and *S*3-HBA [(*S*)-3-hydroxybutyric acid] at resolutions of 2.3, 2.2 and 2.2 Å, respectively (Table 1). For 2-HIBA, the electron density clearly reveals the formation of the substrate adenylate 2-HIB-AMP (Figs. 3a, b and S9A), which was refined with 100% occupancy.

The substrate specificity of HCL corresponds well to the adenylate-forming active-site architecture. Identical to the PaaKs and other acid-CoA ligases, the acyl substrate pocket is located at the interface of the N- and C-terminal domains. However, in direct comparison with PaaK1 and PaaK2, the acyl pocket of HCL is significantly smaller in size (Fig. 4b, d). The amino acids interacting with the acyl substrate moiety of 2-HIB-AMP comprise only the N-terminal domain residues Y164, Y208, S239 and F272. In addition, K455 (catalytic residue for adenylating half-reaction) from the C-terminal domain forms an H-bond to the carbonyl oxygen of the substrate anhydride linkage and a salt bridge to the phosphate residue of the adenylate (Fig. 4b). Thus, binding of phenylacetic acid or other aryl substrates as well as aliphatic carboxylic acids with more than five carbon atoms is impeded by the limited size of the acyl

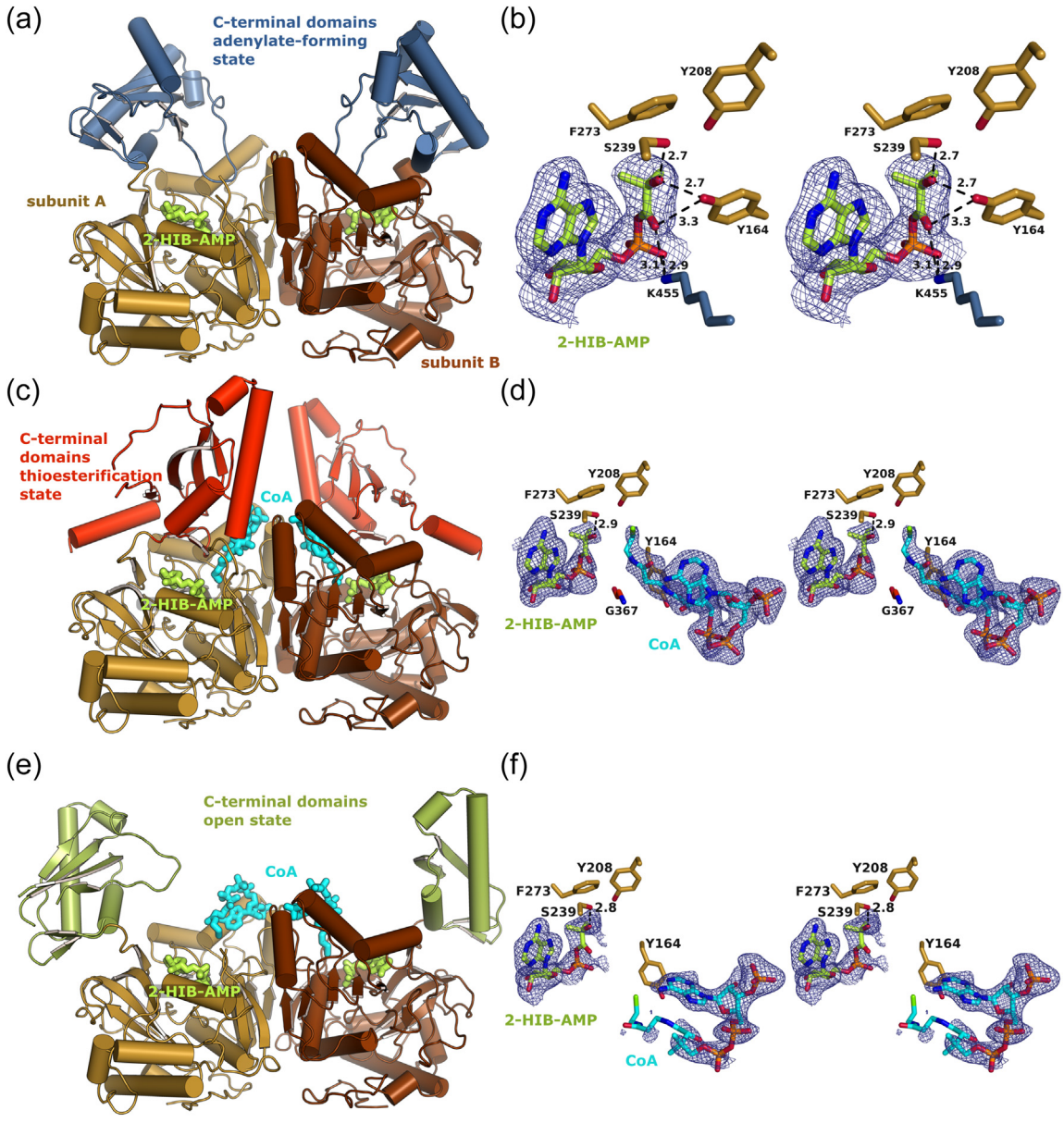

Fig. 3. Different conformations of the HCL dimer. The overall view and the stereoscopic views of bound ligands are shown for the adenylate-forming state in panels a and b, respectively, the thioesterification state in panels c and d and the open conformation in panels e and f. The N-terminal domains are colored in yellow and brown. The C-terminal domains are colored blue (a), red (c) or green (e). The ligands 2-HIB-AMP and CoA are colored in green and cyan, respectively. For all ligands (2-HIB-AMP and CoA), the omit electron density map (blue, 2.0 σ_{rms}) is shown.

binding pocket. Another unique feature of this binding pocket in HCL is its polarity established by residues Y164 and S239, which correspond to the nonpolar residues F141 and G213 in PaaK1 (Fig. S3). Substrate 2-HIB-AMP fits well into this polar substrate pocket and its OH-group forms H-bonds with Y164 and S239 (Fig. 4b). Y164 in HCL corresponds to the hydrophobic residue mentioned in the introduction that participates in substrate positioning in the ANL enzymes and that repositions with the domain reorientation to make space for the nucleophile in the thioesterification step.

Furthermore, we analyzed adenylate-forming complex structures with *R*3-HIB-AMP and *S*3-HB-AMP (Fig. S9B, C). After incubation with ATP and acid substrates, alternative binding of the product mixed anhydride acyl-AMP (refined with 50% occupancy) and of AMP (resulting from the hydrolysis of ATP or the mixed anhydride) together with the free carboxylic acid was observed (both refined with 50% occupancy).

Table 1. Crystallographic data and refinement statistics.

	Adenylation 2-HIB-AMP	Adenylation *R*3-HIB-AMP	Adenylation *S*3-HB-AMP	Thioesterification 2-HIB-AMP + CoA	Open conformation 2-HIB-AMP + CoA
Ligands bound (and occupany)	2-HIB-AMP (1.0)	*R*3-HIB-AMP (0.5), *R*3-HIB (0.5), AMP (0.5)	*S*3-HB-AMP (0.5), *S*3-HB (0.5), AMP (0.5)	2-HIB-AMP (0.5), AMP (0.39), CoA (1.0)	2-HIB-AMP (0.5), AMP (0.4), CoA (1.0)
Data collection					
Beamline	Bessy 14.1	Bessy 14.1	Bessy 14.1	Bessy 14.1	Bessy 14.1
Space group	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P3_121$	$P2_12_12_1$
Cell dimensions					
a, b, c (Å)	83.9, 83.9, 219.8	83.9, 83.9, 219.7	84.2, 84.2, 219.8	95.1, 95.1, 100.7	93.9, 101.9, 109.5
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 120.0	90.0, 90.0, 90.0
Solvent content (%)	64.4	64.4	64.7	47.6	47.4
Resolution (Å)	24.58–2.30 (2.42–2.30)	24.58–2.20 (2.32–2.20)	24.65–2.20 (2.32–2.20)	40.00–2.31 (2.43–2.31)	40.00–2.30 (2.42–2.30)
R_{merge} (%)	10.7 (96.3)	10.7 (89.7)	8.7 (80.9)	7.8 (33.1)	8.7 (47.2)
R_{pim} (%)	5.3 (47.8)	4.9 (41.4)	4.4 (41.2)	2.9 (12.9)	4.0 (21.3)
$<I/\sigma I>$	13.2 (1.7)	11.9 (2.0)	13.2 (2.0)	16.2 (5.0)	11.9 (2.0)
Completeness (%)	98.2 (99.7)	99.8 (99.4)	99.9 (99.7)	99.8 (99.0)	99.9 (100.0)
Redundancy	5.0 (5.0)	5.7 (5.6)	4.8 (4.8)	8.2 (7.5)	5.6 (5.7)
Refinement					
R_{work}/R_{free}	21.3/24.0	20.6/23.7	20.8/24.1	17.4/22.0	18.8/22.2
Ramachandran plot					
Most favored (%)	91.6	91.3	91.3	90.1	90.3
Allowed (%)	7.7	8.2	8.0	9.4	9.0
Generously allowed (%)	0.5	0.2	0.5	0.2	0.3
Disallowed (%)	0.2	0.2	0.2	0.2	0.4
No. atoms					
Protein	3741	3737	3737	3764	7010
Water	102	135	150	164	291
Ligands	29	29	29	100	200
B-factors					
Protein (Å^2)	53.6	47.5	50.2	39.9	43.2
Water (Å^2)	43.9	42.1	44.4	38.5	40.5
Ligands (Å^2)	35.7	35.7	37.6	30.7	52.2
R.m.s. deviations					
Bond lengths (Å)	0.01	0.01	0.01	0.01	0.01
Bond angles (°)	0.98	1.03	1.02	1.06	1.02
Molprobity clash score[a]	2	2	3	3	1
PDB Code	6HDW	6HDX	6HDY	6HE0	6HE2

[a]The Molprobity all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms).

As in the case of 2-HIBA, the OH-groups of *R*3-HIBA and *S*3-HBA form H-bonds with active-site residues S239 and Y164. However, their carbon backbones are longer (Fig. S7) probably causing less-favorable stereochemical arrangements. This is indicated by the significantly higher K_m and lower ligase activity values (Table 2) and by the binding of the acid-AMP product in equilibrium with the free acid and AMP, in contrast to full occupancy binding observed for 2-HIB-AMP.

Pivalic acid, although structurally closely related to 2-HIBA in size and degree of branching (Fig. S7), cannot form the important H-bond interaction with Y164 and S239, as the acyl OH- is replaced by another methyl group. Consequently, enzyme activity and substrate affinity are significantly reduced compared to the main substrate 2-HIBA, resulting in an about 60 times lower catalytic efficiency (Table 2).

In the comparison of the structurally related 3-HIBA and isovaleric acid substrates, the latter compound is again turned over with about 50-fold reduced catalytic efficiency (Table 2). This is likely also due to the lack of an acyl OH-group in isovaleric acid (Fig. S7), resulting in a completely nonpolar acyl moiety unable to form H-bonds with Y164 and S239 of the acyl binding pocket.

HCL in the thioesterification state

We cocrystallized HCL with AMP and 2-hydroxyisobutyryl-CoA, the products of the thioesterification step (Fig. 1), to obtain the thioesterification conformation (Table 1). Electron density maps indicate the presence of AMP and the substrates 2-HIB-AMP and CoA of the thioesterification step (Fig. 3c,d and Fig. 5). AMP and 2HIB-AMP bind to

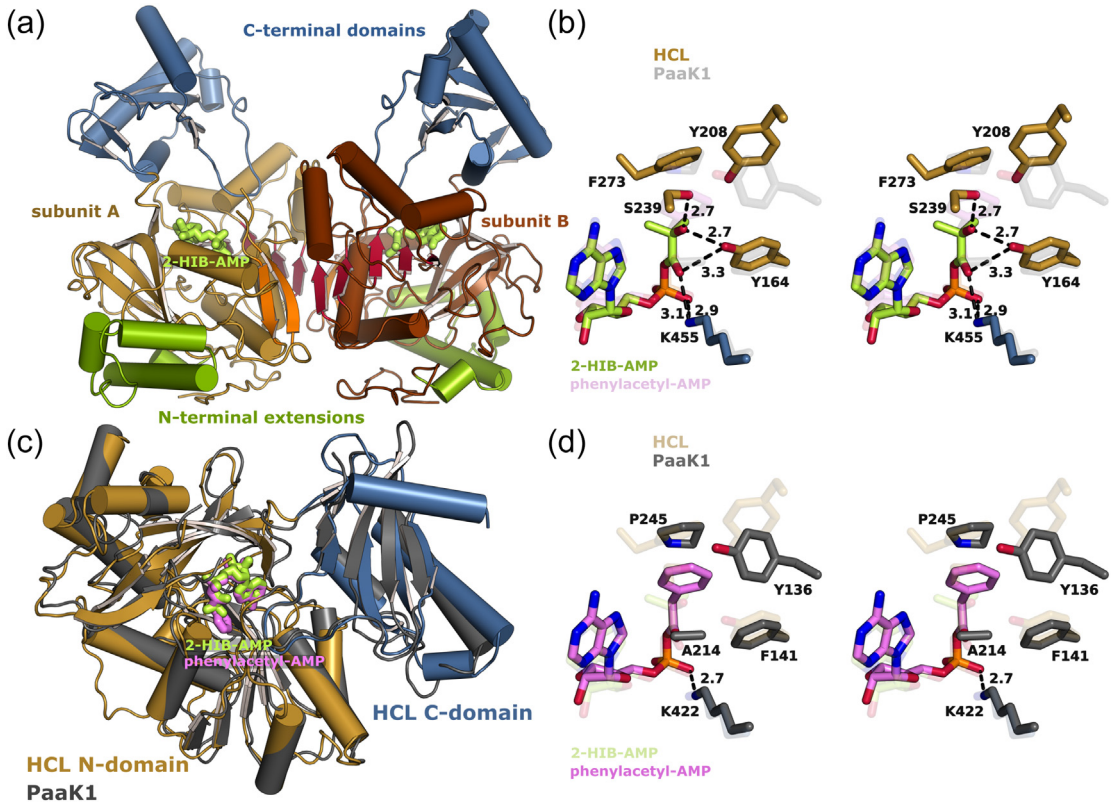

Fig. 4. Adenylate-forming states of HCL and PaaK1. (a) X-ray crystal structure of dimeric HCL in the adenylate-forming state. The N-terminal domains are colored in yellow and brown. The N-terminal extension is colored in green. The C-terminal domains are colored in blue. The ligand 2-HIB-AMP is colored in green. The intersubunit β-sheets are colored in orange and red. (b) Stereoscopic view of the 2-HIB-AMP interactions. Interacting residues of the N-terminal distance to the thioester low, and those of the C-terminal domain are colored in blue. 2-HIB-AMP is colored in green. Equivalent superimposed PaaK1 (PDB ID: 2Y4N) residues and ligand phenylacetyl-AMP are shown in pale colors. (c) Superposition of the adenylate-forming conformations of HCL (N-terminal domain colored in yellow, C-terminal domain colored in blue) and PaaK1 (gray, PDB ID: 2Y4N). Ligands are colored in green (2-HIB-AMP) and magenta (phenylacetyl-AMP). (d) Stereoscopic view of the interactions of phenylacetyl-AMP with PaaK1 (magenta). Equivalent aligned HCL residues and 2-HIB-AMP are shown in pale colors.

the same binding site alternatively with occupancies of 39% and 50%, respectively. The thioesterification states of HCL and PaaK2 have a similar domain orientation with an rmsd value of 2.0 Å for superposition of both domains, and a difference in the domain orientation of 10° (Fig. 5e). The thioesterification state of HCL differs from the adenylate-forming state by a domain rotation of 154° (see analysis in the next section for further details). Whereas the H-bond interaction of 2-HIB-AMP with S239 is maintained in the thioesterification conformation, Y164 is reorientated to make space for the incoming ligand CoA. In addition, K455 is replaced by the much smaller G367 for the accommodation of CoA for the nucleophilic

Table 2. Kinetic parameters of reconstituted HCL incubated with various carboxylic acid substrates at pH 7.8 and 35 °C

Substrate	V_{max} [a] (μmol min^{-1} mg^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
2-Hydroxyisobutyric acid	20.3 ± 1.0	201 ± 33	18.9 ± 0.9	93.9 ± 16.0
Pivalic acid	1.32 ± 0.14	771 ± 173	1.23 ± 0.13	1.60 ± 0.40
3-Hydroxyisobutyric acid (racemate)	1.05 ± 0.05	840 ± 94	0.97 ± 0.05	1.16 ± 0.14
2-Hydroxybutyric acid (racemate)	0.52 ± 0.02	800 ± 77	0.49 ± 0.02	0.61 ± 0.07
Isovaleric acid	0.11 ± 0.03	3990 ± 1840	0.10 ± 0.03	0.025 ± 0.014
(*S*)-3-hydroxybutyric acid	0.031 ± 0.002	490 ± 50	0.029 ± 0.002	0.058 ± 0.007

[a] Acyl-CoA formation was not detected (<0.001 μmol min^{-1} mg^{-1} ligase activity when tested with up to 10 mM carboxylic acid substrate) for (*R*)-3-hydroxybutyric acid, isobutyric acid, butyric acid, phenylacetic acid and acetic acid.

attack (Fig. 5d). CoA is well anchored by many interactions with amino acid side chains and is bound in a substantially bent conformation.

At the entrance of the CoA binding cavity, the 3′-phosphate interacts exclusively with residues of the N-terminal domain forming H-bonds with G190, S192 and H216 as well as with A193 mediated by a water molecule. Likewise, the adenine N1 interacts with A159, V160 and T210 via two water molecules. In addition, hydrophobic interactions with the adenine ring are mediated by A186, Y213 and M399. The phosphopantetheinyl arm of the CoA ligand is mainly bound to the enzyme via salt bridges of the α- and β-phosphates with R438 and R366, respectively. In addition, the pantoate carbonyl interacts via water-mediated H-bonds with the repositioned Y164 and with V437.

The cysteamine and β-alanine groups of CoA have significantly higher *B*-factors and weaker density compared to the rest of the cofactor indicating flexibility. In an alternative conformation, the thiol group of CoA may be modeled in <4 Å distance to the anhydride carbonyl carbon atom of 2-HIB-AMP. Although such a catalytically productive binding mode is supported by positive residual density, we refrained from including this conformation into the crystallographic model, as it is not unambiguously indicated by the density maps. Furthermore, all or part of the CoA in this conformation may also exist in the acylated 2-hydroxyisobutyryl-CoA form.

Crystal structure of an open conformation

Interestingly, a second crystal form grew within the same crystallization drop, representing an open conformation, in which the two domains are almost separated from each other (Table 1). AMP, 2-HIB-AMP and CoA were modeled into the electron density (Fig. 3e, f). As in the thioesterification state, AMP and 2-HIB-AMP bind alternatively with occupancies of 40% and 50%, respectively. CoA binds close to its binding site at the N-terminal domain compared to the thioesterification conformation, but it is rotated and bound in a different conformation (Fig. S10). In both

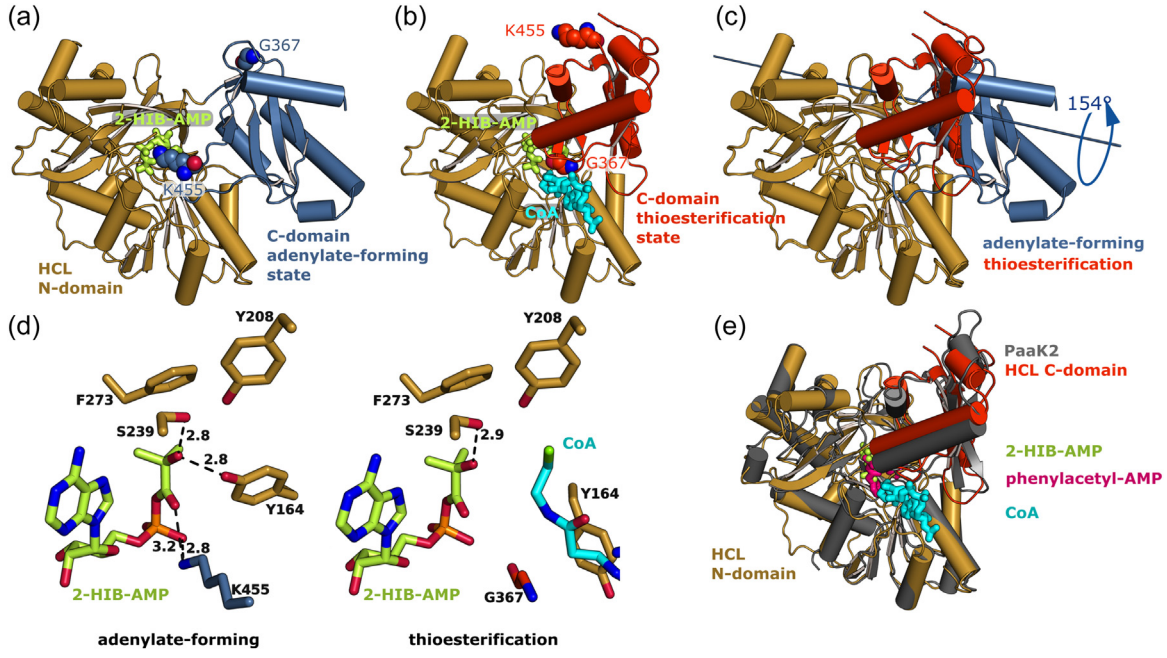

Fig. 5. Adenylate-forming and thioesterification states of HCL. (a) Adenylate-forming state with bound 2-HIB-AMP (green), (b) thioesterification state with bound 2-HIB-AMP (green) and CoA (cyan). The N-terminal domains are colored in yellow. The C-terminal domains are colored in blue and red for the adenylate-forming and thioesterification state, respectively. In the adenylate-forming reaction, K455 points toward the active site. In the thioesterification state, G367 replaces this position. (c) Superposition of the N-terminal domain of HCL of the adenylate-forming and thioesterification states to visualize the rotation of the C-terminal domain relative to the N-terminal domain. The rotation axis is shown as a solid line. (d) Pseudo-stereoscopic view of the interactions of 2-HIB-AMP (green) with HCL in the adenylate-forming state (left) and thioesterification state (right). When viewed in stereo alternating eye-switching results in an optimal impression of the binding modes and interactions of the ligand. Ligand CoA is colored in cyan. Interacting residues of the N-terminal domain are colored in yellow, and those of the C-terminal domain are colored in blue (adenylate-forming) or orange (thioesterification). (e) Superposition of the thioesterification conformations of HCL (N-terminal domain colored in yellow, C-terminal domain colored in red) and PaaK2 (gray, PDB ID: 2Y4O). Ligands are colored in green (2-HIB-AMP), magenta (phenylacetyl-AMP) and cyan (CoA).

protein molecules in the asymmetric unit, the adenosyl 3′-phosphate moiety of CoA is well defined in the electron density and forms the same interactions with H116, G190 and S192 of the N-terminal domain as in the thioester-forming conformation. This indicates a stable binding mode. The rest of the cofactor is flexible, probably due to the absence of the C-terminal domain, which interacts with the pantetheine group of CoA in the thioesterification state. Thus, CoA can bind to the N-terminal domain in the absence of the C-terminal domain in a position that is close to the catalytically competent binding mode of the thioesterification conformation of the enzyme. This open conformation might be important to bind CoA and the substrate prior to the enzyme conformational switch and to release the products.

To analyze the catalytic states with respect to their domain orientation, N- and C-terminal domains were subsequently superposed (Fig. 6). Starting from the open conformation, a rotation of about 50° of the C-terminal domain yields the adenylate-forming state. From this, a 154° rotation is necessary to form the thioester-forming conformation. With a further 130° rotation the open state is restored. The domain reorientations are described by three different rotation axes, which point in distinct directions but intersect at one point at the domain interface, passing near by the hinge residues between the two domains. The non-colinearity of the axes clearly reveals that the open conformation is not an intermediate along a linear rotation trajectory between the adenylate- and the thioester-forming conformations. Nevertheless, since the nature of the movement of the domains between the adenylation and thioesterification states is unknown, a distinction between putative open and intermediate states cannot be made. We denote the observed third conformation of HCL as “open,” whereas other authors have designated additional conformations as “intermediate.”

The average *B*-values of the N-terminal domains are 38.5 and 39.9 $Å^2$ for chains A and B of the asymmetric unit, respectively. In contrast, these values are 50.3 and 81.9 $Å^2$ for the C-terminal domains. The high average *B*-value of the C-terminal domain of chain B compared to a low value of the N-terminal domain indicates a rigid body movement of the complete C-terminal domain in the crystal. A superposition of the two subunits of the dimeric HCL in the asymmetric unit further shows that the C-terminal domains differ by a rotation angle of 11° in their orientation. These findings further demonstrate the high domain mobility of the enzyme.

There is no obvious indication that the open form of HCL in these crystals is stabilized by unusually strong crystal contacts. The strength of crystal packing forces is expected to correlate with improved resolution and with lower solvent content. All three crystal forms diffract to intermediate resolutions around 2.3 Å. The solvent content of the crystals of the open form is 47.4%, representing a normal average value for protein crystals.

Discussion

Position of the PCL group within the ANL superfamily and evolutionary implications

The presence of internal symmetry was recognized in the first crystal structure analysis of an ANL enzyme by Conti and coworkers [23] for firefly luciferase. The authors suggested that this symmetry results from a gene duplication and fusion event in the evolution of these enzymes. Almost all ANL enzymes structurally characterized so far have such a fold with internal symmetry within the subunit(s). Since there are no ANL enzyme structures known that consist of only the structural elements corresponding to one subunit of the homodimeric enzymes HCL or PaaK, we assume that a functional enzyme of this superfamily needs the core elements from both HCL subunits. Thus, the common ancestor enzyme fold most likely resembled HCL and the other PCL enzymes in having a homodimeric core fold forming the common architecture of this superfamily. The gene duplication and fusion event constitutes a distant key branching point in the evolution of the ANL superfamily, which is also the cause of the structural differences between acid-CoA ligases (i.e., between HCL and PaaK with the homodimeric core fold and the other ligases with the duplicated monomeric core fold). In line with this, HCL and the PaaKs share a characteristic N-terminal extension (residues 24–81 in HCL) consisting of three α-helices, which is not present in this form in other ligases (Figs. S3 and 4a). The function of this extension is unknown. Interestingly, the protein EhpF, which is involved in the biosynthesis of the antimicrobial compound D-alanylgriseoluteic acid, has a similar homodimeric architecture to HCL and the PaaKs, and it also contains the N-terminal extension (Fig. S11) [24]. EhpF has no C-terminal domain, and its function has not been clarified. Nevertheless, it appears as if the presence of the N-terminal extension is related to the homodimeric ANL core fold. Coenzyme F390 synthetase from *Methanobacterium thermoautotrophicum* Marburg [25] is also an ANL enzyme with a homodimeric non-fused core fold and an N-terminal extension. Interestingly, this enzyme lacks a large part of the C-terminal domain.

To analyze the structural features that distinguish the now predominant canonical ANL enzymes with pseudo-symmetric fused core structures from the homodimeric PCL enzymes (and EhpF), which probably resemble a common ancient ancestor ANL enzyme, we compared ANL enzymes of known

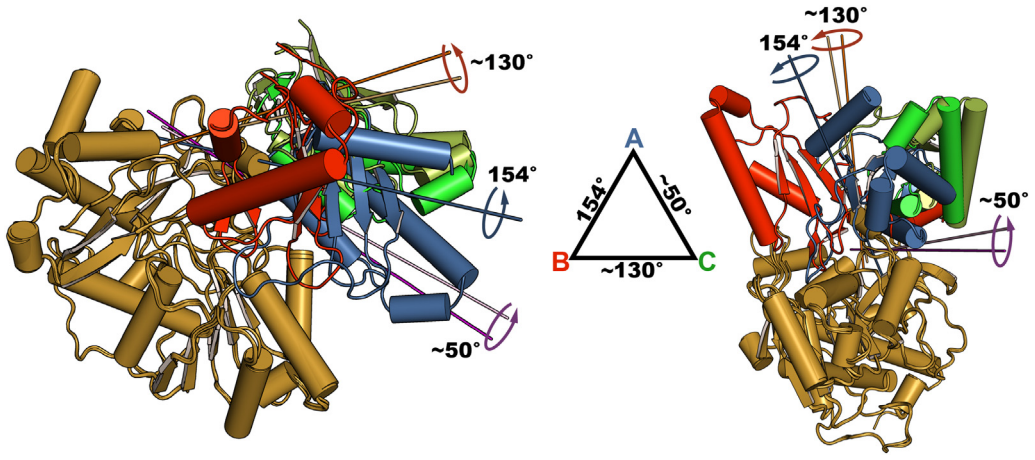

Fig. 6. Domain reorientations between the two catalytic conformations and the open conformation of HCL. The N-terminal domains of all three states have been superimposed and are colored in yellow. The C-terminal domains for the adenylate-forming state (blue), the thioesterification state (red) and the two open conformations (different green colors for the two subunits in the asymmetric unit) have different orientations. The respective rotation axes are labeled by the rotation angle (the triangle defines the rotation angles for the three conformational changes).

structure to HCL (see legend Fig. S4 for further information). In addition to the loss of the N-terminal extension, all analyzed canonical structures exhibited the following main structural differences concerning the secondary structure topology compared to PCL enzymes (Figs. S4 and S5). (1) The N-terminal helix (A′, refer to Fig. S5 for the annotation of secondary structure elements) has a different position in the canonical folds interacting with helices C and C′. (2) In the connection between the two halves (corresponding to the two subunits of PCL enzymes), helix F′ and the complete C-terminal subdomain of sheets b and c are lost. In some canonical ANL enzymes also, strand 7′ is no longer present, and in most enzymes, a short additional helix has formed upstream of the linker that forms the connection to the second half of the pseudo-symmetrical fold (green in Fig. S4B). (3) Helix A is lost. (4) An additional long α-helix has formed between two strands of sheets b and c. Interestingly, this helix has a similar length, position and orientation as the lost helix A of this subunit of the PCL enzymes, but it runs in different N → C-direction. In forming this additional helix in the canonical folds, the main chain passes strand 2 of β-sheet a′, in some enzymes forming a usually short antiparallel β-strand to strand 2, such that this central β-sheet becomes 8-stranded.

Substrate specificity and relationship of HCL to other PCL enzymes

Substrate kinetics as well as phylogenetic and structural analyses clearly define HCL of *A. tertiaricarbonis* L108 as a new subgroup within the PCL family. HCL is highly specific for the aliphatic tertiary OH-group-bearing 2-HIBA. In addition, the phenotype of the mutant strain L108*(Δhcl)*K4 establishes the previously predicted role of HCL in the fuel oxygenate degradation pathway via 2-HIBA to the common metabolite 3-hydroxybutyryl-CoA [14].

Several acid-CoA ligases phylogenetically belonging to the PCL family have been kinetically characterized, such as the enzymes from *Azoarcus evansii* [26], *B. cenocepacia* [12] and *Thermus thermophilus* [27]. In all these cases, short- and medium-chain non-aromatic acids were not transformed into the corresponding CoA thioesters, but preferentially phenylacetic acid and at lower rates often other aromatic monocarboxylic acids, such as benzoic and 4-hydroxyphenylacetic acid, could be used for the thioesterification reaction. Nevertheless, HCL preferentially uses 2-HIBA as acyl substrate, indicating that the PCL family might be much more diverse than assumed thus far, which would also go with its proposed proximity to the origin of the ANL enzymes. In this connection, another PCL subtype has recently been identified in the bacterium *Syntrophorhabdus aromaticivorans* UI, catalyzing the CoA activation of the aromatic dicarboxylic acid isophthalate [28].

The specificity of HCL for 2-HIBA, and to a lesser extent also for other short-chain hydroxylated carboxylic acids, is clearly determined by the architecture of its active site which is significantly smaller and more polar than found in other ligases belonging to the PCL family. The unique H-bond formation between the acyl OH-group and active-site amino acids S239 and Y164 during adenylation seems to be most important for substrate binding, and these residues are well conserved among HCLs (Fig. S12). In line with this, enzyme activity and substrate affinity are dramatically decreased for the substrates isovaleric and pivalic acid because of missing polar OH-groups in the respective acyl moieties.

Despite the evolutionary modification of the acyl binding pocket, structural comparison reveals that the general PCL architecture is still well conserved in HCL, as both homodimeric enzymes build up an intersubunit β-sheet and possess a common N-terminal extension not found in other acid-CoA ligases [12]. Likely, after closer inspection of their substrate kinetics, also other ligases annotated as PCLs may turn out to catalyze CoA activation of aliphatic carboxylic acids. This evolution of PCL enzymes toward short-chain carboxylic acid substrates seems to be quite straightforward, as PCL-dependent degradation of phenylacetic acid and other aromatics is widespread among bacteria and the pathway-encoding genes are present in about 16% of all sequenced bacterial genomes [6]. In addition, paralogous PCL enzymes with divergent substrate specificity, but still preferring aromatic substrates, have been found in some bacterial strains, such as PaaK1 and PaaK2 in *B. cenocepacia* J2315 [12]. In a next evolutionary step, the structure of the aryl substrate pocket of one of these paralogs may adapt toward aliphatic substrates. Thus, in combination with the acquisition of other enzymatic functions, a new degradation pathway might evolve. Another example for this adaptation toward aliphatic acyl residues might be BvaA, which is a PaaK-like ligase supposed to catalyze the synthesis of β-valinyl-CoA in the β-valine-degrading strain *Pseudomonas* sp. DSM 29543 [29]. As outlined in the Introduction section, the domain movement of the ANL enzyme fold might have facilitated the evolutionary development of the diverse substrate specificities described here.

The branched 2-HIBA is rarely found in nature. It is a metabolite in the oxidation of the volatile hydrocarbon isobutene via isobutylene epoxide [30]. In addition, 2-HIBA may be formed in the catabolism of the amino acid valine and is linked to microbial activity of the gut bacterium *Faecalibacterium prausnitzii* [31]. It is used as biomarker for adiposity [32,33] and associated metabolic diseases such as diabetes mellitus [34]. Recently, lysine 2-hydroxyisobutyrylation has also been identified as a posttranslational modification of histones as well as other proteins [35,36], indicating that biosynthesis of this carboxylic acid is more widespread. However, bacterial 2-HIBA metabolism has only been well documented in connection with xenobiotic compounds [15,37], such as the fuel oxygenate MTBE and industrial production of poly(methyl methacrylate). This suggests that the amounts of 2-HIBA formed from the above mentioned natural compounds are negligible and hardly serve as microbial energy and carbon sources. In line with this, BLAST analysis shows that HCL, in contrast to PaaKs, is not widespread among the currently sequenced genomes. Nevertheless, it can be found in several bacterial strains (Fig. S12), while it seems to be absent in the *Archaea* domain. Interestingly, the HCL-bearing bacteria have been isolated from very different habitats, such as MTBE-contaminated groundwater (strain *A. tertiaricarbonis* L108), marine dinoflagellates (strain *Marinobacter algicola* DG893) and geothermal solfatara ponds (strain *K. tusciae* DSM 2912), indicating that HCL and the corresponding 2-HIBA metabolism are present in various ecosystems [17]. This indicates that bacterial 2-HIBA metabolism has obviously evolved earlier than significant anthropogenic sources for it occurred, likely driven by a not-yet-considered biogenic origin of this highly branched carboxylic acid.

CoA binding mode in comparison to other ANL ligases

For the HCL structure trapped in the thioesterification state, CoA binds in a substantially bent conformation. This is in contrast to other ANL ligase structures, where the CoA or acyl-CoA molecules adopt more stretched conformations, for example, in the thioester-forming conformations of bacterial [38] and fungal acetyl-CoA synthetase (PDB ID: 5K85), human medium-chain acyl-CoA synthetase [39] and plant 4-coumarate-CoA ligase [40]. As a result, the CoA binding modes of these enzymes lack the many contacts between the adenine ring and the pantetheine arm in HCL. Furthermore, the adenine ring of CoA in HCL has a different binding site, which overlaps partially with the pantetheine group of CoA in the related enzymes. It has already been pointed out by Reger and co-workers [41] when comparing structures of bacterial 4-chlorobenzoate-CoA ligase and acetyl-CoA synthetases that CoA binding sites are not well conserved among AMP-forming ligases and each ANL subfamily may have a characteristic binding mode resulting in a different anchoring of this ligand. In the case of HCL, the CoA binding mode is even more diverged from the enzymes mentioned above.

ANL domain motion in catalysis

Within the ligase family, the domain rotation enables the catalysis of the two-step reaction of acyl-CoA synthesis (Fig. 1). The active site is quite buried in both catalytic states. Accordingly, ligase enzymes have been captured in the adenylate- and thioester-forming states as well as in so-called open conformations, which may enable substrate binding and product leaving [9]. However, for a single ligase, neither a complete three- nor multi-conformational catalytic cycle has been documented so far. In this study, HCL has been analyzed in the adenylate-forming conformation complexed with acyl-AMP after incubating with ATP and carboxylic acid substrates. In addition, the thioester-forming and an open conformation were obtained in the presence of 2-hydroxyisobutyryl-CoA

and AMP. With the HCL-related PCLs PaaK1 and PaaK2, on the other hand, crystallization with ATP and phenylacetic acid resulted either in the adenylate-forming state (PaaK1) or in the thioester-forming conformation (PaaK2) [12]. Moreover, in the case of the PCL Bt_0428, both catalytically active conformers have been captured even in the same asymmetric unit when incubating the enzyme with AMP and CoA in the absence of carboxylic acid substrates. With these enzymes, however, open or intermediate conformations were not obtained.

We compared the open conformation of HCL to other conformers of ANL enzymes (PDB IDs: 3L8C, 2VSQ, 4FUQ, 1ULT, 3R44, 3A9U, 3FCC, 3IPL, 3LNV, 3KXW, 3NYR, 3NYQ, and 2V7B, as described in chapter 5 of Ref. [13]), which differ significantly from the two catalytically reactive conformers. The most similar conformers are those of 1ULT chain B (33.7° deviation of the C-terminal domain orientation, long chain fatty acyl-CoA synthetase from *T. thermophilus* HB8 [42]) and of 4FUQ chain C (32.7°, methylmalonyl-CoA synthetase MatB from *Rhodopseudomonas palustris* CGA009 [43]) (Fig. S13). In conclusion, binding and release of the different substrates and products in the catalytic cycle of carboxylic acid-CoA ligases may proceed through one or even more stable open conformations.

Materials and Methods

Chemicals, strains and cultivation conditions

Purity and supply sources of chemicals used in this study are listed in the supporting information. CoA thioesters used as standards in HPLC analysis were either prepared from their anhydrides [44] or from the free acids via thiophenyl thioesters [45]. Synthesized CoA thioesters were analyzed by ESI-MS/MS as previously described [14]. *E. coli* ArcticExpress (DE3) (Novagen) was cultivated in LB medium containing 20 mg L^{-1} gentamycin. The bacterial strain *A. tertiaricarbonis* L108 previously isolated from an MTBE-contaminated aquifer at Leuna, Germany [46], was routinely cultivated in a liquid mineral salt medium [14] containing MTBE at a concentration of 0.3 g L^{-1}. An *hcl* knockout mutant strain L108(*Δhcl*)K4 was obtained by using the EZ-Tn5 < KAN-2 > Tnp transposome kit (Epicenter Biotechnologies) as previously described [22]. Briefly, mid-exponential bacterial cells of *A. tertiaricarbonis* L108 were transformed by electroporation and rescued in mineral salt medium amended with 10 mM fructose. Then, dilutions were plated on fructose agar containing 50 µg mL^{-1} kanamycin, and all colonies formed were analyzed for loss of their capability to grow on 2-HIBA. Strain K4 completely failed to grow on 2-HIBA. Direct sequencing of the transposon integration site revealed that the *hcl* gene was disrupted in this strain (Fig. S1).

Resting-cell experiments with mutant strain L108(*Δhcl*)K4

For resting cell experiments, cultures of strain L108(*Δhcl*)K4 were pre-grown on *tert*-amyl alcohol in close glass bottles in 1 L culture medium at 30 °C and harvested by centrifugation applying 13,000*g* at 4 °C for 10 min. After washing twice and adjusting cell concentration to 1.4 g biomass (dry weight) per liter by diluting with mineral salt medium, cells were immediately used as inoculum for resting-cell experiments on TBA. Biomass was followed by measuring the optical density at 700 nm, using a multiplication factor of 0.54 for calculating dry biomass in g per liter [47]. TBA was quantified by headspace GC using flame ionization detection on a GC system (Agilent Technologies, USA) [14]. 2-Methyl-1,2-propandiol and 2-HIBA were quantified using ion exchange HPLC with refractive index detection on an HPLC system (Shimadzu Corp., USA) [47] applying an eluent of 0.005 M sulfuric acid at 0.6 mL min^{-1} and a Nucleogel ION 300 OA column (300 × 7.7 mm, Macherey-Nagel, Switzerland). The data shown in this study represent the mean values and SD of five replicate experiments.

Expression of *hcl* and protein purification

The *hcl* gene was amplified from genomic DNA of strain L108 by applying the forward primer 5′-AGCGGCTCTT CAATGGAAGAGTGGAACTTTCCG-3′ and reverse primer 5′-AGCGGCTCTTCTCCCGGCGCTCGATTG CAGTTG-3′. PCR was accomplished with *Pfu* DNA polymerase (Promega) for 30 cycles, including denaturation at 94 °C for 1 min, annealing at 55 °C for 1 min, and extension at 72 °C for 2 min. The PCR product was cloned into expression vector pASG-IBA43 (IBA Goettingen), which was used for transformation of *E. coli* ArcticExpress (DE3) at 300 mV for 4 ms in chilled 0.1-cm cuvettes in a MicroPulser (Bio-Rad). After growth at 30 °C in LB medium containing 20 mg L^{-1} gentamycin and 100 mg L^{-1} ampicillin to an OD_{550} of 0.35 and further incubation at 12 °C to an OD_{550} of 0.5, induction was performed for 24 h at 12 °C after addition of 200 µg L^{-1} anhydrotetracycline. Then, cells were centrifuged, suspended in Tris buffer and disrupted with glass beads (212 to 300 µm, Sigma) using a mixer mill (MM 400, Retsch GmbH, Germany) at 30 s^{-1} for 30 min. The recombinant protein was purified with the help of its His tag. Crude extract of *E. coli* ArcticExpress (DE3) pASG-IBA43 [*hcl*] were loaded on a nickel-nitrilotriacetic acid Superflow 10-ml column (IBA Goettingen). After washing with 20 column volumes of imidazole buffer (50 mM sodium phosphate, 300 mM NaCl, 20 mM imidazole, pH 8.0), HCL was eluted with the same buffer containing 250 mM imidazole. For determination of

kinetic parameters, fractions containing HCL were concentrated via Viva spin columns (30 kDa, GE Healthcare) and diluted with conservation buffer (50 mM potassium phosphate, 10% glycerol, pH 7.4). For crystallization, the protein was further purified by size-exclusion chromatography using a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) in 20 mM Tris, 150 mM NaCl and 5% glycerol (pH 7.5) as buffer (protein storage buffer) and concentrated to 10 mg mL^{-1} afterward.

Acid-CoA ligase assay

Activity of purified HCL was measured by directly quantifying CoA thioester formation from carboxylic acid substrates employing HPLC ion-pair chromatography. For determining pH and temperature optimum, 2-hydroxyisobutyryl-CoA formation activity was measured incubating recombinant HCL in 50 mM potassium phosphate buffer, amended with 3 mM ATP, 5 mM CoA, 5 mM $MgCl_2$ and 10% glycerol at pH values between 5.8 and 8.6 and at temperatures between 15 and 45 °C. For determining kinetic parameters, 2 μM of recombinant HCL was incubated in 50 mM phosphate buffer adjusted to pH 7.8 and amended with 5 mM $MgCl_2$ and 10% glycerol at 35 °C. As K_m values for ATP and CoA with 2-HIBA as carboxylic acid substrate (Fig. S14) were 270 ± 20 and 410 ± 20 μM, respectively, ligase activities were routinely determined in the presence of 3 mM ATP and 5 mM CoA. The reaction was started by adding carboxylic acid substrate. To quench the reaction, samples were mixed with an equal volume of 100 mM sodium acetate buffer (pH 3.5) and heated at 60 °C for 5 min. Free CoA and CoA thioesters were analyzed by HPLC (Shimadzu Corp., USA) by photometric detection at 260 nm applying a Nucleosil 100–5 C18 column (250 × 3 mm, 5 μm, Macherey-Nagel, Switzerland) and a mobile phase of 20 vol% acetonitrile, 10 mM tetrabutylammonium hydrogen sulfate and 100 mM sodium phosphate at pH 4.5 [14]. Eluent flow was adjusted to 0.8 mL min^{-1} and column oven temperature was 35 °C. Detection limit at 260 nm was 0.5 μM CoA thioester. Kinetic parameters were calculated by nonlinear regression analysis applying the Michaelis–Menten equation (OriginPro 9.0).

Crystallization and structure determination

For the crystallization of the adenylate-forming state, HCL was cocrystallized with 3 mM ATP and various substrates (2-HIBA, 3-HIBA, *S*3-HBA) at 5 mM concentration in 2.1 M lithium sulfate, 10 mM magnesium chloride and 50 mM MES pH 6.0). For obtaining the thioesterification conformation, HCL was cocrystallized with 5 mM AMP and 5 mM 2-hydroxyisobutyryl-CoA in 24% PEG 3350 and 100 mM Tris (pH 8.0). Crystals were cryoprotected using 20% glycerol. Diffraction data were collected using synchrotron radiation (beamline BL14.1 at HZB BESSY II in Berlin, Germany). Data processing was accomplished using XDS [48] for integration and SCALA [49] for scaling within the CCP4 program package. The structure was solved by molecular replacement using the program Molrep [50] and PDB ID: 2Y4N as a search model. Model building was performed using Coot [51], and the structure was refined with BUSTER [52]. The programs MolProbity [53] and Procheck [54] were used to evaluate the final protein model and PyMOL [55] for visualization of the protein structure. Details of data collection and refinement are summarized in Table 1.

Domain orientations of HCL were analyzed by aligning first the N-terminal (M1 to D361) and then the C-terminal domains (M362 to S467) using the program LSQKAB [56]. For comparison of crystal structures from different enzymes, domains rotating as rigid bodies were determined and aligned with the RAPIDO-server [57]. Subsequently, again first the rigid bodies of the N-terminal and then of the C-terminal domain were superposed. Comparison of protein architectures of HCL and PaaK and calculation of rmsd values were performed with the Dali server [58]. Sequences of HCL and related enzymes have been aligned using Clustal Omega [59] and ESPript [60].

Acknowledgments

We would like to thank C. Dilßner (UFZ) and M. Neytschev (UFZ) for excellent technical assistance. In addition, the authors thank the Joint Berlin MX Laboratory at the Helmholtz Zentrum Berlin (Bessy II) for beam time and assistance during synchrotron data collection, as well as for traveling support. We thank Clemens Vonrhein, Andrew Sharff and Gerard Bricogne of the Buster Development Team for help in refinement of the ligand occupancies.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2019.05.027.

Received 21 October 2018;
Received in revised form 14 May 2019;
Available online 28 May 2019

Keywords:
X-ray crystal structure;
catalytic cycle;
domain alternation;
domain rotation;
ANL superfamily

†M.Z. and N.K.-Y. contributed equally to this work.

Abbreviations used:

ANL, superfamily of adenylating enzymes comprising *A*cyl-CoA synthetases, the *N*RPS (non-ribosomal peptide synthetases) adenylation domains, and the *L*uciferase enzymes; PCL, phenylacetate-CoA ligase; 2-HIBA, 2-hydroxyisobutyric acid; MTBE, methyl *tert*-butyl ether; HCL, 2-HIBA-CoA ligase; TBA, *tert*-butyl alcohol; 3-HIBA, 3-hydroxyisobutyryic acid; *S*3-HBA, (*S*)-3-hydroxybutyric acid; 2-HBA, 2-hydroxybutyric acid; PPi, inorganic pyrophosphate.

References

[1] D.H. Kwan, F. Schulz, The stereochemistry of complex polyketide biosynthesis by modular polyketide synthases, Molecules 16 (2011) 6092–6115.

[2] T. Weber, M.A. Marahiel, Exploring the domain structure of modular nonribosomal peptide synthetases, Structure 9 (2001) R3–R9.

[3] E.J. Drake, B.R. Miller, C. Shi, J.T. Tarrasch, J.A. Sundlov, C. L. Allen, G. Skiniotis, C.C. Aldrich, A.M. Gulick, Structures of two distinct conformations of holo-non-ribosomal peptide synthetases, Nature 529 (2016) 235–238.

[4] R. Banerjee, Radical carbon skeleton rearrangements: catalysis by coenzyme B_{12}-dependent mutases, Chem. Rev. 103 (2003) 2083–2094.

[5] P. Estrada, M. Manandhar, S.H. Dong, J. Deveryshetty, V. Agarwal, J.E. Cronan, S.K. Nair, The pimeloyl-CoA synthetase BioW defines a new fold for adenylate-forming enzymes, Nat. Chem. Biol. 13 (2017) 668–674.

[6] R. Teufel, V. Mascaraque, W. Ismail, M. Voss, J. Perera, W. Eisenreich, W. Haehnel, G. Fuchs, Bacterial phenylalanine and phenylacetate catabolic pathway revealed, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 14390–14395.

[7] T.P. Begley, C. Kinsland, E. Strauss, The biosynthesis of coenzyme A in bacteria, Vitam. Horm. 61 (2001) 157–171.

[8] S. Schmelz, J.H. Naismith, Adenylate-forming enzymes, Curr. Opin. Struct. Biol. 19 (2009) 666–671.

[9] A.M. Gulick, Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase, ACS Chem. Biol. 4 (2009) 811–827.

[10] P. Khurana, R.S. Gokhale, D. Mohanty, Genome scale prediction of substrate specificity for acyl adenylate superfamily of enzymes based on active site residue profiles, BMC Bioinformatics. 11 (2010) 57.

[11] M.A. Marahiel, T. Stachelhaus, H.D. Mootz, Modular peptide synthetases involved in nonribosomal peptide synthesis, Chem. Rev. 97 (1997) 2651–2674.

[12] A. Law, M.J. Boulanger, Defining a structural and kinetic rationale for paralogous copies of phenylacetate-CoA ligases from the cystic fibrosis pathogen *Burkholderia cenocepacia* J2315, J. Biol. Chem. 286 (2011) 15577–15585.

[13] C.A. Mitchell, Structural, Functional, and Computational Insights into the ANL Superfamily of Enzymes, Dissertation State University of New York ProQuest LLC, Ann Arbor, 2013.

[14] N. Yaneva, J. Schuster, F. Schäfer, V. Lede, D. Przybylski, T. Paproth, H. Harms, R.H. Müller, T. Rohwerder, Bacterial acyl-CoA mutase specifically catalyzes coenzyme B_{12}-dependent isomerization of 2-hydroxyisobutyryl-CoA and (*S*)-3-hydroxybutyryl-CoA, J. Biol. Chem. 287 (2012) 15502–15511.

[15] T. Rohwerder, U. Breuer, D. Benndorf, U. Lechner, R.H. Müller, The alkyl *tert*-butyl ether intermediate 2-hydroxyisobutyrate is degraded via a novel cobalamin-dependent mutase pathway, Appl. Environ. Microbiol. 72 (2006) 4128–4135.

[16] N. Kurteva-Yaneva, M. Zahn, M.T. Weichler, R. Starke, H. Harms, R.H. Müller, N. Sträter, T. Rohwerder, Structural basis of the stereospecificity of bacterial B_{12}-dependent 2-hydroxyisobutyryl-CoA mutase, J. Biol. Chem. 290 (2015) 9727–9737.

[17] M.T. Weichler, N. Kurteva-Yaneva, D. Przybylski, J. Schuster, R.H. Müller, H. Harms, T. Rohwerder, Thermophilic coenzyme B_{12}-dependent acyl coenzyme A (CoA) mutase from *Kyrpidia tusciae* DSM 2912 preferentially catalyzes isomerization of (*R*)-3-hydroxybutyryl-CoA and 2-hydroxyisobutyryl-CoA, Appl. Environ. Microbiol. 81 (2015) 4564–4572.

[18] M.T. Rohde, S. Tischer, H. Harms, T. Rohwerder, Production of 2-hydroxyisobutyric acid from methanol by *Methylobacterium extorquens* AM1 expressing (*R*)-3-hydroxybutyryl coenzyme A-isomerizing enzymes, Appl. Environ. Microbiol. 83 (2017) e02622-16.

[19] A.R. Horswill, J.C. Escalante-Semerena, Characterization of the propionyl-CoA synthetase (PrpE) enzyme of *Salmonella enterica*: residue Lys592 is required for propionyl-AMP synthesis, Biochemistry. 41 (2002) 2379–2387.

[20] R.M. Morgan-Kiss, J.E. Cronan, The *Escherichia coli fadK* (*ydiD*) gene encodes an anerobically regulated short chain acyl-CoA synthetase, J. Biol. Chem. 279 (2004) 37324–37333.

[21] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.

[22] J. Schuster, F. Schäfer, N. Hübler, A. Brandt, M. Rosell, C. Härtig, H. Harms, R.H. Müller, T. Rohwerder, Bacterial degradation of *tert*-amyl alcohol proceeds via hemiterpene 2-methyl-3-buten-2-ol by employing the tertiary alcohol desaturase function of the Rieske nonheme mononuclear iron oxygenase MdpJ, J. Bacteriol. 194 (2012) 972–981.

[23] E. Conti, N.P. Franks, P. Brick, Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes, Structure 4 (1996) 287–298.

[24] A.K. Bera, V. Atanasova, S. Gamage, H. Robinson, J.F. Parsons, Structure of the D-alanylgriseoluteic acid biosynthetic protein EhpF, an atypical member of the ANL superfamily of adenylating enzymes, Acta Crystallogr D Biol Crystallogr 66 (2010) 664–672.

[25] P. Vermeij, R.J. van der Steen, J.T. Keltjens, G.D. Vogels, T. Leisinger, Coenzyme F390 synthetase from *Methanobacterium thermoautotrophicum* Marburg belongs to the superfamily of adenylate-forming enzymes, J. Bacteriol. 178 (1996) 505–510.

[26] M. El-Said Mohamed, Biochemical and molecular characterization of phenylacetate-coenzyme A ligase, an enzyme catalyzing the first step in aerobic metabolism of phenylacetic acid in *Azoarcus evansii*, J. Bacteriol. 182 (2000) 286–294.

[27] T.J. Erb, W. Ismail, G. Fuchs, Phenylacetate metabolism in thermophiles: characterization of phenylacetate-CoA ligase, the initial enzyme of the hybrid pathway in *Thermus thermophilus*, Curr. Microbiol. 57 (2008) 27–32.

[28] M. Junghare, D. Spiteller, B. Schink, Anaerobic degradation of xenobiotic isophthalate by the fermenting bacterium

Syntrophorhabdus aromaticivorans, ISME J. 13 (2019) 1252–1268.
[29] M. Otzen, C.G. Crismaru, C.P. Postema, H.J. Wijma, M.M. Heberling, W. Szymanski, S. de Wildeman, D.B. Janssen, Metabolism of β-valine via a CoA-dependent ammonia lyase pathway, Appl. Microbiol. Biotechnol. 99 (2015) 8987–8998.
[30] S. Kottegoda, E. Waligora, M. Hyman, Metabolism of 2-methylpropene (isobutylene) by the aerobic bacterium *Mycobacterium* sp. strain ELW1, Appl. Environ. Microbiol. 81 (2015) 1966–1976.
[31] M. Li, B. Wang, M. Zhang, M. Rantalainen, S. Wang, H. Zhou, Y. Zhang, J. Shen, X. Pang, M. Zhang, H. Wei, Y. Chen, H. Lu, J. Zuo, M. Su, Y. Qiu, W. Jia, C. Xiao, L.M. Smith, S. Yang, E. Holmes, H. Tang, G. Zhao, J.K. Nicholson, L. Li, L. Zhao, Symbiotic gut microbes modulate human metabolic phenotypes, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 2117–2122.
[32] R. Calvani, A. Miccheli, G. Capuani, A. Tomassini Miccheli, C. Puccetti, M. Delfini, A. Iaconelli, G. Nanni, G. Mingrone, Gut microbiome-derived metabolites characterize a peculiar obese urinary metabotype, Int. J. Obes. 34 (2010) 1095–1098.
[33] P. Elliott, J.M. Posma, Q. Chan, I. Garcia-Perez, A. Wijeyesekera, M. Bictash, T.M. Ebbels, H. Ueshima, L. Zhao, L. van Horn, M. Daviglus, J. Stamler, E. Holmes, J.K. Nicholson, Urinary metabolic signatures of human adiposity, Sci. Transl. Med. 7 (2015) 285ra62.
[34] X. Li, Z. Xu, X. Lu, X. Yang, P. Yin, H. Kong, Y. Yu, G. Xu, Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabonomics: biomarker discovery for diabetes mellitus, Anal. Chim. Acta 633 (2009) 257–262.
[35] L. Dai, C. Peng, E. Montellier, Z. Lu, Y. Chen, H. Ishii, A. Debernardi, T. Buchou, S. Rousseaux, F. Jin, B.R. Sabari, Z. Deng, C.D. Allis, B. Ren, S. Khochbin, Y. Zhao, Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark, Nat. Chem. Biol. 10 (2014) 365–370.
[36] J. Huang, Z. Luo, W. Ying, Q. Cao, H. Huang, J. Dong, Q. Wu, Y. Zhao, X. Qian, J. Dai, 2-Hydroxyisobutyrylation on histone H4K8 is regulated by glucose homeostasis in *Saccharomyces cerevisiae*, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) 8782–8787.
[37] A. François, H. Mathis, D. Godefroy, P. Piveteau, F. Fayolle, F. Monot, Biodegradation of methyl *tert*-butyl ether and other fuel oxygenates by a new strain, *Mycobacterium austroafricanum* IFP 2012, Appl. Environ. Microbiol. 68 (2002) 2754–2762.
[38] A.M. Gulick, V.J. Starai, A.R. Horswill, K.M. Homick, J.C. Escalante-Semerena, The 1.75 Å crystal structure of acetyl-CoA synthetase bound to adenosine-5′-propylphosphate and coenzyme A, Biochemistry 42 (2003) 2866–2873.
[39] G. Kochan, E.S. Pilka, F. von Delft, U. Oppermann, W.W. Yue, Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase ACSM2A, J. Mol. Biol. 388 (2009) 997–1008.
[40] Z. Li, S.K. Nair, Structural basis for specificity and flexibility in a plant 4-coumarate:CoA ligase, Structure 23 (2015) 2032–2042.
[41] A.S. Reger, R. Wu, D. Dunaway-Mariano, A.M. Gulick, Structural characterization of a 140 degrees domain movement in the two-step reaction catalyzed by 4-chlorobenzoate: CoA ligase, Biochemistry 47 (2008) 8016–8025.
[42] Y. Hisanaga, H. Ago, N. Nakagawa, K. Hamada, K. Ida, M. Yamamoto, T. Hori, Y. Arii, M. Sugahara, S. Kuramitsu, S. Yokoyama, M. Miyano, Structural basis of the substrate-specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer, J. Biol. Chem. 279 (2004) 31717–31726.
[43] H.A. Crosby, K.C. Rank, I. Rayment, J.C. Escalante-Semerena, Structure-guided expansion of the substrate range of methylmalonyl coenzyme A synthetase (MatB) of *Rhodopseudomonas palustris*, Appl. Environ. Microbiol. 78 (2012) 6619–6629.
[44] E.J. Simon, D. Shemin, The preparation of *S*-succinyl coenzyme A, J. Am. Chem. Soc. 75 (1953) 2520.
[45] R. Padmakumar, S. Gantla, R. Banerjee, A rapid method for the synthesis of methylmalonyl-coenzyme A and other CoA-esters, Anal. Biochem. 214 (1993) 318–320.
[46] U. Lechner, D. Brodkorb, R. Geyer, G. Hause, C. Härtig, G. Auling, F. Fayolle-Guichard, P. Piveteau, R.H. Müller, T. Rohwerder, *Aquincola tertiaricarbonis* gen. nov., sp. nov., a tertiary butyl moiety-degrading bacterium, Int. J. Syst. Evol. Microbiol. 57 (2007) 1295–1303.
[47] R.H. Müller, T. Rohwerder, H. Harms, Degradation of fuel oxygenates and their main intermediates by *Aquincola tertiaricarbonis* L108, Microbiology 154 (2008) 1414–1421.
[48] W. Kabsch, XDS, Acta Crystallogr D Biol Crystallogr 66 (2010) 125–132.
[49] The Ccp4 group, The CCP4 suite: programs for protein crystallography, Acta Crystallogr D Biol Crystallogr 50 (1994) 760–763.
[50] A. Vagin, A. Teplyakov, Molecular replacement with MOLREP, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 22–25.
[51] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr D Biol Crystallogr 60 (2004) 2126–2132.
[52] E. Blanc, P. Roversi, C. Vonrhein, C. Flensburg, S.M. Lea, G. Bricogne, Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT, Acta Crystallogr D Biol Crystallogr 60 (2004) 2210–2221.
[53] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, Acta Crystallogr D Biol Crystallogr 66 (2010) 12–21.
[54] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[55] W.L. DeLano, The PyMOL Molecular Graphics System, version 1.7.2, Schrödinger, LLC, New York, 2014.
[56] W. Kabsch, A solution for the best rotation to relate two sets of vectors, Acta Cryst A32 (1976) 922–923.
[57] R. Mosca, T.R. Schneider, RAPIDO: a web server for the alignment of protein structures in the presence of conformational changes, Nucleic Acids Res. 36 (2008) W42–W46.
[58] L. Holm, P. Rosenström, Dali server: conservation mapping in 3D, Nucleic Acids Res. 38 (2010) W545–W549.
[59] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D. G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. 7 (2011) 539.
[60] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, Nucl Acids Res 42 (2014) W320–W324.