Crystal structure of *cis*-aconitate decarboxylase reveals the impact of naturally occurring human mutations on itaconate synthesis

Fangfang Chen[a,b,c,1], Peer Lukat[a,1], Azeem Ahmed Iqbal[b,c], Kyrill Saile[a], Volkhard Kaever[d], Joop van den Heuvel[a], Wulf Blankenfeldt[a,e], Konrad Büssow[a,2,3], and Frank Pessler[b,c,f,2,3]

[a]Department Structure and Function of Proteins, Helmholtz Centre for Infection Research, 38124 Braunschweig, Germany; [b]Research Group Biomarkers for Infectious Diseases, TWINCORE Centre for Experimental and Clinical Infection Research, 30625 Hannover, Germany; [c]Research Group Biomarkers for Infectious Diseases, Helmholtz Centre for Infection Research, 38124 Braunschweig, Germany; [d]Research Core Unit Metabolomics, Hannover Medical School, 30625 Hannover, Germany; [e]Institute for Biochemistry, Biotechnology and Bioinformatics, Technische Universität Braunschweig, 38106 Braunschweig, Germany; and [f]Centre for Individualised Infection Medicine, 30625 Hannover, Germany

Edited by Philippa Marrack, National Jewish Health, Denver, CO, and approved September 3, 2019 (received for review May 24, 2019)

***cis*-Aconitate decarboxylase (CAD, also known as ACOD1 or Irg1) converts *cis*-aconitate to itaconate and plays central roles in linking innate immunity with metabolism and in the biotechnological production of itaconic acid by *Aspergillus terreus*. We have elucidated the crystal structures of human and murine CADs and compared their enzymological properties to CAD from *A. terreus*. Recombinant CAD is fully active in vitro without a cofactor. Murine CAD has the highest catalytic activity, whereas *Aspergillus* CAD is best adapted to a more acidic pH. CAD is not homologous to any known decarboxylase and appears to have evolved from prokaryotic enzymes that bind negatively charged substrates. CADs are homodimers, the active center is located in the interface between 2 distinct subdomains, and structural modeling revealed conservation in zebrafish and *Aspergillus*. We identified 8 active-site residues critical for CAD function and rare naturally occurring human mutations in the active site that abolished CAD activity, as well as a variant (Asn152Ser) that increased CAD activity and is common (allele frequency 20%) in African ethnicity. These results open the way for 1) assessing the potential impact of human CAD variants on disease risk at the population level, 2) developing therapeutic interventions to modify CAD activity, and 3) improving CAD efficiency for biotechnological production of itaconic acid.**

macrophage | enzymology | itaconic acid | *cis*-aconitate | decarboxylase

Itaconic acid was first described in 1836 as an unknown by-product of heating citric acid (1) and subsequently given its name by Crasso (2), but its relevance to the human condition remained elusive until it was found that it can serve as a highly reactive building block for polymer synthesis (3) and can be produced by fermentation in amounts large enough for industrial utilization (4, 5). Recently, itaconic acid moved into the limelight of biomedical research when it was detected during activation of murine macrophages (6, 7) and, subsequently, *cis*-aconitate decarboxylase (CAD; encoded by the *ACOD1* gene) was identified as the enzyme responsible for its formation from the tricarboxylic acid cycle intermediate *cis*-aconitate (8) (Fig. 1). Since then, an astoundingly broad spectrum of itaconic acid functions in immune responses and host–pathogen interactions have been described. These include linking cell metabolism and innate immunity (9–11) (reviewed in ref. 12), limiting inflammatory responses particularly by dampening IFN signaling (13–15), reducing host susceptibility to bacterial (16, 17) and viral (18) infections at the cellular and organismal level, and directly inhibiting growth of bacteria that use the glyoxylate shunt as energy source (8). In addition, itaconic acid appears to exert cytoprotective effects, for example, by limiting synthesis of reactive oxygen species under certain conditions (19, 20). Furthermore, it has been reported to be a growth factor for peritoneal tumors (21), suggesting that CAD inhibitors could potentially be used as antineoplastic agents.

The need for itaconic acid for the synthesis of polymers (e.g., lubricants, thickeners, rubber, resins, and coatings) continues to increase. It is included in the US Department of Energy list of the top 12 biobased (i.e., renewable) chemicals (22), and industry depends on itaconic acid production by fermentation by the fungus *Aspergillus terreus*, which naturally expresses a CAD enzyme with homology to the CAD found in higher organisms (reviewed in refs. 23 and 24). Considering this high importance of itaconic acid to medicine, biology, and biotechnology alike, it is critical to understand the mechanism of its enzymatic synthesis and the functionally important residues in CADs in order to 1) modulate itaconic acid synthesis for therapeutic purposes, 2) assess the possible impact of sequence variants in the *ACOD1* gene upon human disease risk, and 3) genetically engineer improved CADs for use in biotechnology. In addition, considering

Significance

Itaconic acid was first described in the 19th century and was later appreciated mostly as a fungal metabolite of interest to polymer synthesis. Its surprising recent discovery as a key metabolite during the activation of inflammatory macrophages led to the recognition that it plays critical roles in linking metabolism and innate immunity. However, the lack of a crystal structure of *cis*-aconitate decarboxylase (CAD, the enzyme that synthesizes itaconate) has made it impossible to address many questions central to the chemistry, biology, evolution, and medical importance of itaconic acid synthesis. We have now determined the crystal structure of CAD and have identified amino acids that make up the active center, as well as human mutations with strong effects on itaconate synthesis.

Author contributions: W.B., K.B., and F.P. designed research; F.C., P.L., A.A.I., K.S., V.K., J.v.d.H., W.B., and K.B. performed research; F.C., P.L., J.v.d.H., W.B., K.B., and F.P. analyzed data; and P.L., K.B., and F.P. wrote the paper.

The authors declare no competing interests.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 6R6T and 6R6U). Plasmids pCAD16, pCAD29, and pCAD39 and the Asn152Ser, Arg273His, and His103Ala mutants of pCAD29 and human Irg1 pCMV6Entry have been deposited at Addgene, www.addgene.com (ID codes 124842, 124843, 124873–124879).

[1]F.C. and P.L. contributed equally to this work.

[2]K.B. and F.P. contributed equally to this work.

[3]To whom correspondence may be addressed. Email: konrad.buessow@helmholtz-hzi.de or frank.pessler@twincore.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1908770116/-/DCSupplemental.

First published September 23, 2019.

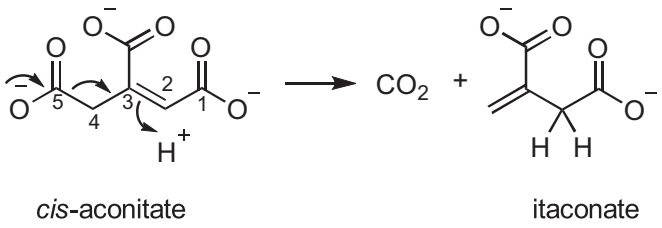


Fig. 1. The formation of itaconate catalyzed by CAD. Bentley and Thiessen (43) showed that CAD removes the C5 carboxyl group and that a proton from the solvent is added to position C2.

its broad conservation from fungi to humans, understanding changes in CAD structure across evolution might improve our understanding of how the structure of such a critical enzyme evolves in order to ensure adaptation to the changing metabolic environment during the evolution from fungi to humans. However, until now the structural basis of itaconic acid synthesis from *cis*-aconitate (Fig. 1) has remained elusive. We have, therefore, determined the crystal structures of human and mouse CAD, identified amino acid residues required for function of the active center, and have begun to assess distribution and impact of naturally occurring variants of the human *ACOD1* gene on itaconic acid synthesis.

Results

Recombinant CADs Form Homodimers. We designed plasmids for recombinant production of CAD from humans (hCAD), mouse (mCAD), and *A. terreus* (aCAD) in *Escherichia coli* and High Five insect cells (BTI-Tn-5B1-4) (25). To this end, structurally disordered protein regions were predicted with the DISOPRED server (26) and then truncated in the constructs for expression in *E. coli.* aCAD was truncated to residues 12 to 490, hCAD to residues 4 to 461, and mCAD to residues 4 to 462. Truncated mCAD expressed in *E. coli* was purified by affinity chromatography but could not be used for crystallization due to formation of aggregates. We found that adding 10% (vol/vol) glycerol to the buffers prevented aggregation. All 3 CADs were thus expressed in *E. coli* and purified by affinity chromatography and gel filtration in the presence of 10% glycerol (*SI Appendix*, Fig. S1 *A* and *B*). Molar masses of the purified proteins were measured by multiangle light scattering (MALS), and it was found that all 3 proteins are homodimers in solution (*SI Appendix*, Fig. S1*C*).

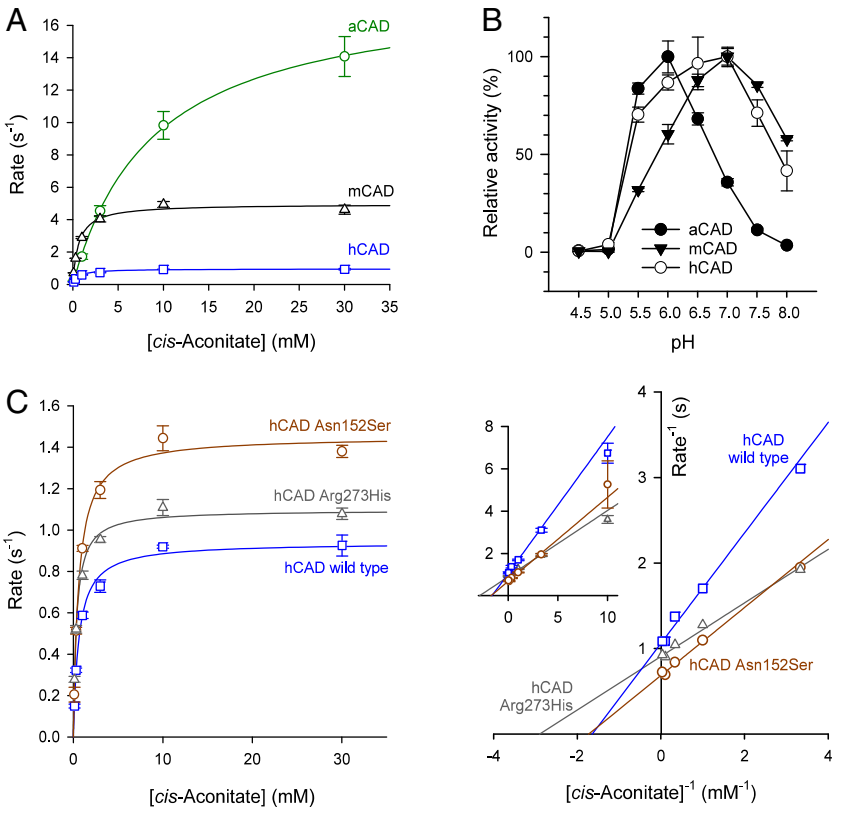
Differences in Enzymatic Activity between Murine, Human, and *Aspergillus* CAD. Previous reports demonstrated that itaconic acid synthesis is substantially higher in murine than in human macrophages upon activation (8). To test whether this is due to higher enzymatic activity of mCAD, Michaelis constants (K_M) and catalytic rate constants (k_{cat}) of the 3 enzymes were determined at pH 6.5 and 37 °C (Fig. 2*A*), and the effect of pH on enzyme activity was tested (Fig. 2*B*). Murine and hCAD had similar K_M values and a similar pH optimum around 7.0. However, the catalytic rate of mCAD was more than 3 times higher, suggesting that higher itaconic acid synthesis in murine macrophages is due to higher turnover of *cis*-aconitate by mCAD. In comparison, aCAD had much higher K_M and k_{cat} values and a markedly lower pH optimum. Taken together, these results suggest 1) that human and mouse CAD are best adapted to physiological pH values found in mammals, whereas the activity of aCAD would increase upon intracellular acidification of the fungus, and 2) that the mouse enzyme (even though its k_{cat} is much lower than that of aCAD) is best adapted to low substrate concentrations.

Crystal Structure of hCAD. The crystal structure of hCAD (comprising residues 4 to 461) was determined by molecular replacement with the structure of iminodisuccinate epimerase (PDB ID code 2HP3, 25% sequence identity) (27) and refined at a resolution of 1.7 Å (Fig. 3*A* and Table 1). Consistent with the results of MALS (*SI Appendix*, Fig. S1*C*), the protein forms a homodimer of ∼70 Å × 70 Å × 100 Å. Each hCAD monomer folds into 2 distinct domains. The larger one consists of residues 1 to 267 and 413 to 462 and is purely α-helical. A smaller domain comprising residues 273 to 410 is inserted into this domain and folds into an antiparallel β-sheet consisting of 4 strands and 3 helical elements. The dimer interface of ∼2,000 Å^2, as calculated by PDBe PISA (28), is located exclusively on the larger domain, opposite to the smaller domain. The putative location of the active site was mapped onto the protein by comparison with the crystal structure of *Agrobacterium tumefaciens* iminodisuccinate (IDS) epimerase, the closest homolog in the Protein Data Bank (PDB ID code 2HP3) (27) with a sequence identity of 25% (Fig. 4). The residues supposed to be important for catalysis are located in the interface between the 2 domains.

As *cis*-aconitate contains 3 carboxyl groups, residues with polar/charged groups in this region were considered to be likely involved in substrate binding or catalysis. Eight potentially important residues were identified, 5 in the large domain (Asp93, Thr97, His103, His159, Lys207) and 3 in the small domain (Lys272, His277, Tyr318). A residual electron density in the cavity (marked green in Fig. 3*A*) between these residues supports this notion. Although it could not be modeled, this electron density can probably be attributed to the use of citrate as precipitant in crystallization, which is structurally similar to *cis*-aconitate as it also contains 3 carboxylate moieties. So far, it has not been possible to obtain the crystal structure of *cis*-aconitate– or itaconate-bound hCAD, neither by ligand-soaking nor by cocrystallization. As the entrance to the putative active site is closed in the crystal structure, it can be assumed that opening of the active site by movement of the small domain would be required for substrate binding. This is corroborated by crystal structures of 3 distantly related bacterial 2-methylcitrate dehydratases (19 to 21% sequence identity with hCAD, but same fold), which have been crystallized in a closed (PDB ID codes 5MUX and 5MVI) but also in an open conformation (PDB ID code 1SZQ). Structural superposition of hCAD with the open form of this enzyme shows that the small domain is tilted away by approximately 27° from the larger domain, thereby opening the active site cavity (*SI Appendix*, Fig. S2). As the B-factors in the regions connecting both domains are not significantly higher than those in rigid elements of the protein (*SI Appendix*, Fig. S3), it can be assumed that movement of the 2 domains relative to each other is hindered by crystal packing in the hCAD structure described here.

mCAD Crystal Structure. Using full-length recombinant mCAD expressed in High Five cells, we were able to also obtain a crystal structure of murine CAD, albeit at a somewhat lower resolution compared to hCAD (2.5 Å vs. 1.7 Å) (Table 1). Both proteins share a sequence identity of 85% and their crystal structures superimpose perfectly (rmsd: 0.85 Å) (Fig. 5*A*). The putative active-site residues are conserved between both proteins, but a difference in the conformation of Tyr318 might indicate a potential role of this residue in gating the active site.

Comparison with *A. terreus* and Zebrafish CAD and *Ustilago maydis trans*-Aconitate Decarboxylase. hCAD and mCAD share only a sequence identity of 25% and 23%, respectively, with aCAD (Fig. 4). A homology model of aCAD was obtained from the Phyre2 server (29): It possesses the same fold as the mammalian proteins, but carries several substitutions in the active site and only a few conserved residues (Fig. 3*B*). This low degree of conservation likely explains different kinetic properties of



D

Protein	k_{cat} ± SE (95% CI) (s^{-1})	K_M ± SE (95% CI) (mM)	k_{cat}/K_M (M^{-1} s^{-1})	R^2
mCAD	4.94 ± 0.13 (4.67-5.22)	0.65 ± 0.08 (0.48-0.82)	7600	0.97
aCAD	18.4 ± 1.3 (15.6-21.1)	9.0 ± 1.7 (5.4-12.5)	2000	0.97
hCAD wild type	0.94 ± 0.02 (0.90-0.98)	0.61 ± 0.06 (0.48-0.74)	1600	0.98
hCAD Asn152Ser	1.45 ± 0.03 (1.39-1.51)	0.58 ± 0.05 (0.46-0.69)	2500	0.98
hCAD Arg273His	1.10 ± 0.02 (1.06-1.14)	0.34 ± 0.03 (0.28-0.41)	3200	0.98

Fig. 2. Enzyme kinetics of wild-type and mutant CADs. (*A*) Michaelis–Menten plots showing the differences between CAD from human, mouse, and *Aspergillus*. (*B*) Effect of pH on enzyme activity. Enzymes were incubated with 8 mM *cis*-aconitate for 10 min at 37 °C and formation of itaconic acid was measured by HPLC. Error bars indicate the SD of triplicate assays. (*C*) Michaelis–Menten and Lineweaver–Burk plots showing the effect of the Asn152Ser and Arg273His mutations on hCAD. The diagram on the right is a zoom-in of the complete diagram shown in the *Inset*. Error bars in *A* and *C* indicate SE of triplicate assays at pH 6.5. Data were fitted to the Michaelis–Menten equation using SigmaPlot, resulting in turnover number (k_{cat}), Michaelis constant (K_M), k_{cat}/K_M and R^2, indicating the goodness of fit (*D*); 95% CI, 95% confidence intervals.

aCAD and may reflect adaption of the *Aspergillus* enzyme to different physiological conditions.

The zebrafish *Acod1* gene is highly induced during infection with *Salmonella* or injection of monosodium urate crystals (11, 20). It encodes a protein that is homologous to mammalian CADs, with nearly complete conservation in the active site (Fig. 4).

Another fungus, the corn pathogen *U. maydis*, also produces itaconic acid. However, it uses as substrate *trans*-aconitate, which in turn is produced from *cis*-aconitate by an epimerase. It has thus been proposed that the ability to synthesize itaconate has evolved independently in these 2 fungal species (30). A structural homology model of *U. maydis trans*-aconitate decarboxylase (TAD, UniProt TAD1_USTMD) generated by Phyre2 revealed a completely different fold than hCAD (*SI Appendix*, Fig. S4): It resembles bacterial fumarate lyase and related enzymes, with which it shares up to 46% sequence identity (PDB ID code 3C8T). Thus, the ability of these 2 fungi to synthesize itaconic acid clearly results from convergent evolution.

Validation of Active-Site Residues by Site-Directed Mutagenesis. We then used site-directed mutagenesis to verify the putative active site of hCAD suggested by the crystal structures. Eight residues lining the pocket of the putative active site (Fig. 3*A*), all of which are conserved between murine and human CAD (Fig. 4), were changed to alanine in hCAD expression plasmids. The activities of purified mutated and wild-type enzymes were assessed in cell-free enzyme assays (Fig. 6*A*). The effect of the mutations on itaconate synthesis in human cells was also determined. Respiratory epithelial A549 cells produce virtually no itaconic acid and were therefore used for transient transfection with plasmids expressing wild-type or mutant hCAD, followed by measuring intracellular itaconic acid concentrations by LC-MS/MS (Fig. 6*B*). There was remarkable agreement between both approaches in that all 8 amino acid changes led to substantial or complete loss of CAD activity in both the cell-free assay and transient transfection, confirming the location of the active site. In particular, no itaconate production at all was detected upon mutating His103 and Lys207. These 2 residues (as well as His159 and Lys272) are conserved between the mammalian CADs, aCAD and the related enzyme IDS epimerase (27), supporting their importance for catalysis (Fig. 4). The other 4 of the 8 mutated residues

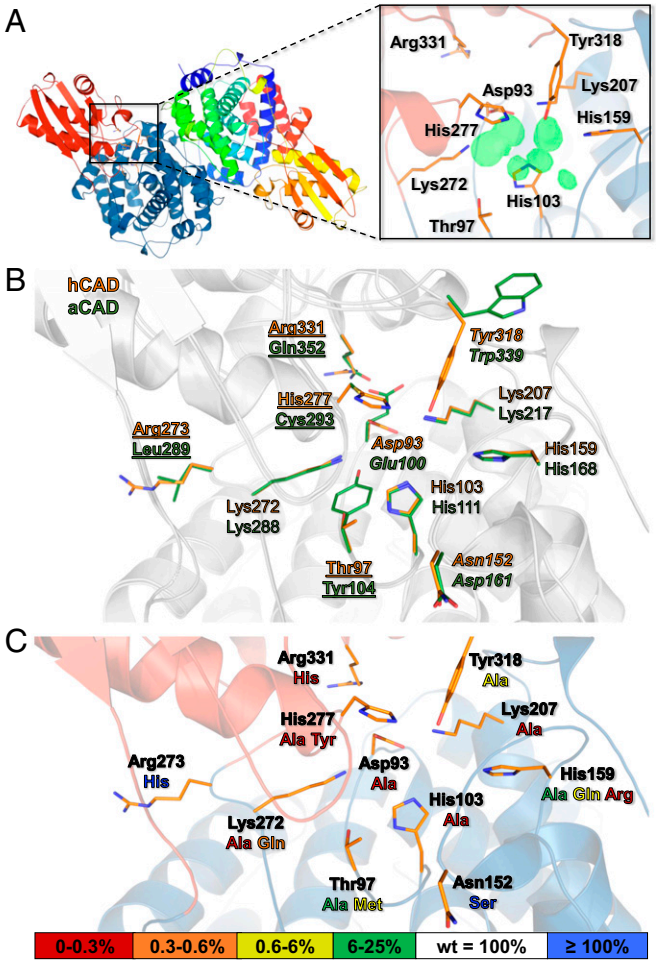


Fig. 3. Crystal structure of CAD. (*A*) Surface and cartoon representation of a dimer of hCAD (*Left*) with 1 polypeptide chain colored to highlight the 2 domains (blue/red) and 1 chain colored from the N terminus (blue) to the C terminus (red). (*Right*) The active site with residues potentially involved in substrate binding or catalysis shown as orange sticks. Positive difference electron density indicating the presence of an unidentified ligand is shown in green (contoured at +3 σ). (*B*) Superposition of hCAD (light gray/orange) with a homology model of *A. terreus* CAD generated by Phyre2 (dark gray/green). The 2 proteins share 25% sequence identity. The residues of the active site and at the positions of frequent polymorphisms are shown as sticks. Conservative substitutions are highlighted in bold and italics, nonconservative substitutions are bold and underlined. Of 11 residues, 4 are conserved between the enzymes, 3 are conservative, and 4 nonconservative substitutions. (*C*) hCAD active site residues and residues at positions of frequent polymorphisms shown as sticks. The effect of single-point mutations relative to activity of the wild-type enzyme is shown by colored labels ranging from red (no activity) via orange (0.3 to 0.6%) and yellow (0.6 to 6%) to green (6 to 25%). Mutations leading to unchanged/increased enzymatic activity are highlighted in blue.

are different in aCAD, which could be the reason for the differences in enzyme kinetics between mammalian CADs and aCAD.

Naturally Occurring Polymorphisms of the Human *ACOD1* Gene. Having identified residues critical for hCAD function, we then interrogated the gnomAD database (31) for sequence variants in the human *ACOD1* gene that could impact on the activity of the enzyme. This analysis identified 8 missense polymorphisms close to the active center (Table 2). Of the 8 polymorphisms, 6 are very rare (gnomAD allele count 1 to 4 in >140,000 genomes) and directly affect active site residues (Fig. 3*C*). To test the effects of 8 polymorphisms on hCAD function, the corresponding mutant proteins were produced in *E. coli*, enzyme activities were determined, and denaturation temperatures were measured by ThermoFluor thermal shift assay. All 6 rare variants (Thr97Met, His159Arg, His159Gln, Lys272Gln, His277Tyr, and Arg331His in Fig. 6) abolished or markedly reduced enzyme activity in the cell-free assay and in transfected A549 cells. As shown in the immunoblot analysis in Fig. 6*B*, intracellular levels of 5 of the 6 rare hCAD mutants were low after transfection. Denaturation temperatures of the corresponding purified mutant proteins were distinctly lower than of wild-type hCAD, suggesting that the reduced levels after transfection were due to decreased stability (Fig. 6*C*). In contrast, mutating active-site residues to alanine had only negligible effects on protein stability, demonstrating that the affected residues are indeed critical for catalytic activity of CAD. The natural mutations thus abolished CAD activity by both disrupting the active center and reducing intracellular stability.

We also analyzed 2 relatively frequent polymorphisms that affect residues close to the active site (i.e., at the interface of the 2 CAD domains) (Table 3). The polymorphism with dbSNP ID rs61741168 (gnomAD 13:77531492) affects position 273, following active-site residue Lys272, and changes arginine to histidine (Fig. 3*C*). This amino acid change affected enzyme activity only to a minor extent in that k_{cat} of the corresponding mutated protein was somewhat higher and K_M somewhat lower compared

Table 1. Data collection and refinement statistics

Structure	hCAD	mCAD
PDB ID code	6R6U	6R6T
Data collection		
Beamline	SLS X06DA (PXIII)	BESSY BL 14.2
Wavelength (Å)	1.00	0.92
Space group	$P2_12_12$	$P4_32_12$
Cell dimensions		
a, *b*, *c* (Å)	102.0, 110.0, 75.0	174.1, 174.1, 71.5
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)*	102.01–1.71 (1.74–1.71)	123.09–2.54 (2.58–2.54)
R_{merge} (%)*	8.4 (123.1)	12.2 (158.3)
R_{pim} (%)*	2.4 (35.1)	3.3 (41.7)
$I/\sigma I$*	22.6 (2.2)	18.0 (2.0)
Completeness (%)*	100 (100)	97.3 (100)
Redundancy*	13.3 (13.2)	14.6 (15.3)
$CC_{1/2}$ (64) (%)*	100.0 (81.0)	99.9 (63.1)
Refinement		
Resolution (Å)	1.71	2.54
No. reflections	92,397	36,050
R_{work}/ R_{free} (%)	14.5/16.2	21.1/24.7
No. atoms (non-H)	8,579	7,730
Protein	7,939	7,683
Ligand/ion	30	0
Water	610	47
B-factors (Å^2)	27	63
Protein	26	63
Ligand/ion	41	0
Water	36	50
Rmsd		
Bond lengths (Å)	0.004	0.002
Bond angles (°)	0.80	0.51
Clashscore	1.6	1.5
Ramachandran statistics (%)		
Favored	98.1	96.5
Allowed	1.9	3.5

*Values for the highest resolution shell are shown in parentheses.

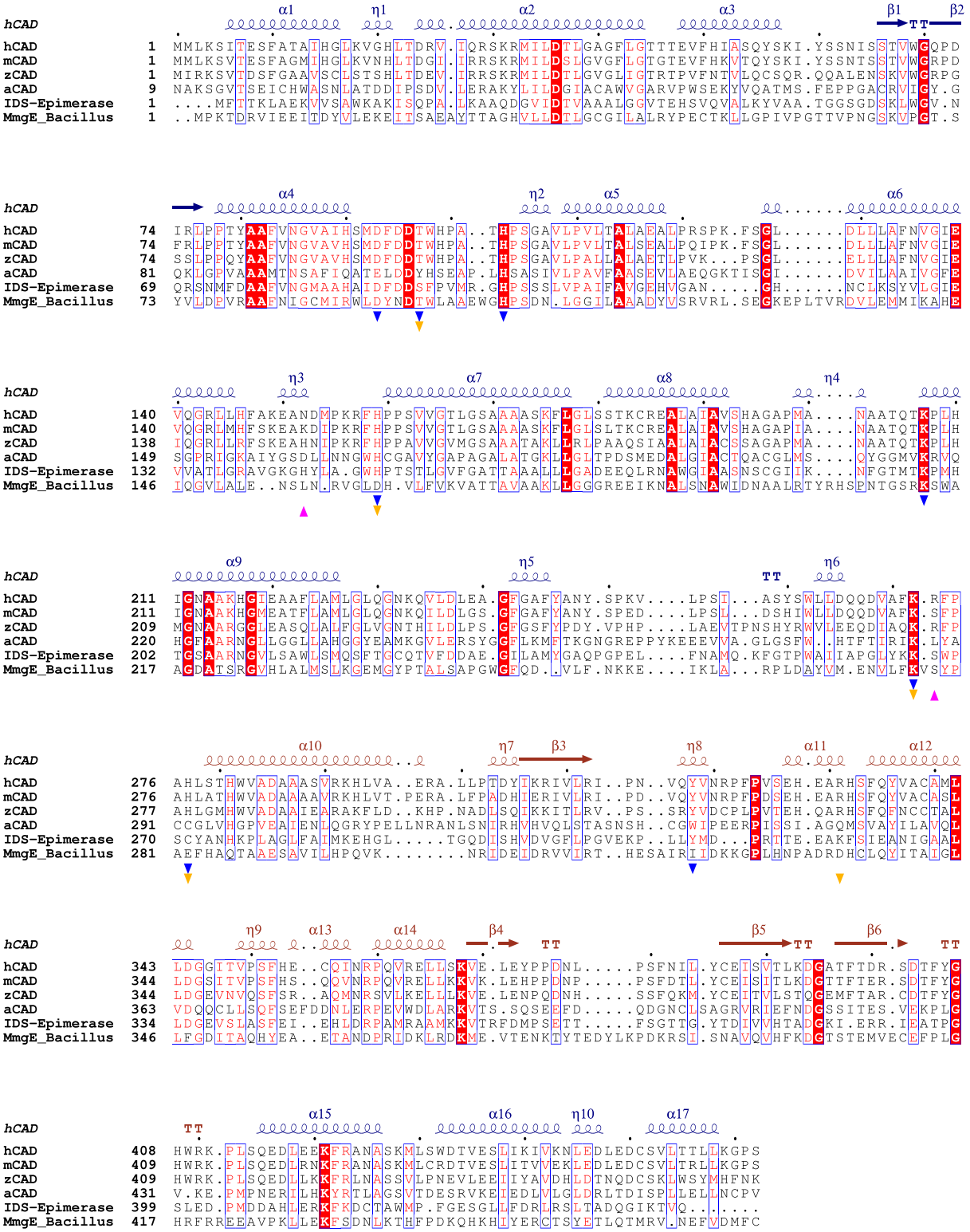


Fig. 4. Sequence alignment of CAD and related enzyme sequences. The Swiss-Prot/UniProt sequences used are: hCAD (IRG1_HUMAN), mCAD (IRG1_MOUSE), aCAD (CAD_ASPTE), IDS epimerase (*A. tumefaciens*, Q1L4E3_RHIRD), MmgE_Bacillus (*Bacillus subtilis* 2-methylcitrate dehydratase, MMGE_BACSU), and zCAD (zebrafish CAD, B0UYM1_DANRE). Active-site residues that were mutated to alanine in hCAD are labeled with blue triangles. Orange triangles indicate hCAD polymorphisms that reduced enzyme activity, while pink upward triangles indicate those with neutral/increasing effect. Secondary structure elements are presented on top: helices with squiggles, β-strands with arrows, turns with TT letters. Blue and brown indicate the large and the small domains, respectively. The alignment was prepared with T-Coffee (69) and ESPript (70). Red boxes and red text highlight identical and similar residues, respectively.

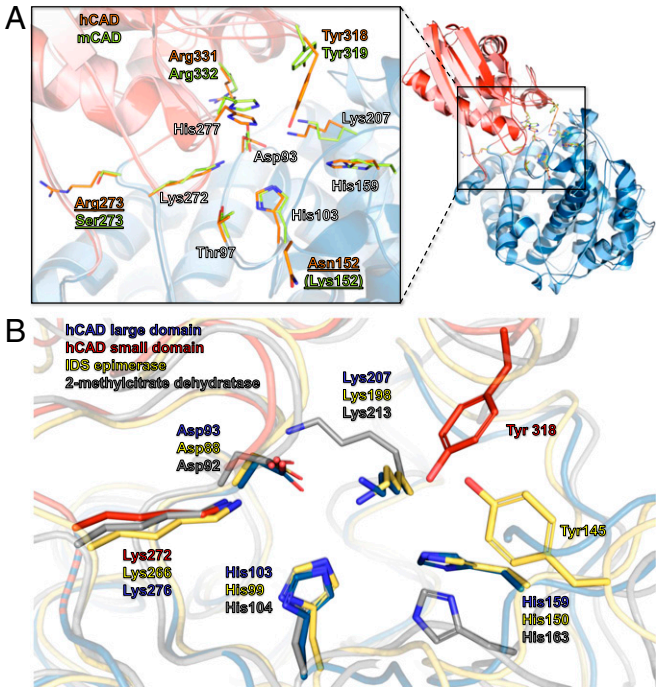


Fig. 5. Active sites of hCAD and related enzymes. (*A*) Superposition of hCAD (1.7 Å) with mCAD (2.5 Å). The proteins have 84% sequence identity. The rmsd of the C_α positions was 0.85 Å. Active-site residues and positions of frequent polymorphisms are shown as sticks. hCAD was rendered in dark red, dark blue, and orange, whereas light red, blue, and lime were used for mCAD. Differences between both proteins are underlined. The mCAD residue Lys152 was not modeled. (*B*) Conserved residues in the active site of hCAD, IDS epimerase from *A. tumefaciens* (PDB ID code 2HP3) and 2-methylcitrate dehydratase (MCDH) from *B. subtilis* (PDB ID code 5MUX). Asp93, His103, Lys207, and Lys272 are strictly conserved among all 3 enzymes. His163 of MCDH is shifted by 1 residue compared to His159 of hCAD or His150 of IDS epimerase, but the flexibility of the side-chain could compensate for this. Tyr145 of IDS epimerase is a potentially important residue in the enzyme's reaction mechanism (27). The hydroxyl groups of this residue and Tyr318 of hCAD are positioned very closely in the superposition of the crystal structures.

to the wild-type (Fig. 2*C*). No change was measured after transfection of A549 cells (Fig. 6*B*).

The most frequent missense polymorphism of hCAD changes Asn152, a residue in vicinity to the active site (Fig. 3*C*), to serine (gnomAD 13:77529644, dbSNP ID rs640192). k_{cat} of the corresponding mutant protein was ∼50% higher while K_M was unchanged (Fig. 2*C*), and itaconic acid synthesis in A549 cells was more than twice as high as that of the wild type (Fig. 6*B*). According to data of the 1000 Genomes Project (32), this polymorphism has an allele frequency of 19.7% in genomes of African and African American ancestry and is essentially absent in other populations (Table 3). Interestingly, mCAD and aCAD (both of which have higher activity than hCAD) differ from wild-type hCAD at the corresponding position (Fig. 4). mCAD has a lysine residue there, but its side-chain could not be modeled in the structure due to insufficient electron density at this position. It is possible that the asparagine found in wild-type hCAD contributes to the somewhat lower enzyme activity of hCAD compared to mCAD.

Taking these data together, this genotype/enzyme activity analysis showed that hCAD loss-of-function mutations are very rare, suggesting that they are selected against, whereas a mutation leading to increased activity may have conferred a selective advantage in the African gene pool.

Discussion

We have determined the crystal structures of hCAD and mCAD, identified residues critical for their function, assessed their enzymatic properties in functional assays, and compared the results to CAD from the evolutionarily distant fungus *A. terreus*.

Potential Origin of CAD in Evolution. Unlike many other decarboxylases (33), human and mouse CAD did not require a cofactor for catalytic activity, which is consistent with the results by Dwiarti et al. (34), who first described enzyme properties of purified aCAD. The CADs therefore establish a new class of the diverse but otherwise relatively rare cofactor-free decarboxylases (33, 35–41). In addition to IDS epimerase, prokaryotic 2-methylcitrate dehydratase also shares structural similarity with CAD. Although the mechanisms of IDS epimerase and the dehydratase remain to be characterized, they share some intriguing common properties with CAD. The 2-methylcitrate dehydratase from *E. coli* catalyzes, as the back reaction, hydratation of 2-methyl-*cis*-aconitate, and it also hydrates, more slowly, *cis*-aconitate (42). The substrate of IDS epimerase is also rich in carboxylate groups. This suggests that CADs (which appear to occur naturally only in eukaryotes) have evolved from a prokaryotic enzyme that recognizes *cis*-aconitate and in addition acquired a functional feature that enables the decarboxylation reaction.

Implications for a Putative Mechanism of CAD Catalysis. It can be assumed that during their catalytic cycle, CAD, IDS epimerase, and 2-methylcitrate dehydratase require the protonation/deprotonation of a carbon atom. Comparison of the sequences and crystal structures of the 3 enzymes revealed that despite the low sequence identity, several histidine and lysine residues, potential proton donors, are conserved among these enzymes (Figs. 4 and 5*B*). Although the residual electron density in the active site could neither be assigned to a substance from the crystallization buffer nor to the substrate or reaction product of CAD, 2 areas of high electron density located next to positively charged or polar residues can probably be attributed to 2 carboxylate groups in a molecule containing a linear C4-moiety. We have used this to model 2 possible binding modes of the substrate *cis*-aconitate, one locating the C1- and C6-carboxylates in the electron density, the other placing C5 and C6 in the respective positions. The better fit resulted from the first of these 2 possibilities (Fig. 7*A*). In this, the C1- and C6-carboxylates would be tightly coordinated by several hydrogen bonds, but the leaving C5-carboxylate would not be involved in any polar interactions and instead be exposed toward a hydrophobic pocket (Fig. 7 *B* and *C*). The location of the negatively charged group within this nonpolar environment would make the formation of an uncharged carbon dioxide and the C3-C4 ethylene moiety favorable, while hydrogen bonds of the C1- and C6-carboxylates to several positively charged or polar residues would provide additional polarization and thus destabilization of the substrate. Together, these contributions would provide a driving force for the decarboxylation reaction. It has been shown that during the reaction a proton from the solvent is transferred to C2 (43). This could be facilitated by His103, which would in this model be located in an appropriate position for the respective proton transfer and was identified to be crucial for enzyme activity in our mutagenesis experiments.

Adaptation of CAD Sequence and Structure throughout Evolution. Even though the overall structural organization of CAD was preserved from *Aspergillus* to mice and humans, we found considerable differences in enzyme kinetics and pH optima in *Aspergillus*, and a much higher activity in mCAD than hCAD. In a larger investigation into CAD structure and function across key organisms along the evolutionary tree it will now be of great interest which amino acid substitutions account for changes in enzymatic properties during key steps in evolution. Consistent

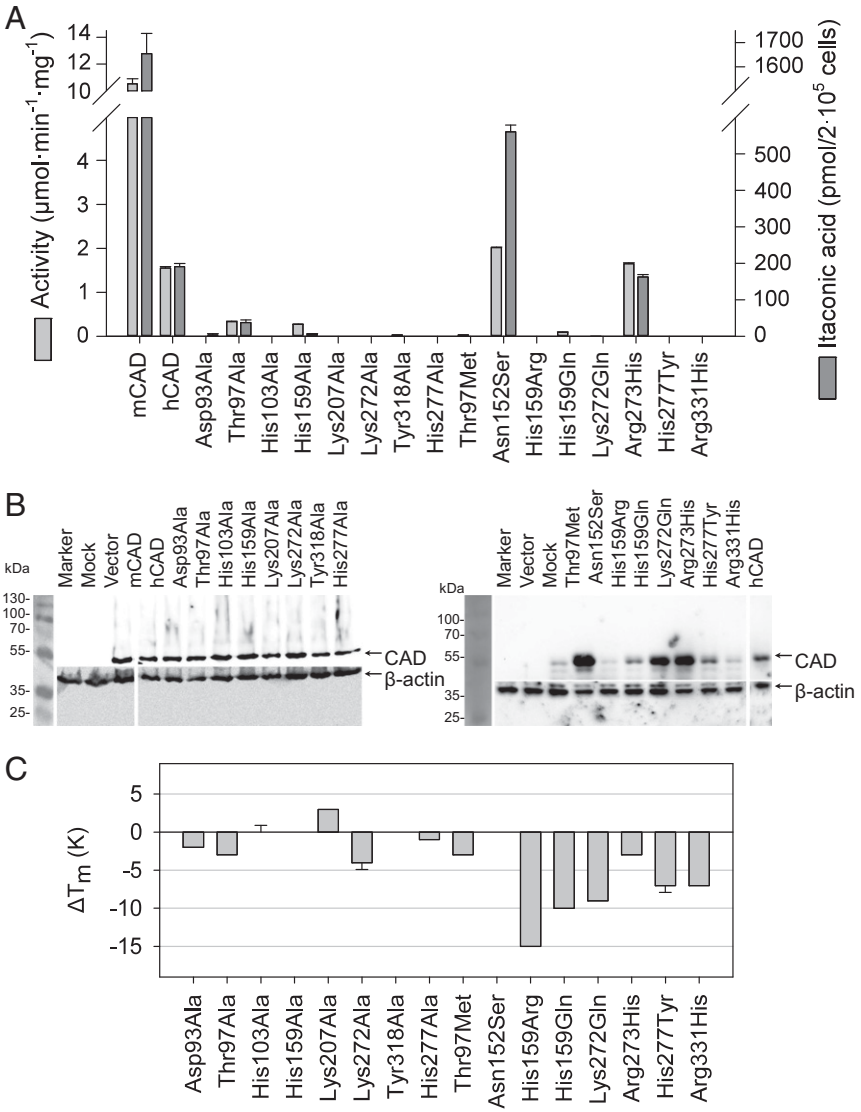


Fig. 6. Enzymatic activity of human CAD mutants. (*A*) The activities of wild-type CAD from human (hCAD) and mouse (mCAD) and of hCAD mutants were tested. Activities of purified enzymes (*Left y* axis) and itaconic acid production by transfected A549 cells (*Right y* axis) are shown. The activities of human CAD mutants were determined by incubation of purified proteins with 8 mM *cis*-aconitate for 10 min at pH 6.5 and 37 °C, followed by itaconic acid measurement by HPLC. Wild-type hCAD was compared to 8 alanine mutants of active site residues and to 8 mutants corresponding to natural polymorphisms of the human *ACOD1* gene. Error bars indicate the SD of triplicate assays. (*B*) Western blot analysis of mCAD and wild-type and mutant mCAD overexpression in transfected A549 cells. β-Actin was used as loading control. Mock, transfection without DNA; vector, transfection with the empty pCMV6-Entry vector. (*C*) Melting temperatures (T_m) at which 50% of the protein is unfolded were determined by a thermal-shift assay. T_m differences (ΔT_m) between mutants of hCAD and the wild-type protein (T_m = 63 °C) are shown. Error bars indicate SDs from triplicate measurements.

with a previous report that itaconate synthesis is substantially higher in murine than in human macrophages upon activation (8), we found that mCAD is markedly more active than hCAD. The active sites of the 2 enzymes are well preserved and all residues predicted to interact with the substrate are identical. There are 5 differences near the active sites of hCAD and mCAD, namely Asn152Lys, Met154Ile, Met199Ile, Arg273Ser, and Ser279Ala. This includes the 2 positions, Asn152 and Arg273, that are changed by frequent hCAD polymorphisms affecting enzyme activity. Arg273 is the first residue of the small domain. A large backbone rotation occurs here in 2-methylcitrate-dehydratase during the opening of the active center (*SI Appendix*, Fig. S2). hCAD Met199 and the loop containing Asn152 and Met154, all located on the large domain, contact the small domain in the closed conformation. An open conformation of hCAD would require separation of these contacts. These observations suggest that the differing activities between mCAD, hCAD, and the frequent hCAD variants could be caused by differences in the efficiency of opening and closing of the active site. In addition, Met154 and Met199 line a hydrophobic pocket in the active center that was hypothesized to accommodate the leaving carboxylate group of the substrate (Fig. 7) and the presence of isoleucine instead of methionine in mCAD may also influence enzyme acitivity. While it remains to be explained why CAD activity has diverged during mammalian evolution, it can be expected that a similar level of transcription of CAD mRNA would lead to much higher levels of itaconic acid in murine than in human cells. This should be considered when inferring from studies in mice about CAD function in humans.

Significance of the Identification of the Active Site to Studies of Human Disease. Recent studies have suggested that common sequence variants in the *ACOD1* gene impact upon clinically relevant

Table 2. Natural active site variants in the human *ACOD1* gene

GnomAD ID	SNP ID	Residue change*	Allele count†	Allele number†	Frequency†
13:77529479 C/T	rs767323284	Thr97Met	4	174,458	0.000023
13:77529644 A/G	rs640192	Asn152Ser	2,602	174,616	0.014901
13:77531150 A/G	rs1471882722	His159Arg	1	140,584	0.000007
13:77531151 T/A	rs1179279159	His159Gln	1	140,356	0.000007
13:77531488 A/C	rs1018207074	Lys272Gln	4	174,384	0.000023
13:77531492 G/A	rs61741168	Arg273His	415	174,382	0.002380
13:77531503 C/T	rs1289708092	His277Tyr	1	143,394	0.000007
13:77531666 G/A	rs755737247	Arg331His	4	174,458	0.000023

*GenBank NP_001245335.
†According to gnomAD (30), version of September 2018.

human immune responses (e.g., efficacy of hepatitis B vaccination and innate immune tolerance) (44, 45). However, it was not possible in those studies to specifically test sequence variants affecting the active site. Our results show that loss-of-function mutations are extremely rare. The available evidence of CAD function in infection and immunity suggests that reduced CAD activity would result in a phenotype characterized by increased propensity to inflammation and susceptibility to infections, including those caused by *Mycobacterium tuberculosis* (16), *Salmonella* (11), and *Legionella* (17), all of which may contribute to limiting the spread of CAD loss-of-function mutants in a human population. The Asn152Ser variation, which led to increased CAD activity, was common specifically in populations of African descent. On one hand, it may constitute a neutral variant (or one of unknown significance) that became fixed in these populations known for their greater genetic diversity compared to non-African populations (32). Alternatively, it may constitute an adaption that is associated with decreased susceptibility to infectious or inflammatory diseases. This is plausible in the case of mycobacteria, as they likely coevolved with early humans on the African continent (46) for an estimated time of at least 2.6 million y (47), which led to the speciation of *M. tuberculosis* about 40,000 y ago (48, 49). The importance of CAD/itaconic acid in limiting *M. tuberculosis* has been demonstrated in vitro and in vivo (8, 16, 50). It is therefore tempting to speculate that the Asn152Ser mutation became enriched in the African gene pool because it conferred a selective advantage by reducing susceptibility to mycobacterial infection during an extensive period of coevolution. Our results also now enable investigating the possible impact of SNPs in CAD at the population level on chronic inflammation and inflammation-associated common disorders, such as cardio- and cerebrovascular disease, neurodegenerative disorders, and inflammatory joint disease. Indeed, CAD expression and itaconic acid synthesis are induced in mouse models of gout (51) and chronic arthritis (52). Here, we hypothesize that diminished CAD function would be associated with increased, and gain-of-function with decreased disease risk.

In summary, elucidation of the structure of CAD will now enable further studies geared toward detailed characterization of the mechanism of CAD, structure/function studies across a broad range of eukaryotic organisms expressing CAD under diverse ecological conditions, the possible impact of sequence variants on risk of human disease, and genetic engineering of more efficient CADs for improved biotechnological production of itaconic acid.

Materials and Methods

Molecular weights were determined by size-exclusion chromatography-coupled MALS. Melting temperatures were determined by thermofluor, a fluorescence-based thermal shift assay. The ACOD1 antibody D6H2Y, a rabbit mAb (Cell Signaling #77510), and an HRP-conjugated anti–β-actin antibody (Abcam #ab49900) were used for immunoblotting. For more information on molecular weight determination, thermal shift assay and immunoblotting, see *SI Appendix, Materials and Methods*.

Plasmids. A codon-optimized sequence encoding *A. terreus* CAD (GenBank AB326105) was gene-synthesized (Eurofins). Sequences encoding residues 4 to 461 of hCAD (GenBank NM_001258406), 4 to 462 of mCAD (GenBank NM_008392), and 12 to 490 of aCAD were PCR-amplified from Irg1 pCMV6-Entry plasmids (OriGene) and from the synthetic gene with primer pairs CAD38/29, CAD39/28, and CAD41/24, respectively (*SI Appendix*, Table S1). PCR products were cloned between the NotI and KpnI sites of a modified pCOLADuet-1 vector containing an N-terminal StrepTagII with tobacco etch virus (TEV) cleavage site (MASWSHPQFEKVDENLYFQGGG) using a QuickFusion kit (Absource Diagnostics), generating plasmids pCAD29 (human), pCAD39 (mouse), and pCAD16 (*A. terreus*). For cloning into pOpIE2, a PCR product encoding mCAD was PCR-amplified with primers CAD12/21 and was cloned into a modified pOpIE2 vector (53) with a C-terminal thrombin cleavage site and StrepTagII (GTLVPRGSAA GKGSA WSHPQFEK GGGSGGGSGGSA WSHPQFEK), generating plasmid CAD09. Site-directed mutagenesis of pCAD29 and pCMV6-Entry plasmids by QuikChange mutagenesis was done with PfuUltra II Fusion HS (Agilent) for the "CAD50-59" primers (*SI Appendix*, Table S2) or Phusion flash mix (ThermoFisher) for the "CADpoly" primers, as described previously (54).

Protein Production with *E. coli* Cells. Proteins were expressed in the *E. coli* strain BL21 (DE3) CodonPlus RIL (Agilent) by autoinduction at 25 °C (55). Twenty milliliters of preculture grown in MDAG-135 medium containing 100 μg/mL kanamycin and 34 μg/mL chloramphenicol at 37 °C was used to

Table 3. Frequencies of human *ACOD1* genotypes

	Arg273His (rs61741168)		Asn152Ser (rs640192)		
Population	Allele frequency*, %	Heterozygotes*, %	Allele frequency*, %	Heterozygotes*, %	Homozygotes*, %
African	2.3	4.5	19.7	31.0	4.2
American	0.3	0.6	1.3	2.0	0.3
Asian	0.0	0.0	0.0	0.0	0.0
European	0.0	0.0	0.2	0.4	0.0
All	0.6	1.3	5.4	8.5	1.2

*1000 Genomes Project data (31).

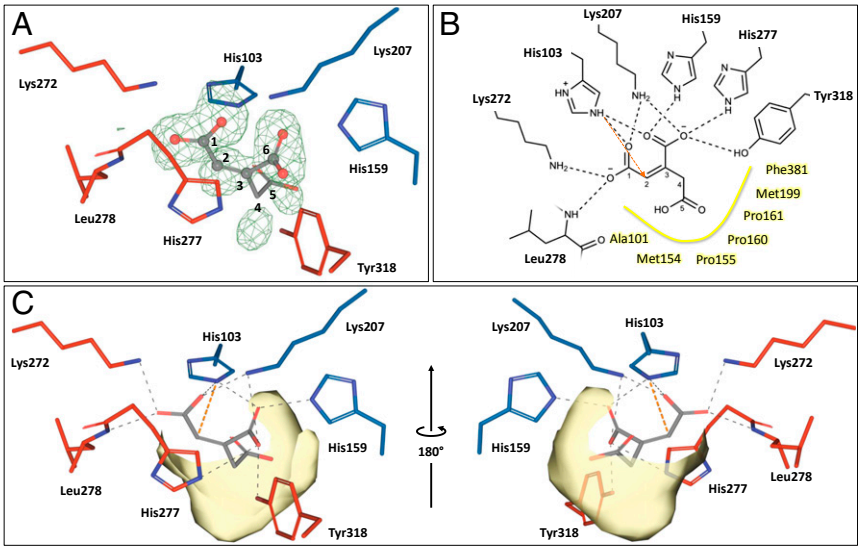


Fig. 7. Modeling of *cis*-aconitate binding and implications on the potential reaction mechanism of hCAD. (*A*) Fitting of *cis*-aconitate (gray) into the positive $F_{obs}-F_{calc}$ difference electron density in the active site of hCAD (green mesh, contoured at $\sigma = +3$). The atoms used for the fit are also shown as spheres. The C1 and C6 carboxylates can be modeled well into the electron density, while the leaving C5 carboxylate cannot be attributed to the map. (*B*) Ligand interaction diagram and (*C*) corresponding 3D representation of *cis*-aconitate fitted into the active site as in *A*. The C1 and C6 carboxylates of the substrate are tightly coordinated by several hydrogen bonds to the enzyme, but the leaving C5 carboxylate is not involved in any interactions. Any possible conformation of this carboxylate moiety would be positioned it in a hydrophobic pocket (yellow), making it likely that this carboxylate would not be charged but protonated. This hydrophobic environment together with the polarizing hydrogen bonds at the C1 and C6 carboxylates would make the formation of an uncharged carbon dioxide favorable. The protonation that has been shown to occur at C2 (43) could be facilitated by His103, which was shown to be crucial for the reaction of hCAD. In this model, the side chain of this residue seems to be oriented appropriately for transferring a proton to C2 of *cis*-aconitate (dashed orange arrow/line).

inoculate 1 L ZYM-5052 medium containing the same antibiotics in a 2.8-L Fernbach flask. Cultures were shaken at 130 rpm for 24 h at 25 °C to saturation at an OD_{600} of 13 to 14. Cells were harvested (15 min, 5,000 rpm) and about 90 g wet cell mass was typically obtained from 4 L of culture. Cells were frozen in liquid nitrogen and were stored at −80 °C. Cells were resuspended in 200 mL lysis buffer (20 mM Tris·HCl, pH 8.0, 0.5 M NaCl, 10% [vol/vol] glycerol, 1 mM DTT) and were lysed by high-pressure homogenization. Polyethylenimine (PEI) of molecular weight 50,000 to 100,000 (No. 195444, MP Biomedicals) was added to 0.5% (wt/vol) and the mixture was centrifuged at 16,500 rpm for 45 min, followed by filtration of the supernatant through a 0.45-μm filter. The lysate was loaded onto a 6-mL Strep-Tactin XT column at 3 mL/min. The column was washed with lysis buffer and bound protein was eluted with lysis buffer containing 50 mM biotin. The protein was incubated with TEV protease overnight to remove the affinity tag and was further purified with a Superdex 200 column using GF buffer (10 mM Hepes, pH 7.4, 150 mM NaCl, 0.1 mM TCEP, 10% [vol/vol] glycerol). Peak fractions were pooled, TCEP was added to 1 mM and the protein was concentrated to 8 to 10 mg/mL with a VivaSpin 20 ultrafiltration concentrator, molecular weight cutoff (MWCO) 10,000 (Sartorius). A protein amount of 5 to 20 mg was obtained from 4-L culture. Concentrated proteins were frozen in liquid nitrogen and stored at −80 °C.

Protein Production with Insect Cells. mCAD was produced by transient gene expression in High Five insect cells with the pCAD09 plasmid, followed by purification by 2 steps of chromatography. One liter of High Five insect cell culture was transfected using PEI as described previously (56), resulting in 27 g of cell pellet. The cells were resuspended in lysis buffer (50 mM Hepes, pH 8.0, 0.5 M NaCl, 0.5% [vol/vol] IGEPAL CA-630, 1 μL Benzonase, 1 tablet cOmplete protease inhibitor [Roche], 10% [vol/vol] glycerol, 2 mM TCEP, 1 mM EDTA) and lysed by mixing and passing through a 0.9-mm needle. Following centrifugation for 2 × 20 min at 16,000 rpm, the supernatant was diluted to 100 mL binding buffer (50 mM Hepes, pH 8.0, 0.5 mM NaCl, 10% [vol/vol] glycerol, 2 mM TCEP, 1 mM EDTA). The protein was purified with StrepTactin Superflow beads (IBA GmbH) by elution with 10 mM d-desthiobiotin added to the binding buffer. The eluted CAD protein was further purified with a Superdex 75 26/60 column using GF2 buffer (10 mM Hepes, pH 7.4, 150 mM NaCl, 2 mM TCEP, 10% [vol/vol] glycerol). Peak fractions were pooled and the protein was concentrated to 3 mg/mL with a VivaSpin 20 ultrafiltration concentrator, MWCO 10,000 (Sartorius). Two milligrams protein was obtained from 1-L culture. Concentrated proteins were shock-frozen using liquid nitrogen and stored at −80 °C.

Enzyme Assays. Enzyme activity was measured according to Dwiarti et al. (34) Assays were performed in triplicate. In general, 125 μL 0.2 M sodium phosphate buffer or sodium acetate buffer was mixed on ice with 15 μL enzyme in GF buffer and 10 μL *cis*-aconitate (Sigma-Aldrich #A3412), pH 6.5 (NaOH). Incubation for 10 min at 37 °C was immediately followed by heat inactivation of the enzyme at 95 °C for 15 min. The protein precipitate was pelleted by centrifugation. Supernatants were filtered using Spin-X 0.22-μm centrifuge tube filters and were acidified with 100 μL 100 mM H_3PO_4. Itaconic acid was measured by HPLC (Licrosphere C-8 reverse-phase column, 1 mL/min 10 mM H_3PO_4, detection at 210 nm).

Enzyme Kinetics. Proteins were diluted to 1 mg/mL in GF buffer and exact concentrations were measured by spectrophotometry at 280 nm. The appropriate amounts of protein were mixed with 125 μL sodium phosphate buffer, pH 6.5, and GF buffer to 140 μL. Ten microliters of *cis*-aconitate (pH 6.5) at 1.5, 4.5, 15, 45, 150, and 450 mM was added to obtain 0.1- to 30-mM final concentration. Assays were performed as described above. Enzyme amounts per 150-μL assay were 0.75 μg (with 0.1 or 0.3 mM *cis*-aconitate), 1.5 μg (with 1 mM *cis*-aconitate), 3 μg (3 mM), or 6 μg (10, 30 mM) of hCAD or 0.5 μg (0.1 to 1 mM), 0.75 μg (3 mM), or 1 μg (10, 30 mM) of mCAD or aCAD.

Effect of pH and Mutations on Enzyme Activity. Proteins were diluted in GF buffer to 2.5 mg/mL (hCAD), 1.5 mg/mL (mCAD), or 1 mg/mL (aCAD) and the exact concentrations were measured photometrically. Next, 0.2 M sodium acetate buffers at pH 4.0 and 4.5 and 0.2 M sodium phosphate buffers at pH 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0 were prepared. Then 125 μL buffer was mixed with 15 μL hCAD, 5 μL mCAD, or 1 μL aCAD and GF buffer to 140 μL, resulting in 33 μg hCAD, 5 μg mCAD, and 1.5 μg aCAD per assay. Ten microliters of 120 mM *cis*-aconitate, pH 6.5 (NaOH) was added (8 mM final concentration). Activity of hCAD mutants was measured in the same way at pH 6.5.

Cell Transfection. Human epithelial A549 adenocarcinoma cells (DSMZ no. ACC 107) were cultivated in 1× RPMI Medium 1640 (Gibco #31870-025) with 1% GlutaMAX-I (100×) (Gibco #35050-038), supplemented with 10% FBS (Biochrom GmbH #S0115) and 0.1 mM MEM Non-Essential Amino Acids Solution 100× (Sigma-Aldrich #M7145). Next, 2 × 10^5 cells per well were

transfected in a 6-well plate with human or mouse Irg1 pCMV6-Entry plasmids (OriGene). pCMV6-Empty vector and mock transfection (no DNA) were used as negative controls. Transfections were done using Lipofectamine LTX and PLUS Reagent (ThermoFisher #15338100) following the manufacturer's handbook and then incubated for 48 h at 37 °C with 5% CO_2.

Itaconic Acid Quantification. A549 cells were grown in 6-well plates and harvested for itaconic acid concentration measurements 48 h posttransfection. For harvesting and simultaneous metabolite extraction, medium was removed and cells were placed on ice. Metabolites were extracted by adding 300 μL ice-cold organic extraction solvent per well (acetonitrile/methanol/water, 2/2/1 [vol/vol/v]) with internal standard (ISTD: 2 μM citric acid-2,4-$^{13}C_2$, Merck #492078 or $^{13}C_5$-itaconic acid: Toronto Research Chemicals #1931004), scraping the cells and then transferring the cell suspension to a 2-mL Eppendorf reaction tube. Scraper and well were rinsed twice using 400 μL extraction solvent without ISTD. Metabolite extracts were frozen for at least 2 h at −20 °C to complete protein precipitation. After centrifugation (10 min, 4 °C, 20,800 × *g*), the supernatant was evaporated and the dried metabolite extract was resuspended in 75 μL HPLC-grade water, resulting in a final ISTD concentration of 8 μM. The sample was transferred to an MS-vial (Wicom #WIC41160) with insert (Macherey-Nagel #702813). Itaconic acid was measured by HPLC-MS/MS on a liquid-chromatography system (SIL HTc) coupled to a triple quadrupole mass spectrometer (API4000, Sciex). Data acquisition and further quantification was performed using the Analyst software 1.5.2 (Sciex).

Protein Crystallization. Crystallization trials were set up at room temperature with a HoneyBee 961 crystallization robot (Digilab Genomic Solutions) in Intelli 96-3 plates (Art Robbins Instruments) with 200 nL protein solution at different concentrations and 200 nL reservoir solution. Several commercial sparse matrix screens were used to identify suitable crystallization conditions and an optimization screen was designed containing a gradient of either Na_3-citrate (0.7 to 1.2 M) or itaconate (0.7 to 1.8 M) and 0.1 M of Tris·HCl, Bis-Tris·HCl or Hepes with a pH range of 6.5 to 10. Optimized crystals of hCAD were obtained from this screen (6.8 mg/mL protein, Tris·HCl pH 8.7, 1.2 M Na_3-citrate) and crystals of mCAD (expressed in insect cells) grew in condition 29 (1.5 mg/mL protein, 60% tacsimate pH 7.0) of the INDEX sparse matrix screen (Hampton Research). The crystals were cryoprotected by addition of 10% (2*R*, 3*R*)-2,3-butanediol before flash-freezing.

Data Collection and Processing. Datasets were collected at beamline X06DA (PXIII) of the SLS (Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland) and at beamline 14.2 of the BESSY II (Helmholtz-Zentrum Berlin, Germany) (57). All datasets were recorded at a temperature of 100 K. The datasets were processed using the AutoPROC (58) toolbox (Global Phasing) executing XDS (59), Pointless (60), and Aimless (61).

Structure Determination, Refinement, and Model Building. The crystal structure of hCAD was solved by molecular replacement executing BALBES (62) from the CCP4 suite (63). Iminodisuccinate epimerase (PDB ID code 2HP3, sequence identity: 25%) (27) was found as a working search model. The crystal structure of mCAD was solved using the structure of the human protein as search model for Phaser (64) from the Phenix suite (65). The structural models were initially generated running AutoBuild (66) from the Phenix software package (65). These models were further analyzed and manually completed in Coot (67) and crystallographic refinement was performed with Phenix.refine (68), including the addition of hydrogens in riding positions and TLS-refinement. Five percent of random reflections were flagged to calculate R_{free}. hCAD was at a resolution of 1.7 Å refined to R/R_{free} of 14/16% in space group $P2_12_12$. mCAD was at a resolution of 2.5 Å refined to R/R_{free} of 20/25% in space group $P4_32_12$. Data collection and refinement statistics are summarized in Table 1.

Data Availability. Structural data have been deposited in the Protein Data Bank (https://www.rcsb.org) with the PDB ID codes 6R6T for mCAD and 6R6U for hCAD. Plasmids pCAD16, pCAD29, and pCAD39 and the Asn152Ser, Arg273His, and His103Ala mutants of pCAD29 and human Irg1 pCMV6Entry have been deposited at Addgene.

ACKNOWLEDGMENTS. We thank Kevin Walkling, Daniela Gebauer, Annette Garbe, and Claudia Wylegalla for expert technical assistance; Luca Codutti (Biomolekulares Wirkstoffzentrum, Hannover) and Stefan Schmelz for size-exclusion chromatography–multiangle light-scattering measurements; and Kwanghoon Sung for molecular docking calculations. The study was funded by the Helmholtz Association's Initiative on Individualised Medicine (iMed), by the Helmholtz Association's Future Topic "Aging and Metabolic Programming" (AMPro), and by internal funds of the Helmholtz Centre for Infection Research.

1. S. Baup, Ueber eine neue Pyrogen-Citronensäure, und über Benennung der Pyrogen-Säuren überhaupt. *Ann. Phar.* **19**, 29–38 (1836).
2. G. L. Crasso, Untersuchungen über das Verhalten der Citronensäure in höherer Temperatur und die die daraus hervorgehenden Produkte. *J. Prakt. Chem.* **20**, 322–339 (1840).
3. H. Stobbe, A. Lippold, Einfluss des Lichtes auf die Polymerisation des Itaconsäure-äthylesters. *J. Prakt. Chem.* **90**, 336–344 (1914).
4. T. Cordes, A. Michelucci, K. Hiller, Itaconic acid: The surprising role of an industrial compound as a mammalian antimicrobial metabolite. *Annu. Rev. Nutr.* **35**, 451–473 (2015).
5. J. H. Kane, A. C. Finlay, P. F. Amann, *Production of Itaconic Acid*, US Patent 2385283 (1945).
6. C. L. Strelko *et al*., Itaconic acid is a mammalian metabolite induced during macrophage activation. *J. Am. Chem. Soc.* **133**, 16386–16389 (2011).
7. M. Sugimoto *et al*., Non-targeted metabolite profiling in activated macrophage secretion. *Metabolomics* **8**, 624–633 (2012).
8. A. Michelucci *et al*., Immune-responsive gene 1 protein links metabolism to immunity by catalyzing itaconic acid production. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7820–7825 (2013).
9. V. Lampropoulou *et al*., Itaconate links inhibition of succinate dehydrogenase with macrophage metabolic remodeling and regulation of inflammation. *Cell Metab.* **24**, 158–166 (2016).
10. T. Cordes *et al*., Immunoresponsive gene 1 and itaconate inhibit succinate dehydrogenase to modulate intracellular succinate levels. *J. Biol. Chem.* **291**, 14274–14284 (2016).
11. C. J. Hall *et al*., Immunoresponsive gene 1 augments bactericidal activity of macrophage-lineage cells by regulating β-oxidation-dependent mitochondrial ROS production. *Cell Metab.* **18**, 265–278 (2013).
12. L. A. J. O'Neill, M. N. Artyomov, Itaconate: The poster child of metabolic reprogramming in macrophage function. *Nat. Rev. Immunol.* **19**, 273–281 (2019).
13. C. Tang *et al*., 4-Octyl itaconate activates Nrf2 signaling to inhibit pro-inflammatory cytokine production in peripheral blood mononuclear cells of systemic lupus erythematosus patients. *Cell. Physiol. Biochem.* **51**, 979–990 (2018).
14. E. L. Mills *et al*., Itaconate is an anti-inflammatory metabolite that activates Nrf2 via alkylation of KEAP1. *Nature* **556**, 113–117 (2018).
15. M. Bambouskova *et al*., Electrophilic properties of itaconate and derivatives regulate the IκBζ-ATF3 inflammatory axis. *Nature* **556**, 501–504 (2018).
16. S. Nair *et al*., *Irg1* expression in myeloid cells prevents immunopathology during *M. tuberculosis* infection. *J. Exp. Med.* **215**, 1035–1045 (2018).
17. J. Naujoks *et al*., IFNs modify the proteome of legionella-containing vacuoles and restrict infection via IRG1-derived itaconic acid. *PLoS Pathog.* **12**, e1005408 (2016).
18. B. P. Daniels *et al*., The nucleotide sensor ZBP1 and kinase RIPK3 induce the enzyme IRG1 to promote an antiviral metabolic state in neurons. *Immunity* **50**, 64–76.e4 (2019).
19. H. Liu *et al*., Four-octyl itaconate activates Keap1-Nrf2 signaling to protect neuronal cells from hydrogen peroxide. *Cell Commun. Signal.* **16**, 81 (2018).
20. C. J. Hall *et al*., Blocking fatty acid-fueled mROS production within macrophages alleviates acute gouty inflammation. *J. Clin. Invest.* **128**, 1752–1771 (2018).
21. J. M. Weiss *et al*., Itaconic acid mediates crosstalk between macrophage metabolism and peritoneal tumors. *J. Clin. Invest.* **128**, 3794–3805 (2018).
22. T. Werpy, G. Petersen, *Top Value Added Chemicals From Biomass, Volume I: Results of Screening for Potential Candidates from Sugars and Synthesis Gas* (National Renewable Energy Laboratory, Golden, CO, 2004), vol. I.
23. T. Willke, K. D. Vorlop, Biotechnological production of itaconic acid. *Appl. Microbiol. Biotechnol.* **56**, 289–295 (2001).
24. M. Okabe, D. Lies, S. Kanamasa, E. Y. Park, Biotechnological production of itaconic acid and its biosynthesis in Aspergillus terreus. *Appl. Microbiol. Biotechnol.* **84**, 597–606 (2009).
25. T. J. Wickham, T. Davis, R. R. Granados, M. L. Shuler, H. A. Wood, Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system. *Biotechnol. Prog.* **8**, 391–396 (1992).
26. J. J. Ward, L. J. McGuffin, K. Bryson, B. F. Buxton, D. T. Jones, The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20**, 2138–2139 (2004).
27. B. Lohkamp, B. Bäuerle, P.-G. Rieger, G. Schneider, Three-dimensional structure of iminodisuccinate epimerase defines the fold of the MmgE/PrpD protein family. *J. Mol. Biol.* **362**, 555–566 (2006).
28. J. J. Paxman, B. Heras, Bioinformatics tools and resources for analyzing protein structures. *Methods Mol. Biol.* **1549**, 209–220 (2017).
29. L. A. Kelley, M. J. Sternberg, Protein structure prediction on the Web: A case study using the Phyre server. *Nat. Protoc.* **4**, 363–371 (2009).
30. E. Geiser *et al*., Ustilago maydis produces itaconic acid via the unusual intermediate trans-aconitate. *Microb. Biotechnol.* **9**, 116–126 (2016).
31. M. Lek *et al*.; Exome Aggregation Consortium, Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
32. A. Auton *et al*.; 1000 Genomes Project Consortium, A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
33. T. Li, L. Huo, C. Pulley, A. Liu, Decarboxylation mechanisms in biological system. *Bioorg. Chem.* **43**, 2–14 (2012).

34. L. Dwiarti, K. Yamane, H. Yamatani, P. Kahar, M. Okabe, Purification and characterization of cis-aconitic acid decarboxylase from Aspergillus terreus TN484-M1. *J. Biosci. Bioeng.* **94**, 29–33 (2002).
35. J. J. Almrud *et al*., Crystal structures of the wild-type, P1A mutant, and inactivated malonate semialdehyde decarboxylase: A structural basis for the decarboxylase and hydratase activities. *Biochemistry* **44**, 14818–14827 (2005).
36. T. C. Appleby, C. Kinsland, T. P. Begley, S. E. Ealick, The crystal structure and mechanism of orotidine 5′-monophosphate decarboxylase. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2005–2010 (2000).
37. M. M. Benning, T. Haller, J. A. Gerlt, H. M. Holden, New reactions in the crotonase superfamily: Structure of methylmalonyl CoA decarboxylase from Escherichia coli. *Biochemistry* **39**, 4630–4639 (2000).
38. T. Bock *et al*., AibA/AibB induces an intramolecular decarboxylation in isovalerate biosynthesis by myxococcus xanthus. *Angew. Chem. Int. Ed. Engl.* **56**, 9986–9989 (2017).
39. L. Cendron *et al*., The structure of 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase provides insights into the mechanism of uric acid degradation. *J. Biol. Chem.* **282**, 18182–18189 (2007).
40. L. A. Highbarger, J. A. Gerlt, G. L. Kenyon, Mechanism of the reaction catalyzed by acetoacetate decarboxylase. Importance of lysine 116 in determining the pKa of active-site lysine 115. *Biochemistry* **35**, 41–46 (1996).
41. K. Miyamoto, H. Ohta, Purification and properties of a novel arylmalonate decarboxylase from Alcaligenes bronchisepticus KU 1201. *Eur. J. Biochem.* **210**, 475–481 (1992).
42. M. Brock, C. Maerker, A. Schütz, U. Völker, W. Buckel, Oxidation of propionate to pyruvate in Escherichia coli. Involvement of methylcitrate dehydratase and aconitase. *Eur. J. Biochem.* **269**, 6184–6194 (2002).
43. R. Bentley, C. P. Thiessen, Biosynthesis of itaconic acid in Aspergillus terreus. III. The properties and reaction mechanism of cis-aconitic acid decarboxylase. *J. Biol. Chem.* **226**, 703–720 (1957).
44. X. Liu *et al*., Polymorphisms in IRG1 gene associated with immune responses to hepatitis B vaccination in a Chinese Han population and function to restrain the HBV life cycle. *J. Med. Virol.* **89**, 1215–1223 (2017).
45. J. Dominguez-Andres *et al*., The itaconate pathway is a central regulatory node linking innate immune tolerance and trained immunity. *Cell Metab.* **29**, 211–220.e5 (2019).
46. Z. Djelouadji, D. Raoult, M. Drancourt, Palaeogenomics of Mycobacterium tuberculosis: Epidemic bursts with a degrading genome. *Lancet Infect. Dis.* **11**, 641–650 (2011).
47. M. C. Gutierrez *et al*., Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. *PLoS Pathog.* **1**, e5 (2005).
48. I. Comas *et al*., Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nat. Genet.* **45**, 1176–1182 (2013).
49. I. Hershkovitz *et al*., Tuberculosis origin: The Neolithic scenario. *Tuberculosis (Edinb.)* **95** (suppl. 1), S122–S126 (2015).
50. H. Wang *et al*., An essential bifunctional enzyme in *Mycobacterium tuberculosis* for itaconate dissimilation and leucine catabolism. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 15907–15913 (2019).
51. F. Pessler *et al*., Identification of novel monosodium urate crystal regulated mRNAs by transcript profiling of dissected murine air pouch membranes. *Arthritis Res. Ther.* **10**, R64 (2008).
52. F. Michopoulos *et al*., Targeted metabolic profiling of the Tg197 mouse model reveals itaconic acid as a marker of rheumatoid arthritis. *J. Proteome Res.* **15**, 4579–4590 (2016).
53. M. Bleckmann *et al*., Identification of essential genetic baculoviral elements for recombinant protein expression by transactivation in Sf21 insect cells. *PLoS One* **11**, e0149424 (2016).
54. L. Zheng, U. Baumann, J.-L. Reymond, An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32**, e115 (2004).
55. F. W. Studier, Stable expression clones and auto-induction for protein production in E. coli. *Methods Mol. Biol.* **1091**, 17–32 (2014).
56. K. Karste, M. Bleckmann, J. van den Heuvel, Not limited to E. coli: Versatile expression vectors for mammalian protein expression. *Methods Mol. Biol.* **1586**, 313–324 (2017).
57. M. Gerlach, U. Mueller, M. S. Weiss, The MX beamlines BL14.1-3 at BESSY II. *J. Large-Scale Res. Facil.* **2**, A47 (2016).
58. C. Vonrhein *et al*., Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 293–302 (2011).
59. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
60. P. Evans, Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006).
61. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
62. F. Long, A. A. Vagin, P. Young, G. N. Murshudov, BALBES: A molecular-replacement pipeline. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 125–132 (2008).
63. M. D. Winn *et al*., Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
64. A. J. McCoy *et al*., Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
65. P. D. Adams *et al*., PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
66. T. C. Terwilliger *et al*., Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 61–69 (2008).
67. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
68. P. V. Afonine *et al*., Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
69. C. Notredame, D. G. Higgins, J. Heringa, T-coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).
70. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).