BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 2687–2696 2687
DOI: 10.1021/bi901766z

Substrate Specificity Determinants of the Methanogen Homoaconitase Enzyme: Structure and Function of the Small Subunit[†,‡]

Jeyaraman Jeyakanthan,[§,▼] Randy M. Drevland,[∥,▼] Dasara Raju Gayathri,[⊥] Devadasan Velmurugan,[⊥] Akeo Shinkai,[#] Seiki Kuramitsu,[#,△] Shigeyuki Yokoyama,[▲,◇] and David E. Graham*[,∥,♦,▽]

[§]*Life Science Group, National Synchrotron Radiation Research Center, 101 Hsin-Ann Road, Hsinchu Science Park, Hsinch 30076, Taiwan,* [∥]*Chemistry and Biochemistry Department, The University of Texas at Austin, Austin, Texas 78712,* [⊥]*Centre of Advanced Study in Crystallography and Biophysics, University of Madras, Guindy Campus, Chennai 600 025, India,* [#]*RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan,* [△]*Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan,* [▲]*RIKEN Systems and Structural Biology Center, Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan,* [◇]*Graduate School of Science, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan,* [♦]*Department of Microbiology, University of Tennessee, Knoxville, Tennessee 37996, and* [▽]*Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831.* [▼]*These two authors contributed equally to this paper.*

Received October 13, 2009; Revised Manuscript Received February 1, 2010

ABSTRACT: The aconitase family of hydro-lyase enzymes includes three classes of proteins that catalyze the isomerization of α-hydroxy acids to *β*-hydroxy acids. Besides aconitase, isopropylmalate isomerase (IPMI) proteins specifically catalyze the isomerization of α,*β*-dicarboxylates with hydrophobic *γ*-chain groups, and homoaconitase (HACN) proteins catalyze the isomerization of tricarboxylates with variable chain length *γ*-carboxylate groups. These enzymes' stereospecific hydro-lyase activities make them attractive catalysts to produce diastereomers from unsaturated precursors. However, sequence similarity and convergent evolution among these proteins lead to widespread misannotation and uncertainty about gene function. To find the substrate specificity determinants of homologous IPMI and HACN proteins from *Methanocaldococcus jannaschii*, the small-subunit HACN protein (MJ1271) was crystallized for X-ray diffraction. The structural model showed characteristic residues in a flexible loop region between α2 and α3 that distinguish HACN from IPMI and aconitase proteins. Site-directed mutagenesis of MJ1271 produced loop-region variant proteins that were reconstituted with wild-type MJ1003 large-subunit protein. The heteromers formed promiscuous hydro-lyases with reduced activity but broader substrate specificity. Both R26K and R26V variants formed relatively efficient IPMI enzymes, while the T27A variant had uniformly lower specificity constants for both IPMI and HACN substrates. The R26V T27Y variant resembles the MJ1277 IPMI small subunit in its flexible loop sequence but demonstrated the broad substrate specificity of the R26V variant. These mutations may reverse the evolution of HACN activity from an ancestral IPMI gene, demonstrating the evolutionary potential for promiscuity in hydro-lyase enzymes. Understanding these specificity determinants enables the functional reannotation of paralogous HACN and IPMI genes in numerous genome sequences. These structural and kinetic results will help to engineer new stereospecific hydro-lyase enzymes for chemoenzymatic syntheses.

The methanogen homoaconitase (HACN)[1] enzyme catalyzes the isomerization of (*R*)-homocitrate to (2*R*,3*S*)-homoisocitrate in the biosynthesis of coenzyme B (Scheme 1) (*1*). This enzyme evolved from an isopropylmalate isomerase (IPMI) ancestor, a ubiquitous enzyme required for leucine biosynthesis. HACN and IPMI, together with aconitase (ACN; citric acid cycle), homoaconitate hydratase (α-aminoadipate pathway for lysine biosynthesis), 2,3-dimethylmalate dehydratase (nicotinate catabolism), and 2-methylcitrate dehydratase, belong to the aconitase family of [4Fe-4S]-dependent hydro-lyases (*2*−*7*). Only three members of this family, HACN, IPMI, and ACN, catalyze sequential dehydration and hydration reactions, creating new isomers. Despite extensive structural and mechanistic characterization of ACN proteins, the substrate specificity determinants of IPMI and HACN enzymes were unknown (*8*, *9*). As a result, sequence databases often list incorrect functional annotations for these proteins, hindering metabolic reconstruction projects. Efforts to design new hydro-lyase enzymes for the chemoenzymatic synthesis of diastereomeric acids will require understanding the specificity determinants reported here.

In the leucine biosynthetic pathway, IPMI catalyzes the reversible dehydration of α-isopropylmalate to *cis*-dimethylcitraconate and the subsequent *trans* addition of water, yielding (2*R*,3*S*)-3-isopropylmalate (*β*-isopropylmalate) (*10*). IPMI also catalyzes

[†]The present work was supported by the RIKEN Structural Genomic/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science, and Technology of Japan, National Science Foundation Grant MCB-0817903, and U.S. Department of Energy, Office of Science Biological and Environmental Research. R.M.D. was supported in part by a Faraday Fellowship. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract DE-AC05-00OR22725.

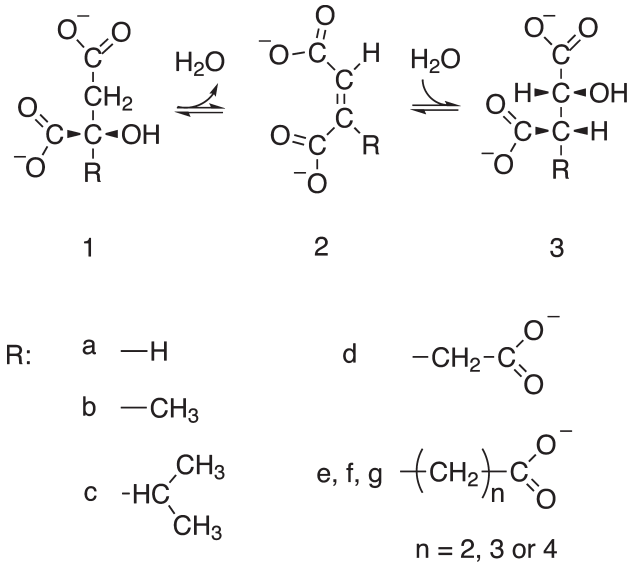
[‡]The structural model of the MJ1271 protein has been deposited in the RCSB Protein Data Bank with identifier 2PKP.

*Address correspondence to this author. Tel: 865-574-0559. Fax: 865-576-8646. E-mail: grahamde@ornl.gov.

[1]Abbreviations: HACN, homoaconitase; IPMI, isopropylmalate isomerase; ACN, aconitase; mACN, mitochondrial aconitase; SOE-PCR, splicing overlap extension polymerase chain reaction; DTT, dithiothreitol; CHES, 2-(cyclohexylamino)ethanesulfonic acid; MES, 2-(*N*-morpholino)ethanesulfonic acid; PEG, polyethylene glycol; HICDH, homoisocitrate dehydrogenase; IPMDH, isopropylmalate dehydrogenase.

Published on Web 02/19/2010

Scheme 1: Hydro-lyase Enzymes Catalyze the Isomerization of α- and β-Hydroxy Acids[a]



[a]R-group substituents are shown below the reaction. Malease (EC 4.2.1.31) catalyzes the hydration of maleate (**2a**) to form D-malate (**3a**). The isopropylmalate isomerase (IPMI, EC 4.2.1.33) catalyzes the dehydration of 2-isopropylmalate (**1c**), forming isopropylmaleate (**2c**) that is hydrated to produce 3-isopropylmalate (**3c**). This enzyme can also catalyze the dehydration of citramalate (**1b**) and the subsequent hydration of citraconate (**2b**) to produce 2-methylmalate (**3b**). The mitochondrial aconitase (mACN, EC 4.2.1.3) catalyzes the dehydration of citrate (**1d**) to produce *cis*-aconitate (**2d**) that is hydrated to form isocitrate (**3d**). The homoaconitate hydratase enzyme (EC 4.2.1.36) catalyzes only the hydration of *cis*-homoaconitate (**2e**) to form homoisocitrate (**3e**). Finally, the methanogen homoaconitase (HACN, EC 4.2.1.114) catalyzes the dehydration of $(homo)_{1-3}$-citrate (**1e**,**f**,**g**) and the hydration of *cis*-$(homo)_{1-3}$-aconitate (**2e**,**f**,**g**) to form $(homo)_{1-3}$-isocitrate (**3e**,**f**,**g**).

the isomerization of (*R*)-citramalate to β-methylmalate in the pyruvate pathway for isoleucine biosynthesis that replaces threonine dehydratase in some microbes (*11*). In the α-aminoadipate pathways of *Thermus thermophilus*, *Saccharomyces cerevisiae*, and *Aspergillus nidulans* (*12*), homoaconitate hydratase catalyzes only the second half-reaction in the isomerization of (*R*)-homocitrate: the hydration of *cis*-homoaconitate to (2*R*,3*S*)-homoisocitrate (*3*, *4*, *13*). A second enzyme, predicted to be ACN, was proposed to catalyze the initial dehydration of (*R*)-homocitrate to *cis*-homoaconitate (*4*). It is therefore remarkable that the homoaconitase from *Methanocaldococcus jannaschii* ($HACN_{Mj}$) catalyzes the full conversion of (*R*)-homocitrate to homoisocitrate, as well as the full isomerizations of (*R*)-$homo_2$-citrate and (*R*)-$homo_3$-citrate, in the chain elongation reactions of coenzyme B biosynthesis (*1*).

The HACN and IPMI homologues are monomeric (fungi) or heterodimeric (bacteria and archaea) iron−sulfur cluster proteins, similar to the well-characterized monomeric mitochondrial aconitase (mACN). In *M. jannaschii*, the MJ1271 and MJ1003 proteins comprise $HACN_{Mj}$, and the paralogous MJ1277 and MJ0499 proteins comprise $IPMI_{Mj}$. The large subunits of the archaeal proteins (MJ1003 or MJ0499, ∼48 kDa) are homologous to domains 1, 2, and 3 of mACN, including the $CX_{63}CXXC$ motif that provides ligands for the [4Fe-4S] cluster. These large subunits interact with small subunits (MJ1271 or MJ1277, ∼18 kDa), homologous to domain 4 of mACN. Sequence alignments of archaeal HACN and IPMI homologues indicate that the majority of the 21 active site residues in mACN are highly conserved, implying similar structural organization and catalytic mechanisms. Structural and mutational analysis of mACN implicated Arg^{580} from a flexible loop of domain 4 as the key residue responsible for recognizing citrate, *cis*-aconitate, and isocitrate in the aconitase mechanism (*14*, *15*). The guanidinium nitrogen atoms of Arg^{580} hydrogen bond to the substrate's γ-carboxylate group. Replacing Arg^{580} with lysine resulted in a 30-fold increase in K_M with isocitrate (*15*). However, this arginine residue is not conserved in the archaeal HACN or IPMI sequences. The crystal structure of a putative bifunctional HACN/IPMI small subunit from the euryarchaeon *Pyrococcus horikoshii* shows that the loop region between α-helices 1 and 2 corresponds to the flexible loop region of mACN (*16*). Complementation analysis demonstrated the homoaconitase (or homoaconitate hydratase) function of the *P. horikoshii* protein (*17*), while metabolic reconstruction and phylogenetic analysis made the IPMI function doubtful (*11*). An alignment of the mACN and *P. horikoshii* HACN sequences suggested thatArg^{580} is replaced with a polar threonine residue in the loop region (*16*), inconsistent with the mACN substrate-binding model (*18*). Without biochemical characterization, both the function and mechanism of this *P. horikoshii* protein are unresolved.

The similarity between small subunits of the *M. jannaschii* proteins $IPMI_{Mj}$ (MJ1277) and $HACN_{Mj}$ (MJ1271) underscores the importance of defining the consensus sequence responsible for substrate recognition. These two protein sequences are 53% identical, and both were previously annotated as IPMI subunits based on primary sequence alone, only to be distinguished by heterologous expression, purification, reconstitution, and *in vitro* kinetic analysis (*1*, *11*). Sequences of the predicted flexible loop regions of these enzymes, identified by homology models and sequence alignments, may be the best indicators of substrate specificity. The $HACN_{Mj}$ small subunit has a consensus sequence of $Y^{24}LRT$ while the homologous $IPMI_{Mj}$ has the sequence $Y^{26}LVY$. The polar Arg^{26} and Thr^{27} residues of $HACN_{Mj}$ could potentially form hydrogen bonds with the γ-carboxylates of *cis*-homoaconitate, *cis*-$homo_2$-aconitate, and *cis*-$homo_3$-aconitate. The corresponding hydrophobic Val^{28} and Tyr^{29} residues of the $IPMI_{Mj}$ loop region may be structural determinants for recognizing hydrophobic γ-chains. These consensus sequences are conserved in most of the Euryarchaeota.

In this work, we present the first structural study of a characterized homoaconitase. The $HACN_{Mj}$ small subunit protein, MJ1271, was crystallized and refined to 2.1 Å. Site-directed mutagenesis of the proposed flexible loop region of MJ1271, combined with wild-type MJ1003 large subunit protein, revealed residues affecting substrate specificity. Replacing residues Arg^{26} and Thr^{27} in the $Y^{24}LRT$ loop region of MJ1271 with corresponding residues from the $IPMI_{Mj}$ small subunit created a promiscuous enzyme with both HACN and IPMI activities. We propose that Arg^{26} is the equivalent of mACN Arg^{580}, coordinating the γ-carboxylate of the HACN substrates. The results provide the first structure−function analysis of HACN substrate specificity, which will be used to correct the functional annotations of uncharacterized aconitase-like proteins and design new hydro-lyase enzymes for biocatalysis.

EXPERIMENTAL PROCEDURES

The *Escherichia coli* BL21(DE3) strains with plasmids pDG141 (expressing MJ1003) and pDG163 (expressing MJ1277) were described previously (Table 1) (*11*). Splicing overlap extension

Table 1: List of Plasmids and Microorganisms

strain or plasmid (parent plasmid)	description	source or ref
E. coli		
BL21(DE3)	protein expression host	Novagen
BL21 CodonPlus-RIL	expression host with additional tRNAs	Stratagene
XL-1 Blue	general cloning host	Stratagene
plasmids		
pMJ1271 (pET-21a)	MJ1271	this work
pDG141 (pCDF-Duet1)	MJ1003	*11*
pDG160 (pET-19b)	MJ1271	*11*
pDG163 (pET-19b)	MJ1277	*11*
pDG476 (pET-19b)	MJ1271-MJ1277 chimera	this work
pDG625 (pET-19b)	MJ1271-LysU	this work
pRD03 (pDG160)	MJ1271 R26V	this work
pRD06 (pDG160)	MJ1271 R26K	this work
pRD09 (pDG160)	MJ1271 T27A	this work
pRD17 (pRD03)	MJ1271 R26V T27Y	this work

Table 2: Oligonucleotide Primers Used To Construct MJ1271 Mutations

primer name	sequence (5′ to 3′)
T7	TAATACGACTCACTATAGGG
MJ1277Rev-overlap	GCTAACTCGTAAGGGTCTGT-ATAAACTAAATACCTTGCTGG
MJ1277Fwd-overlap	CCAGCAAGGTATTTAGTTTA-TACAGACCCTTACGAGTTAGC
T7-terminator	GCTAGTTATTGCTCAGCGG
3MJ1271-LysU1	CCAACCATGAACGGAGCGTA-TTTTCCTGGAATTATTGCGTC
5MJ1271-LysU1	CGCTCCGTTCATGGTTGGTGA-ATACGAGTTAGCTTCACACTG
R26V[a]	CCAGGACCTTACTTAGTGACTA-CAGACCCTTACGAG
R26K[a]	CCAGGACCTTACTTAAAGAC-TACAGACCCTTACGAG
T27A[a]	CCAGGACCTTACTTAAGGGCTA-CAGACCCTTACGAG
R26V/T27Y[a]	CCAGGACCTTACTTAGTGTATA-CAGACCCTTACGAG

[a]Mutations are underlined in these sequences.

polymerase chain reaction (SOE-PCR) was used to construct two chimeric genes (*19*). The MJ1271-MJ1277 chimeric gene contained codons 1−29 from MJ1277 fused to codons 28−170 from MJ1271. T7 promoter and MJ1271Rev-overlap primers (Table 2) were used to amplify the 5′-fragment of this chimeric gene from pDG163, and MJ1271Fwd-overlap and T7 terminator primers were used to amplify the 3′-fragment from pDG160. The purified fragments were joined by SOE-PCR. The chimeric product was purified and digested with *Nco*I and *Bam*HI restriction enzymes and then ligated into the same sites of vector pET-19b to create vector pDG476. The MJ1271-LysU chimera replaced codons 23−30 of MJ1271 with codons 18−23 from the *T. thermophilus lysU* gene (*4*). T7 promoter and 3MJ1271-LysU1 primers were used to amplify the 5′-fragment from pDG163, and 5MJ1271-LysU and T7-terminator primers were used to amplify the 3′-fragment from pDG163. SOE-PCR produced a full-length chimeric gene that was ligated into the *Nco*I and *Bam*HI sites of pET-19b to form vector pDG476. For crystallography experiments, the MJ1271 gene was amplified from *M. jannaschii* chromosomal DNA and ligated into vector pET-21a (Novagen) to produce plasmid pMJ1271.

Plasmid pDG160, encoding the wild-type MJ1271 gene, was mutated using the QuikChange II site-directed mutagenesis kit (Stratagene) and the mutagenic primers listed in Table 2 (together with their reverse complements). The resulting mutations in the plasmids listed in Table 1 were confirmed by DNA sequencing. Plasmid pRD03 was used as a template to generate the double mutation in pRD17 by the same strategy. Electroporation transformed *E. coli* BL21(DE3) pDG141 cells with the new plasmids encoding small subunit proteins.

To express proteins for kinetic studies, *E. coli* strains were grown in LB medium, supplemented with 100 μg mL^{-1} ampicillin and 50 μg mL^{-1} streptomycin, at 37 °C with shaking at 250 rpm until the culture reached an optical density at 600 nm of 0.6−0.8. Expression was then induced by the addition of 50 μM isopropyl 1-β-D-galactopyranoside, and the culture was then incubated for an additional 3−4 h. The wild-type MJ1003 and MJ1271 variant proteins were purified by heating the cell lysate, followed by anion-exchange chromatography, dialysis, and concentration by centrifugal ultrafiltration (*1*). Protein purity was determined by SDS−PAGE with Coomassie Blue dye staining. Interactions between MJ1003 and MJ1271 mutants were tested by measuring apparent masses of protein complexes using analytical size exclusion chromatography (*20*).

Copurified apoenzymes containing wild-type MJ1003 and variant MJ1271 proteins (1−2 mg mL^{-1}) were reconstituted in the presence of dithiothreitol (DTT), $Fe(NH_4)_2(SO_4)_2$, and Na_2S as described previously (*1*). A mock reconstitution solution consisted of the above mixture without protein. Enzymatic activity was measured using previously described continuous spectrophotometric assays (*1*), where a decrease in UV absorbance by *cis*-unsaturated substrates was measured under anoxic conditions. Reactions (1 mL) were conducted in quartz semimicrocells with screw cap septa (Starna) containing 50 mM 2-(cyclohexylamino)ethanesulfonic acid (CHES)−KOH (pH 9.0), 200 mM KCl, 10−100 μg mL^{-1} holoenzyme, and various substrate concentrations.

For protein crystallization, *E. coli* BL21 CodonPlus (DE3)-RIL (pMJ1271) cells were grown at 37 °C overnight in a medium containing 1.0% polypeptone, 0.5% yeast extract, 0.5% NaCl, and 100 μg mL^{-1} ampicillin (pH 7.0). The cells were lysed by sonication in 20 mM Tris-HCl (pH 8.0) containing 500 mM NaCl, 5 mM β-mercaptoethanol, and 1 mM phenylmethanesulfonyl fluoride. The lysate was incubated at 70 °C for 10 min and centrifuged at 15000 rpm for 30 min at 4 °C. After exchanging the buffer with 20 mM Tris-HCl (pH 8.0), the protein sample was loaded onto a TOYOPEARL SuperQ-650 M column (Tosoh) preequilibrated with 20 mM Tris-HCl (pH 8.0). The flow-through, containing most MJ1271 protein, was subjected to a buffer exchange with 20 mM 2-(*N*-morpholino)ethanesulfonic acid (MES)−NaOH (pH 6.0) and loaded onto a RESOURCE S column (GE Healthcare Bio-Sciences) preequilibrated with the same buffer. The protein was bound to the column and eluted in a linear gradient from 0 to 400 mM NaCl. Fractions containing MJ1271 protein were subjected to a buffer exchange with 10 mM potassium phosphate buffer (pH 7.0) and loaded on a Bio-Scale CHT20-I column (Bio-Rad) preequilibrated with the same buffer. The protein was bound to the column and eluted in a linear gradient from 10 to 500 mM potassium phosphate (pH 7.0). The MJ1271 protein was subjected to gel filtration on a HiLoad 16/60 Superdex 200 pg column (GE Healthcare Bio-Sciences) preequilibrated with 20 mM Tris-HCl and 200 mM NaCl (pH 8.0). The fraction containing purified protein was

concentrated using a VIVASPIN 10000 molecular weight cutoff ultrafiltration device (Sartorius). The final preparation was dissolved in 20 mM Tris-HCl buffer containing 200 mM NaCl and 1 mM DTT (pH 8.0), at a concentration of 18 mg mL^{-1}.

Initial crystallization conditions were screened using the TERA (automatic crystallization) system (*21*) from 144 conditions. The crystallization of MJ1271 protein was performed using the micro batch method by setting up drops consisting of 1 μL of protein solution and 1 μL of reservoir solution at 295 K. The reservoir solution consisted of 50% (w/v) polyethylene glycol (PEG) 200 and 0.1 M Tris-HCl, pH 4.6. The crystals were transferred to mother liquor containing 20% (v/v) glycerol cryoprotectant.

The crystals were cooled in an N_2 gas stream at 100 K using the SPring-8 precise automatic cryosample exchanger (SPACE), which was controlled using the beamline-scheduling software BSS (*22*, *23*). The intensity data were integrated, scaled, and reduced to 2.1 Å resolution using DENZO and SCALEPACK programs (*24*). The data were indexed in the centered orthorhombic space group $C222_1$, with unit cell parameters a = 90.9 Å, b = 108.3 Å, and c = 45.4 Å. The value of the Matthews coefficient (*25*) was 2.9 $Å^3$ Da^{-1}, and the solvent content was 58.9% assuming a homodimer in the asymmetric unit.

The MJ1271 protein structure was determined by the molecular replacement method, using the orthorhombic form of the *P. horikoshii* small subunit protein (PDB id code 1V7L). The *P. horikoshii* protein contains 163 amino acids, with 44% sequence identity to MJ1271. Residues 1−137 (only core regions or truncations) were used for molecular replacement (phasing) using the AMoRe program (*26*), and the resolution limit of 10.0−6.0 Å was used for molecular replacement. The remaining MJ1271 residues, from 137 to 167, were manually traced based on the electron density map, and then the residues were placed using the MJ1271 protein sequence. A unique solution was obtained with this search model, giving an R_{factor} of 54% and a correlation factor of 66%. The structure was refined using the program CNS 1.1 (*27*). The programs FRODO (*28*) and COOT (*29*) were used for model building and viewing electron density maps. Initially 60 cycles of rigid body refinement were carried out, followed by 50 cycles of positional refinement. Without the mutated residues, the R_{work} dropped to 44% (R_{free} = 50.5%) for all the reflections in the resolution range 20−2.1 Å. The map was clear enough to mutate the residues into the electron density map, and then the model was subjected to simulated annealing by employing a slow-cooling protocol, followed by 100 cycles of positional refinement that dropped the R_{work} to 33.2% (R_{free} = 37.5%). After four rounds of manual rebuilding and subsequent refinement, the model had an R_{free} of 28.5% and R_{work} of 32.5% using all observed reflections. Solvent molecules were gradually included into the structure at stereochemically sensible positions and with difference density higher than 3.0σ and $2|F_o| - |F_c|$ density higher than 0.8σ. The final R_{work} and R_{free} are 20.1% and 24.8%, respectively. During the progress of the refinement, based on the $|F_o| - |F_c|$ density, a Zn ion and PEG molecule were located. The stereochemistry of the refined structure was analyzed with the program PROCHECK (*30*). A summary of structure determination and refinement statistics is given in Table 3.

The SWISS-MODEL automated comparative protein modeling server (version 8.03) constructed a MJ1277 homology model using the MJ1271 structure as the template (*31*). Figures were prepared using the PYMOL program (version 1.2r1; DeLano Scientific). Archaeal HACN and IPMI small subunit sequences were aligned with the T-COFFEE program (version 7.71) (*32*) and edited using the CINEMA5 editor. Sequence logos were produced using the WebLogo 3 program (*33*).

Table 3: Data Collection and Refinement Statistics for the MJ1271 Structure

	native
data collection	
X-ray source	BL26B2, SPring-8
wavelength (Å)	1.0
detector	Jupiter 210cs CCD
temperature (K)	100
distance (mm)	170
space group	$C222_1$
unit cell parameters (Å)	a = 90.9, b = 108.3, c = 45.4
resolution range (Å)	50−2.1 (2.18−2.10)
total reflections	90990
unique reflections	13385 (1286)
completeness (%)	99.7 (97.7)
R_{merge} (%)[a]	6.9 (32.9)
redundancy	6.8 (5.7)
V_M ($Å^3$ Da^{-1})	2.9
solvent content (%)	58.9
refinement	
resolution range (Å)	20.0−2.1
R_{work}[b]	20.1
R_{free}[c]	24.8
no. of protein atoms	1286
no. of solvent atoms	149
zinc ion and PEG di(hydroxyethyl) ether	1 and 1
Ramachandran plot[d] (%)	
most favored	96.4
allowed	3.6
PDB identifier	2PKP

[a] $R_{merge} = \sum|I_i - \langle I\rangle|/\sum\langle I\rangle$. The values within the parentheses refer to the last shell. [b] $R_{work} = \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. [c] R_{free} is the R factor for a subset of 5% of the reflections that were omitted from refinement. [d] As calculated by PROCHECK.

A 4-year-old commercial preparation of *cis*-aconitic acid (stored dried at room temperature) contained 31% *trans*-aconitic acid. For *cis*-aconitic acid: ^{1}H NMR (400 MHz, D_2O + TSP, δ) 6.40 (s, 1H) and 3.51 (d, 2H, J = 1.2 Hz); ^{13}C NMR (D_2O, 100 MHz, δ) 173.1, 169.6, 169.2, 137.2, 129.0, and 34.5. For *trans*-aconitic acid: ^{1}H NMR (400 MHz, D_2O + TSP, δ) 7.03 (s, 1H) and 3.88 (d, 2H, J = 0.6 Hz). A recently purchased batch of *cis*-aconitic acid (Sigma) contained 7% *trans*-aconitic acid. A fresh sample of the latter batch was dissolved in water, immediately adjusted to pH 7 using NaOH, stored briefly on ice, and used for the kinetic experiments reported here. The synthesis and purification of *cis*-homoaconitate analogues were described previously (*1*).

RESULTS

Previous studies showed that $IPMI_{Mj}$ specifically catalyzed the hydration of dicarboxylate substrates with nonpolar γ-chain substituents: maleate, citraconate, and isopropylmaleate (*11*). While $HACN_{Mj}$ also catalyzed the hydration of the minimal substrate maleate, the enzyme specifically recognized tricarboxylate substrates with variable γ-chain lengths: *cis*-$(homo)_1$-aconitate, *cis*-$(homo)_2$-aconitate, *cis*-$(homo)_3$-aconitate, and even the nonphysiological *cis*-$(homo)_4$-aconitate (*1*). Surprisingly, no activity was observed in mixtures containing $HACN_{Mj}$ and *cis*-aconitate, whose γ-chain is one methylene group shorter than

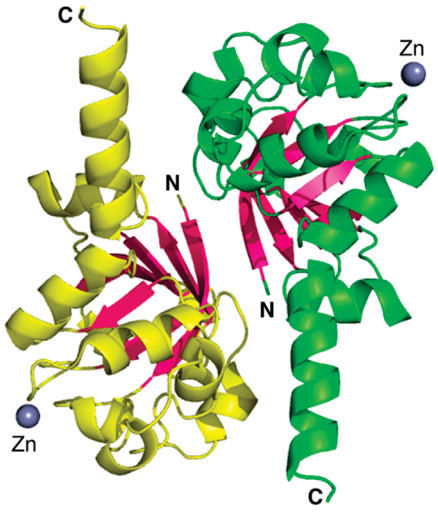


FIGURE 1: Ribbon diagram showing the crystallographic dimer of two MJ1271 subunits. Each protomer coordinates a Zn^{2+} ion (purple).

cis-(homo)$_1$-aconitate. We subsequently discovered that the dry, commercially produced *cis*-aconitic acid sample was contaminated with 31% *trans*-aconitic acid, a potent aconitase, and homoaconitase inhibitor (*1*, *34*). Concentrated *cis*-aconitic acid rapidly isomerizes to *trans*-aconitic acid in aqueous solutions, although the dried compound and neutral solutions are considered stable (*35*). Using a new batch of *cis*-aconitate (adjusted to pH 7 with NaOH, containing 7% *trans*-aconitate), we determined that $HACN_{Mj}$ catalyzed *cis*-aconitate hydration with a K_M of $300 \pm 90\ \mu M$ and a $k_{cat}/K_M = 2.5 \times 10^3\ M^{-1}\ s^{-1}$. This specificity constant is 10-fold lower that values for *cis*-(homo)$_{1-4}$-aconitate due to the high K_M value for *cis*-aconitate. Therefore, $HACN_{Mj}$ is specific for *cis*-unsaturated tricarboxylates, while $IPMI_{Mj}$ recognizes *cis*-unsaturated dicarboxylates.

The structure of the $HACN_{Mj}$ small subunit, MJ1271, was solved by molecular replacement using the *P. horikoshii* small subunit structure that shares 44% amino acid identity and 73% similarity. The MJ1271 protein crystallized as a dimer, with Asp^{13} and Cys^{63} side chains from each subunit coordinating Zn^{2+} ions (Figure 1). Residues 1−167 (out of 170) were modeled using X-ray diffraction data at 2.1 Å resolution. These proteins have a "swiveling" $\beta/\beta/\alpha$ domain consisting of a β-sheet with parallel strands (β1−4) and a second β-sheet with antiparallel strands (β1, β5−7) in an α/β fold (Figure 2). Together, these two β-sheets form a β-barrel-type structure that is capped by α7 and α8. The conserved Ser^{65} residue, corresponding to the alkoxide base Ser^{642} in mACN (*36*), is located in a loop region between β2 and α5 (Figure 3). The main-chain bond angle of this residue is 10.2° greater than the ideal small molecule value. The adjacent Arg^{67} (corresponding to Arg^{644} in mACN) stabilizes the serine alkoxide and interacts with substrate carboxylate groups in mACN (*14*). This strained loop is approximately 15 Å from a flexible loop between α2 and α3, which includes Arg^{580} in mACN. A structural alignment shows that Arg^{580} has no direct equivalent in the smaller loop of MJ1271; instead, we predicted the nearby Arg^{26}-Thr^{27} residues could interact with the substrates' γ-chain substituents (Figure 3).

The $IPMI_{Mj}$ small subunit, MJ1277, shares 53% sequence identity with the $HACN_{Mj}$ small subunit, MJ1271. Therefore, a homology model of MJ1277 was built using the MJ1271 structure coordinates as a template, resulting in a structural alignment with

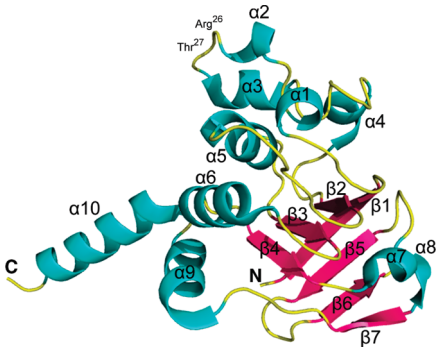


FIGURE 2: Ribbon diagram showing the MJ1271 subunit in the asymmetric unit. Positions of the Arg^{26} and Thr^{27} residues are labeled on the flexible loop between α-helices 2 and 3.

an rmsd of 0.1 Å. The catalytic Ser^{67} and Arg^{69} residues in the MJ1277 model are equivalent to Ser^{65} and Arg^{67} in MJ1271, but residues Val^{28}-Tyr^{29} replace the polar Arg^{26}-Thr^{27} residues in the flexible loop region between α2 and α3 (Figure 4). An alignment of numerous small subunit sequences from homologous archaeal HACN and IPMI proteins confirmed the correlation between residues in this flexible loop and the protein's predicted substrate specificity (Figure 5). Putative HACN subunits usually have a basic residue (Lys or Arg) and Thr in this loop. IPMI subunits show more diversity in this loop: Thr/Ile/Val and Ile/Tyr replace their more polar counterparts in HACN.

To determine whether the Y^{24}LRT sequence of MJ1271 confers substrate specificity for $HACN_{Mj}$, six mutations in the MJ1271 gene were constructed, and the altered proteins were coexpressed with wild-type MJ1003 in *E. coli*, as described previously (*1*). In aerobic analytical size exclusion chromatography experiments, the six purified $HACN_{Mj}$ variants eluted as complexes with apparent molecular masses of 144−155 kDa. Masses of the large and small subunit monomers are approximately 46 and 18 kDa, respectively. Therefore, these variant forms of $HACN_{Mj}$ form heterotetramers consisting of two large and two small subunits, consistent with wild-type $HACN_{Mj}$, which eluted with a molecular mass of 143 kDa.

Sequence and structural alignments indicated that Arg^{26} in the MJ1271 flexible loop could act similarly to Arg^{580} of mACN (*14*). Zheng et al. reported that replacing Arg^{580} with Lys resulted in a 30-fold increase in K_M for mACN with isocitrate, a decrease in activity, and a loss of tight substrate binding (*15*). The analogous R26K substitution in MJ1271 broadened the enzyme's substrate specificity: the enzyme catalyzed citraconate hydration and the dehydration of 3-isopropylmalate, and the substitution caused a 6-fold increase in K_M for *cis*-homo$_1$-aconitate (Tables 4 and 5). In contrast to Arg, the smaller Lys side chain may not interfere with the hydrophobic γ-chains of the citraconate and 3-isopropylmalate. Consistent with this model, the steady-state kinetic parameters K_M and k_{cat} are almost identical for the R26K-catalyzed hydration of citraconate (315 μM, 8.4 s^{-1}) and maleate (310 μM, 8.1 s^{-1}) (Figure 6).

Replacing the MJ1271 Arg^{26} with Val, the analogous residue in MJ1277, had a greater effect on substrate recognition. This variant catalyzed citraconate and maleate hydration with kinetic rate constants similar to those of the R26K protein; however, 3-isopropylmalate dehydratase activity increased substantially and homoaconitase activity decreased (Table 4 and Figure 6). The R26V variant catalyzed 3-isopropylmalate dehydration with a K_M less than 2-fold higher than the wild-type $IPMI_{Mj}$ and a similar

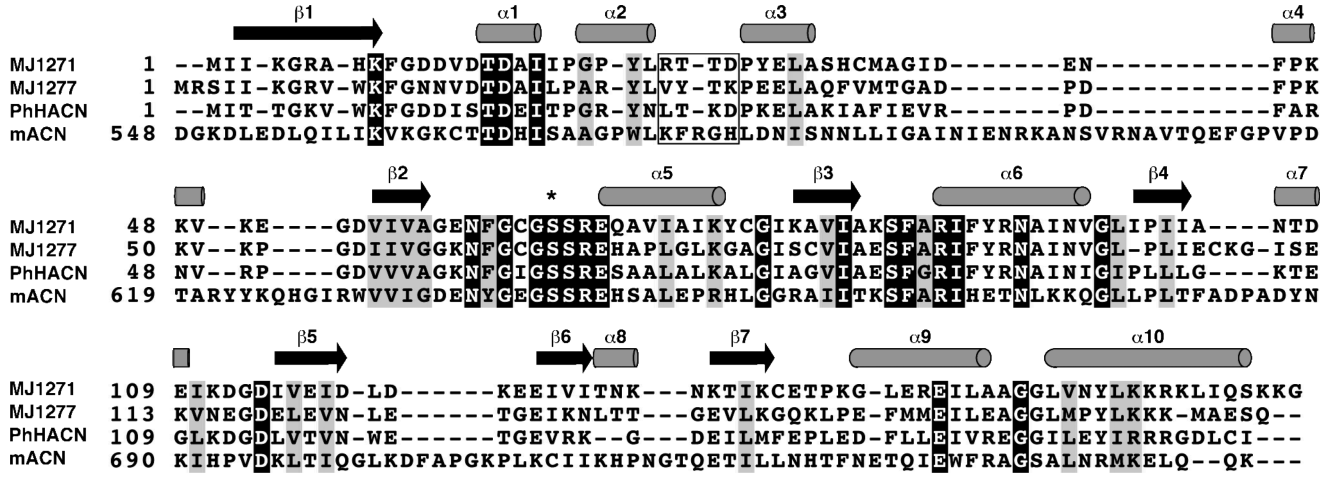


FIGURE 3: Amino acid sequence alignment of the MJ1271, MJ1277, *P. horikoshii* HACN small subunit, and mitochondrial aconitase (domain 4) proteins. The MATRAS program (*47*) aligned structural models of MJ1271 (PDB id 2PKP), PhHACN (1V7L), and mACN (7ACN). A pairwise sequence alignment of the MJ1277 and MJ1271 sequences was manually added to the alignment. Identical amino acid positions are shown in white text on a black background, while similar positions are shown in black on a gray background. β-Strands identified in the MJ1271 structure are shown as black arrows above the sequence, and helices are shown as gray cylinders. A box highlights the flexible loop region between α2 and α3, which includes MJ1271 Arg^{26} and Thr^{27} residues. An asterisk above MJ1271 Ser^{65} identifies the alkoxide base catalyst.

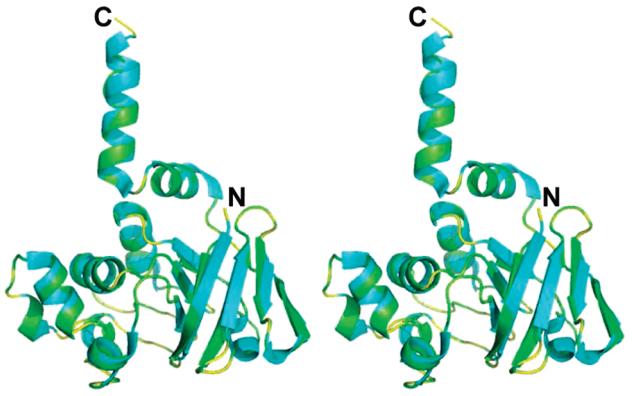


FIGURE 4: Stereoview of the MJ1271 structure (cyan) aligned with the homology model of the MJ1277 protein (green).

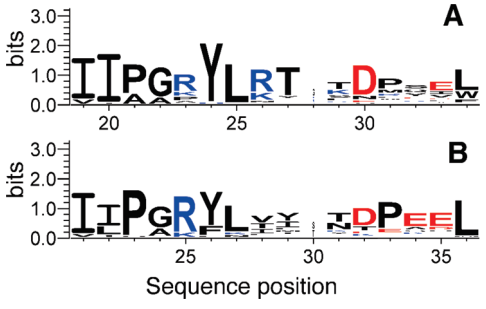


FIGURE 5: Sequence logos show the correlation between the flexible loop region of archaeal small-subunit proteins and their predicted specificities. HACN small subunits contain conserved Arg/Lys^{26} and Thr^{27} residues (part A), while IPMI paralogues contain less polar Val/Thr/Ile^{28} and Tyr/Ile^{29} residues at the homologous positions (part B). Conserved residues from alignments of 26 putative HACN and 23 putative IPMI proteins are represented by single letter amino acid abbreviations, where letter heights (bits of information) correspond to their frequency in the aligned column (*33*).

turnover (1.1 and 1.9 s^{-1}, respectively). In a trade-off, the K_M for *cis*-$(homo)_1$-aconitate, *cis*-$(homo)_2$-aconitate, and *cis*-$(homo)_3$-aconitate increased between 10- and 30-fold for the R26V protein. Therefore, Arg^{26} of MJ1271 plays a key role in homoaconitate substrate recognition, while discriminating against the hydrophobic methyl or isopropyl γ-chains of citraconate and 3-isopropylmalate.

In the conserved loop regions of $HACN_{Mj}$ and $IPMI_{Mj}$, residues adjacent to the critical arginine or valine also differ. Although no other substrate-binding residues have been identified in the corresponding loop of mACN, any change in the loop's orientation or flexibility could affect substrate specificity. To determine the catalytic role of Thr^{27} in $HACN_{Mj}$, the MJ1271 Thr^{27}Ala variant was constructed. In a holoenzyme complex, this T27A variant catalyzed the hydration of citraconate and maleate substrates with a 10-fold higher K_M than wild-type $IPMI_{Mj}$ (Table 4). Also, the K_M values for *cis*-homoaconitate substrates increased 10−20-fold relative to the wild-type $HACN_{Mj}$, which was offset by the variant's increased turnover to produce similar specificity constants (Figure 6). Unlike the R26V variant, the T27A variant had no detectable dehydratase activity with 3-isopropylmalate, probably due to steric hindrance between the Arg^{26} guanidinium group and the substrate's isopropyl group.

A double replacement, MJ1271 $R^{26}V$ $T^{27}Y$, was constructed to remodel the loop region of $HACN_{Mj}$ to resemble the $IPMI_{Mj}$ loop region. The double mutation resulted in a K_M of 489 μM for maleic acid, similar to wild-type $HACN_{Mj}$ and $IPMI_{Mj}$, and a turnover that was close to the Thr^{27}Ala and $IPMI_{Mj}$ proteins' values. However, compared to the single Arg^{26}Val protein, the K_M and k_{cat} values for both citraconate and 3-isopropylmalate increased slightly. The K_M values for *cis*-$(homo)_1$-aconitate, *cis*-$(homo)_2$-aconitate, and *cis*-$(homo)_3$-aconitate increased 20−50-fold compared to wild-type $HACN_{Mj}$. Although the double replacement did result in an increase in turnover for *cis*-$(homo)_1$-aconitate and *cis*-$(homo)_2$-aconitate, all three tricarboxylate substrates had specificity constants that were an order of magnitude lower than those of the wild-type enzyme (Table 5 and Figure 6). The difference in rate constants for citraconate and 3-isopropylmalate between the double replacement and Arg^{26}Val replacement proteins probably results from structural changes in the loop region caused by the threonine to tyrosine substitution, as discussed for MJ1271 T27A above.

A chimeric MJ1271 protein was constructed with the first 28 amino acids of MJ1277 to test whether the holoenzyme's substrate specificity could be completely altered from $HACN_{Mj}$

Table 4: Steady-State Kinetic Parameters for IPMI Substrates

	maleate		citraconate		3-isopropylmalate	
enzyme	K_M (μM)[a]	k_{cat} (s^{-1})	K_M (μM)	k_{cat} (s^{-1})	K_M (μM)	k_{cat} (s^{-1})
$IPMI_{Mj}$[b]	400 ± 50	36	80 ± 20	14	39 ± 7	1.9
MJ1271 $R^{26}V$	230 ± 10	9.9	190 ± 20	5.5	68 ± 5	1.1
MJ1271 $R^{26}K$	310 ± 32	8.1	315 ± 71	8.36	180 ± 30	0.8
MJ1271 $T^{27}A$	3400 ± 200	24.2	920 ± 230	4.1	ND[c]	ND
MJ1271 $R^{26}V$ $T^{27}Y$	489 ± 86	23.1	290 ± 70	9.8	133 ± 37	1.6

[a]Initial rate data for the first-order reactions were fitted with the Michaelis−Menten−Henri equation by nonlinear regression. The mean values for both kinetic parameters are shown, along with the standard error for K_M values. [b]Values for wild-type $IMPI_{Mj}$ were reported previously (*11*). The parameters for $HACN_{Mj}$ in maleate hydratase reactions were $K_M = 330 \pm 50\,\mu M$ and $k_{cat} = 6\,s^{-1}$. $HACN_{Mj}$ has no detectable activity with citraconate or isopropylmalate (*1*). [c]ND, not detected.

Table 5: Steady-State Kinetic Parameters for HACN Substrates

	cis-homo$_1$-aconitate		*cis*-homo$_2$-aconitate		*cis*-homo$_3$-aconitate	
enzyme	K_M (μM)[a]	k_{cat} (s^{-1})	K_M (μM)	k_{cat} (s^{-1})	K_M (μM)	k_{cat} (s^{-1})
$HACN_{Mj}$[b]	22 ± 3	0.75	30 ± 5	0.66	36 ± 6	2.5
MJ1271 $R^{26}V$	220 ± 30	0.48	870 ± 150	5.8	660 ± 170	2.8
MJ1271 $R^{26}K$	135 ± 27	1.43	NT[c]	NT	NT	NT
MJ1271 $T^{27}A$	220 ± 20	2.5	269 ± 68	2.2	650 ± 80	4.1
MJ1271 $R^{26}V$ $T^{27}Y$	460 ± 80	1.7	1600 ± 500	6.6	640 ± 90	1.9

[a]Kinetic parameters were determined as described for Table 4. [b]Values for wild-type $HACN_{Mj}$ were reported previously (*1*). [c]NT, not tested.

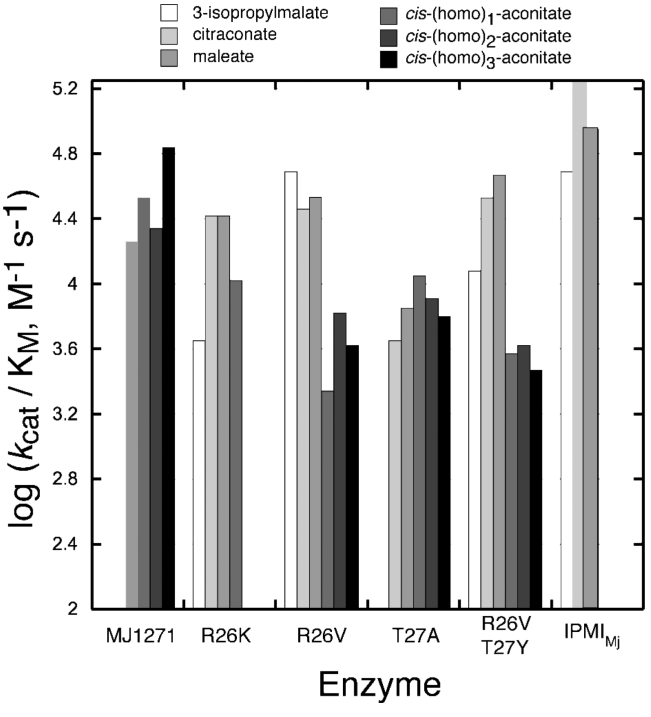


FIGURE 6: Substrate specificities for wild-type MJ1271 (HACN), MJ1271 variants, and wild-type IPMI proteins were estimated from steady-state rate constants using the six substrates shown in the legend above the graph. Together with the MJ1003 large-subunit protein, wild-type MJ1271 catalyzes the hydration of maleate and *cis*-(homo)$_{1-3}$-aconitate substrates with similar specificity constants but has no activity with 3-isopropylmalate or citraconate (*1*). The MJ1271 R26K variant catalyzes 3-isopropylmalate dehydration, as well the hydration of citraconate, maleate, and *cis*-homoaconitate. The R26V variant is promiscuous, with enhanced IPMI activity and reduced HACN activity. The T26A variant is also promiscuous, with uniform specificity constants for both IPMI and HACN reactions. The specificity of the R26V T27Y variant resembles the R26V variant. Wild-type $IPMI_{Mj}$ (MJ1277 and MJ0499) efficiently acts on 3-isopropylmalate, citraconate, and maleate but has no detectable activity with *cis*-(homo)-aconitate analogues (*11*).

to $IPMI_{Mj}$. The chimera interacted with MJ1003, forming a heterotetramer in size exclusion chromatography. Although this association suggested that the chimeric protein folded correctly, no activity was detected for any of the tested substrates after reconstitution. No further experiments were performed to examine the chimeric protein's structure. This inactive protein is reminiscent of the stable MJ1003-MJ1277 heterotetramer, which also lacked catalytic activity (*1*).

In contrast to HACN, the homoaconitate hydratase protein catalyzes only the hydration of *cis*-homoaconitate to (2*R*,3*S*)-homoisocitrate. Sequence alignments and homology models of the *T. thermophilus* homoaconitate hydratase (LysU) protein reveal a flexible loop region with the sequence Y^{21}APFMV but do not provide insight into the mechanistic discrepancy. A chimeric form of MJ1271 was created, replacing the loop region with the corresponding region of the *T. thermophilus* LysU protein, in an effort to alter the reactions catalyzed by the $HACN_{Mj}$ to that of the homoaconitate hydratases. Although the chimeric protein still interacted with MJ1003 and maintained relatively low activity with maleic acid ($k_{cat}/K_M = 1.4 \times 10^3\ M^{-1}\ s^{-1}$), no activity was detected for *cis*-(homo)$_1$-aconitate or for any other tested substrates.

DISCUSSION

The $HACN_{Mj}$ small subunit protein, MJ1271, is homologous to the fourth domain of mACN. Structural alignment of MJ1271 with the relaxed (without substrate) or tense (substrate bound) structures of mACN resulted in an rmsd of 1.9 Å (*Z* scores between 19.0 and 19.2), as indicated by the DALI server (Figure 7). In both the relaxed and tense states of mACN, the α-helical region of domain 4 forms the active site cleft at the interface with domains 1, 2, and 3. In HACN and IPMI proteins, the large subunit contains domains 1−3 and provides cysteine ligands for the [4Fe-4S] cluster. The catalytically important Arg^{580} and Ser^{642} residues, located within the loop regions of mACN domain 4, are considered essential for aconitase activity. Therefore, we predict that the corresponding

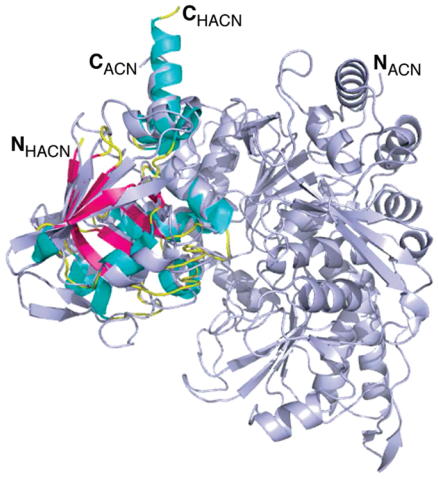


FIGURE 7: Alignment of MJ1271 (cyan and pink) with mACN in the relaxed conformation (gray, PDB id 1AMJ) (*18*). The MJ1271 subunit is homologous to domain 4 of mACN. N- and C-termini are indicated for both proteins.

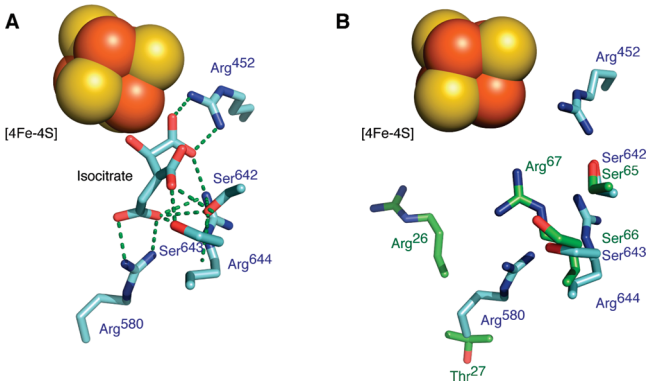


FIGURE 8: Side chains of serine and arginine residues form hydrogen bonds with isocitrate in the porcine mACN active site (part A, PDB id 7ACN). An alignment of the MJ1271 subunit structure with mACN in the relaxed state (part B, PDB id 1AMJ), made by the DaliLite server, shows a substantial shift in the corresponding arginine residues (Arg^{26}/Arg^{580} and Arg^{67}/Arg^{644}). Carbon atoms for the MJ1271 residues are shown in green, while those from mACN are shown in light blue.

regions of MJ1271, including Arg^{26} and Ser^{65}, are positioned at the interface with the MJ1003 protein in the $HACN_{Mj}$ holoprotein.

The C_α backbones of the conserved GSSRE sequences in relaxed mACN and MJ1271 overlap in a structural alignment (Figures 3 and 7), and the homologous serine side chains have similar orientations. However, Arg^{644} in mACN, which stabilizes the Ser^{642} alkoxide, differs in orientation from Arg^{67} in MJ1271 (Figure 8). Also, the GSSRE sequence of mACN resides completely in a loop region, while the GSSRE sequence of MJ1271 extends into α6. Aligning MJ1271 with tense, substrate-bound mACN (PDB id 1AMI) shows a similar alignment of C_α backbones and side chain residues of the GSSRE sequence. Ser^{642} of tense mACN is rotated almost 180° from the relaxed mACN and MJ1271 serine side chains. Therefore, the MJ1271 structure most resembles the relaxed state of mACN.

The Y^{24}LRT residues are close to the GS^{65}SRE loop in MJ1271, indicating that this sequence could interact with the γ-carboxylates of the HACN substrates. Still, the YLRT sequence aligns poorly with the corresponding loop region of mACN that contains Arg^{580}. Instead, Arg^{26} of MJ1271 aligns with Leu^{577} of mACN, and there is an approximately 120° difference in the orientation of the arginine residues (Figure 8). Arg^{26} is oriented perpendicular to the GS^{65}SRE sequence and does not appear to be in line to contact the substrate's γ-carboxylate, as in mACN Arg^{580}. Substrate binding at the mACN active site results in conformational changes 30 Å from the active site; similar changes could reposition MJ1271 Arg^{26}. The orientation of the Y^{24}LRT loop region may also be influenced by interaction with MJ1003. Aconitase proteins require a [4Fe-4S] cluster for substrate binding and orientation; therefore, we did not attempt to cocrystallize the MJ1271 protein with ligands. Structures of the MJ1003-MJ1271 ternary complex with substrate will be required to accurately assess the orientation of Arg^{26} of MJ1271 in $HACN_{Mj}$.

Site-directed mutagenesis of MJ1271 suggests that Arg^{26} is a critical residue for $HACN_{Mj}$ substrate specificity. Replacing this residue with valine or lysine dramatically increased the Michaelis constants for *cis*-$(homo)_1$-aconitate, *cis*-$(homo)_2$-aconitate, and *cis*-$(homo)_3$-aconitate compared to the wild-type $HACN_{Mj}$. At the same time, these variants had broader substrate specificity: the Arg^{26}Val substitution resulted in K_M and k_{cat} values comparable to those of $IPMI_{Mj}$, suggesting that Val^{28} of the MJ1277 Y^{26}LVY sequence is necessary to accommodate the hydrophobic γ-chains of citraconate and dimethylcitraconate.

The residues adjacent to mACN Arg^{580} have not been shown to function in the aconitase mechanism. The adjacent residue to MJ1271 Arg^{26} is Thr^{27}, which corresponds to Tyr^{29} of MJ1277. Replacing Thr^{27} with alanine disrupted the activity of $HACN_{Mj}$ for all substrates, most notably by increasing the K_M for maleate 10-fold compared to wild-type $HACN_{Mj}$. The MJ1271 structure shows the Thr^{27} side chain oriented away from the predicted active site. This residue is highly conserved in HACN small subunit sequences, but additional mutagenesis experiments will be required to distinguish polar interactions with the hydroxyl group from steric effects on conformation. Corresponding residues in this position are overwhelmingly aromatic (Tyr) or β-branched amino acids (Thr or Ile), which favor an extended conformation.

The homoisocitrate dehydrogenase (HICDH) enzyme evolved from isopropylmalate dehydrogenase through a similar series of active site amino acid substitutions. HICDH binds products of HACN or homoaconitate hydratase and catalyzes their oxidative decarboxylation, forming 2-oxoadipate and longer chain analogues. Members of this family catalyze the $NAD(P)^+$-dependent decarboxylation of α-hydroxy acids with different γ-chain substituents (*37*). Both isocitrate dehydrogenase and homoisocitrate dehydrogenase recognize α-hydroxytricarboxylates, while isopropylmalate dehydrogenase (IPMDH) recognizes α-hydroxydicarboxylates. Random mutagenesis identified isocitrate dehydrogenase mutants with low levels of IPMDH activity (*38*). Site-directed mutagenesis of the *T. thermophilus* HICDH identified Arg^{85} as the key specificity determinant: an R85V variant was transformed into an IPMDH (*39*). Subsequently, a crystal structure model of the *T. thermophilus* HICDH identified the Arg^{85} guanidinium group near the putative substrate-binding site, poised to form an ion pair with the γ-carboxylate of isocitrate or homoisocitrate (*40*). The *P. horikoshii* HICDH sequence, however, lacks a basic residue homologous to Arg^{85} although that enzyme's substrate specificity closely resembles that of *T. thermophilus* HICDH (*41*). $HACN_{Mj}$ and HICDH enzymes from modern, currently living species are highly specific for tricarboxylate substrates. Yet a small number of amino acid substitutions converts them into promiscuous enzymes (*42*, *43*). Homologous

proteins in ancestral organisms could have had broader specificies. From these examples, we infer that the evolutionary potential for catalytic promiscuity fostered the diversification of both the isomerase and oxidative decarboxylase families (*44*, *45*), enabling the evolution of the diverse 2-oxo acid elongation pathways (*46*).

Although the substrate specificity of the MJ1271 protein was altered to accommodate the substrates of MJ1277, activity was not completely lost for tricarboxylate substrates, forming a promiscuous enzyme. The interacting large subunits could also contribute to substrate specificity by modulating conformational changes in the small subunits and by interacting with the substrates' α- and *β*-carboxylate groups. However, the steady-state kinetic analysis of $HACN_{Mj}$ and $IPMI_{Mj}$, the crystallography and homology modeling of MJ1271 and MJ1277, and the site-directed mutagenesis of MJ1271 all indicate that the conserved Y^{24}LRT and Y^{26}LVY sequences are consistent indicators of protein specificity in the Archaea. Therefore, uncharacterized proteins with the Y^{24}LRT sequence are likely HACNs, while proteins with the YLV(Y/I/M) sequence are IPMIs, validating the motifs deduced in Figure 5. These studies should facilitate annotation of these uncharacterized proteins.

The $HACN_{Mj}$, $IPMI_{Mj}$, and mACN proteins catalyze the isomerization of α-hydroxy acids to form *β*-hydroxy acids. Given the similarity of the homoaconitate hydratases to these enzymes, it is surprising the initial dehydration of (*R*)-homocitrate has not been observed. We replaced the region of MJ1271 containing the Y^{24}LRT sequence with the corresponding sequence from the *T. thermophilus* HACN in an attempt to alter the catalytic mechanism to accommodate only the hydration of *cis*-homoaconitate to (2*R*,3*S*)-homoisocitrate. Although the MJ1003 and MJ1271/LysU variant subunits interacted to form a heterotetramer with malease activity, no activity was observed with any of the other $HACN_{Mj}$ or $IPMI_{Mj}$ substrates. Preserving the position and orientation of residues in this flexible loop may be essential for reproducing activity with *cis*-homoaconitate. Therefore, a crystal structure of HACN with bound substrate, as well as mutagenesis of homoaconitate hydratase active site amino acids, will be required to resolve this difference in reactivity.

Although most crenarchaea use the α-aminoadipate and isopropylmalate pathways for lysine and leucine biosyntheses, respectively, they appear to only have one HACN/IPMI homologue with a conserved YL(K/V)Y sequence in the flexible loop region. These putative promiscuous enzymes are expected to accept a broad pool of hydroxy acid substrates. No crenarchaeal homologues have been purified and characterized, although the results of the MJ1271 mutagenesis show that mutating Arg^{26} of YLRT to either lysine or valine allows for the recognition of both $HACN_{MJ}$ and $IPMI_{MJ}$ substrates. Therefore, future studies will involve the characterization, crystallization, and mutagenesis of these crenarchaeal proteins to determine how ancestral hydrolyases recognize their substrates.

ACKNOWLEDGMENT

We thank the staff at the SPring-8 beamlines BL12B2 and BL26B2 for use of their excellent facilities and assistance with X-ray data collection.

REFERENCES

1. Drevland, R. M., Jia, Y., Palmer, D. R. J., and Graham, D. E. (2008) Methanogen homoaconitase catalyzes both hydrolyase reactions in coenzyme B biosynthesis. *J. Biol. Chem. 283*, 28888–28896.
2. Gruer, M. J., Artymiuk, P. J., and Guest, J. R. (1997) The aconitase family: three structural variations on a common theme. *Trends Biochem. Sci. 22*, 3–6.
3. Weidner, G., Steffan, B., and Brakhage, A. A. (1997) The *Aspergillus nidulans lysF* gene encodes homoaconitase, an enzyme involved in the fungus-specific lysine biosynthesis pathway. *Mol. Gen. Genet. 255*, 237–247.
4. Jia, Y., Tomita, T., Yamauchi, K., Nishiyama, M., and Palmer, D. R. J. (2006) Kinetics and product analysis of the reaction catalysed by recombinant homoaconitase from *Thermus thermophilus*. *Biochem. J. 396*, 479–485.
5. Grimek, T. L., and Escalante-Semerena, J. C. (2004) The *acnD* genes of *Shewenella oneidensis* and *Vibrio cholerae* encode a new Fe/S-dependent 2-methylcitrate dehydratase enzyme that requires *prpF* function in vivo. *J. Bacteriol. 186*, 454–462.
6. Alhapel, A., Darley, D. J., Wagener, N., Eckel, E., Elsner, N., and Pierik, A. J. (2006) Molecular and functional analysis of nicotinate catabolism in *Eubacterium barkeri*. *Proc. Natl. Acad. Sci. U.S.A. 103*, 12341–12346.
7. Blank, L., Green, J., and Guest, J. R. (2002) AcnC of *Escherichia coli* is a 2-methylcitrate dehydratase (PrpD) that can use citrate and isocitrate as substrates. *Microbiology 148*, 133–146.
8. Beinert, H., Kennedy, M. C., and Stout, C. D. (1996) Aconitase as iron-sulfur protein, enzyme, and iron-regulatory protein. *Chem. Rev. 96*, 2335–2374.
9. Kennedy, M. C., and Beinert, H. (2007) Aconitase, in Biological Inorganic Chemistry: Structure and Reactivity (Bertini, I., Gray, H. B., Stiefel, E. I., and Valentine, J. S., Eds.) pp 209−215, University Science Books, Sausalito, CA.
10. Gross, S. R., Burns, R. O., and Umbarger, H. E. (1963) The biosynthesis of leucine. II. The enzymic isomerization of *β*-carboxy-*β*-hydroxyisocaproate and α-hydroxy-*β*-carboxyisocaproate. *Biochemistry 2*, 1046–1052.
11. Drevland, R. M., Waheed, A., and Graham, D. E. (2007) Enzymology and evolution of the pyruvate pathway to 2-oxobutyrate in *Methanocaldococcus jannaschii*. *J. Bacteriol. 189*, 4391–4400.
12. Xu, H., Andi, B., Qian, J., West, A. H., and Cook, P. F. (2006) The α-aminoadipate pathway for lysine biosynthesis in fungi. *Cell Biochem. Biophys. 46*, 43–64.
13. Strassman, M., and Ceci, L. N. (1966) Enzymatic formation of *cis*-homoaconitic acid, an intermediate in lysine biosynthesis in yeast. *J. Biol. Chem. 241*, 5401–5407.
14. Lauble, H., Kennedy, M. C., Beinert, H., and Stout, C. D. (1992) Crystal structures of aconitase with isocitrate and nitroisocitrate bound. *Biochemistry 31*, 2735–2748.
15. Zheng, L., Kennedy, M. C., Beinert, H., and Zalkin, H. (1992) Mutational analysis of active site residues in pig heart aconitase. *J. Biol. Chem. 267*, 7895–7903.
16. Yasutake, Y., Yao, M., Sakai, N., Kirita, T., and Tanaka, I. (2004) Crystal structure of the *Pyrococcus horikoshii* isopropylmalate isomerase small subunit provides insight into the dual substrate specificity of the enzyme. *J. Mol. Biol. 344*, 325–333.
17. Lombo, T., Takaya, N., Miyazaki, J., Gotoh, K., Nishiyama, M., Kosuge, T., Nakamura, A., and Hoshino, T. (2004) Functional analysis of the small subunit of the putative homoaconitase from *Pyrococcus horikoshii* in the *Thermus* lysine biosynthetic pathway. *FEMS Microbiol. Lett. 233*, 315–324.
18. Lauble, H., and Stout, C. D. (1995) Steric and conformational features of the aconitase mechanism. *Proteins 22*, 1–11.
19. Horton, R. M., Cai, Z., Ho, S. N., and Pease, L. R. (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *BioTechniques 8*, 528–535.
20. Helgadóttir, S., Rosas-Sandoval, G., Söll, D., and Graham, D. E. (2007) Biosynthesis of phosphoserine in the *Methanococcales*. *J. Bacteriol. 189*, 575–582.
21. Sugahara, M., and Miyano, M. (2002) Development of high-throughput automatic protein crystallization and observation system. *Tanpakushitsu Kakusan Koso 47*, 1026–1032.
22. Ueno, G., Hirose, R., Ida, K., Kumasaka, T., and Yamamoto, M. (2004) Sample management system for a vast amount of frozen crystals at SPring-8. *J. Appl. Crystallogr. 37*, 867–873.
23. Ueno, G., Kanda, H., Kumasaka, T., and Yamamoto, M. (2005) Beamline Scheduling Software: administration software for automatic operation of the RIKEN structural genomics beamlines at SPring-8. *J. Synchrotron Radiat. 12*, 380–384.
24. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
25. Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol. 33*, 491–497.

26. Navaza, J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A 50*, 157–163.
27. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905–921.
28. Jones, T. A. (1985) Interactive computer graphics: FRODO. *Methods Enzymol. 115*, 157–171.
29. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126–2132.
30. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *PROCHECK*: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283–291.
31. Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res. 31*, 3381–3385.
32. Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol. 302*, 205–217.
33. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res. 14*, 1188–1190.
34. Lauble, H., Kennedy, M. C., Beinert, H., and Stout, C. D. (1994) Crystal structures of aconitase with *trans*-aconitate and nitrocitrate bound. *J. Mol. Biol. 237*, 437–451.
35. Krebs, H. A., and Eggleston, L. V. (1944) Micro-determination of isocitric and *cis*-aconitic acids in biological material. *Biochem. J. 38*, 426–437.
36. Lloyd, S. J., Lauble, H., Prasad, G. S., and Stout, C. D. (1999) The mechanism of aconitase: 1.8 Å resolution crystal structure of the S642A:citrate complex. *Protein Sci. 8*, 2655–2662.
37. Aktas, D. F., and Cook, P. F. (2009) A lysine-tyrosine pair carries out acid-base chemistry in the metal ion-dependent pyridine dinucleotide-linked α-hydroxyacid oxidative decarboxylases. *Biochemistry 48*, 3565–3577.
38. Doyle, S. A., Fung, S.-Y. F., and Koshland, D. E. (2000) Redesigning the substrate specificity of an enzyme: isocitrate dehydrogenase. *Biochemistry 39*, 14348–14355.
39. Miyazaki, J., Kobashi, N., Nishiyama, M., and Yamane, H. (2003) Characterization of homoisocitrate dehydrogenase involved in lysine biosynthesis of an extremely thermophilic bacterium, *Thermus thermophilus* HB27, and evolutionary implication of beta-decarboxylating dehydrogenase. *J. Biol. Chem. 278*, 1864–1871.
40. Miyazaki, J., Asada, K., Fushinobu, S., Kuzuyama, T., and Nishiyama, M. (2005) Crystal structure of tetrameric homoisocitrate dehydrogenase from an extreme thermophile, *Thermus thermophilus*: involvement of hydrophobic dimer-dimer interaction in extremely high thermotolerance. *J. Bacteriol. 187*, 6779–6788.
41. Miyazaki, K. (2005) Bifunctional isocitrate-homoisocitrate dehydrogenase: a missing link in the evolution of beta-decarboxylating dehydrogenase. *Biochem. Biophys. Res. Commun. 331*, 341–346.
42. O'Brien, P. J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol. 6*, R91–R105.
43. Copley, S. D. (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr. Opin. Chem. Biol. 7*, 265–272.
44. Schmidt, D. M. Z., Mundorff, E. C., Dojka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., and Gerlt, J. A. (2003) Evolutionary potential of $(\beta/\alpha)_8$-barrels: functional promiscuity produced by single substitutions in the enolase superfamily. *Biochemistry 42*, 8387–8393.
45. Khersonsky, O., Roodveldt, C., and Tawfik, D. S. (2006) Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr. Opin. Chem. Biol. 10*, 498–508.
46. Graham, D. E., and White, R. H. (2002) Elucidation of methanogenic coenzyme biosyntheses: from spectroscopy to genomics. *Nat. Prod. Rep. 19*, 133–147.
47. Kawabata, T. (2003) MATRAS: a program for protein 3D structure comparison. *Nucleic Acids Res. 31*, 3367–3369.