Supplementary Materials

Spectroscopic properties of synthesized compounds

Trimethyl *cis*-homoaconitate 5a (trimethyl (1*Z*)-but-1-ene-1,2,4-tricarboxylate). Yellow oil. ^{1}H NMR (400 MHz, $CDCl_3$, δ): 5.90 (s, CH=C, 1H), 3.83 (s, OCH_3, 3H), 3.73 (s, OCH_3, 3H), 3.70 (s, OCH_3, 3H), 2.69 (t, *J*= 7.5 Hz, CH=C-CH_2, 2H), 2.55, (t, *J*= 7.3 Hz, CH_2- CH_2-CO-, 2H).

^{13}C NMR (100 MHz, $CDCl_3$+TMS, δ): 172.2 (OC=O), 168.6 (OC=O), 165.3 (OC=O), 147.8 (C-2), 120.9 (C-1), 52.5 (OMe), 51.95 (OMe), 51.90 (OMe), 31.5 (CH_2), 29.2 (CH_2).

GC-MS (CI+, CH_4) *m*/*z* (% relative intensity, ion): 139 (100%), 167 (71%), 199 (47%), 140 (8%), 259 (2%, M + $C_2H_5^+$). GC-MS/MS (199 *m*/z): 139 (100%), 167 (23%), 111 (6%), 127 (5%).

HRMS-CI (*m*/*z*): M^- calcd for $C_{10}H_{14}O_6$ 230.0790; found, 230.0786.

NOE difference spectroscopy: Irradiation at 5.90 ppm shows enhanced resonances at 3.73, 2.69 and 2.55 ppm. Irradiation at 2.69 ppm shows enhanced resonances at 5.90 and 3.83 ppm. High-resolution ^{1}H-NMR indicates the vinylic proton is coupled to a methylene group (*J*=1.3 Hz).

Sodium *cis*-homoaconitate.
^{1}H NMR (400 MHz, D_2O+TSP, δ): 5.74 (s, CH=C, 1H), 2.59 (t, *J*= 7.6 Hz, CH_2, 2H), 2.46, (t, *J*= 7.4 Hz, CH_2, 2H).

LC-MS/MS (ESI-, CID): Fragmentation of the 187 *m/z* $[M-H]^-$ molecular ion gave product ions with peaks at 143 *m/z* and 125 *m/z*.

Trimethyl *trans*-homoaconitate (trimethyl (1*E*)-but-1-ene-1,2,4-tricarboxylate).
^{1}H NMR (400 MHz, $CDCl_3$, δ): 6.82 (s, CH=C, 1H), 3.82 (s, OCH_3, 3H), 3.78 (s, OCH_3, 3H), 3.67 (s, OCH_3, 3H), 3.09 (t, *J*= 7.7 Hz, CH=C-CH_2, 2H), 2.54, (t, *J*= 7.8 Hz, CH_2- CH_2-CO-, 2H).

GC-MS (CI+, CH_4) *m*/*z* (% relative intensity, ion): 139 (100%), 167 (85%), 199 (55%), 140 (7%), 231 (6%, MH^+), 259 (6%, M + $C_2H_5^+$), 111 (3%).

GC-MS/MS (167 *m*/*z*): 139 (100%), 123 (21%), 110 (9%), 108 (6%).

NOE difference spectroscopy: Irradiation at 6.82 ppm shows no enhanced resonances. Irradiation at 3.09 ppm showed enhanced resonance at 2.54 ppm.

Sodium *trans*-homoaconitate.
^{1}H NMR (400 MHz, D_2O+TMS, δ): 6.51 (s, CH=C, 1H), 2.74 (t, *J*= 8.1 Hz, CH_2, 2H), 2.54, (t, *J*= 7.9 Hz, CH_2, 2H).

Dimethyl 2-oxoadipate (dimethyl 2-oxohexanoate)**.**
^{1}H NMR (400 MHz, $CDCl_3$+TMS, δ): 3.88 (s, OCH_3, 3H), 3.68 (s, OCH_3, 3H), 2.94 (t, *J*= 7.1 Hz, CH_2, 2H), 2.42-2.37 (m, 2H), 1.98 (quintet, 2H).

Trimethyl *cis*-homo$_2$-aconitate (trimethyl (1*Z*)-pent-1-ene-1,2,5-tricarboxylate).
^{1}H NMR (400 MHz, $CDCl_3$+TMS, δ): 5.80 (s, CH=C, 1H), 3.83 (s, OCH_3, 3H), 3.73 (s, OCH_3, 3H), 3.68 (s, OCH_3, 3H), 2.44 (m, 4H), 1.84, (quintet, 2H).

^{13}C NMR (100 MHz, $CDCl_3$+TMS, δ): 173.2 (OC=O), 169.0 (OC=O), 165.3 (OC=O), 149.4 (C-2), 120.2 (C-1), 52.4 (OMe), 51.9 (OMe), 51.6 (OMe), 33.5 (CH_2), 32.8 (CH_2), 22.3 (C-4).

GC-MS (CI+, CH_4) *m*/*z* (% relative intensity, ion): 152 (100%), 245 (51%, MH^+), 124 (42%), 213 (34%, $[M\text{-}OCH_3]^+$), 180 (22%), 185 (14%, $[M\text{-}C(=O)OCH_3]^+$).

GC-MS/MS (213 *m*/z): 152 (100%), 185 (76%). GC-MS/MS (180 *m*/z): 152 (100%), 124 (15%).

HRMS-CI (*m*/*z*): MH^+ calcd for $C_{11}H_{17}O_6$ 245.1025; found, 245.1027.

Sodium *cis*-homo$_2$-aconitate.
^{1}H NMR (400 MHz, D_2O+TSP, δ): 5.75 (s, CH=C, 1H), 2.47-2.34 (m, 4H), 1.77, (quintet, 2H).

Trimethyl *trans*-homo$_2$-aconitate (trimethyl (1*E*)-pent-1-ene-1,2,5-tricarboxylate).
^{1}H NMR (400 MHz, $CDCl_3$+TMS, δ): 6.79 (s, CH=C, 1H), 3.81 (s, OCH_3, 3H), 3.77 (s, OCH_3, 3H), 3.67 (s, OCH_3, 3H), 2.85 (t, *J*= 7.6 Hz, CH=C-C*H*$_2$, 2H), 2.44-2.34 (m, 4H), 1.83 (quintet, 2H).

GC-MS (CI+, CH_4) *m*/*z* (% relative intensity, ion): 152 (100%), 245 (74%, MH^+), 124 (34%), 180 (28%), 213 (22%, $[M\text{-}OCH_3]^+$), 185 (14%, $[M\text{-}C(=O)OCH_3]^+$).

HRMS-CI (*m*/*z*): MH^+ calcd for $C_{11}H_{17}O_6$ 245.1025; found, 245.1023.

Sodium *trans*-homo$_2$-aconitate.
^{1}H NMR (400 MHz, D_2O+TSP, δ): 6.39 (s, CH=C, 1H), 2.26-2.15 (m, 4H), 1.63, (quintet, 2H).

Dimethyl 2-oxopimelate (dimethyl 2-oxoheptanedioate).
^{1}H NMR (400 MHz, $CDCl_3$+TMS, δ): 3.92 (s, OCH_3, 3H), 3.87 (s, OCH_3, 3H), 2.94 (t, *J*= 6.9 Hz, CH_2, 2H), 2.36-2.32 (m, 2H), 1.70-1.65 (m, 4H).

Trimethyl *cis*-homo$_3$-aconitate (trimethyl (1*Z*)-hex-1-ene-1,2,6-tricarboxylate). ^{1}H NMR (400 MHz, $CDCl_3$+TMS, δ): 5.84 (s, CH=C, 1H), 3.83 (s, OCH_3, 3H), 3.73 (s, OCH_3, 3H), 3.67 (s, OCH_3, 3H), 2.38 (t, *J*=6.9 Hz, 2H), 2.35 (t, *J*=7.3 Hz, 2H), 1.68 (quintet, 2H), 1.55 (quintet, 2H).

^{13}C NMR (100 MHz, $CDCl_3$+TMS, δ): 173.7 (OC=O), 169.2 (OC=O), 165.3 (OC=O), 150.2 (C-2), 119.5 (C-1), 52.3 (OMe), 51.8 (OMe), 51.6 (OMe), 34.0 (CH_2), 33.6 (CH_2), 26.4 (CH_2), 24.1 (CH_2).

GC-MS (CI+, CH_4) *m*/*z* (% relative intensity, ion): 227 (100%, $[M\text{-}OCH_3]^+$), 199 (51%, $[M\text{-}C(=O)OCH_3]^+$), 287 (44%, $[M+C_2H_5^+]$), 135 (23%), 195 (20%), 167 (12%), 299 (10%, $[M+C_3H_5^+]$), 259 (6%, MH^+).

HRMS-CI (*m*/*z*): MH^+ calcd for $C_{12}H_{19}O_6$ 259.1182; found, 259.1180.

Trimethyl *trans*-homo$_3$-aconitate (trimethyl (1*E*)-hex-1-ene-1,2,6-tricarboxylate).

^{1}H NMR (400 MHz, $CDCl_3$+TMS, δ): 6.76 (s, CH=C, 1H), 3.81 (s, OCH_3, 3H), 3.77 (s, OCH_3, 3H), 3.67 (s, OCH_3, 3H), 2.36-2.31 (m, 4H), 1.71-1.63 (m, 4H).

GC-MS (CI+, CH_4) *m*/*z* (% relative intensity, ion): 227 (100%, $[M\text{-}OCH_3]^+$), 199 (78%, $[M\text{-}C(=O)OCH_3]^+$), 287 (60%, $[M+C_2H_5^+]$), 259 (50%, MH^+), 135 (42%), 195 (30%), 167 (24%), 299 (10%, $[M+C_3H_5^+]$).

HRMS-CI (*m*/*z*): MH^+ calcd for $C_{12}H_{19}O_6$ 259.1182; found, 259.1179.

***cis*-Homo$_3$-aconitic acid (free acid).**
^{1}H NMR (400 MHz, D_2O+TSP, δ): 5.94 (s, CH=C, 1H), 2.45-2.39 (m, 4H), 1.69-1.61 (m, 2H), 1.59-1.53 (m, 2H).

***trans*-Homo$_3$-aconitic acid (free acid).**
^{1}H NMR (400 MHz, D_2O+TSP, δ): 6.75 (s, CH=C, 1H), 2.46-2.38 (m, 4H), 1.68-1.58 (m, 2H), 1.56-1.48 (m, 2H).

Dimethyl 2-oxosuberate (dimethyl 2-oxooctanedioate).
^{1}H NMR (400 MHz, $CDCl_3$+TMS, δ): 3.92 (s, OCH_3, 3H), 3.87 (s, OCH_3, 3H), 2.86 (t, *J*= 7.4 Hz, 2H), 2.34-2.28 (m, 2H), 1.69-1.61 (m, 4H), 1.41-1.34 (m, 2H).

GC-MS (CI+, CH_4) *m*/*z* (% relative intensity, ion): 125 (100%), 217 (50%, MH^+), 157 (42%, $[M\text{-}C(=O)OCH_3]^+$), 185 (40%, $[M\text{-}OCH_3]^+$), 245 (18%, $[M+C_3H_5^+]$).

2-Oxosuberic acid (2-oxooctanedioic acid).
^{1}H NMR (400 MHz, D_2O+TSP adjusted to pD = 7 with sodium deuteroxide, δ): 2.76 (t, *J*=7.2 Hz, 2H), 2.45 (t, *J*=7.4 Hz, 2H), 1.58 (m, 4H), 1.34 (pentet, 2H).

^{13}C NMR (100 MHz, D_2O+TSP adjusted to pD = 7 with sodium deuteroxide, δ): 208.0 (C-2), 182.7 (C-8), 171.1 (C-1), 39.2, 36.5, 28.3, 25.3, 22.7 (C-3 to C-7).

Trimethyl *cis*-homo$_4$-aconitate (trimethyl (1*Z*)-hept-1-ene-1,2,7-tricarboxylate).

^{1}H NMR (400 MHz, $CDCl_3$+TMS, δ): 5.82 (s, CH=C, 1H), 3.83 (s, OCH_3, 3H), 3.72 (s, OCH_3, 3H), 3.67 (s, OCH_3, 3H), 2.37 (t, *J*=7.1 Hz), 2.32 (t, *J*=7.4 Hz), 1.65 (quintet, 2H), 1.53 (quintet, 2H), 1.38 (m, 2H).

^{13}C NMR (100 MHz, $CDCl_3$+TMS, δ): 173.9 (OC=O), 169.3 (OC=O), 165.4 (OC=O), 150.7 (C-2), 119.4 (C-1), 52.3 (OMe), 51.8 (OMe), 51.5 (OMe), 34.2 (CH_2) 33.9 (CH_2), 28.4 (CH_2), 26.7 (CH_2), 24.6 (CH_2).

GC-MS (CI+, CH_4) *m/z* (% relative intensity, ion): 241 (100%, $[M\text{-}OCH_3]^+$), 213 (55%, $[M\text{-}C(=O)OCH_3]^+$), 301 (50%, $[M+C_2H_5^+]$), 149 (30%), 181 (15%), 209 (12%), 313 (10%, $[M+C_3H_5^+]$), 273 (5%, MH^+).

HRMS-CI (*m/z*): MH^+ calcd for $C_{13}H_{21}O_6$ 273.1338; found, 273.1340.

***cis*-Homo$_4$-aconitic acid (free acid).**
^{1}H NMR (400 MHz, D_2O+TSP, δ): 5.92 (s, CH=C, 1H), 2.39 (m, 4H), 1.62 (quintet, 2H), 1.53 (quintet, 2H), 1.37 (quintet, 2H).

***(R)*-Homocitrate.**
LC-MS/MS (ESI-, CID): Fragmentation of the 205 *m/z* $[M\text{-}H]^-$ molecular ion gave product ions with peaks at 187 *m/z* ($[M\text{-}OH]^-$), 143 *m/z* ($[M\text{-}OH\text{-}CO_2]^-$) and 125 *m/z* ($[M\text{-}OH\text{-}CO_2\text{-}H_2O]^-$.

Supplementary Table 1: Oligodeoxyribonucleotide primers

Primer name	Sequence[1]
5MJ1596BN	GGTGGATCCAT*ATG*ATGAAGGTGTGTG
3MJ1596B	GGTGGATCCTCAATATCCCTTTAACTTC

[1] Restriction sites are underlined and the initiator codon is shown in italics.

Supplementary Table 2: List of plasmids and microorganisms

Strain or plasmid	Description	Reference or source
Methanocaldoccus jannaschii JAL-1	Wild-type	DSM 2661
Escherichia coli		
DH5α	General cloning host	Invitrogen
BL21 (DE3)	Protein expression host	Novagen
ArcticExpress (DE3)-RIL	Host for low-temperature protein expression with chaperones	Stratagene
Plasmids		
pET-19b	Expression vector for proteins with an N-terminal decahistidine tag	Novagen
pET-43.1c	Expression vector for proteins with an N-terminal NusA fusion and hexahistidine tag	Novagen
pDG141	MJ1003 cloned into NdeI/KpnI sites of pCDF-Duet1	(1)
pDG160	MJ1271 cloned into NcoI/BamHI sites of pET-19b	(1)
pDG162	MJ1277 cloned into NcoI/BamHI sites of pET-19b	(1)
pDG130	MJ1596 PCR product obtained using primers 5MJ1596BN and 3MJ1596B cloned into BamHI site of pET-43.1c.	This study
pDG131	pDG130 with the 1659 bp NdeI fragment deleted.	This study
pDG382	MJ1596 PCR producted cloned into NdeI/BamHI sites of pET-19b	This study
pT7-MJ1596	MJ1596 cloned in expression vector pT7-7	(2)

Phylogenetic analysis of archaeal small-subunit hydro-lyase homologs

An alignment of 34 homologs of the small IPMI subunit was prepared using the T-Coffee program (ver. 5.57) (3). From the full alignment 164 positions that were deemed to be confidently aligned were chosen for phylogenetic analysis. The phylogeny was inferred using the proml program or the protdist and neighbor programs (with 100 bootstrap replicates) from the Phylip package (ver. 3.66) (4). Both programs used the Jones-Taylor-Thornton model of amino acid changes and assumed a γ-distribution of rates (α=2.4) approximated by three states.

Sequences of small subunit homologs (and their RefSeq accession numbers) were from *Archaeoglobus fulgidus* (NP_070589.1 and NP_069463.1), *Caldivirga maquilingensis* (YP_001540990.1), *Metallosphaera sedula* (YP_001190928.1), Candidatus *Methanoregula boonei* (YP_001403761.1 and YP_001405233.1), *Methanobrevibacter smithii* (YP_001273420.1 and YP_001273872.1), *Methanocaldococcus jannaschii* (NP_248267.1 and NP_248273.1), *Methanococcoides burtonii* (YP_566232.1 and YP_566284.1) *Methanococcus maripaludis* (NP_987501.1 and NP_987256.1), *Methanococcus vannielii* (YP_001323877.1 and YP_001323653.1), *Methanocorpusculum labreanum* (YP_001029677.1 and YP_001030049.1), *Methanosaeta thermophila* (ZP_01153511.1 and ZP_01153233.1), *Methanosarcina acetivorans* (NP_618624.1, NP_616162.1 and NP_615175.1), *Methanosphaera stadtmanae* (YP_447420.1 and YP_448498.1), *Methanospirillum hungatei* (YP_503239.1and YP_503785.1), *Methanothermobacter thermautotrophicus* (NP_275966.1 and NP_276503.1), *Pyrobaculum aerophilum* (NP_559684.1), *Pyrococcus furiosus* (NP_579409.1 and NP_578668.1), *Sulfolobus solfataricus* (NP_343817.1), *Thermococcus kodakarensis* (YP_182694.1). For each genome with multiple paralogs, the predicted HACN subunit accession number is listed first, followed by the IPMI subunit.

References for Supplementary Materials

1. Drevland, R. M., Waheed, A., and Graham, D. E. (2007) *J. Bacteriol.* **189**, 4391-4400
2. Howell, D. M., Graupner, M., Xu, H., and White, R. H. (2000) *J. Bacteriol.* **182**, 5013-5016
3. Notredame, C., Higgins, D. G., and Heringa, J. (2000) *J. Mol. Biol.* **302**, 205-217
4. Felsenstein, J. (2005) PHYLIP (phylogeny inference package). 3.67 Ed., Department of Genetics, University of Washington, Seattle