Supporting Information

Characterization of Choline Trimethylamine-Lyase Expands the Chemistry of Glycyl Radical Enzymes

Smaranda Craciun, Jonathan Marks, and Emily P. Balskus*

Department of Chemistry & Chemical Biology, Harvard University,
Cambridge, Massachusetts 02138, USA

*Correspondence to: balskus@chemistry.harvard.edu.

Supporting Information

1. General materials and methods

All chemicals and solvents were purchased from Sigma-Aldrich, except where otherwise noted. Luria-Bertani Lennox (LB) medium was purchased from EMD Millipore, d_9-trimethylamine (TMA) from CDN Isotopes, choline chloride-1-^{13}C from Sigma-Aldrich, and d_4-acetaldehyde and NMR solvents from Cambridge Isotope Laboratories. Methanol and water used for liquid chromatography-mass spectrometry (LC-MS) were B&J Brand high-purity solvents (Honeywell Burdick & Jackson).

Oligonucleotide primers were synthesized by Integrated DNA Technologies. Recombinant plasmid DNA was purified with a Qiaprep Kit from Qiagen. Gel extraction of DNA fragments and restriction endonuclease clean up were performed using an Illustra GFX PCR DNA and Gel Band Purification Kit from GE Healthcare. DNA sequencing was performed by Genewiz or Beckman Coulter Genomics. Multiple sequence alignments were performed with Clustal W,[1] and the results analyzed with Geneious Pro 6.1.6.[2]

Nickel-nitrilotriacetic acid agarose (Ni-NTA) resin was purchased from Qiagen. SDS-PAGE gels were purchased from BioRad. Protein concentrations were determined according to the method of Bradford using bovine serum albumin (BSA) as a standard,[3] or using a NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific) for CutC variants (ε = 128 870 M^{-1} cm^{-1}). The extinction coefficient for CutC variants was obtained using the ExPASy protparam tool (http://www.expasy.org/tools/protparam.html).

All anaerobic experiments were conducted in an MBraun glovebox (MBraun) under an atmosphere consisting of 99.997% N_2 with less than 5 ppm O_2, or in an anaerobic chamber (Coy Laboratory Products) under an atmosphere of 98% nitrogen and 2% hydrogen.

Gel filtration experiments were performed on a BioLogic DuoFlow Chromatography System (Bio-Rad) equipped with a Superdex 200 10/300 GL column (GE Healthcare). Bovine thyroglobulin (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), horse myoglobin (17 kDa), and vitamin B_{12} (1350 Da) were used as molecular weight markers (Bio-Rad, #151-190). The molecular weights of the proteins analyzed by gel filtration were calculated

from their elution volume, using a second-degree polynomial for the relationship between log(molecular weight) and retention time.

Circular dichroism (CD) spectra were taken with a Jasco J-710 Spectropolarimeter in the wavelength range 185–300 nm. The scan speed was 50 nm min^{-1}, the data pitch 0.1 nm, the bandwidth 0.5 nm, and four scans were accumulated per sample. The spectropolarimeter was equipped with a Jasco PTC-348W temperature controller set to 22 °C. The cuvettes used for CD measurements were Type 21 quartz 0.4 mL cells with a PTFE stopper (1 mm path length, Starna Cells; 21-Q-1).

Perpendicular mode X-band electron paramagnetic resonance (EPR) spectra were recorded on a Bruker ElexSysE500 EPR instrument fitted with a quartz dewar (Wilmad Lab-Glass) for measurements at 77 K, or a cryostat for measurements at 9–40 K. All samples were loaded into EPR tubes with 4 mm o.d. and 8" length (Wilmad Lab-Glass, 734-LPV-7). Data acquisition was performed with Xepr software (Bruker). The magnetic field was calibrated with a standard sample of α,γ-bisdiphenylene-β-phenylallyl (BDPA), g = 2.0026 (Bruker). The experimental spectra for the glycyl radical were modeled with EasySpin for Matlab to obtain g values, hyperfine coupling constants, and line widths.[4] Spin concentration measurements were performed for CutC activation assays by numerically calculating the double integral of the simulated spectra and comparing the area with that of a $K_2(SO_3)_2NO$ standard, taking into consideration the difference in receiver gain.[5] This standard is not stable for a long time[6], so it was prepared before each set of EPR measurements. Solid $K_2(SO_3)_2NO$ was dissolved under anaerobic conditions in anoxic 0.5 M $KHCO_3$ and diluted to a final concentration of 0.3–0.5 mM. To account for any decomposition during dissolution, the concentration was measured at 248 nm (ε = 1690 M^{-1} cm^{-1})[6], using a NanoDrop 2000 UV-Vis Spectrophotometer. The standard used for quantification of the signal arising from the iron-sulfur clusters of CutD was prepared by combining 800 μM $CuSO_4$ with 8 mM EDTA. For glycyl radical-containing samples and the $K_2(SO_3)_2NO$ standard, EPR spectra represent the average of 16 scans and were recorded under the following conditions: temperature, 77 K; center field, 3370 Gauss; range, 200 Gauss; microwave power, 20 μW; microwave frequency, 9.45 MHz; modulation amplitude, 0.4 mT; modulation frequency, 100 kHz; time constant, 20.48 ms; conversion time, 20.48 ms; scan time, 20.97 s; receiver gain, 60 dB (for enzymatic assays) or 30 dB (for standards). For iron-sulfur

cluster-containing samples and the Cu-EDTA standard, EPR spectra represent the average of 4 scans, and were recorded under similar conditions except for: temperature, 9 – 40 K, center field, 3500 Gauss; range, 2000 Gauss; microwave power, 2 mW.

Analytical HPLC was performed on a Dionex Ultimate 3000 instrument (Thermo Scientific). The spectrophotometric coupled assay for CutC activity was conducted at 20 ºC in 96-well plates using a PowerWave HT Microplate Spectrophotometer (BioTek). The absorbance of each well was monitored at 340 nm for 10 min, reads being 20 s apart. Path length correction was employed such that absorbance values were reported for a path length of 1 cm. Data was analyzed using Gen5 Data Analysis Software and kinetic parameters were calculated using GraphPad Prism6. The UV-Vis spectra for CutD were collected using the same instrument and a quartz 96-well plate.

LC-MS/MS analysis was performed in the Saghatelian research labs in the Department of Chemistry and Chemical Biology, Harvard University, on an Agilent 6410 Triple Quadrupole LC/MS instrument (Agilent Technologies), or in the Small Molecule Mass Spectrometry Facility at Harvard University on an Agilent 6460 Triple Quadrupole Mass Spectrometer with Agilent 1290 uHPLC (Agilent Technologies). The same LC-MS/MS method was used as previously described,[7] with the exception that for the Agilent 6460 Triple Quadrupole Mass Spectrometer, the drying gas temperature was maintained at 350 °C with a flow rate of 12 L min^{-1} and a nebulizer pressure of 25 psi.

All GC-MS experiments were conducted using a Waters Quattro micro GC Mass Spectrometer (Waters) equipped with a Combi PAL headspace autosampler (CTC Analytics) and a split/splitless injector. The column used was a DB-624UI, 30 m x 0.32 mm x 1.80 μm (Agilent) and the inlet liner was a 1 mm straight single taper Ultra Inert liner (Agilent). The carrier gas was helium, held at 2.3 mL min^{-1} constant flow. Samples were incubated at 37 °C for 30 min, with an agitator speed of 500 rpm. The GC injection port was set at 220 °C, the syringe temperature to 120 °C, and the needle flush time was 120 s. A headspace sample volume of 1 mL was injected into the inlet with a split ratio of 20:1. The column temperature was initially maintained at 30 °C for 3 min, then increased to 250 °C at a rate of 50 °C min^{-1} with a total run time of 7.4 min. The GC was coupled to a mass spectrometer, equipped with an electron ionization source and a triple

quadrupole mass analyzer with acquisition in positive mode. The GC interface temperature was maintained at 220 °C, and the ion source temperature at 200 °C. In order to identify the products present in enzymatic reactions, the mass analyzer was set to scan over a 45–100 m/z range for 0–1.3 min and a 20–100 m/z range for 1.3–6.2 min of the run. The Selected Ion Recording (SIR) function was used for detection and quantification of acetaldehyde over the first 6.2 min of the run. The monitored SIR ions were m/z 29 for unlabeled acetaldehyde and m/z 30 for d_4-acetaldehyde. GC-MS data analysis was performed with the MassLynx software package. Peak intensities for the ions indicated above, at a retention time of ~1.44 min, were used for the final quantification of acetaldehyde.

Proton nuclear magnetic resonance (^{1}H NMR) spectra and carbon nuclear magnetic resonance (^{13}C NMR) spectra were recorded in the Magnetic Resonance Laboratory in the Harvard University Department of Chemistry and Chemical Biology on a Varian Inova-500 (500 MHz, 125 MHz) or Varian Mercury-400 (400 MHz, 100 MHz) NMR spectrometer. Chemical shifts are reported in parts per million downfield from tetramethylsilane using the solvent resonance as an internal standard for ^{1}H (D_2O = 4.75 ppm) and ^{13}C (DMSO-d_6= 39.52 ppm). Data are reported as follows: chemical shift, integration multiplicity (s = singlet, d = doublet, t = triplet), coupling constant, integration, and assignment. NMR spectra were visualized using ACD/NMR Processor Academic Edition.

High-resolution mass spectral (HRMS) data for the synthetic compounds were obtained in the Magnetic Resonance Laboratory in the Harvard University Department of Chemistry and Chemical Biology on a Bruker Micro QTOF-QII fitted with a dual-spray electrospray ionization (ESI) source. The capillary voltage was set to 4.5 kV and the end plate offset to –500 V, the drying gas temperature was maintained at 190 ^{o}C with a flow rate of 8 L min^{-1} and a nebulizer pressure of 21.8 psi. The liquid chromatography (LC) was performed using an Agilent Technologies 1100 series LC with 50% H_2O and 50% acetonitrile as solvent.

2. Cloning, overexpression and purification of CutC wild type, truncated CutC variants, CutC active site mutants, and CutD

Table S1. Oligonucleotides used for cloning. Restriction sites are underlined.

Oligonucleotide	Target	Sequence (5' to 3')	Restriction site
Dde-3282-NdeI-start	CutC	GCATCATATGGATCTCCAGGACTTTTCACATAAGC	NdeI
Dde-3282-XhoI-stop	CutC CutC (–18 aa) CutC (–30 aa)	GATCCTCGAGTTAGAAACCATGCAGCATGG	XhoI
CutC-truncated-18aa-NdeI-start	CutC (–18 aa)	GCATCATATGCCGGCAGAACGTGCTTCG	NdeI
CutC-truncated-30aa-NdeI-start	CutC (–30 aa)	GCATCATATGGGCGTATCCGCCGAGGTCTTC	NdeI
CutC-truncated-52aa-NdeI-start	CutC (–52 aa)	GCATCATATGGGCATTCCCGACGGGCCTAC	NdeI
Ddes3282-NotI-stop	CutC (–52 aa)	GATTAA GCG GCC GCTTAGAAACCATGCAGC	NotI
Dde-3281-NdeI-start	CutD	GCATCATATGAGAACCGCAACACACAGAGACG	NdeI
Dde-3281-XhoI-nostop	CutD	GATTCTCGAGGTGGCGGATCACCGAAACC	XhoI
Dde3282-QC-D216N-for	CutC-D216N	GTGAACGGCGGCGGCAACTCCAACCCCGGTTAC	–
Dde3282-QC-D216N-rev	CutC-D216N	GTAACCGGGGTTGGAGTTGCCGCCGCCGTTCAC	–
Dde3282-QC-T334S-for	CutC-T334S	GTTGAAGAAAACCAGAGCGGTATGTCCATCGGC	–
Dde3282-QC-T334S-rev	CutC-T334S	GCCGATGGACATACCGCTCTGGTTTTCTTCAAC	–
Dde3282-QC-F395L-for	CutC-F395L	GCCGGGTACCAGCCTTTAGTGAACATGTGCGTG	–
Dde3282-QC-F395L-rev	CutC-F395L	CACGCACATGTTCACTAAAGGCTGGTACCCGGC	–
Dde3282-QC-C489A-for	CutC-C489A CutC (–52 aa) C489A	GACTACTGCCTGATGGGTGCCGTGGAACCGCAG	–
Dde3282-QC-C489A-rev	CutC-C489A	CTGCGGTTCCACGGCACCCATCAGGCAGTAGTC	–
Dde3282-QC-E491Q-for	CutC-E491Q	CTGATGGGTTGCGTGCAACCGCAGAAATCAGGCCGT	–
Dde3282-QC-E491Q-rev	CutC-E491Q	ACGGCCTGATTTCTGCGGTTGCACGCAACCCATCAG	–
Dde3282-QC-T502S-for	CutC-T502S	CGTCTGTACCAGTGGAGCTCCACCGGCTATACC	–
Dde3282-QC-T502S-rev	CutC-T502S	GGTATAGCCGGTGGAGCTCCACTGGTACAGACG	–
Dde3282-QC-G821A-for	CutC-G821A	GTGGTGCGCGTGGCCGCATACAGCGCCTTCTTC	–
Dde3282-QC-G821A-rev	CutC-G821A	GAAGAAGGCGCTGTATGCGGCCACGCGCACCAC	–

Cloning of *cutC* (wild type and truncated variants) and *cutD*

Amplification of *cutC* (Dde_3282) and *cutD* (Dde_3281) from *Desulfovibrio alaskensis* G20 genomic DNA was performed according to a previously published procedure using the primers listed in Table S1.[7] Variants of CutC lacking the first 18 (*cutC*–18 aa), 30 (*cutC*–30 aa), and 52 (*cutC*–52 aa) amino acids were PCR amplified using the same method, except that an annealing

temperature of 63.8 °C was used.

Amplified fragments were digested with *NdeI* and *XhoI* (*cutC*, *cutD, cutC*–18 aa, *cutC*–30 aa) or *NotI* (*cutC*–52 aa) *(*New England Biolabs, NEB), purified and ligated into linearized expression vectors as previously described,[7] to afford pET-29b-CutD (C-terminal His_6-tagged construct) and pET-28a-CutC (N-terminal His_6-tagged construct). These constructs were transformed into chemically competent *E. coli* BL21 (DE3) cells (Invitrogen) and stored at –80 °C as frozen LB/glycerol stocks.

Site-directed mutagenesis of CutC

Site-directed mutagenesis of CutC (full length enzyme) was performed using the corresponding oligonucleotides listed in Table S1. PCR reactions of 20 µL contained 10 µL of Phusion High-Fidelity PCR Master Mix (New England Biolabs), 50 ng of pET-28a-CutC template, and 100 pmoles of each primer. Thermocycling was carried out in a C1000 Gradient Cycler (Bio-Rad) using the following parameters: denaturation for 1 min at 95 °C, followed by 18 cycles of 30 s at 95 °C, 1 min at 67 °C, and a final extension time of 8 min at 72 °C. Digestion of the methylated template plasmid was performed with DpnI (NEB), and 5 µL of each digestion was used to transform a single tube of chemically competent *E. coli* TOP10 cells. The identities of the resulting pET-28a-CutC-D216N, pET-28a-CutC-T334S, pET-28a-CutC-F395L, pET-28a-CutC-C489A, pET-28a-CutC-E491Q, pET-28a-CutC-T502S, pET-28a-CutC-G821A, and pET-28a-CutC-52-C489A constructs were confirmed by sequencing purified plasmid DNA. Chemically competent *E. coli* BL21(DE3) cells were transformed with plasmid DNA and stored at –80 °C as frozen LB/glycerol stocks.

Large scale overexpression and purification of wild type CutC, truncated variants, and active site mutants

An LB agar plate with 50 µg mL^{-1} kanamycin (Kan) was streaked with a frozen stock of *E. coli* BL21(DE3) cells transformed with pET-28a-CutC wild type (WT), truncated CutC variant, or mutant. A single colony from each plate was inoculated into 50 mL of LB medium with 50 µg/ml Kan, which was grown overnight at 37 °C. These starter cultures were diluted 1:100 into 2 L of LB medium containing 50 µg mL^{-1} Kan, then incubated at 37 °C with shaking at 175 rpm,

induced with 500 μM IPTG (Teknova) at OD_{600} = 0.5–0.6, and incubated at 37 °C for 4–5 h (wild type CutC and active site mutants) or at 15 °C overnight (truncated CutC variants). Cells from 2 L of culture were harvested by centrifugation (6,000 rpm x 10 min), resuspended in 45-50 mL of lysis buffer. The cells were lysed by passage through a cell disruptor (Avestin EmulsiFlex-C3) twice at 8,000–10,000 psi, and the lysate was clarified by centrifugation (10,000 rpm x 30 min). The supernatant was incubated with 2 mL of Ni-NTA resin and 5 mM imidazole for 2 h at 4 °C. The mixture was centrifuged (3,000 rpm x 5 min) and the unbound fraction discarded. The Ni-NTA resin was resuspended in 10 mL of elution buffer, loaded into a glass column, and washed with 10 mL of elution buffer. Protein was eluted from the column using a stepwise imidazole gradient in elution buffer (25 mM, 50 mM, 100 mM, 200 mM, 500 mM), collecting 2 mL fractions. SDS–PAGE analysis (4–15% Tris-HCl gel) was employed to ascertain the presence and purity of protein in each fraction. Fractions containing the desired protein were combined and dialyzed twice against 2 L of storage buffer. Upon removal of any precipitated protein through centrifugation at 3,000 rpm x 2 min, the solutions were concentrated using a Spin-X® UF 20 mL centrifugal concentrator with a 30,000 MWCO membrane (Corning®), then frozen in liquid N_2 and stored at –80 °C.

All CutC mutants and the preparations of CutC WT used for MS assays were purified using this protocol and the following buffers: lysis buffer contained 20 mM Tris-HCl pH 8, 500 mM NaCl, 10 mM MgCl2, and half of a SIGMAFAST™ protease inhibitor cocktail tablet (Sigma-Aldrich), elution buffer consisted of 20 mM Tris-HCl pH 8, 500 mM NaCl, 10 mM $MgCl_2$, and 5 mM imidazole, and storage buffer contained 20 mM Tris-HCl pH 8, 50 mM NaCl, 10% (v/v) glycerol. This procedure afforded yields of 1–3 mg L^{-1} of bacterial culture for CutC WT, 2 mg L^{-1} for CutC-D216N, 6 mg L^{-1} for CutC-T334S, 16 mg L^{-1} for CutC-F395L, 1.8 mg L^{-1} for CutC-C489A, 0.9 mg L^{-1} for CutC-E491Q, 9 mg L^{-1} for CutC-T502S, and 1.5 mg L^{-1} for CutC-G821A. Proper folding of CutC WT and mutant proteins was verified by analyzing a 10 μM solution of each protein by circular dichroism, as outlined in ‘general materials and methods’.

Since the yields obtained by purifying CutC WT with Tris-HCl pH 8 buffer were too low to allow its characterization by EPR, we optimized a purification procedure employing potassium phosphate buffer pH 8. This protocol was also used to purify the truncated variants of CutC:

CutC (–18 aa), CutC (–30 aa), CutC (–52 aa), and the mutant CutC (–52aa) C489A. The overall procedure was similar to the one described above with the following exceptions: the lysis buffer contained 50 mM potassium phosphate pH 8 and 300 mM KCl; no protease inhibitor was added; the elution buffer contained 50 mM potassium phosphate pH 8, 300 mM KCl, and 5 mM imidazole; and the storage buffer consisted of 50 mM potassium phosphate pH 8, 50 mM KCl, and 10% (v/v) glycerol. This procedure afforded yields of 6-10 mg L^{-1} of bacterial culture for CutC WT, 17 mg L^{-1} for CutC (–18 aa), 20 mg L^{-1} for CutC (–30 aa), 24 mg L^{-1} for CutC (–52 aa), and 25 mg L^{-1} for CutC (–52 aa) C489A.

Determination of the molecular mass of native CutC and truncated CutC variants

A ~15 μM solution of CutC (full length or truncated variant) purified with potassium phosphate buffer pH 8 was analyzed by gel filtration as described in 'general materials and methods' (250 μL injection volume). The proteins were eluted over 62 min with 25 mM potassium phosphate buffer pH 8, 50 mM KCl at 0.5 mL min^{-1}. A solution of molecular weight markers was analyzed under the same conditions. Table S2 indicates that the full length CutC and the truncated variants exist as dimers in solution, a result consistent with the oligomeric state found for homologous glycyl radical enzymes[8].

Table S2. Molecular weights for full length CutC and truncated variants determined from analytical size exclusion chromatography.

Protein	Calculated molecular weight (kDa)[a]	Observed molecular weight (kDa)
CutC (full length)	96.8	191.4
CutC (–18 aa)	94.9	201.2
CutC (–30 aa)	93.4	204.4
CutC (–52 aa)	91.5	180.1

[a] Molecular weights were calculated for the N-terminal His_6-tagged constructs using Geneious Pro 6.1.6.

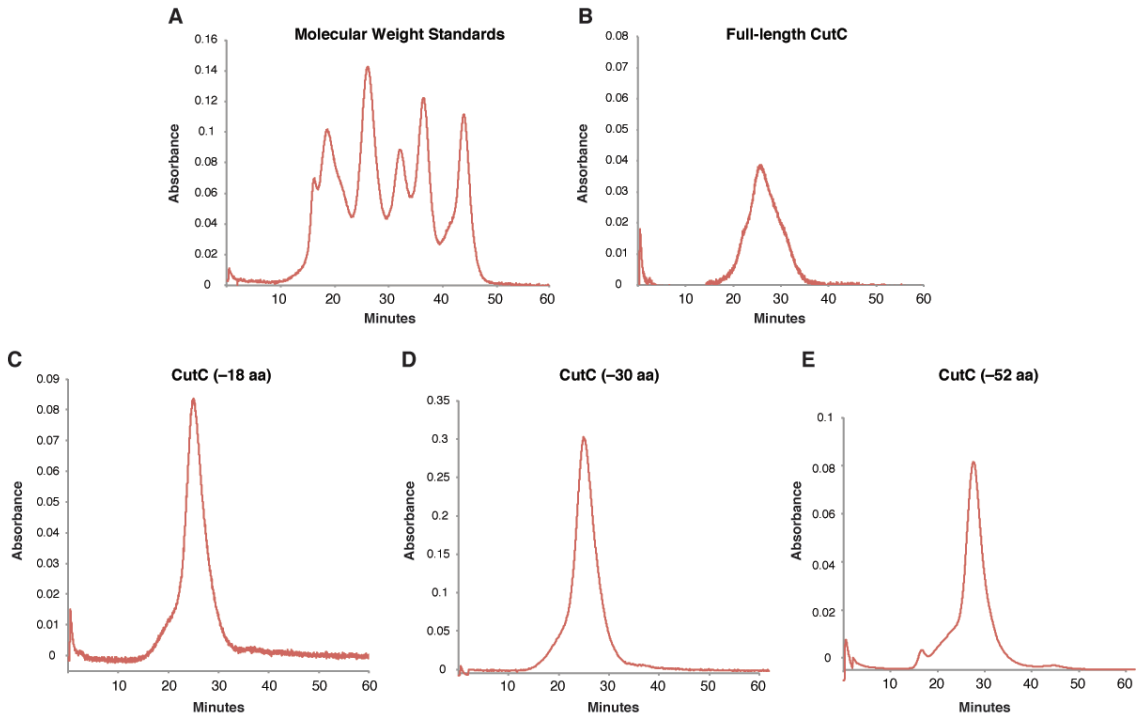


Figure S1. Determination of CutC molecular mass via size exclusion chromatography. (A) FPLC chromatogram of molecular weight standards, (B) full-length CutC, (C) CutC (–18 aa), (D) CutC (–30 aa), (E) CutC (–52 aa).

Large scale overexpression and purification of CutD

The best CutD yields were achieved when plasmid pPH149[9] containing the *E. coli IscSUA-HscBA-Fd* genes (ISC system) was co-transformed into chemically competent *E. coli* BL21 (DE3) cells (Invitrogen) together with pET-29b-CutD. An LB agar plate with 50 μg mL^{-1} Kan and 50 μg mL^{-1} chloramphenicol (Cam) was streaked with a frozen stock of *E. coli* BL21(DE3) cells transformed with pET-29b-CutD and the ISC system. A single colony was used to inoculate 50 mL of LB medium with 50 μg mL^{-1} Kan and 50 μg mL^{-1} Cam, which was grown overnight at 37 °C. This starter culture were diluted 1:100 into a screw-capped 2.8 L flask with 2 L of LB medium supplemented with 50 μg mL^{-1} Kan, 50 μg mL^{-1} Cam, 2 mM iron (III) ammonium citrate, and 0.5 % (w/v) glucose. The culture was incubated at 37 °C with shaking at 175 rpm until reaching an OD_{600} = 0.35–0.4, at which point it was placed at room temperature and sparged with argon gas for ~30 min. Upon addition of L-cysteine (2 mM final concentration) and sodium fumarate (20 mM final concentration), the culture was further sparged until reaching an OD_{600} = 0.5–0.6, at which point protein expression was induced with 500 μM IPTG (Teknova),

the flask was capped tightly and wrapped with electrical tape, and the culture was incubated at 15 °C overnight (~16 to 18 h) with shaking at 100–150 rpm.

All purification steps were performed in an anaerobic chamber at 11 °C (MBraun glovebox) or 17 °C (Coy Labs glovebox), with the exception of cell lysis and centrifugation. Cells from 2 L of culture were harvested by centrifugation in 250 mL polypropylene centrifuge tubes (6,000 rpm x 10 min) and resuspended in 45-50 mL of anoxic lysis buffer (50 mM HEPES, 200 mM NaCl, 10 mM $MgCl_2$, pH 8) supplemented with half of a SIGMAFAST™ protease inhibitor cocktail tablet (Sigma-Aldrich), 8 mg of chicken egg lysozyme, and 5 mM DL-dithiothreitol (DTT). The resuspension was brought out of the anaerobic chamber and the cells were lysed by passage through a cell disruptor (Avestin EmulsiFlex-C3) twice at 8,000–10,000 psi while maintaining a nitrogen line above the receiving tube and one above the cell disruptor reservoir. The lysate was clarified by centrifugation (10,000 rpm x 30 min), and the supernatant was incubated with 2 mL of Ni-NTA resin (previously sparged with nitrogen) and 5 mM imidazole for 2 h at 4 °C. The mixture was brought out of the anaerobic chamber, centrifuged (3,000 rpm x 5 min), and the unbound fraction discarded. The Ni-NTA resin was resuspended inside the anaerobic chamber in 10 mL of elution buffer (50 mM HEPES, 200 mM NaCl, 10 mM $MgCl_2$, 5 mM imidazole, pH 8) and loaded into a glass column. Protein was eluted from the column using a stepwise imidazole gradient in elution buffer (10 mM, 25 mM, 50 mM, 100 mM, 200 mM), collecting 2 mL fractions. SDS–PAGE analysis (4–15% Tris-HCl gel) was employed to ascertain the presence and purity of protein in each fraction. Fractions containing CutD were combined and stirred slowly for 13–16 h with 2 mM DTT, 0.2 mM Na_2S, and 0.25 mM $Fe(NH_4)_2(SO_4)_2$ (the iron was added from a 10 mM stock 5 min after the DTT and Na_2S solutions were added).

Reconstituted CutD was filtered through a 25 mm, 0.22 μm pore-size Acrodisc syringe filter with HT Tuffryn Membrane (Pall Life Sciences) to remove particulates and was dialyzed with three 1 L batches of 50 mM HEPES pH 8, 50 mM NaCl, and 10 % (v/v) glycerol. Upon concentration, the protein solution was aliquoted into 0.5 mL cryogenic vials, frozen in liquid N_2, and placed in 18 x 150 mm Hungate tubes (Chemglass) sealed with butyl stoppers and aluminum seals. This procedure afforded up to 9 mg of CutD per L of bacterial culture, while expression without the ISC system provided a 2-fold lower yield, without impacting activity.

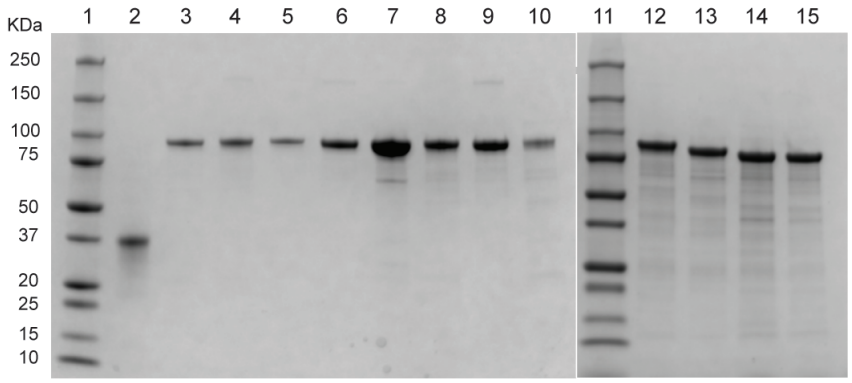


Figure S2: SDS-PAGE (4-15% Tris-HCl gel) of the proteins used in this work. Precision Plus Protein™ All Blue Standards (Bio Rad) (lanes 1 and 11), CutD (lane 2), CutC WT (lane 3), CutC mutants: CutC-D216N (lane 4), CutC-T334S (lane 5), CutC-F395L (lane 6), CutC-C489A (lane 7), CutC-E491Q (lane 8), CutC-T502S (lane 9), CutC-G821A (lane 10), CutC truncated variants: CutC (–18 aa) (lane 12), CutC (–30 aa) (lane 13), CutC (–52 aa) (lane 14), CutC (–52 aa)-C489A (lane 15).

Determination of the molecular mass of native CutD

A ~60 µM solution of CutD was analyzed by gel filtration as described in 'general materials and methods' (250 µL injection volume). The protein was eluted over 62 min with 50 mM HEPES buffer pH 8, 50 mM NaCl at 0.5 mL min^{-1}. A solution of molecular weight markers was analyzed under the same conditions. The observed molecular weight for CutD was 37.4 kDa, whereas the calculated molecular weight for the C-terminal His_6-tagged construct is 35.8 kDa. This result indicates that CutD exists as a monomer in solution, which is consistent with the oligomeric state of pyruvate formate-lyase activating enzyme.[10]

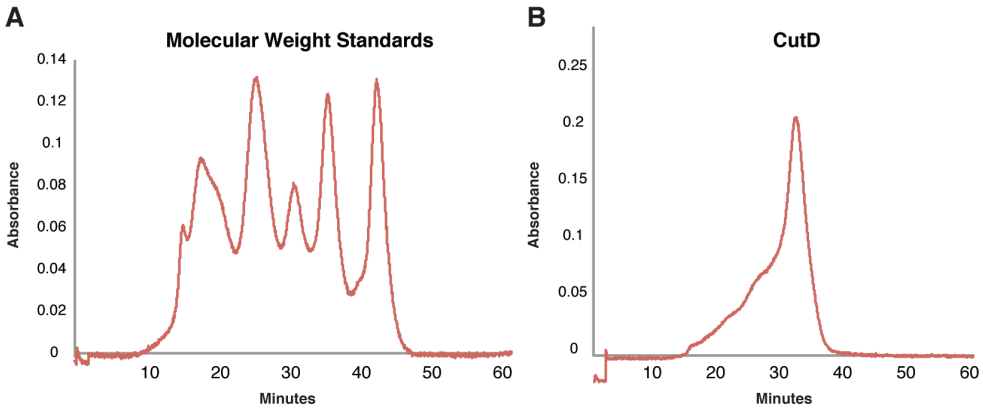


Figure S3. Determination of CutD molecular mass via size exclusion chromatography. A) FPLC chromatogram of molecular weight standards. B) FPLC chromatogram of CutD.

Overexpression and purification of flavodoxin and flavodoxin reductase from *E. coli* K-12

The plasmids for flavodoxin (pET-23b-Fld) and flavodoxin reductase (pET-28b-Fpr) from *E. coli* K-12 were constructed as previously described.[11,12] Overexpression and purification of these two enzymes employed a published method[12] with the following modifications: 10 µM of riboflavin was added to the growth medium prior to induction with IPTG, and the storage buffer used was 20 mM sodium phosphate pH 7.4, 50 mM NaCl, 10 % (v/v) glycerol. We obtained yields of 28 mg L^{-1} of bacterial culture for Fld, and 60 mg L^{-1} for Fpr.

3. Biochemical characterization of CutD

UV-Vis assay for assembly of the [4Fe-4S] cluster

After reconstitution of CutD, the protein solution was diluted inside the anaerobic chamber to 10 µM with anoxic buffer (50 mM HEPES pH 8, 50 mM NaCl), and its absorbance was measured over a 250–800 nm range using a quartz 96-well plate as described in 'general materials and methods'. To obtain the spectrum for reduced CutD, 10 µM of the protein was incubated with a 10-fold excess of sodium dithionite (NaDT) for 15 min prior to the absorbance measurement.

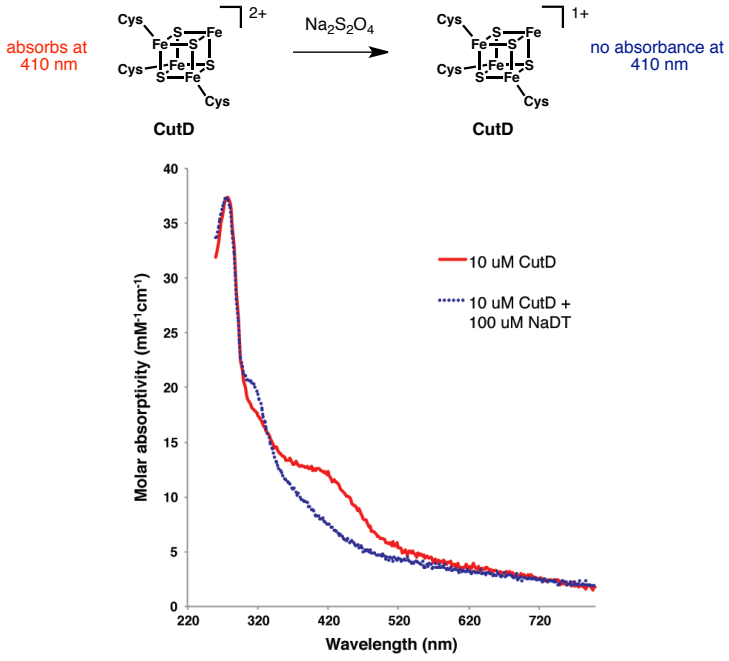


Figure S4: UV-Vis spectrum of purified CutD before (red) and after (blue) reduction with sodium dithionite (NaDT).

Determination of iron content

The iron content of a 6 μM solution of CutD was determined using Ferene (3-(2-pyridyl)-5,6-di(2-furyl)-1,2,4-triazine-5 and 5''-disulfonic acid disodium salt), according to a previously published procedure,[13] with the exceptions that the assay volume was tripled and the standard curve was prepared with ammonium iron (II) sulfate hexahydrate (0-70 μM).

Determination of sulfide content

The sulfide content of a 6 μM solution of CutD was measured using a previously published procedure,[14] with the following modifications: the assays were performed in microcentrifuge tubes, vortexing was used instead of stirring, and the mixture was incubated for 20 min at room temperature after addition of NaOH.

EPR characterization of CutD

The samples for EPR analysis of the reduced iron-sulfur cluster contained 50 mM HEPES pH 8, 50 mM NaCl, 40 μM CutD, and 4 mM NaDT, and were incubated at room temperature (22 °C) under anaerobic conditions for 20 min prior to freezing in liquid N_2. In order to determine the effect of S-adenosylmethionine (SAM) on the EPR signal, 4 mM SAM was added after the reduction, and the sample was frozen within less than 5 min. Quantification of the signal at 9 K for each sample was performed with a Cu-EDTA standard as described in the 'general materials and methods', and indicated the presence of ~0.4 spins per CutD monomer in each case. Despite the apparent splitting noticeable in the A) and B) spectra from Figure S5 (which could be due to a superimposition of signals), the signal with $g_{\|}$ of 2.045 and $g_{\perp}$ of 1.94 can be unambiguously attributed to a $[4Fe\text{-}4S]^{+}$ cluster, because this signal broadens and decreases in intensity at higher (40 K) temperatures, as shown in spectrum C). The shape and amplitude of the EPR signal do not seem to change upon addition of excess SAM, an observation that has been made for other glycyl radical activating proteins.[15–17]

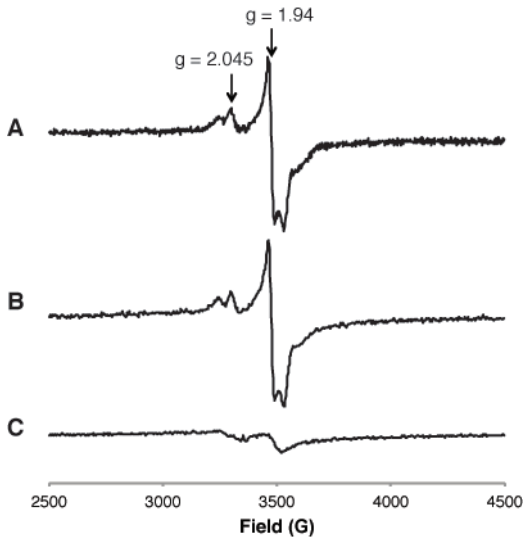


Figure S5. EPR characterization of the iron-sulfur cluster in CutD. A) Spectrum at 9K of NaDT-reduced CutD. B) Spectrum at 9 K of NaDT-reduced CutD incubated for < 5 min with excess SAM. C) Spectrum at 40 K of NaDT-reduced CutD. Acquisition parameters are described in 'general materials and methods'.

HPLC assays for SAM cleavage by CutD

Detection of 5'-deoxyadenosine (dA)

The 5'-dA detection assay contained 25 mM Tris HCl buffer pH 8, 50 mM NaCl, 16 mM NaDT, 40 μM CutD, and 1 mM SAM in a total volume of 100 μL. All components except for SAM were combined inside an anaerobic chamber in order to allow CutD to reduce and SAM was added after 15 min. The reactions were incubated at room temperature in the chamber for 4 h, quenched with 2.56 μL formic acid, and centrifuged (13,200 rpm x 10 min) to remove precipitated protein. The supernatant was analyzed by HPLC (20 μL injection volume) on a Phenomenex Luna analytical C18 reverse phase column (5 μm, 4.6 x 250 mm, 100 Å), preceded by a HAIPEEK guard cartridge holder (Higgins Analytical) equipped with a CLIPEUS C18 (5 μm, 2 cm x 3.2 mm) cartridge (Higgins Analytical). The solvent system consisted of 40 mM ammonium acetate pH 6 (A) and acetonitrile (B), and the elution conditions were: 0% B for 2 min, 0-25% B over 25 min, 25-0% B over 2 min, 0% B for 5 min. The flow rate was set to 1 mL min^{-1}, the products were detected by UV at 216 nm and 260 nm, and commercial 5'-deoxyadenosine was used as a standard.

Detection of methionine

The methionine detection assay was performed in a similar manner as the 5'-dA detection assay, except that it contained 12 mM NaDT and 30 μM CutD, and it was incubated for 18 h. The reactions were quenched with 200 μL methanol, kept on ice for 10 min, and centrifuged (13,200 rpm x 10 min). The organic solvent was removed *in vacuo*, concentrating the mixtures to ~100 μL. A 20 μL aliquot of each assay was mixed with o-phthaldialdehyde (OPA) reagent (Sigma-Aldrich; containing 2 μL of 2-mercaptoethanol per mL of incomplete reagent). After an incubation of 5-6 min, the samples were analyzed by HPLC (40 μL injection volume) on a Dionex Acclaim C18 reverse phase column (3 μm, 2.1 x 150 mm, 120 Å), preceded by the same guard cartridge as for the 5'-dA assay. The solvent system consisted of water (A) and acetonitrile (B). The products eluted at 0.25 mL min^{-1} with the following elution conditions: 0% B for 1 min, 0-70% B over 15 min, 70-0% B over 5 min, 0% B for 2 min. OPA-derivatized L-methionine was used as a standard and the products were detected by UV at 338 nm.

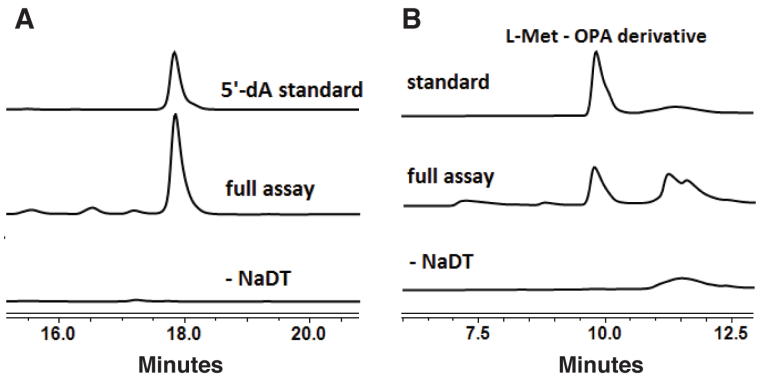


Figure S6. HPLC assays for characterization of SAM cleavage products generated in the presence of CutD. A) Detection of 5'-deoxyadenosine (5'-dA). B) Detection of L-methionine as an OPA derivative.

EPR detection of CutC glycyl radical formation by CutD

Quantification of the glycyl radical content of CutC truncated variants

All activation assays contained 25 mM Tris HCl buffer pH 8, 50 mM NaCl, 3 mM NaDT, 30 μM CutD, 7.5 μM CutC dimeric truncated variant (–18 aa, –30 aa, or –52 aa), and 3 mM SAM in a total volume of 300 μL. For all assays involving CutC, the enzyme was made anoxic by sparging

with argon for 15-30 min. All components except for CutC and SAM were combined inside the anaerobic chamber in order to allow CutD to reduce, then SAM was added after 20 min. The reactions were incubated at room temperature for approximately 60 min before being transferred to an EPR tube and frozen in liquid nitrogen. For each CutC truncated variant, the activation assay was performed in triplicate. Quantification of the amount of glycyl radical present in the assay was done with a $K_2(SO_3)_2NO$ standard as described in the 'general materials and methods'. The extent of activation was calculated for each variant assuming a maximum of 1 radical site per CutC dimer, a result previously reported for PFL[18] and class III RNR[19]: 8.2 ± 0.2 % for CutC (–18 aa), 8.6 ± 0.3 % for CutC (–30 aa), 9.3 ± 0.4 % CutC (–52 aa). So far, our attempts to increase the extent of CutC activation have been unsuccessful. The amount of glycyl radical observed by EPR was lower when the CutC:CutD ratio employed was 1:0.4 (with either 1 h or 5.5 h incubation), when choline (0.3 mM) was included in the activation assay, and when NaDT was replaced with 1.7 mM DTT, 4 μM Fld, 2 μM Fpr and 0.8 mM NADPH (the NADPH:flavodoxin reductase – flavodoxin system from *E. coli*). Preliminary experiments also indicated that the maximum amount of glycyl radical is obtained after 1 h of incubation and that the radical has a half-life of ~6 h.

Activation assays conducted in buffered H_2O and D_2O

In order to exchange CutC (–52 aa) (wild type or C489A mutant) and CutD from buffered H_2O to D_2O, a 25 mM Tris DCl pH 8, 50 mM NaCl buffer was prepared anoxically in 99.9% D_2O and used to swell 0.5 g of Sephadex G-25 fine resin. A mixture of CutC (–52 aa) and CutD was passed through the resin, and the proteins were eluted with the D_2O-containing buffer described above. In order to determine which of the 0.2 mL-fractions collected contained protein, 10 μL of each fraction was mixed with 90 μL Bradford reagent. The ones that afforded a color change were pooled and concentrated to ~280 μL with Spin-X® UF 6 mL centrifugal concentrators with a 5,000 MWCO membrane (Corning®) placed in 50 mL conical-bottom centrifuge tubes with plug seal caps. The activation assays were conducted as presented above and contained 25 mM Tris HCl (or DCl) buffer pH 8, 50 mM NaCl, 5 mM NaDT, 50 μM CutD, 12.5 μM CutC (–52 aa) (wild type or C489A mutant), and 1 mM SAM in a total volume of 300 μL. EPR spectra were recorded and modeled as described in the 'general materials and methods'.

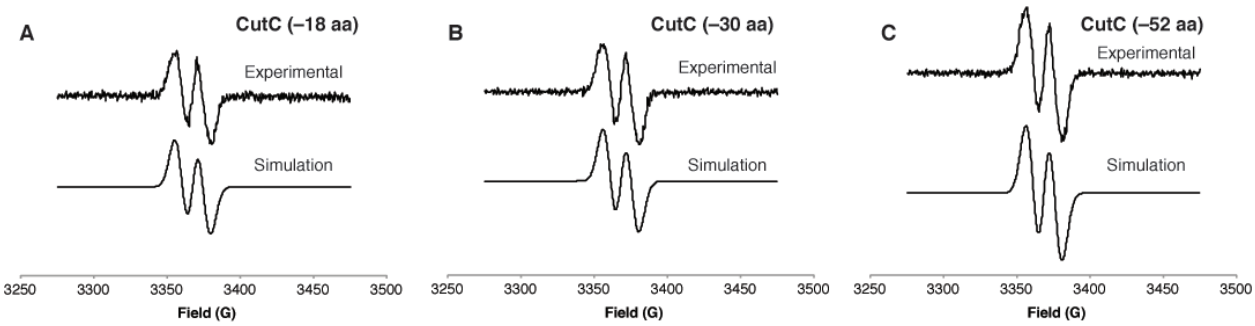


Figure S7. EPR spectra of activated CutC truncated variants. A) Experimental spectrum of activated CutC (–18 aa) and simulation assuming an isotropic signal with g = 2.0033, linewidth 1 mT, and hyperfine coupling A = 1.46 mT. B) Experimental spectrum of activated CutC (–30 aa) and simulation assuming an isotropic signal with g = 2.004, linewidth 1 mT, and hyperfine coupling A = 1.46 mT. C) Experimental spectrum of activated CutC (–52 aa) and simulation assuming an isotropic signal with g = 2.0035, linewidth 1.01 mT, and hyperfine coupling A = 1.44 mT. Acquisition parameters are described in 'general materials and methods'.

4. Biochemical characterization of wild type CutC, truncated variants, and active site mutants

LC-MS/MS assay for trimethylamine detection

A typical assay for LC-MS/MS detection of trimethylamine (TMA) contained 25 mM Tris HCl buffer pH 8, 50 mM NaCl, 2 mM NaDT, 5 μM CutD, 0.5 μM CutC (WT, truncated variant, or mutant), 0.2 mM substrate (choline, betaine aldehyde, glycine betaine or DL-carnitine), and 1 mM SAM in a total volume of 250 μL. No difference in activity was observed between CutC purified with Tris-HCl pH 8 buffer and with potassium phosphate pH 8 buffer. To a 2 mL screw top amber vial (Agilent Technologies, 5188-6535) located inside the anaerobic chamber were added buffer, NaDT, and CutD, and the resulting mixture was incubated at room temperature for 15 min in order to allow CutD to reduce. CutC, substrate, and SAM were added in this order, the vials were capped immediately with a screw top cap with PTFE/red silicone septa (Agilent Technologies, 5182-0717), and the reactions were incubated at room temperature for 3 h (exception: the assay with DL-carnitine as substrate was incubated for 13 h). Upon removal from the anaerobic chamber, the TMA concentration in each assay was determined using LC-MS/MS after derivatization with ethyl bromoacetate[7] and using 100 μL of d_9-TMA (200 μM in water) as the internal standard. A 10 μL aliquot of the quenched derivatization mixture was diluted 50-fold with infusion solution [acetonitrile/ water/ formic acid, 50/50/0.025 (v/v/v)] and analyzed by LC-

MS (3 µL injection volume) as described in 'general materials and methods'. For the assays that did not afford TMA, the concentration of internal standard used was 0.3 µM, and the samples were not diluted before analysis.

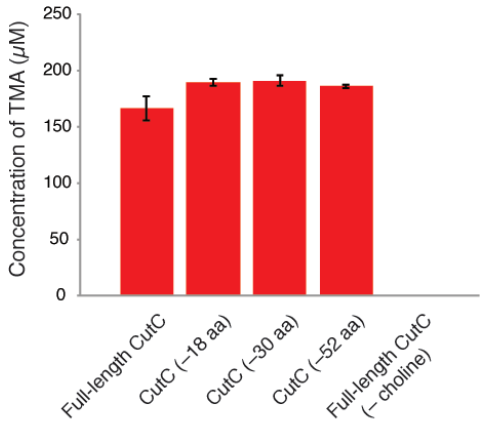


Figure S8. LC-MS assays detecting TMA produced from cleavage of choline by full length CutC (purified with either Tris-HCl or potassium phosphate buffers) and truncated CutC variants. Each bar represents the mean ± standard error of the mean (SEM) of three replicates.

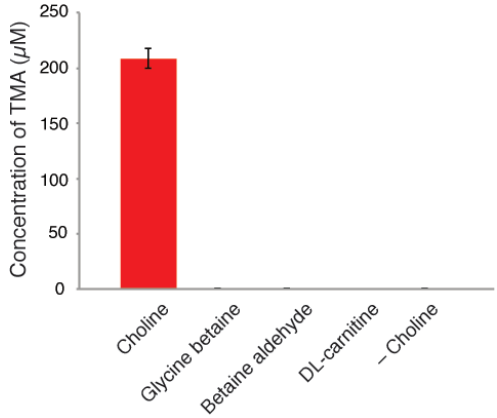


Figure S9. LC-MS assays detecting TMA produced from cleavage of choline, glycine betaine, betaine aldehyde, or DL-carnitine by CutC. Each bar represents the mean ± standard error of the mean (SEM) of three replicates.

GC-MS/MS assay for acetaldehyde, acetaldehyde-1-^{13}C, propionaldehyde, and 3-hydroxy-propionaldehyde detection

The enzymatic reactions were conducted as described for the LC-MS/MS assays, except that 1 mM substrate (choline chloride, choline chloride-1-^{13}C, dimethylaminoethanol, methylaminoethanol, ethanolamine, or 1,2-propandiol) was used and the incubation time was ~12 h. CutC WT used for these assays was purified with potassium phosphate buffer pH 8. Since

both CutC and CutD are stored with 10% (v/v) glycerol (179 mM final concentration), no additional glycerol was added to the assays testing it as a potential substrate. Upon removal from the anaerobic chamber, reaction mixtures were kept at 4 °C for 1 h, uncapped, and the contents were quickly transferred to 10 mL headspace vials with screw tops (Sigma-Aldrich, SU860099) containing 1 mM d_4-acetaldehyde (only for acetaldehyde quantification assays), 1.8 g NaCl, and water up to a final volume of 2.5 mL. The headspace of each vial was analyzed by GC-MS/MS as outlined in 'general materials and methods'. For assays with choline chloride-1-^{13}C, the % ^{13}C incorporation was calculated based on the relative intensities of the fragments with m/z 29 and 30 from the MS/MS spectrum of ^{13}C-acetaldehyde, also taking into consideration the abundance ratio of the fragments with m/z 28 and 29 in the MS/MS spectrum of unlabeled acetaldehyde.

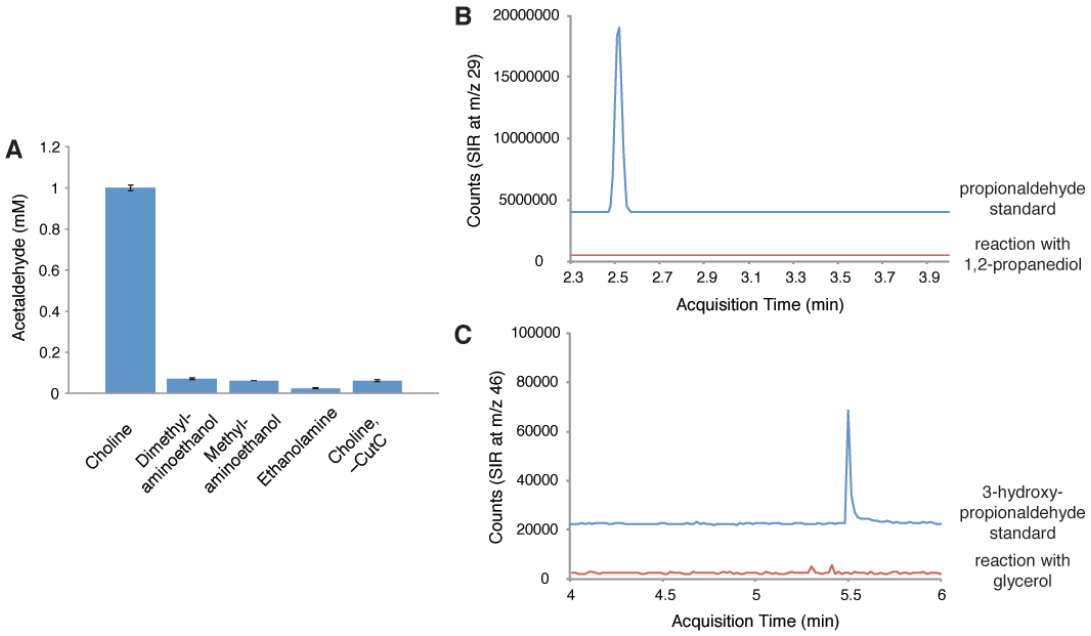


Figure S10. GC-MS assays for CutC activity with choline and other potential substrates. A) Quantification of acetaldehyde production with alternate substrates. Bar graphs represent the mean ± standard error of the mean (SEM) of three independent experiments. B) Selected ion recording (SIR) chromatogram for CutC incubated with 1,2-propanediol. C) SIR chromatogram for CutC incubated with glycerol.

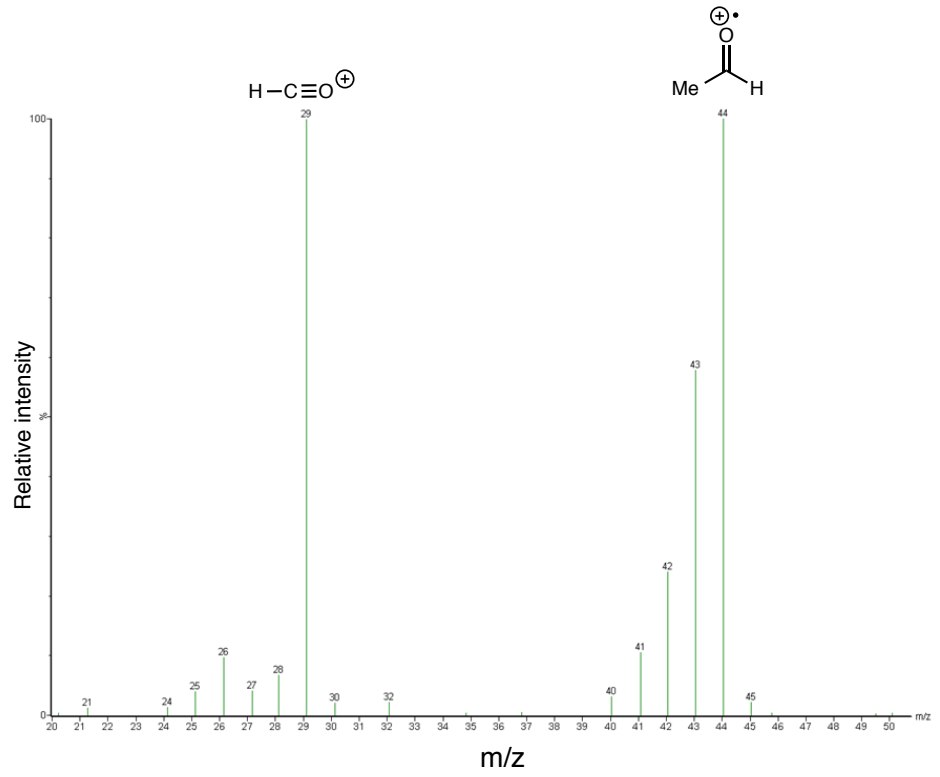


Figure S11. MS/MS spectrum of acetaldehyde generated from reaction of CutC with unlabeled choline.

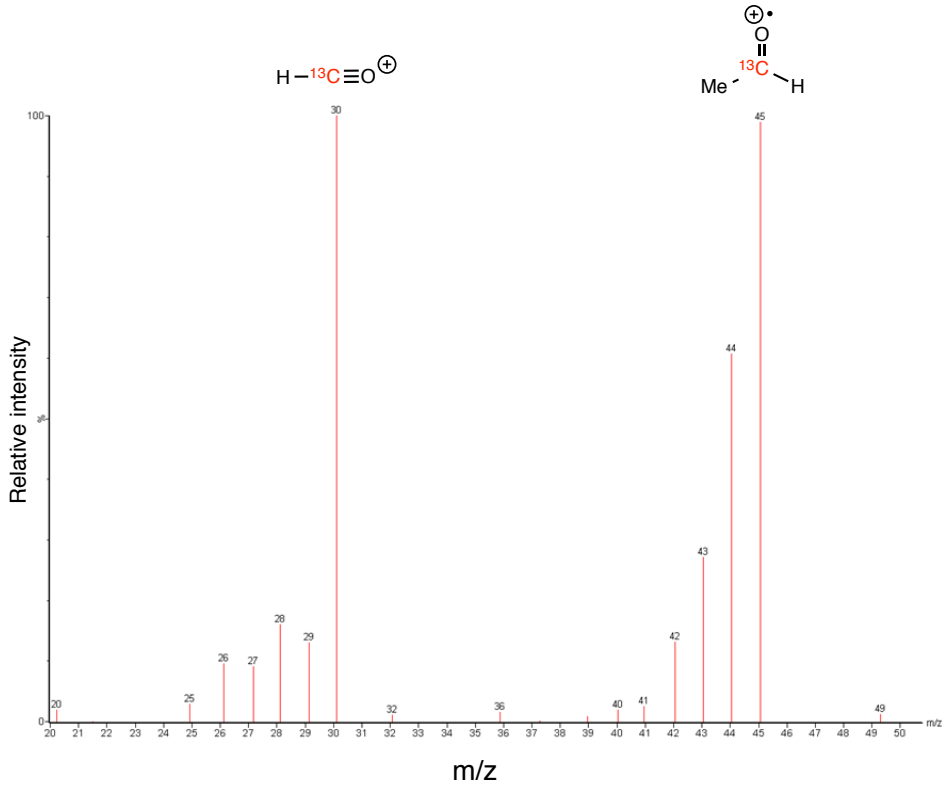


Figure S12. MS/MS spectrum of acetaldehyde generated from reaction of CutC with choline-1-^{13}C.

Spectrophotometric coupled assay for kinetic analysis of choline cleavage

Before each assay, CutC (–52 aa) was activated without substrate using the same conditions described for the EPR quantification of the glycyl radical. After 1 h of incubation, this activation mixture was diluted 75-fold and added to the kinetic assay. Each coupled assay contained 25 mM Tris HCl buffer pH 8, 50 mM NaCl, 200 µM NADH, 320 nM yeast alcohol dehydrogenase, 8 nM CutC dimer (–52 aa), 32 nM CutD, 3.2 µM NaDT, 3.2 µM SAM, and 0-2 mM choline, in a total volume of 200 µM. The reaction was initiated by adding choline. The turnover number was calculated as the µmoles of product formed per second per µmol of active CutC dimer, based on a molecular mass of 182.94 kDa for CutC dimer (–52 aa), the fact that only 9.3 ± 0.4 % of CutC (–52 aa) is activated by CutD, and assuming that each dimer possesses up to 1 glycyl radical site.

Table S3. Kinetic parameters of CutC and comparison to those of other glycyl radical enzymes.

Enzyme	Substrate	K_m (μM)	V_{max} (μmol min^{-1} mg^{-1})	k_{cat} (s^{-1})[a]	k_{cat}/K_m (μM^{-1} s^{-1})
CutC (–52 aa)	choline	302.5 ± 7.9	22.7 ± 0.2	747 ± 27	2.47
4-HPAD[15]	HPA	649 ± 90	15.0 ± 0.2	110[b]	0.17[c]
PFL[20]	pyruvate	2050	16.6	1100[b]	0.536[c]
BSS[21]	toluene	< 100	0.0074	N/A	N/A
GDH[8]	1,2-propanediol	N/A	1560	N/A	N/A
class III RNR[22,23]	CTP (with 0.1 mM ATP as modulator)	450	1.4	6.8[c]	0.015[c]

[a] Turnover number shown is based on the amount of active protein. [b] Value previously reported. [c] Value calculated from published data.

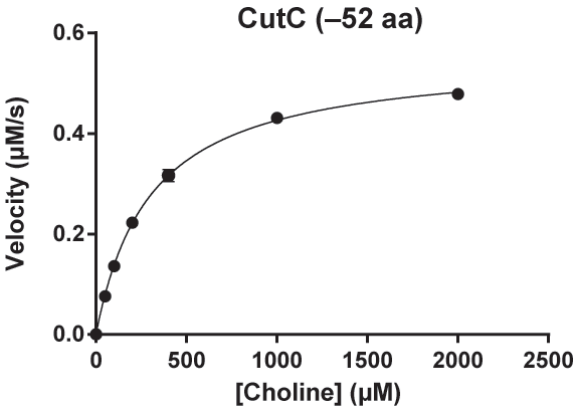
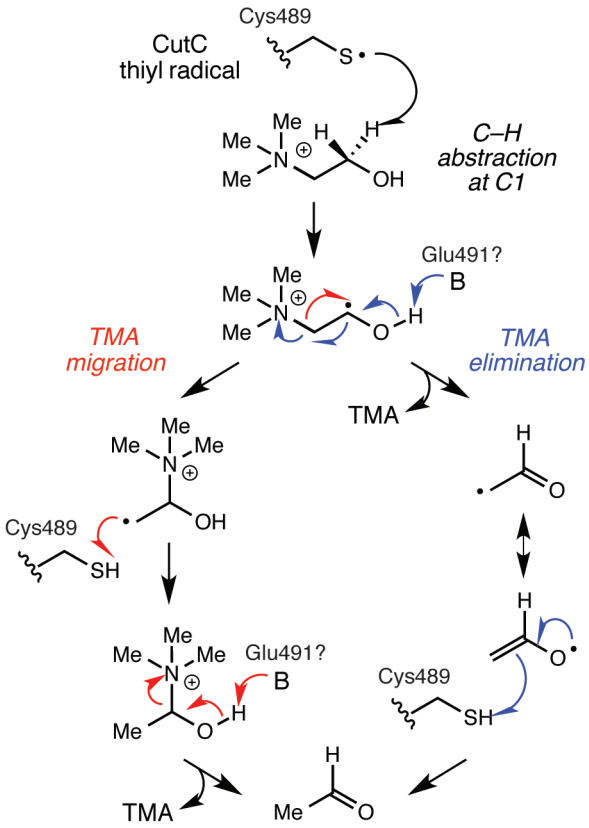


Figure S13. Michaelis–Menten kinetics of CutC (–52 aa).
The error bars represent the standard deviation of three individual experiments.

Mechanistic hypotheses for choline cleavage by CutC

Scheme S1.



5. Bioinformatics

Multiple sequence alignment of CutD homologs with characterized glycyl radical activating enzymes

A multiple sequence alignment of CutD from *D. alaskensis* G20 (Dde_3281) was performed using sequences of biochemically characterized glycyl radical activating enzymes, as well as putative CutD homologues from different bacterial genera previously found to possess a *cut* gene cluster.[7] Besides the conserved SAM-binding, GRE-binding, and [4Fe-4S] cluster binding sites,[10] Figure S14 also reveals that CutD homologs possess two "Clostridial"-type[24] $CX_{2-5}CX_{2-4}CX_3C$ motifs, suggesting that they might bind auxiliary [4Fe-4S] clusters.

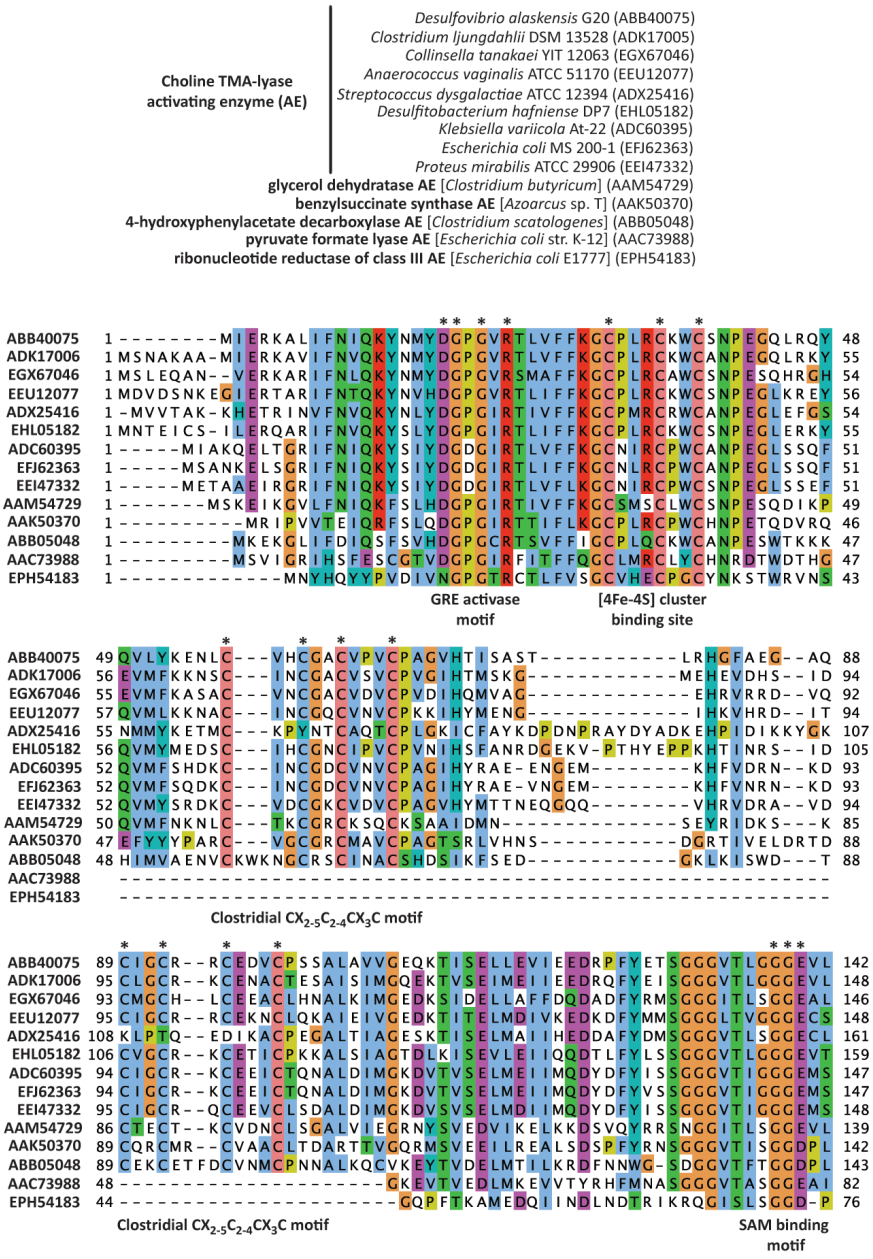


Figure S14. Multiple sequence alignment of putative choline TMA-lyase activating enzymes and known glycyl radical activating enzymes. Asterisks indicate conserved binding motifs. GenBank accession number identifies amino acid sequences used.

Multiple sequence alignment of the N-terminal regions of representative choline TMA-lyase homologs

Most CutC homologs are clustered together with genes encoding putative microcompartment structural proteins, suggesting that these enzymes are microcompartment-associated in vivo.

Given our observation that wild type CutC from *D. alaskensis* G20 tends to precipitate during purification, we hypothesized that its instability could be due to aggregation in the absence of the microcompartment. Previous work with proteins from the 1,2-propanediol[25] and ethanolamine[26] utilization pathways revealed that short N-terminal sequences are responsible for packaging proteins into bacterial microcompartments. Removal of this sequence led to an improved in vitro solubility of ethanolamine ammonia lyase subunit EutC, without negatively impacting activity.[26]

We performed a multiple sequence alignment of CutC from *D. alaskensis* G20 with several CutC homologs, three of which were not microcompartment-associated based on genomic context (Figure S15). We observed that the N-terminal regions of these proteins have a lower degree of similarity (32%) than the rest of the sequences (74%). Thus, we hypothesize that the aggregation of full length CutC in vitro is due to a short (up to ~50 amino acid) N-terminal sequence. As described above, truncated variants of CutC lacking the first 18, 30 or 52 amino acids possessed significantly improved solubility and stability in comparison to the full-length enzyme.

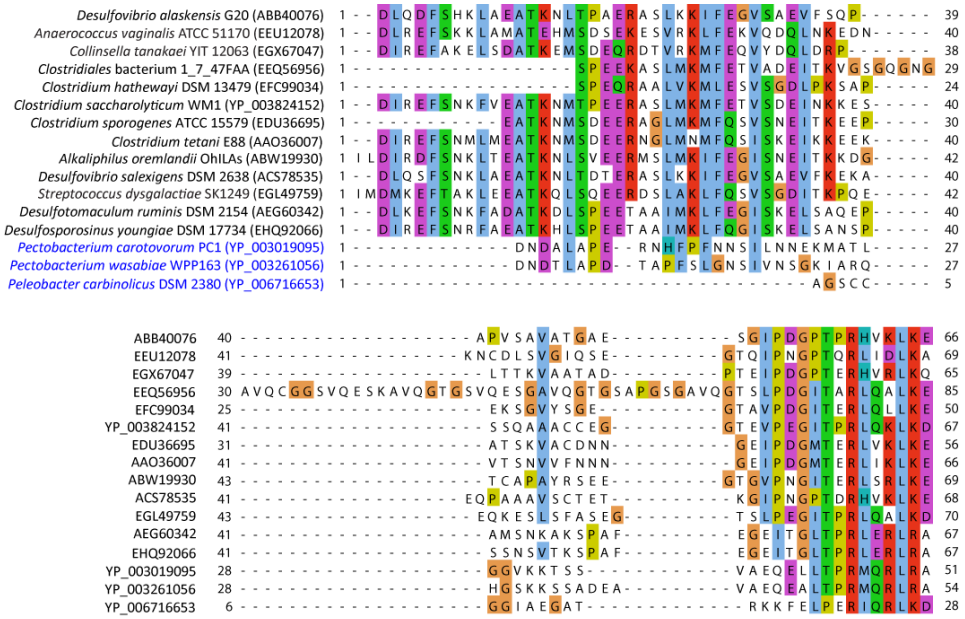


Figure S15. Multiple sequence alignment of the N-terminal regions of representative choline TMA-lyase homologs reveals that this region might target CutC to a protein microcompartment. The last three homologs (marked with blue) are part of *cut* operons that do not contain putative microcompartment proteins. N-terminal methionine residues were omitted to prevent spurious alignment of the divergent termini. The percent similarity of these regions is 32%, as opposed to 74% for the rest of the sequences.

Multiple sequence alignment of CutC homologs with characterized glycyl radical enzymes

A multiple sequence alignment of CutC from *Desulfovibrio alaskensis* G20 (Dde_3282) was performed with biochemically characterized glycyl radical enzymes, as well as putative CutC homologs (60-80.1 % amino acid similarity) from different bacterial genera.[7]

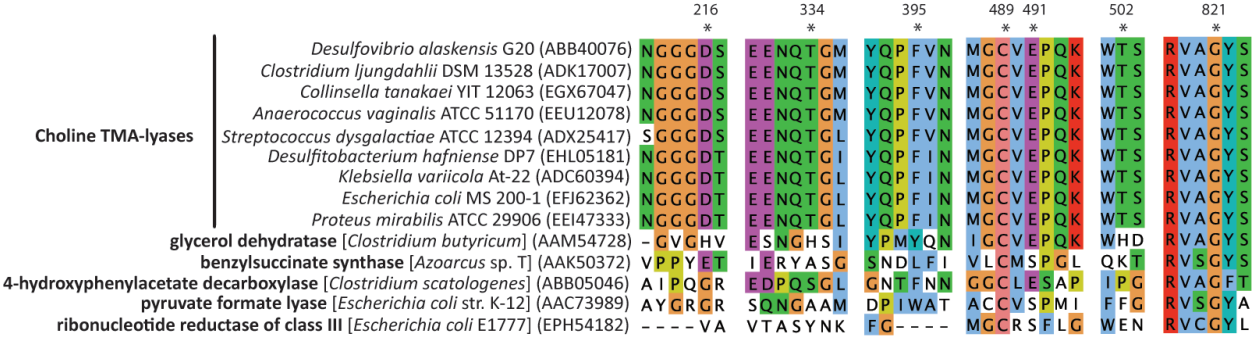


Figure S16. Multiple sequence alignment of choline TMA-lyase homologs and characterized glycyl radical activating enzymes. Asterisks indicate putative CutC active site residues. GenBank accession number identifies amino acid sequences used.

Construction of the CutC homology model

The Basic Local Alignment Search Tool (BLAST)[27] was used to search the protein databank (PDB) to find templates for homology model creation, and templates were selected to maximize sequence identity and query coverage. Potential template structures with the highest identity were all GREs: glycerol dehydratase (GDH) from *Clostridium butyricum* (PDB ID: 2R8W),[8] a putative GDH from *Archaeoglobus fulgidus* (PDB ID: 2F3O),[28] a Kolbe-type decarboxylase from *Clostridium scatologenes* (PDB ID: 2Y8N),[29] and pyruvate formate lyase (PFL) from *E. coli* (PDB ID: 2PFL).[30]

Modeller v9.12 was used to generate five randomly-seeded homology models of CutC with each individual potential template structure and with multiple alignment of the top two, top three, and top four template structures. Heteroatoms and water molecules were excluded from the calculations and multiple template alignment was performed using Salign[31]. Following model generation using the automodel algorithm, models were optimized using Modeller's built-in conjugate gradient optimization algorithm set to the highest setting and optimization was repeated twice per model. Models with the lowest Discrete Optimization Potential Energy

(DOPE) score from each set of homology models were compared based on their DOPE profiles and putative active site geometries to determine the best set of template structures for homology model creation. The homology model based on the multiple alignment of GDH (2R8W) and the putative GDH GRE from *A. fulgidus* (2F3O) was selected due to its low DOPE score and favorable geometry, and 100 randomly-seeded models were generated using the same algorithm as above. The homology model with the lowest DOPE score was analyzed with the verify3D,[32] PROCHECK,[33] QMEAN6,[34] ERRAT,[35] WHATCHECK,[36] and PROVE[37] packages.

The CutC homology model is consistent with known GRE structures and passes multiple structure quality benchmarks (Table S4). The canonical GRE fold with Gly821 and Cys489 surrounded by the characteristic 10-stranded α/β barrel within the homology model aligns well with other members of the GRE family. With an overall DOPE score that is within 3.1-5.5% of the published GDH[8] and putative GDH (*A. fulgidus*) structures[28], respectively, and with a QMean score that is at the 55th percentile of all known structures (Z-score of 0.142), the CutC homology model is compatible with experimentally-determined structures. In the DOPE profile (Figure S17A) and the 3D-1D profile of the model (Figure S17B), there are few notable deviations from the template structure profiles, with the exception of the increased DOPE scores for the 209-220 loop in the predicted CutC structure. The Ramachandran plot statistics for the homology model reveal that 99.2% of the model residues are within allowable regions and fewer than 0.1% occupy disallowed regions, indicating good stereochemical quality (Figure S17C).

Table S4. Comparison of structure quality benchmarks for CutC homology model with those for the crystal structures of GDH (1R8W) and the predicted GDH GRE from *A. fulgidus* (2F3O).

Overall Scoring Metrics	**CutC Model**	**1R8W**	**2F3O**
DOPE (from Modeller)	-101582.21[a]	-104901.53	-107507.26
Verify3D average per residue	0.35417[a]	0.45087	0.45981
DFIRE2 Energy	-1557.36	-1584.73	-1613.4
dDFIRE Energy	-2019.05	-2046.52	-2079.01
QMean6	0.776	0.808	0.774
QMean6 z-score	0.142	0.494	0.125
ERRAT Quality Factor	78.777[a]	97.301	94.51
PROVE Score	0.988	1.055	1.414
Backbone conformation z-score	-4.518	-5.185	-3.377
Ramachandran			
Z-score (WHATCHECK)	1.46	-2.312	-0.969
% of residues in favorable regions	91.9	90.4	88.4
% of residues in allowed regions	8	9.5	11.6
% of residues in disallowed regions	0.1	0.1	0

[a]Denotes metrics that are particularly affected by the presence of the long disordered N-terminal tail of CutC.

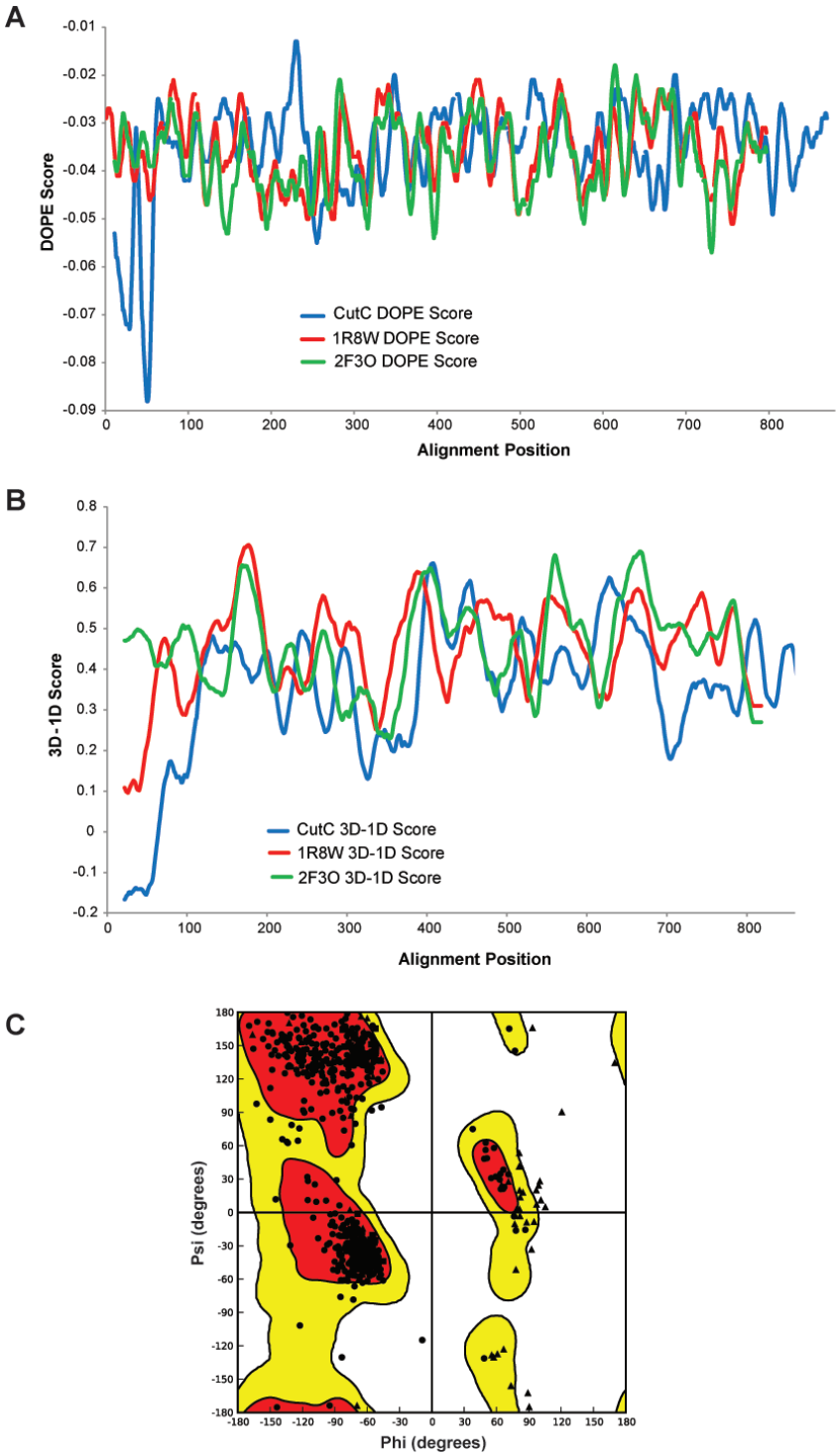


Figure S17. A) Comparison of the DOPE profiles for the CutC homology model (blue), the published GDH structure (1R8W, red), and the predicted GDH (*A. fulgidus*) structure (2F3O, green). B) 3D-1D profile of the CutC homology model (blue), the published GDH structure (red), and the predicted GDH (*A. fulgidus*) structure (green). C) Ramachandran plot of homology modeled CutC structure, generated with Schrödinger Maestro. Most favored regions are shown in red. Additional allowed regions are shown in yellow. Disallowed regions are shown in white.

Induced fit docking of choline into the active site of the CutC homology model

The Glide[38,39] and Prime[40] programs within the Schrödinger suite of molecular dynamics programs have been successfully used in the past to detect and model enzyme active sites and substrate binding orientations.[41,42] In this study, Schrödinger Suite 2012 was used for active site refinement of CutC and induced-fit docking of calculated conformations of the choline substrate.

The standard protein preprocessing protocol of assigning bond orders, adding hydrogens, and capping termini was followed to prepare the selected receptor input file. Hydrogen bond assignment was optimized using PROPKA with a pH of 7.4, and a restrained minimization using the OPLS 2005 force field was used to resolve a single atom-atom clash produced by hydrogen addition. The choline ligand was prepared using the LigPrep algorithm with an approximate pH of 7.4, and ligand minimizations were carried out using the OPLS 2005 force field.

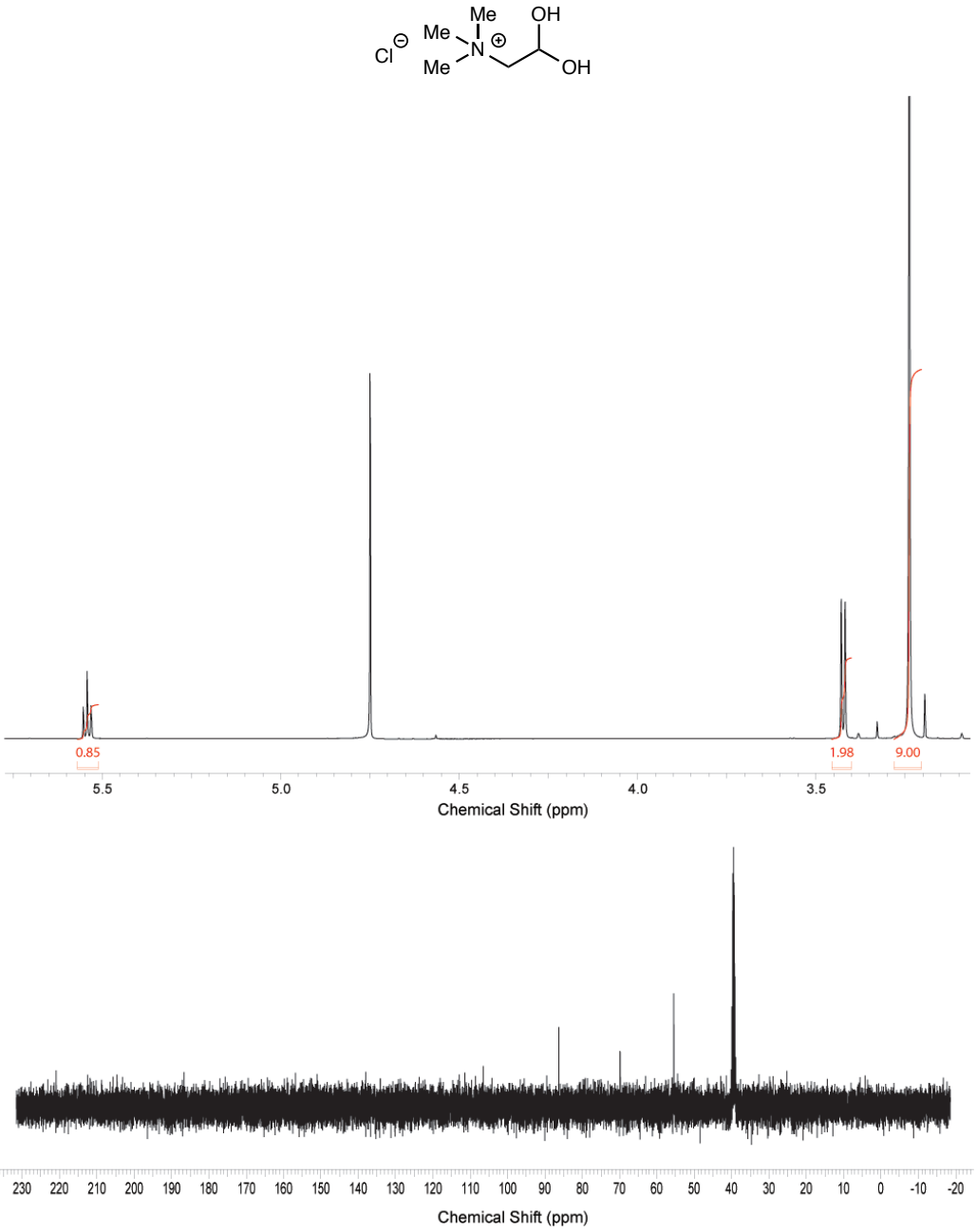
Preliminary docking was performed with Glide using SP precision and van der Waals scaling of 0.5. The receptor grid (18.65Å x 18.65Å x 18.65Å) was centered on the centroid of the active site residues Asp216, Thr334, Phe395, Cys489, Glu491, and Thr502. Following preliminary docking, Prime was used to optimize the active site. All amino acids within 5.0 Å of the docked choline were refined and all putative active site residues were marked for additional refinement. Glide docking with XP precision was then used to re-dock the choline ligand within the refined active site, and calculated poses were compared based on their Glide XP docking scores, Glide Emodel scores, and Prime Energy scores as well as active site geometry. The scores for the best model were Glide: –4.95, Emodel: –38.55, and Prime: 54491.94.

6. Chemical synthesis procedures and characterization data

Betaine aldehyde (2,2-dihydroxy-*N*,*N*,*N*-trimethylethan-1-aminium chloride) was synthesized following a previously published procedure,[43] with the exception that the recrystallization step was replaced with an ethyl acetate extraction to remove any unreacted acetal. Concentration *in vacuo* of the HCl-catalyzed hydrolysis reaction afforded 2,2-dihydroxy-*N*,*N*,*N*-trimethylethan-1-aminium chloride as a light yellow solid (0.389 g, 49%).

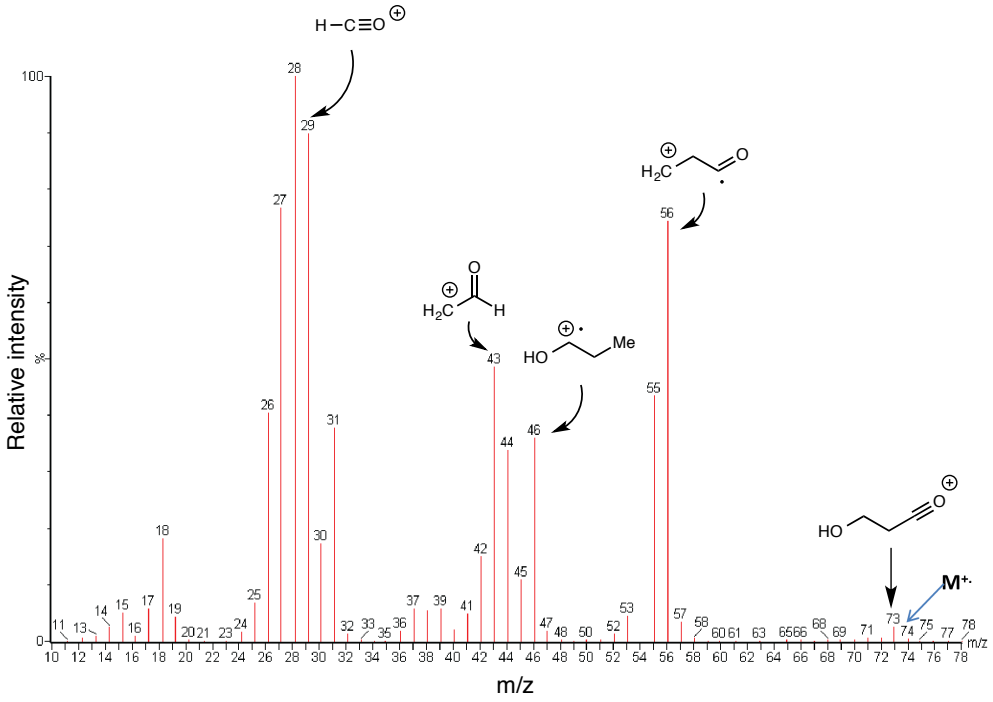
^{1}H NMR: (500 MHz, D_2O) 5.55 (t, 1H, C**H**$(OH)_2$, J = 5.5 Hz), 3.42 (d, 2H, NC**H**$_2$, J = 5.5 Hz), 3.22 (s, 9H, N(C**H**$_3)_3$). ^{13}C NMR (100 MHz, DMSO-d_6): 86.28, 69.86, 55.50. HRMS (ESI): calcd for $C_5H_{12}NO^+$ $[M]^+$, 102.0913; found, 102.1071

^{1}H and ^{13}C NMR spectra:



3-Hydroxypropanal was synthesized following a published procedure,[44] with the following modification: upon neutralization with calcium carbonate, the final mixture was filtered and stored at 4 ºC. The identity of the resulting 3-hydroxypropanal product was confirmed through GC-MS/MS. The ^{13}C NMR data also matched that reported previously, revealing a mixture of 3-HPA and 2-(2-hydroxyethyl)-4-hydroxy-1,3-dioxane (HPA dimer).[44]

MS/MS data:



References

1. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673–4680.

2. Kearse, M., Moir, R, Wilson, A, Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., and Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data, *Bioinformatics 28*, 1647–1649.

3. Stoscheck, C. M. (1980) Quantitation of protein, *Method. Enzymol. 182,* 50–68.

4. Stoll, S., and Schweiger, A. (2006) EasySpin, a comprehensive software package for spectral simulation and analysis in EPR, *J. Magn. Reson. 178*, 42–55.

5. Eaton, G. R., Eaton, S.S., Barr, D. P., and Weber, R. T. (2010) A More in Depth Look at the EPR Signal Response in *Quantitative EPR* 1st ed., pp. 37-61, SpringerWeinNewYork, Germany.

6. Murib, J. M., and Ritter, D. M. (1952) Decomposition of Nitrosyl Disulfonate Ion I. Products and Mechanism of Color Fading in Acid Solution, *J. Am. Chem. Soc*. *74*, 3394–3398.
7. Craciun, S., and Balskus, E. P. (2012) Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme, *Proc. Natl. Acad. Sci., USA 109*, 21307–21312.
8. O'Brien, J. R., Raynaud, C., Croux, C., Girbal, L., Soucaille, P., and Lanzilotta, W. N. (2004) Insight into the mechanism of the B12-independent glycerol dehydratase from Clostridium butyricum: preliminary biochemical and structural characterization, *Biochemistry 43*, 4635–4645.
9. Wecksler, S. R., Stoll, S., Tran, H., Magnusson, O. T., Wu, S.-P., King, D., Britt, R. D., and Klinman, J. P. (2009) Pyrroloquinoline quinone biogenesis: demonstration that PqqE from *Klebsiella pneumoniae* is a radical S-adenosyl-L-methionine enzyme, *Biochemistry 48*, 10151–10161.
10. Vey, J. L., Yang, J., Li, M., Broderick, W. E., Broderick, J. B., and Drennan, C. L. (2008) Structural basis for glycyl radical formation by pyruvate formate-lyase activating enzyme, *Proc. Natl. Acad. Sci. USA 105*, 16137–16141.
11. Hall, D. A., Jordan-Starck, T. C., Loo, R. O., Ludwig, M. L., and Matthews, R. G. (2000) Interaction of flavodoxin with cobalamin-dependent methionine synthase, *Biochemistry 39*, 10711–10719.
12. Szu, P. H., Ruszczycky, M. W., Choi, S. H., Yan, F., and Liu, H. W. (2009) Characterization and mechanistic studies of DesII: a radical S-adenosyl-L-methionine enzyme involved in the biosynthesis of TDP-D-desosamine, *J. Am. Chem. Soc. 131*, 14030–14042.
13. Kennedy, M. C., Kent, T. A., Emptage, M., Merkle, H., Beinert, H., and Münck, E. (1984) Evidence for the formation of a linear [3Fe-4S] cluster in partially unfolded aconitase, *J. Biol. Chem. 259*, 14463–14471.
14. Beinert H. (1983) Semi-micro methods for analysis of labile sulfide and of labile sulfide plus sulfane sulfur in unusually stable iron-sulfur proteins, *Anal Biochem. 131*, 373–378.
15. Yu, L., Blaser, M., Andrei, P. I., Pierik, A. J., and Selmer, T. (2006) 4-Hydroxyphenylacetate decarboxylases: properties of a novel subclass of glycyl radical enzyme systems, *Biochemistry 45*, 9584–9592.
16. Ollagnier, S., Mulliez, E., Schmidt, P. P., Eliasson, R., Gaillard, J., Deronzier, C., Bergman, T., Gräslund, A., Reichard, P., and Fontecave M. (1997) Activation of the anaerobic ribonucleotide reductase from *Escherichia coli*. The essential role of the iron-sulfur center for S-adenosylmethionine reduction, *J. Biol. Chem. 272*, 24216–24223.
17. Külzer, R., Pils, T., Kappl, R., Hüttermann, J., and Knappe J. (1998) Reconstitution and characterization of the polynuclear iron-sulfur cluster in pyruvate formate-lyase-activating enzyme. Molecular properties of the holoenzyme form, *J. Biol. Chem. 273*, 4897–4903.
18. Crain, A. V. and Broderick, J. B. (2013) Pyruvate formate-lyase and its activation by pyruvate formate-lyase activating enzyme, *J. Biol. Chem. 289*, 5723–5729.

19. Mulliez, E., Padovani, D., Atta, M., Alcouffe, C., and Fontecave, M. (2001) Activation of class III ribonucleotide reductase by flavodoxin: a protein radical-driven electron transfer to the iron-sulfur center, *Biochemistry 40*, 3730–3736.
20. Knappe, J.; Blaschkowski, H. P.; Grobner, P.; Schmitt, T. (1974) Pyruvate formate-lyase of Escherichia coli: the acetyl-enzyme intermediate, *Eur. J. Biochem. 50*, 253–263.
21. Li, L., and Marsh, E. N. G. (2006) Deuterium isotope effects in the unusual addition of toluene to fumarate catalyzed by benzylsuccinate synthase, *Biochemistry 45*, 13932–13938.
22. Eliasson, R., Pontis, E., Sun, X., and Reichard, P. (1994) Allosteric control of the substrate specificity of the anaerobic ribonucleotide reductase from *Escherichia coli*, *J. Biol. Chem. 269*, 26052–26057.
23. Mulliez, E., Ollagnier, S., Fontcave, M, Elliason, R., and Reichard, P. (1995) Formate is the hydrogen donor for the anaerobic ribonucleotide reductase from *Escherichia coli*, *Proc. Natl. Acad. Sci. USA 92*, 8759–8762.
24. Selvaraj, B., Pierik, A. J., Bill, E., and Martins, B. M. (2013) 4-Hydroxyphenylacetate decarboxylase activating enzyme catalyses a classical S-adenosylmethionine reductive cleavage reaction, *J. Biol. Inorg. Chem. 18*, 633–643.
25. Fan, C., Cheng, S., Liu, Y., Escobar, C. M., Crowley, C. S., Jefferson, R. E., Yeates, T. O., and Bobik, T. A. (2010) Short N-terminal sequences package proteins into bacterial microcompartments, *Proc. Natl. Acad. Sci. USA 107*, 7509–7514.
26. Akita, K., Hieda, N., Baba, N., Kawaguchi, S., Sakamoto, H., Nakanishi, Y., Yamanishi, M., Mori, K., and Toraya, T. (2010) Purification and some properties of wild-type and N-terminal-truncated ethanolamine ammonia-lyase of *Escherichia coli*, *J. Biochem. 147*, 83–93.
27. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, *J. Mol. Biol. 215*, 403–410.
28. Lehtiö, L., Grossmann, J. G., Kokona, B., Fairman, R., and Goldman, A. (2006) Crystal structure of a glycyl radical enzyme from *Archaeoglobus fulgidus*, *J. Mol. Biol. 357*, 221–235.
29. Martins, B. M., Blaser, M., Feliks, M., Ullmann, G. M., Buckel, W., and Selmer, T. (2011) Structural basis for a Kolbe-type decarboxylation catalyzed by a glycyl radical enzyme, *J. Am. Chem. Soc. 133*, 14666–14674.
30. Becker, A., Fritz-Wolf, K., Kabsch, W., Knappe, J., Schultz, S., and Volker Wagner, A. F. (1999) Structure and mechanism of the glycyl radical enzyme pyruvate formate-lyase, *Nat. Struct. Biol. 6*, 969–975.
31. Marti-Renom, M. A., Madhusudhan, M. S., and Sali, A. (2004) Alignment of protein sequences by their profiles, *Protein Sci. 13*, 1071–1087.
32. Bowie, J. U., Lüthy, R., and Eisenberg, D. (1991) A method to identify protein sequences that fold into a known three-dimensional structure, *Science 253*, 164–170.
33. 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.

34. Benkert, P., Biasini, M., and Schwede, T. (2011) Toward the estimation of the absolute quality of individual protein structure models, *Bioinformatics 27*, 343–350.
35. Colovos, C. and Yeates, T. O. (1993) Verification of protein structures: patterns of nonbonded atomic interactions, *Protein Sci. 2*, 1511–1519.
36. Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996) Errors in protein structures, *Nature 381*, 272.
37. Pontius, J., Richelle, J., and Wodak, S. J. (1996) Deviations from standard atomic volumes as a quality measure for protein crystal structures, *J. Mol. Biol. 264*, 121–136.
38. Glide, version 5.8; Schrödinger, LLC: New York, NY, 2012.
39. Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., and Banks, J. L. (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, *J. Med. Chem. 47*, 1750–1759.
40. Jacobson, M. P., Friesner, R. A., Xiang, Z., and Honig, B. (2002) On the role of the crystal environment in determining protein side-chain conformations, *J. Mol. Biol. 320*, 597–608.
41. McRobb, F. M., Capuano, B., Crosby, I. T., Chalmers, D. K., and Yuriev, E. (2010) Homology modeling and docking evaluation of aminergic G protein-coupled receptors, *J. Chem. Inf. Model. 50*, 626–637.
42. Cortial, S., Chaignon, P., Iorga, B. I., Aymerich, S., Truan, G., Gueguen-Chaignon, V., Meyer, P., Moréra, S., and Ouazzani, J. (2010) NADH oxidase activity of *Bacillus subtilis* nitroreductase NfrA1: insight into its biological role, *FEBS Lett. 584*, 3916–3922.
43. Landfald, B. and Strøm, A. R. (1986) Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*, *J. Bacteriol. 165*, 849–855.
44. Vollenweider, S., Grassi, G., König, I., and Puhan, Z. (2003) Purification and structural characterization of 3-hydroxypropionaldehyde and its derivatives, *J. Agric. Food Chem. 51*, 3287–3293.