SUPPLEMENTARY INFORMATION for:

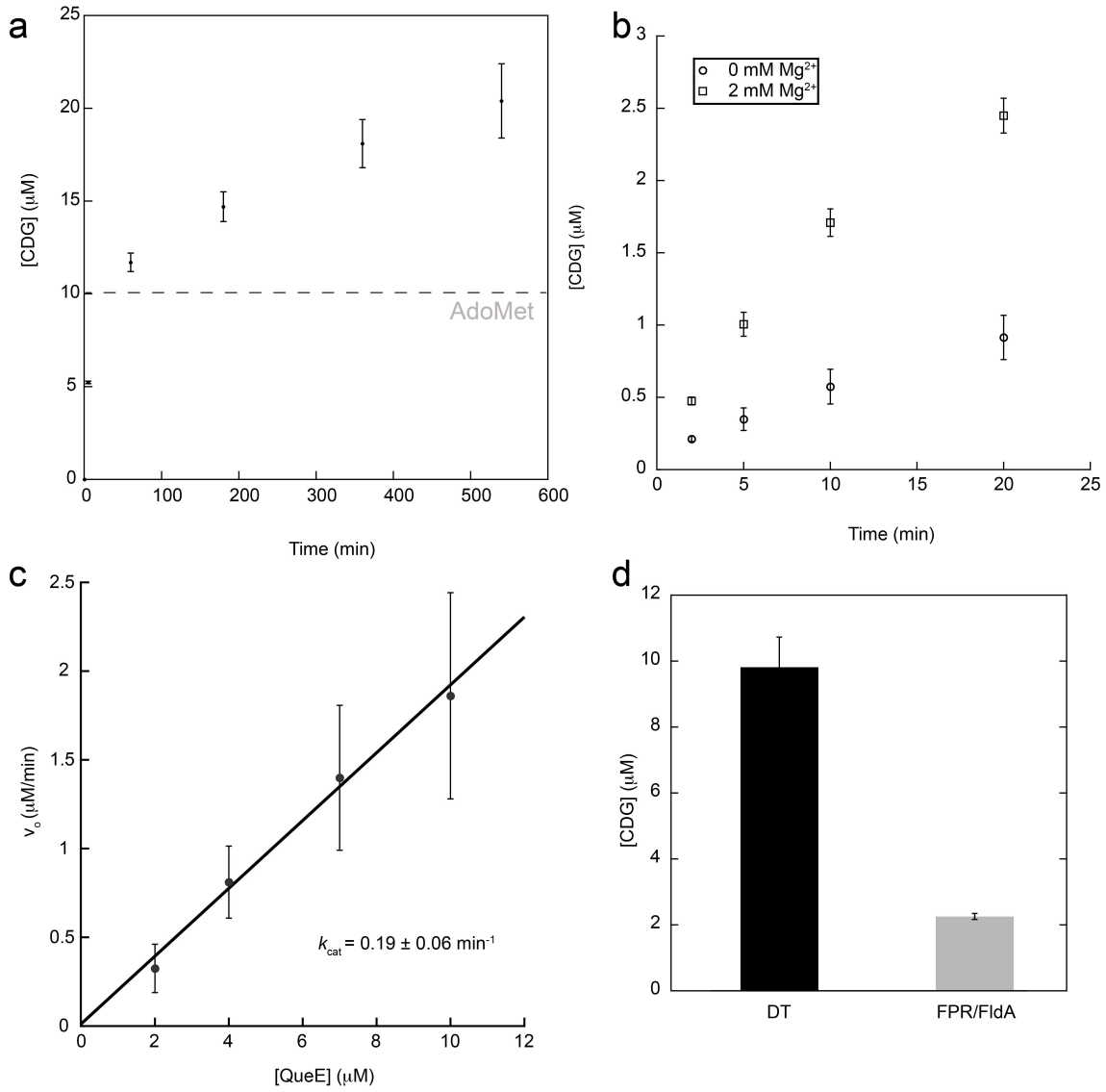
Radical SAM enzyme QueE defines a new minimal core fold and metal-dependent mechanism

Daniel P. Dowling[1,2], Nathan A. Bruender[3], Anthony P. Young[3], Reid M. McCarty[3], Vahe Bandarian[3], and Catherine L. Drennan[1,2,4,$]

[1]Howard Hughes Medical Institute, Departments of Chemistry[2] and Biology[4], Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; cdrennan@mit.edu
[3]Department of Chemistry and Biochemistry, University of Arizona, Tucson, AZ 85721
[$]Corresponding author

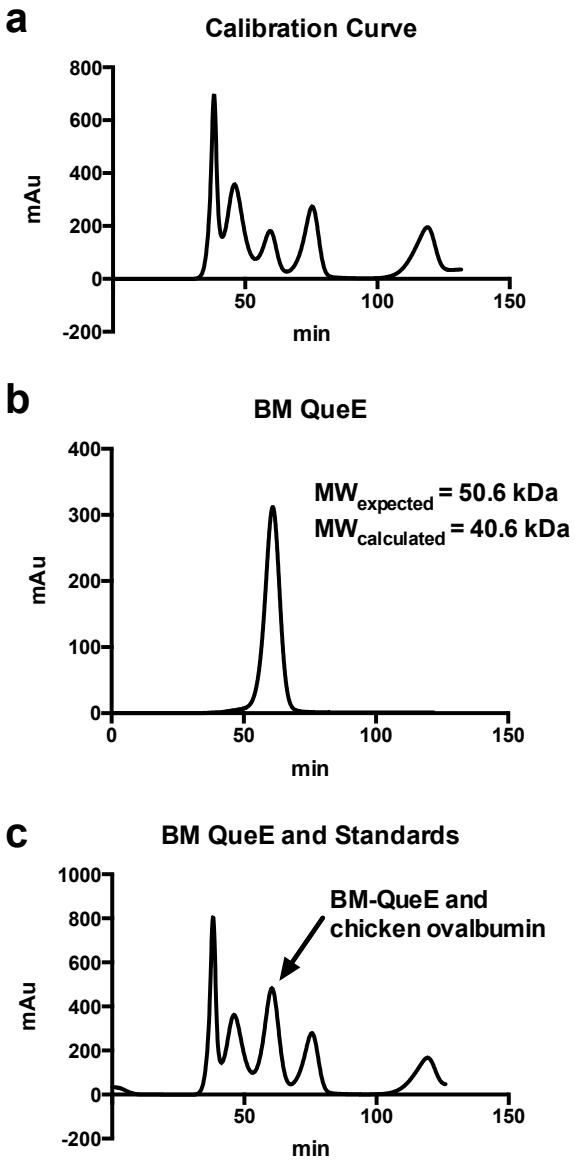
SUPPLEMENTARY RESULTS

Supplementary Fig. 1. Biochemical activity of QueE from *B. multivorans*.
(**a**) CDG formation over time in presence of 10 μM AdoMet (represented by a gray dashed line) showing that AdoMet is used catalytically. (**b**) CDG formation was monitored over time by HPLC analysis with either 0 mM or 2 mM added Mg^{2+}, showing an ~3-fold increase in the rate of product formation with added divalent cation. Basal level of activity without Mg^{2+} is likely due to the presence of Mg^{2+} from the protein purification. (**c**) The k_{cat} for CDG formation was determined using saturating concentrations of CPH_4, Mg^{2+}, and AdoMet and varying concentrations of QueE (2, 4, 7, and 10 μM). (**d**) *B. multivorans* QueE is more active with dithionite (DT) versus the *E. coli* flavodoxin/flavodoxin reductase/NADPH system (FPR/FldA); therefore all assays were performed using dithionite as reductant. All data represent mean value of three replicates ± s.d.

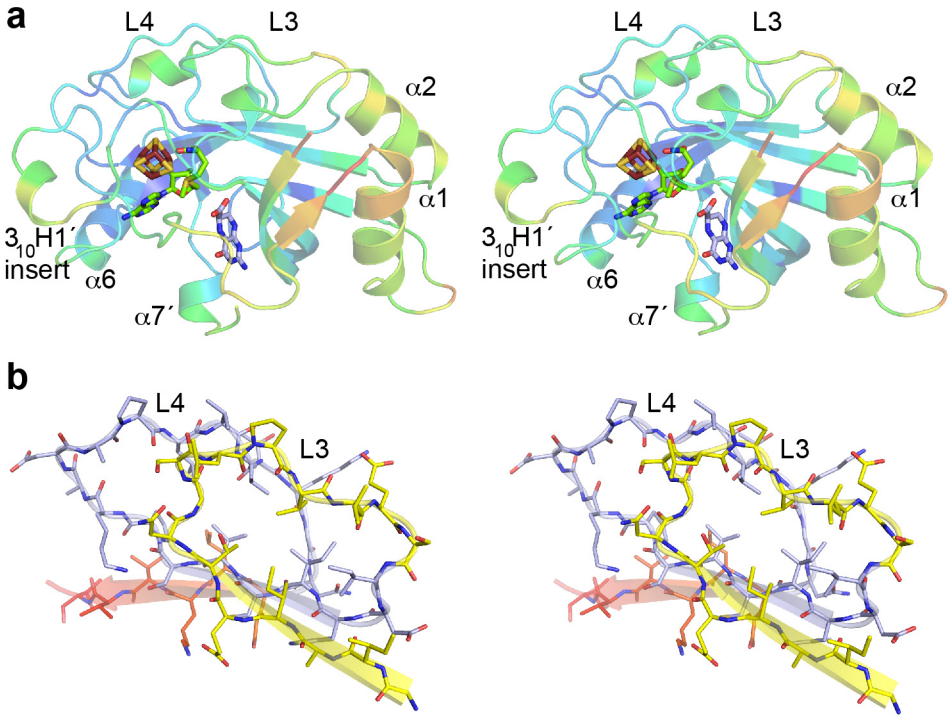


Supplementary Fig. 2. QueE sequence alignments.
Sequences include QueE from *Burkholderia multivorans ATCC 17616* (gi161526284), *Streptomyces griseoflavus Tu4000* (gi302558410), *Ralstonia solanacearum Po82* (gi386333569), *Oceanobacter sp. RED65* (gi94500711), *Rhodothermus marinus DSM 4252* (gi268315892), *Bacteroides thetaiotaomicron BT 4454* (gi29349862), and *Bacillus subtilis* str 168 (gi16078438). β-strands and helices (labeled α or 3_{10} for α-helices or 3_{10}-helices, respectively) within the structural core are highlighted yellow and purple, respectively. Secondary structure of the N- and C-terminal extensions are highlighted gray and orange, respectively, and the cluster binding loop 3_{10}-helical insert is highlighted cyan. Cysteines of the cluster-binding loop are colored red. Residues that contact AdoMet are green with green circles and that contact the substrate are blue with blue circles. Residues that make up the dimerization interface in QueE are indicated by black circles. The location of AdoMet binding motifs are indicated; note "GGE" and "GxIxGxxE" motifs were named for sequences found in BioB and are not strictly conserved.

Supplementary Fig. 3. Solution studies of *B. multivorans* QueE.
Anaerobic size exclusion chromatogram of QueE on a Sephacryl S-200 column in 50 mM PIPES (pH 7.4) and 10 mM DTT. (**a**) A calibration curve was generated with a mixture of thyroglobulin (bovine, 670 kDa), γ-globulin (bovine, 158 kDa), ovalbumin (chicken, 44 kDa), myoglobin (horse, 17 kDa), and vitamin B_{12} (1.35 kDa). (**b**) *B. multivorans* QueE has a retention time consistent with a ~51 kDa homodimer. (**c**) In co-elution experiments with standards in (**a**), QueE coelutes with ovalbumin (44 kDa) at 60.4 min. Each chromatogram is a result of a single injection.

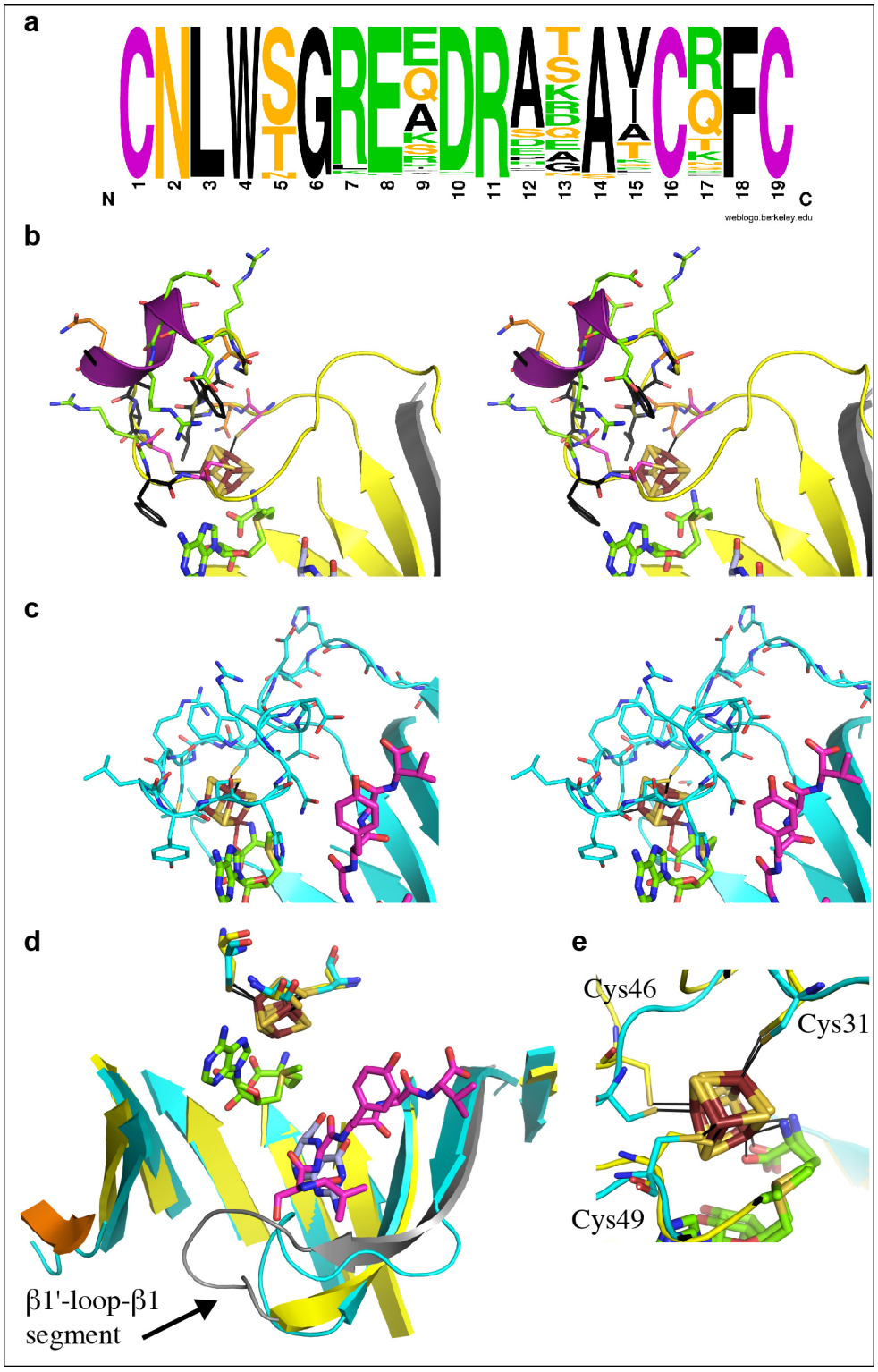


Supplementary Fig. 4. Stereoviews showing loops L3 and L4.
Helices α3 and α4 of the standard AdoMet radical protein core are replaced with two loops in QueE. (a) Ribbon stereoview depiction of one molecule of QueE, colored blue to red for B-factor values ranging from 25 to 59 Å^2. α-Helices and the 3_{10}H1′ helix in the cluster-binding loop are labeled to orient the molecule. (**b**) Stereoview depiction of the L3 and L4 loops, which bury hydrophobic residues against their respective β-strands from the protein core. Protein carbons are colored as follows: β3 and L3, yellow; β4 and L4, light blue; and β5, red.



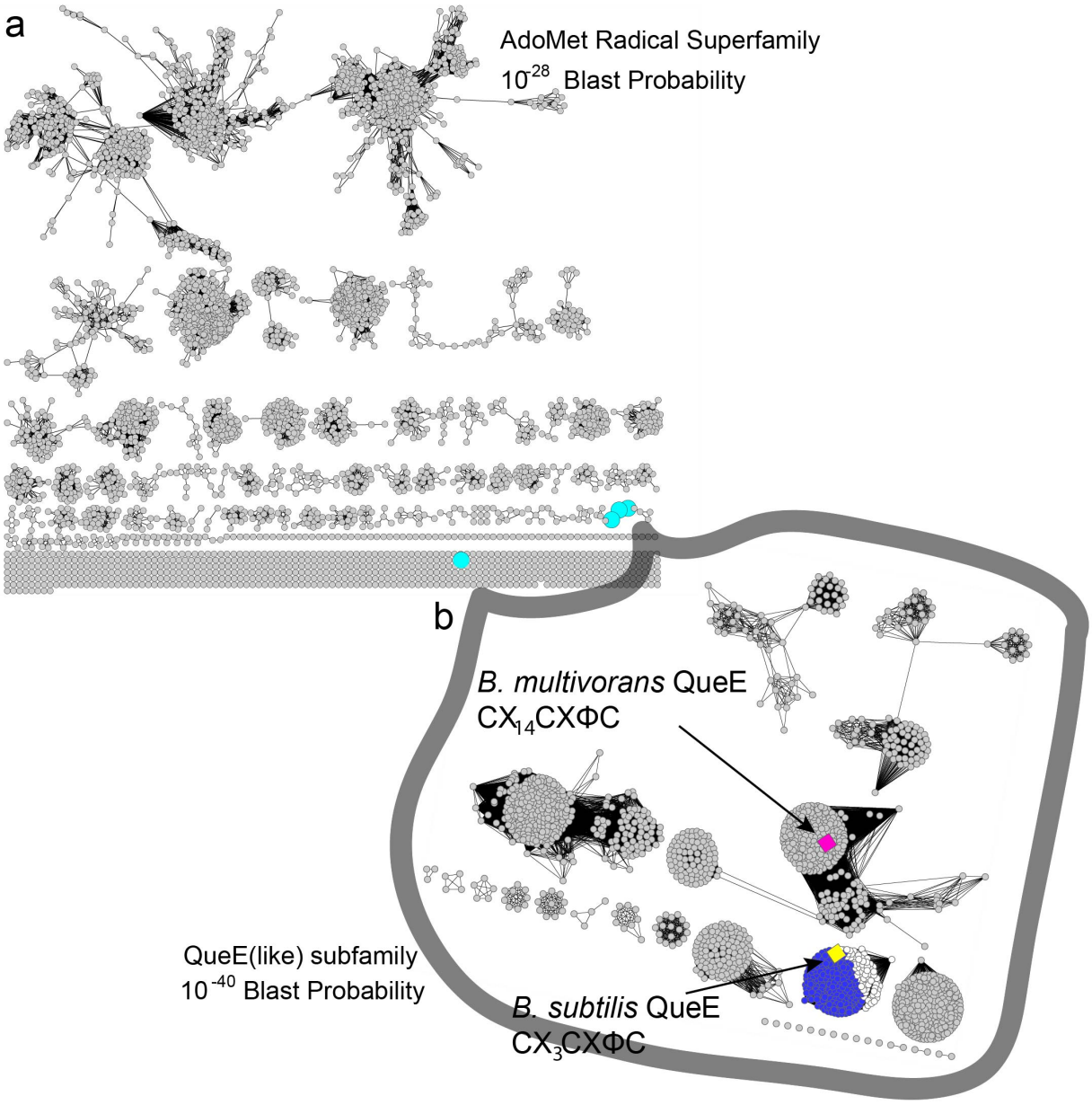
Supplementary Fig. 5. Cluster-binding loop insertion in QueE and comparison to the traditional $CX_3CX\Phi C$ motif as observed in PFL-AE.

(**a**) Representation of the $CX_{14\text{-}20}CX1\text{-}2\Phi C$, determined from 334 proteobacterial protein sequences. Figure was generated using weblogo.berkeley.edu. Stereoview images of the cluster-binding loops from (**b**) QueE and (**c**) PFL-AE (PDB ID 3CB8) [1], accentuating the high degree to which the cluster of QueE is buried in comparison to enzymes with the more common $CX_3CX\Phi C$ motif. (**d**) Overlay of the partial β-barrels of QueE with PFL-AE, depicting the β1′-loop-β1 segments of both proteins. (**e**) Close-up depiction of the [4Fe-4S] coordination environment of QueE and PFL-AE. The cluster loop insertion positions Cys46 differently, but the sulfur of this cysteine superimposes with the cysteine sulfur of PFL-AE[2]. QueE is colored in gray, yellow, and orange for N-terminal, core, and C-terminal residues, respectively, with the 3_{10} helical insert in purple. PFL-AE is colored cyan. Ligand carbons are colored as follows: AdoMet, green; CPH_4, light blue; and PFL-AE substrate peptide, magenta. [4Fe-4S] atoms are ruby and yellow, respectively.

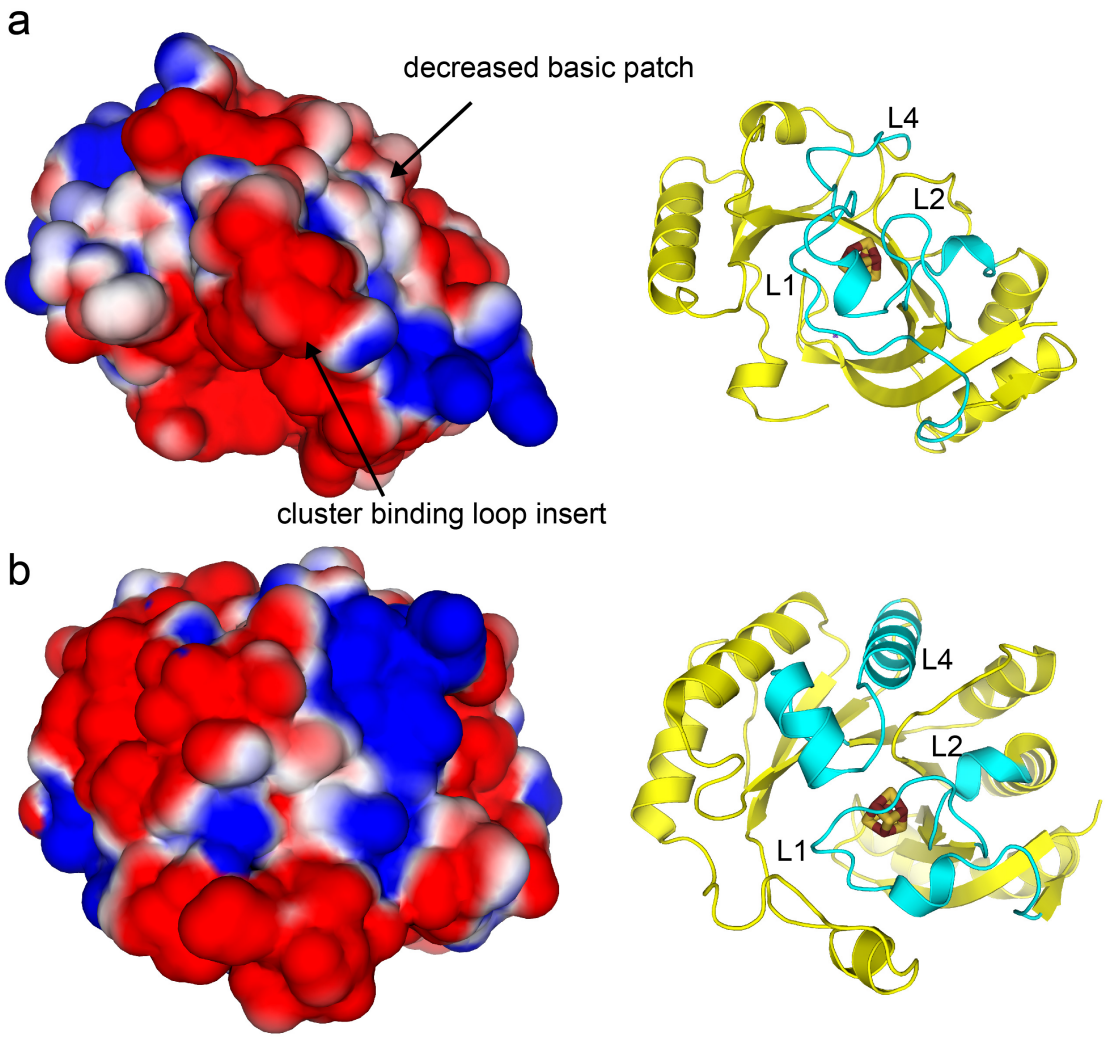

a
N
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
C
weblogo.berkeley.edu
b
c
d
β1'-loop-β1
segment
e
Cys46
Cys31
Cys49

Supplementary Fig. 6. Sequence similarity network of the AdoMet radical superfamily and 7-carboxy-7-deazaguanine synthases.

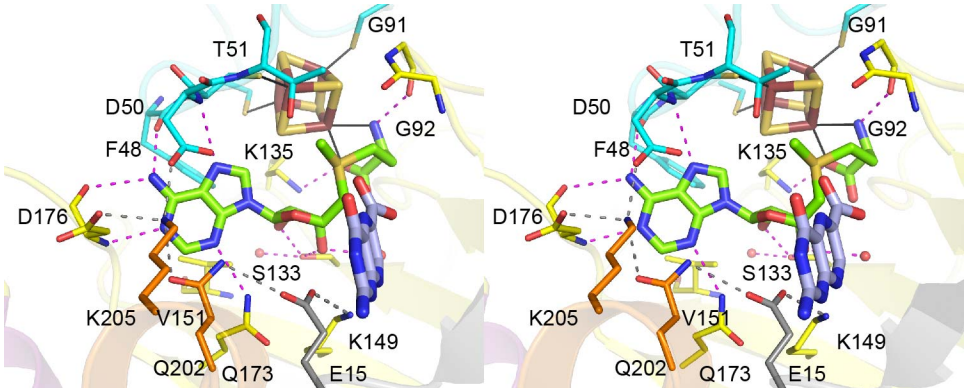
(a) Protein sequence similarity network of the AdoMet radical superfamily, in which separate nodes (gray) represent protein sequences with at least 40% sequence identity; connections are filtered at 10^{-28} Blast Probability. Nodes that contain QueE-like sequences are colored cyan. **(b)** Blow-up of QueE sequence space containing 1589 protein sequences, with node connections filtered at 10^{-40} Blast Probability. Nodes in the QueE-like subfamily are colored as follows: blue, sequences annotated as 7-carboxy-7-deazaguanine synthases; gray, unclassified sequences; white, sequences annotated as ExsD. The gene product ExsD has been predicted to be involved in pterin biosynthesis[3]. The *B. multivorans* and *B. subtilis* QueE sequences are labeled as magenta and yellow diamonds, respectively, depicting the separate grouping of QueE sequences that contain the $CX_{14}CX\Phi C$ motif observed in *B. multivorans* compared to the traditional $CX_3CX\Phi C$ motif.



Supplementary Fig. 7. QueE electrostatic calculations in comparison with PFL-AE. Electrostatic, solvent accessible surface representations of (**a**) QueE and (**b**) PFL-AE (PDB entry 3C8F) [1] are displayed on the left, with the corresponding ribbon depiction shown on the right. Loops proposed to be involved in binding external reductants[1,2,4,5], such as flavodoxin, are colored cyan and labeled. Electrostatic potentials are depicted on a colorimetric scale from red to blue for -1 to +1 kT e^{-1}.

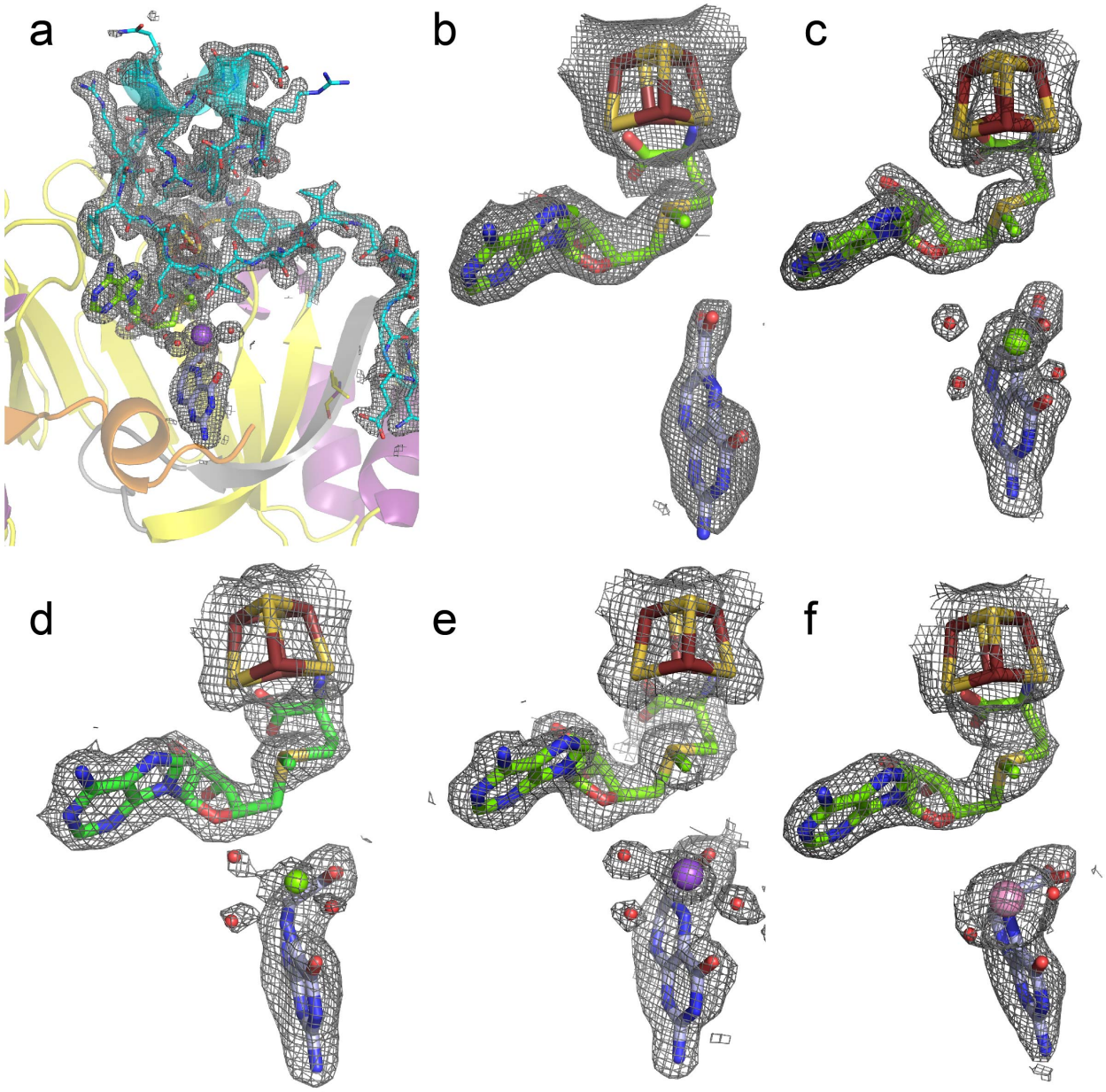


Supplementary Fig. 8. Stereoview of AdoMet binding in *B. multivorans* QueE.
Residues that interact with AdoMet are displayed as sticks. AdoMet and 6CP carbon atoms are colored green and light blue, respectively. The protein backbone is displayed as ribbons; the AdoMet radical core is colored in yellow for loops and strands and purple for helices. The cluster-binding loop is colored cyan. The N− and C−terminal extensions of the protein core are colored gray and orange, respectively. [4Fe-4S] cluster atoms are ruby and yellow, respectively. Coordination bonds to the iron-sulfur cluster are displayed as solid black lines. Protein-ligand interactions are represented as magenta dashed lines, and the protein hydrogen-bonding network located above the plane of the adenine is displayed as dashed gray lines.



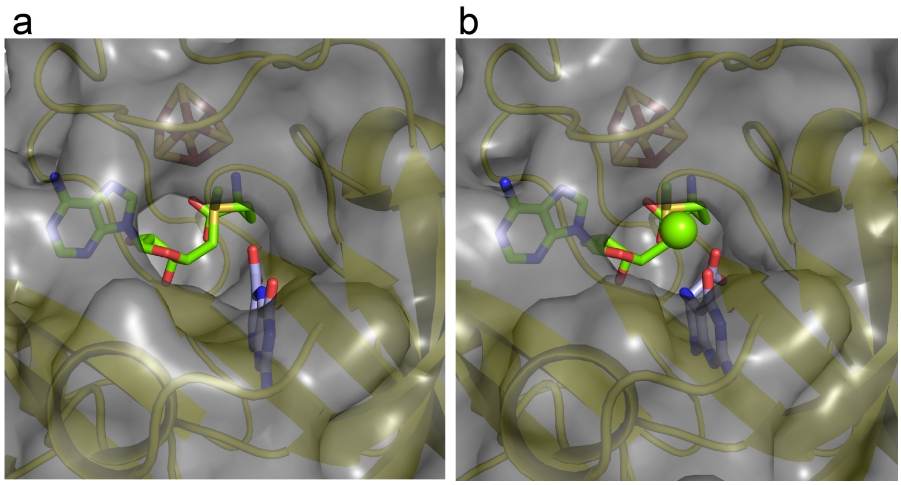
Supplementary Fig. 9. Composite omit electron density for QueE ligands and the cluster loop insertion sequence.

(**a**) $CX_{14}CX\Phi C$ motif and cofactors at 1.9 Å resolution. (**b**) 2.6 Å resolution structure with 6CP and AdoMet bound. (**c**) 1.9 Å resolution structure with CDG, Mg^{2+} and AdoMet bound. (**d**) 2.2 Å resolution structure with CPH_4, Mg^{2+} and AdoMet bound. (**e**) 1.9 Å resolution structure with CPH_4, Na^{+} and AdoMet bound. (**f**) 2.7 Å resolution structure with CPH_4, Mn^{2+} and AdoMet. The protein backbone is displayed as ribbons in (a) only; the AdoMet radical core is colored in yellow for loops and strands and purple for helices. The cluster-binding loop is colored cyan. The N− and C−terminal extensions of the protein core are colored gray and orange, respectively. Ligand carbons are colored as follows: AdoMet, green; 6CP, CDG and CPH_4, light blue; Mg^{2+}, green; Na^{+}, purple; and Mn^{2+}, pink. [4Fe-4S] atoms are ruby and yellow, respectively. Composite omit electron density (2Fo-Fc) at 1σ is displayed as gray mesh.

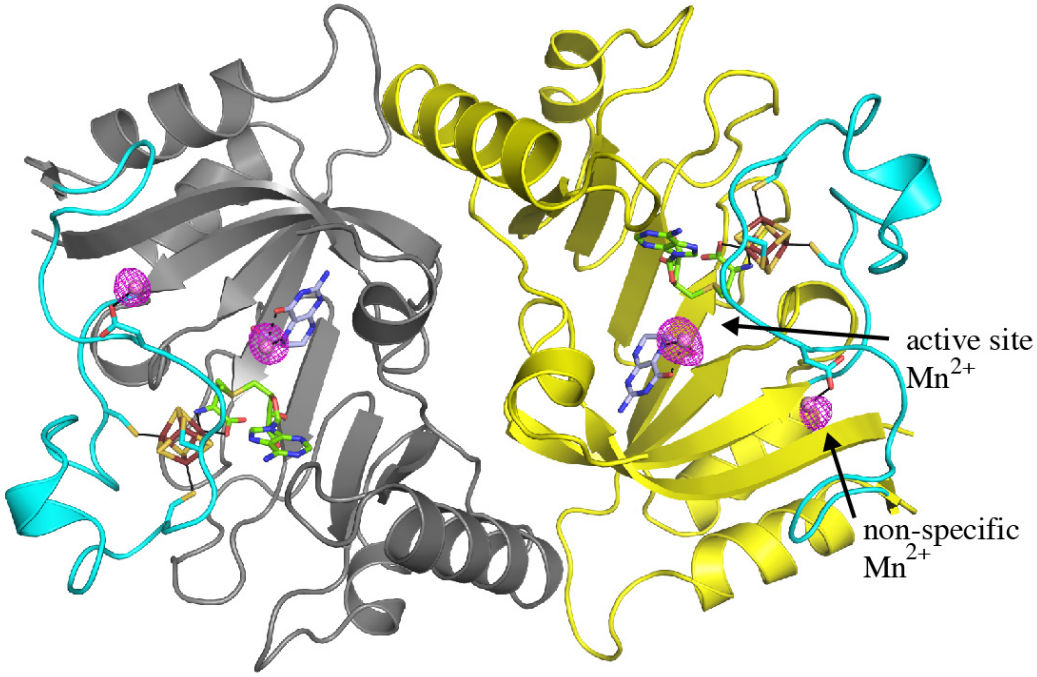


Supplementary Fig. 10. Solvent access to the enzyme active site of QueE from *B. multivorans*.

(**a**) Crystal structure of QueE with AdoMet and 6CP. (**b**) Crystal structure of QueE complexed with AdoMet, CPH_4, and a Mg^{2+} ion. Protein is colored yellow with gray accessible surface area displayed. AdoMet and CPH_4 are displayed with green and blue carbon atoms, respectively, and Mg^{2+} is shown as a green sphere. The [4Fe-4S] cluster is displayed as ruby and yellow sticks for iron and sulfur, respectively.



Supplementary Fig. 11. Anomalous electron density maps for Mn^{2+}-QueE.
Anomalous electron density (based on data collected at wavelength of 1.7399 Å) shows two Mn^{2+} binding sites per monomer in QueE. One Mn^{2+} is bound in the active site; the other is bound on the surface of the protein, where it interacts with the carboxylate of D52 in a monodentate fashion. Anomalous electron density map is colored magenta and contoured at 3σ. The QueE dimer is depicted in ribbons, with separate molecules colored yellow and gray. The cluster-binding loop of each molecule is colored cyan. AdoMet and CPH_4 carbon atoms are colored green and blue, respectively. Iron and sulfur of the clusters are colored ruby and yellow, respectively.



Supplementary Fig. 12. *B. multivorans* QueE gene and protein sequence.
(a) Nucleotide and (b) protein sequence for recombinant codon optimized *B. multivorans* QueE. The nucleotide sequences in bold represent the *Nde*I and *Hind*III restriction endonuclease sites that were used to clone the gene into the corresponding sites in pET28a. The start and stop codons are shown in *italic.* The underlined sequence corresponds to the His_6 tag.

(a)

*ATG*GGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGC**CATATG**A
CCTACGCTGTGAAAGAAATCTTTTATACGCTGCAGGGTGAAGGCGCTAATGCGGGTCGTCCGGC
GGTGTTTTGTCGTTTTGCCGGCTGCAACCTGTGGAGCGGTCGTGAAGAAGATCGCGCACAGGCT
GTGTGCCGTTTTTGTGATACCGACTTCGTTGGCACGGATGGTGAAAATGGCGGTAAATTTAAAG
ATGCAGACGCCCTGGTTGCGACCATTGCAGGTCTGTGGCCGGCAGGTGAAGCTCATCGCTTCGT
GGTTTGCACGGGCGGTGAACCGATGCTGCAGCTGGATCAACCGCTGGTCGACGCACTGCACGCA
GCAGGTTTTGGTATTGCGATCGAAACCAACGGCAGCCTGCCGGTCCTGGAATCTATTGATTGGA
TCTGTGTGAGTCCGAAAGCAGACGCTCCGCTGGTCGTGACGAAGGGTAACGAACTGAAAGTTGT
CATTCCGCAGGATAATCAACGTCTGGCGGATTATGCCAAACTGGACTTTGAATACTTCCTGGTG
CAACCGATGGATGGCCCGTCCCGCGACCTGAATACCAAACTGGCCATCGATTGGTGTAAACGTC
ATCCGCAGTGGCGTCTGTCAATGCAAACCCATAAATACCTGAACATCCCG*TGA*T**AAGCTT**

(b)

MGSSHHHHHHSSGLVPRGSHMTYAVKEIFYTLQGEGANAGRPAVFCRFAGCNLWSGREEDRAQA
VCRFCDTDFVGTDGENGGKFKDADALVATIAGLWPAGEAHRFVVCTGGEPMLQLDQPLVDALHA
AGFGIAIETNGSLPVLESIDWICVSPKADAPLVVTKGNELKVVIPQDNQRLADYAKLDFEYFLV
QPMDGPSRDLNTKLAIDWCKRHPQWRLSMQTHKYLNIP

Supplementary Table 1 Data collection, phasing and refinement statistics for Fe-SAD structure of QueE from *B. multivorans*.

	AdoMet 6CP	
Data collection #		
Space group		*P* 4_3 2_1 2
Cell dimensions		
a, *b*, *c* (Å)		119.0,119.0,102.0
α, β, γ (°)		90, 90, 90
	Native	*Peak*
Wavelength	0.990	1.54178
Resolution (Å)	47.2-2.6	33.0-2.9
R_{sym}	0.067 (0.536) *	0.105 (0.482)
$I / \sigma I$	18.0 (2.5)	21.1 (4.9)
Completeness (%)	99.0 (99.4)	99.8 (100)
Redundancy	3.6 (3.6)	7.6 (7.4)
Refinement		
Resolution (Å)	47.2-2.6	
No. reflections	22,885	
R_{work} / R_{free}	0.176 / 0.208	
No. atoms		
Protein	3285	
Ligand/ion	115	
Water	103	
B-factors (Å^2)		
Protein	54.3	
Ligand/ion	50.4	
Water	48.1	
R.m.s deviations		
Bond lengths (Å)	0.002	
Bond angles (°)	0.6	

Each dataset was collected on a single crystal. *Highest-resolution shell is shown in parentheses.

Supplementary Table 2 Data collection and refinement statistics for QueE from *B. multivorans* (molecular replacement).

	AdoMet CPH_4 Na^+	AdoMet CPH_4 Mg^{2+}	AdoMet CPH_4 Mn^{2+} &	AdoMet CDG Mg^{2+}
Data collection #				
Space group	$P\,4_3\,2_1\,2$	$P\,4_3\,2_1\,2$	$P\,4_3\,2_1\,2$	$P\,4_3\,2_1\,2$
Cell dimensions				
a, *b*, *c* (Å)	119.1, 119.1, 103.4	119.1, 119.1, 104.3	119.2, 119.2, 105.3	118.58, 118.8, 102.7
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	39.0-1.9	28.9-2.2	48.2-2.7	47.2-1.91
R_{sym}	0.093 (0.565) *	0.102 (0.459)	0.076 (0.551)	0.054 (0.575)
$I / \sigma I$	15.7 (2.8)	12.1 (3.7)	27.7 (3.9)	29.2 (2.8)
Completeness (%)	99.0 (99.9)	99.6 (99.7)	99.9 (99.6)	96.8 (94.7)
Redundancy	4.6 (4.6)	4.2 (4.1)	8.6 (8.3)	5.4 (5.5)
Refinement				
Resolution (Å)	39.0-1.9	28.9-2.2	48.2-2.7	47.2-1.91
No. reflections	56302	38352	39665	55489
R_{work} / R_{free}	0.145 / 0.174	0.156 / 0.180	0.165 / 0.203	0.152 / 0.184
No. atoms				
Protein	3304	3291	3286	3302
Ligand/ion	114	113	104	112
Water	558	308	69	446
B-factors (Å^2)				
Protein	29.5	39.7	51.5	38.1
Ligand/ion	26.3	38.7	61.7	32.6
Water	41.4	44.6	45.2	45.5
R.m.s. deviations				
Bond lengths (Å)	0.013	0.010	0.003	0.017
Bond angles (°)	1.4	1.2	0.8	1.6

Each dataset was collected from a single crystal. & Friedel mates not merged for the Mn^{2+} - dataset. *Highest-resolution shell is shown in parentheses.

Supplementary Table 3 Structural comparison of QueE from *B. multivorans* to other AdoMet radical enzyme structures.

RMSD values are calculated using the DALI server[6].

AdoMet radical family member	PDB entry	RMSD (Å)	Cα atoms aligned	% Sequence identity
pyruvate formate-lyase activating enzyme	3C8F	2.7	171	17
lysine 2,3-aminomutase	2A5H	3.3	166	13
RlmN	4RF9	3.3	166	17
spore photoproduct lyase	4FHC	3.5	156	10
TYW1	2YX0	3.3	163	17
MoaA	2FB2	3.4	159	17
HydE	3CIW	4.0	156	11
coproporphyrinogen III oxidase	1OLT	4.0	156	13
biotin sythase	1R30	4.0	158	11
ThiC	3EPM	4.6	134	7
methyl ornithine synthase	3T7V	3.8	147	14

Supplementary Table 4 Metal ion distances and substrate dihedral angles.
Water molecules are labeled in **Figure 4**. Values indicated are in angstroms (Å) for distances or degrees (°) for dihedrals for both molecules of the QueE dimer.

	CPH_4 C4=O	**CPH_4 C6-COOH**	**H_2O-A**	**H_2O-B**	**H_2O-C**	**T51 Oγ1**	**CPH_4 N5**	**CPH_4 C4a-N5-C6-C**
AdoMet CPH_4 Mg^{2+}	3.0 3.0	2.8 3.0	2.4 2.8	2.8 2.7	3.2 3.3	2.9 2.8	2.9 3.0	-99° -96°
AdoMet CPH_4 Na^{+}	3.3 3.4	2.9 2.9	2.7 2.8	2.8 2.7	2.8 2.9	2.8 2.8	2.9 2.9	-145° -145°
AdoMet CPH_4 Mn^{2+}	2.6 2.7	3.1 3.4	-	2.3 2.5	3.0 2.4	3.4 3.7	2.0 2.0	-84° -82°

	CDG C6=O	**CDG C7-COOH**	**H_2O-A**	**H_2O-B**	**H_2O-C**	**T51 Oγ1**
AdoMet CDG Mg^{2+}	2.9 2.9	2.4 2.3	3.0 3.0	2.8 2.6	2.9 3.1	2.9 3.0

References

1. Vey, J.L. et al. Structural basis for glycyl radical formation by pyruvate formate-lyase activating enzyme. *Proc. Natl. Acad. Sci. USA* **105**, 16137-16141 (2008).

2. Vey, J.L. & Drennan, C.L. Structural insights into radical generation by the radical SAM superfamily. *Chem. Rev.* **111**, 2487-2506 (2011).

3. Sofia, H.J., Chen, G., Hetzler, B.G., Reyes-Spindola, J.F. & Miller, N.E. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res.* **29**, 1097-1106 (2001).

4. Layer, G., Moser, J., Heinz, D.W., Jahn, D. & Schubert, W.-D. Crystal structure of coproporphyrinogen III oxidase reveals cofactor geometry of Radical SAM enzymes. *EMBO J.* **22**, 6214-6224 (2003).

5. Berkovitch, F., Nicolet, Y., Wan, J.T., Jarrett, J.T. & Drennan, C.L. Crystal structure of biotin synthase, an *S*-adenosylmethionine-dependent radical enzyme. *Science* **303**, 76-79 (2004).

6. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545-W549 (2010).