Supporting Information

Vey *et al.* 10.1073/pnas.0806640105

Materials and Methods

Two pyruvate formate-lyase activating enzyme (PFL-AE) samples were used to solve the structures of PFL-AE presented. The first PFL-AE sample, referred to as AE, contained as-isolated enzyme with a 10-fold molar excess of *S*-adenosylmethionine (10× AdoMet), whereas the second PFL-AE sample, referred to as pept-AE, contained as-isolated enzyme, 10× AdoMet and a 10-fold molar excess of peptide substrate (10× peptide). The peptide substrate is the 7-mer RVSGYAV (K_m = 0.22 mM, compared with K_m = 1.4 μM for full PFL), corresponding to the area surrounding the catalytic glycine residue 734 of PFL on which PFL-AE can generate a glycyl radical (1).

Purification and Crystallization of AE. PFL-AE was purified according to published procedures (2, 3) except that 1 mM DTT was included in all buffers. The peptide (RVSGYAV, purity ≥98%) was obtained from Celteck Peptides. AdoMet was synthesized by using AdoMet synthetase as described in ref. 2. Crystals of AE were obtained in a Coy anaerobic chamber under an atmosphere of 5% H_2 and 95% argon gas mix at room temperature via hanging-drop vapor diffusion by mixing 1 μl of protein [0.730 mM in 50 mM Hepes, 200 mM NaCl, 1 mM DTT, 7.25 mM AdoMet (pH 7.5)] and 1 μl of crystallization buffer [0.1 M Tris (pH 8.5), 25% PEG 3350]. Drops were equilibrated against a well solution of 2.5 M ammonium sulfate to improve reproducibility and the crystallization time (typically 4 days). To improve diffraction quality, a detergent (0.11 mM octaethylene glycol monododecyl ether, $C_{12}E_8$; Hampton Research) was added to the crystallization drop and to all cryoprotectant solutions. Crystals were cryoprotected with a final concentration of 20% PEG 400 or 20% 2-methylpentane-2,4-diol added to the crystallization solutions, followed by freezing in liquid nitrogen in the anaerobic chamber.

Data Collection and Structure Determination of AE. Numerous native and multiwavelength anomalous dispersion (MAD) datasets were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) on beamlines 9-1, 9-2, and 1-5, and at the Advanced Light Source (ALS) on beamline 5.0.2. MAD datasets were collected in 20°–30° wedges by using the inverse-beam technique at the iron absorption edge, inflection, and a remote wavelength. The highest quality datasets are presented in Table S1. The three MAD datasets summarized here were collected from one crystal, and the native dataset was collected from a separate crystal. The data processed in mosflm (4) and HKL2000 (5) scaled reasonably well in space groups C2 (three molecules per asymmetric unit, ASU), $P3_1$ (two molecules per ASU), and $P3_121$ (one molecule per ASU). The data were of poor quality as judged by χ^2 and mosaicity values (which ranged between 1.0 and 1.8) and were initially difficult to index because of poor spot profiles and multiple lattices. Attempts to phase the data in $P3_121$ were unsuccessful, and the first reasonable maps were generated in $P3_1$. In the $P3_1$ space group, two iron sites (one for each cluster per molecule) were found by SOLVE (6) (www.solve.lanl.gov) by using SAD data obtained with the iron peak wavelength dataset. These sites matched anomalous difference Patterson maps calculated by XtalView (7) and yielded interpretable maps to 3.5 Å resolution after refinement of the iron sites in SHARP (8) (FOM (acent/cent) = 0.65/0.48) and solvent flattening in SOLOMON (9).

Initial Model Building and Refinement of AE by Using MAD Data. A model comprising two molecules (residues 5–230) was built in XtalView into solvent-flattened experimental maps calculated to 3.5 Å resolution. After building the cluster-binding loop and placing the cluster ligands into density, the individual iron sites were fit and refined to calculate maps to 2.87 Å resolution. The higher-resolution maps were used to complete the model further, and ≈40% of the residue side chains were added. Two chain breaks per molecule were observed, the first at residues 46–48 and a second break of unknown length toward the C terminus of each chain. The side chains of residues at the C terminus could not be modeled unambiguously. High *R* factors and overfitting of the model in CNS (10), possible signs of twinning, led to use of the h,-h-k,-l twin law in SHELXL (11) with a twin fraction of 0.5 (perfect merohedral twinning), although the cumulative intensity distribution of the data was normal and did not indicate twinning in any of the twin tests used. Refinement by this strategy yielded encouraging *R* factors (R = 28.82, R_{free} = 34.69) although only to 2.5 Å resolution and not the full (2.25 Å) extent of the data. In addition, no further improvement was possible. These problems (possible twinning and poor data quality) initially prevented the complete refinement of the model. Crystals of the pept-AE sample were then identified, allowing us to surmount this problem.

Crystallization of Pept-AE. Substrate-bound PFL-AE was also crystallized at room temperature by using hanging-drop vapor diffusion in an anaerobic chamber. The initial conditions were identified by sitting-drop vapor diffusion with high-throughput trays and the Hampton Index Screen. Crystals formed in approximately 6 days from drops comprising 1 μl of protein (0.68 mM protein, 6.80 mM AdoMet, 6.95 mM peptide, 50 mM Hepes, 200 mM NaCl, and 1 mM DTT) and 1 μl of crystallization buffer [100 mM Hepes (pH 6.8), 3.5 M sodium formate]. Larger crystals were obtained after optimization of the original condition by addition of detergents (Detergent Screens; Hampton Research). A single drop containing 0.2 mM 2,6-dimethyl-4-heptyl-β-D-maltopyranoside as an additive yielded several large (80 μm × 80 μm × 100 μm) crystals of good enough quality for data collection. These crystals were cryocooled with liquid nitrogen directly from the drop in the anaerobic chamber and stored for synchrotron data collection.

Data Collection and Structure Determination of Pept-AE. Native and MAD datasets were collected on beamlines 9-1 and 9-2 at SSRL. The data were processed with HKL2000 in space group $P6_122$, with 1 molecule per ASU and a solvent content of 54.2% (Table S1). Each dataset presented here (Table S1) was collected on a separate crystal because of crystal decay. Because molecular replacement with the initial AE model did not yield a solution that refined well, MAD data were used to obtain experimental maps. One site corresponding to the iron sulfur cluster was identified with SOLVE and refined in SHARP. The resulting maps were solvent-flattened in SOLOMON (through SHARP) to give interpretable density to 2.77 Å resolution (FOM (acent/cent) = 0.47/0.33).

Model Building and Refinement of Pept-AE. The initial AE model was fit into the new experimental maps manually, followed by rigid-body refinement in CNS. The new maps allowed placement of most side chains and correction and placement of the entire polypeptide backbone (residues 2–245), with no chain breaks, as

well as placement of the cluster, AdoMet, and peptide substrate. Calculation of phase combined $2F_o - F_c$, $F_o - F_c$ and composite omit maps in CNS using the MLHL target and careful manual rebuilding with extensive refinement allowed confident assignment of the peptide substrate orientation. Refinement was carried out in CNS by using the MLHL target with no sigma cutoff to $R_{work} = 22.9$ and $R_{free} = 26.1$ with 5.2% of reflections in the test set (Table 1). The final model contains PFL-AE residues 2–256, 4 iron atoms, 4 sulfur atoms, 27 AdoMet atoms, 6 peptide substrate residues, 16 waters, and 5 formate molecules (10 atoms) from the crystallization conditions.

Final Model Building and Refinement of AE Using Pept-AE Structure. After the pept-AE model was refined to reasonable *R* factors, these coordinates, minus the peptide substrate, were used to refine against data from the first crystal form. Refinement was carried out in CNS and SHELXL in C2, $P3_1$, and $P3_121$ to determine the proper space group. Fortunately, refinement using the new (corrected) model in CNS in $P3_121$ (with no twin law) gave better *R* factors and improved electron density maps. Apparently, the initial experimental maps were simply too poor to build parts of the model correctly, in particular loop C, which extends 25 Å over the cluster from the C terminus of the β sheet and includes 25 residues. Electron density at the 4Fe-4S unique iron has been modeled as just the methionyl moiety and C4′ and C5′ atoms of AdoMet; there is no electron density for the adenine or sugar of AdoMet in these maps. Although we cannot rule out that this density at the unique Fe is caused by water or another component of the crystallization conditions, the presence of AdoMet in the crystallization solution makes it a good candidate. It is unlikely that the partially ordered AdoMet observed here is a result of AdoMet cleavage; rather, the limited electron density, the lack of a suitable reducing agent, and the presence of electron density for atom C4′ suggests that AdoMet is disordered in this structure with only a portion of it being visible. Crystal lattice contacts, which are located far from this site, cannot explain this apparent disorder. Rounds of manual rebuilding and refinement in CNS with no sigma cutoff were done to a final R_{work} and R_{free} of 22.4 and 29.8, respectively, by using a test set of 9.5% of reflections (Table 1). This final model contains residues 1–245 of PFL-AE, 4 iron atoms, 4 sulfur atoms, 11 AdoMet atoms, 10 PEG atoms, and 49 waters.

We have confirmed in several ways that our AE data are indeed not twinned. First, there was no indication of twinning in the cumulative intensity distribution of the AE data in $P3_1$ and $P3_121$ (the following programs and servers were used to confirm: Merohedral Crystal Twinning Server (http://nihserver.mbi.ucla.edu/Twinning/), the Merohedral Twin Detector: Padilla–Yeates Algorithm (http://nihserver.mbi.ucla.edu/pystats/), Scala in ccp4, and CNS). Second, it is unlikely that a strong anomalous signal would be detected and then used successfully to generate reasonable experimental maps in perfectly twinned data. Third, the self-rotation functions calculated for these data are completely consistent with the space group $P3_121$. Fourth, the twin operator (h,-h-k,-l, equivalent to k,h,-l) used for refinement in Shelxl is actually a symmetry operator for space group $P3_121$. Finally, we were able to obtain lower *R* factors with the corrected model with normal refinement in CNS. We therefore attribute the overfitting observed in our final *R* factors and the high R_{free} to poor data quality. Because of this poor data quality, we have limited our discussion of this structure to the conformation of loop A, and for loop A, we show omit electron density maps so that the data quality for this region can be evaluated directly (Fig. S5).

Model Preparation for Docking Studies. The AE and pept-AE models were docked with a fragment of the PFL model [Protein Data Bank (PDB) ID code 2PFL] (12) corresponding to the portion of PFL homologous to the small protein YfiD (13) (residues 700–759). This fragment, termed the PFL radical domain (RD), was further truncated at the N terminus to improve docking results and minimize steric clashes (residues 712–759). When the pept-AE was used in docking studies, the peptide residues were deleted from the model before manipulation.

Manual Docking. A manual docking model of the RD with pept-AE was generated by orientation of the glycine loop of the RD in the active site in as close a conformation as possible to that of the bound peptide (see Fig. S6*d*). This docking model (Fig. S6*a*) was used as a guide to help evaluate the complexes output by ZDOCK (14).

Docking with ZDOCK. The docking algorithm ZDOCK generates the possible binding modes of two proteins based on their shape complementarity, electrostatics, and desolvation energies (14). ZDOCK was used as described on the program's website (http://zlab.bu.edu/zdock/). Residues of pept-AE and the RD were blocked from evaluation by the program if their involvement in the catalytically relevant complex was not expected (i.e., residues located on the opposite side of the AE with respect to the active site; see Fig. S6 *b* and *c*).

Evaluation of Docking Models. The top 20 complexes as scored by ZDOCK from each of six separate trials were evaluated based on the position of PFL G^{734} with respect to the AE active site and distance between G^{734} C_α and AdoMet C5′. Also, the degree of conservation of residues on the surface of PFL-AE, calculated by using ESPript (15), was used to evaluate the docking models. The eight most reasonable models were then refined by using rigid-body refinement and minimization in CNS (with no x-ray term) and reevaluated based on the criteria mentioned above as well as possible hydrogen-bonding contacts, interaction of RD with conserved AE residues, and agreement with biochemical data on the involvement of specific PFL residues in complex formation. In particular, R^{731} of PFL is necessary for the PFL activation reaction, whereas the hydroxyl groups of S^{733} and Y^{735} are not required (1). The model reported here (ZDOCK score of 31.12) was selected as the best complex based on the criteria discussed above, although it was not the highest scored in the automated docking algorithms.

Equipment and Settings. All figures were generated with Pymol (www.pymol.org) (16) with the exception of Fig. S4, which was generated with ClustalW (17) and Weblogo (18). Adobe Photoshop was used to add labels to all figures.

1. Frey M, Rothe M, Wagner AF, Knappe J (1994) Adenosylmethionine-dependent synthesis of the glycyl radical in pyruvate formate-lyase by abstraction of the glycine C-2 pro-S hydrogen atom: Studies of [2H]glycine-substituted enzyme and peptides homologous to the glycine 734 site. *J Biol Chem* 269:12432–12437.
2. Walsby CJ, *et al*. (2002) Electron-nuclear double resonance spectroscopic evidence that *S*-adenosylmethionine binds in contact with the catalytically active $[4Fe\text{-}4S]^+$ cluster of pyruvate formate-lyase-activating enzyme. *J Am Chem Soc* 124:3143–3151.
3. Broderick JB, *et al*. (2000) Pyruvate formate-lyase-activating enzyme: Strictly anaerobic isolation yields active enzyme containing a $[3Fe\text{-}4S]^+$ cluster. *Biochem Biophys Res Commun* 269:451–456.
4. Leslie AGW (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* 26.
5. Otwinowski Z, Minor W (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
6. Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D* 55:849–861.
7. McRee DE (1999) XtalView/Xfit: A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* 125:156–165.

8. Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W (2003) Generation, representation, and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr D* 59:2023–2030.
9. Abrahams JP, Leslie AG (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D* 52:30–42.
10. Brünger AT, *et al*. (1998) Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* 54:905–921.
11. Sheldrick GM, Schneider TR (1997) SHELXL: High-resolution refinement. *Methods Enzymol* 277:319–343.
12. Becker A, *et al*. (1999) Structure and mechanism of the glycyl radical enzyme pyruvate formate-lyase. *Nat Struct Biol* 6:969–975.
13. Wagner AF, *et al*. (2001) YfiD of *Escherichia coli* and Y06I of bacteriophage T4 as autonomous glycyl radical cofactors reconstituting the catalytic center of oxygen-fragmented pyruvate formate-lyase. *Biochem Biophys Res Commun* 285:456–462.
14. Chen R, Li L, Weng Z (2003) ZDOCK: An initial-stage protein-docking algorithm. *Proteins* 52:80–87.
15. Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15:305–308.
16. DeLano WL (2002) The PyMOL Molecular Graphics System (DeLano Scientific, Palo Alto, CA).
17. Claverie J, States D (1993) Information enhancement: Methods for large-scale sequence analysis. *Comput Chem* 17:191–201.
18. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Res* 14:1188–1190.

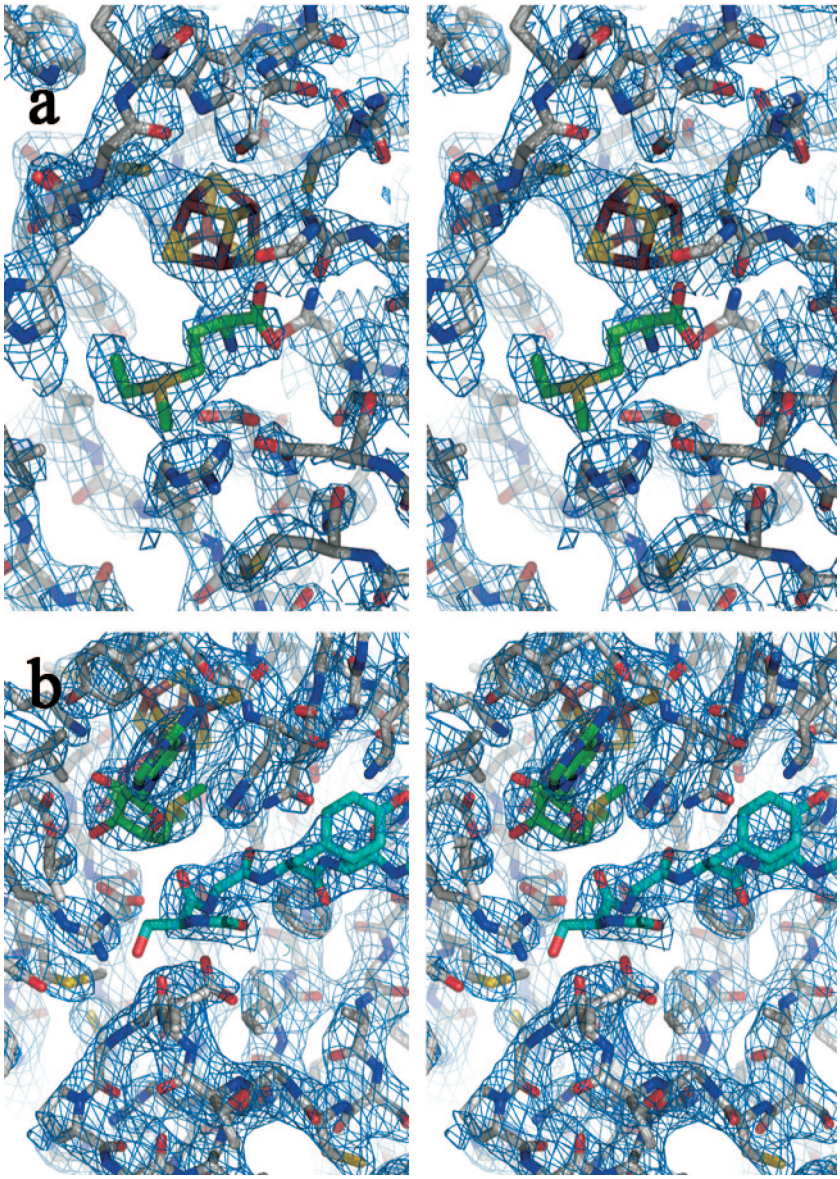

Fig. S1. Active-site electron density. Stereoview of the AE (*a*) and pept-AE (*b*) active sites with the 4Fe-4S cluster, AdoMet peptide, and protein side chains colored as follows: protein carbon atoms are gray; AdoMet carbons, green; peptide carbons, teal; oxygen, red; nitrogen, blue; sulfur, gold; and iron, ruby. (*a*) A $2F_o - F_c$ electron density map calculated by CNS and contoured at 1σ is shown in blue mesh. Extra density near the unique iron can be modeled as the AdoMet methionyl and C4′ and C5′ atoms (green carbon atoms). AdoMet is not fully ordered in this structure. (*b*) A phase-combined $2F_o - F_c$ electron density map calculated by CNS and contoured at 1σ is shown in blue mesh.

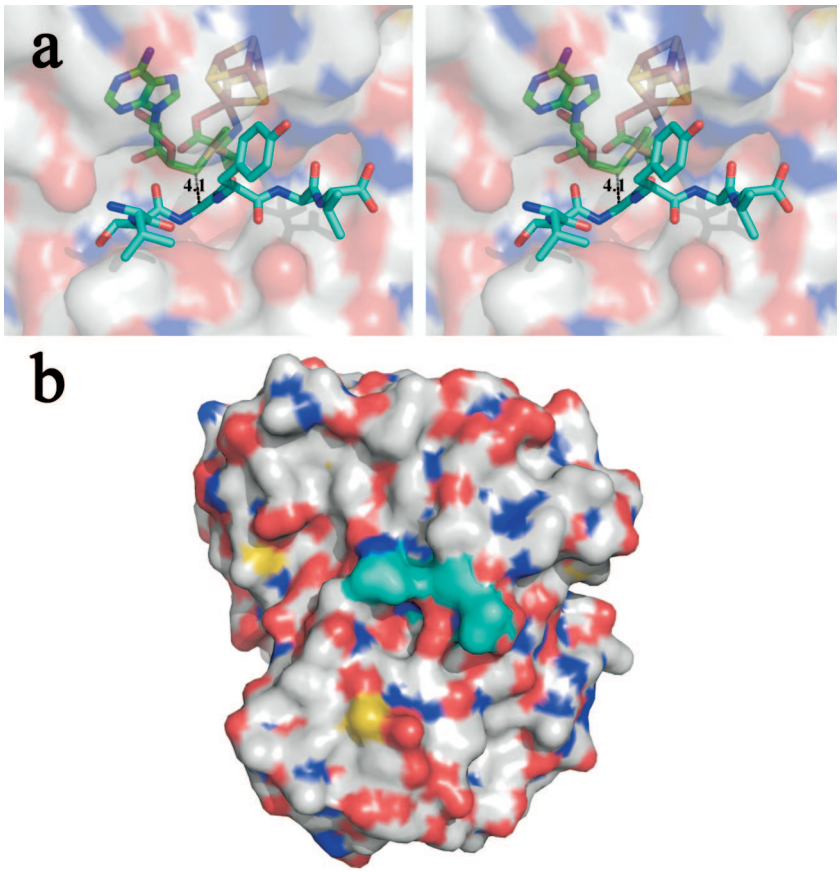

Fig. S2. Location of peptide-binding site. The 4Fe-4S cluster, AdoMet, and the peptide are shown in sticks colored as follows: iron, ruby; sulfur, gold; AdoMet carbons, green; peptide carbons, teal. The protein surface is shown colored by atom type (protein carbons, light gray; oxygens, red; nitrogens, blue). (*a*) Close-up stereoview of the active site. (*b*) Surface representation of the protein–peptide complex shows that no AdoMet (green) or 4Fe-4S cluster atoms (ruby and gold) are visible and therefore solvent-exposed when substrate is bound.

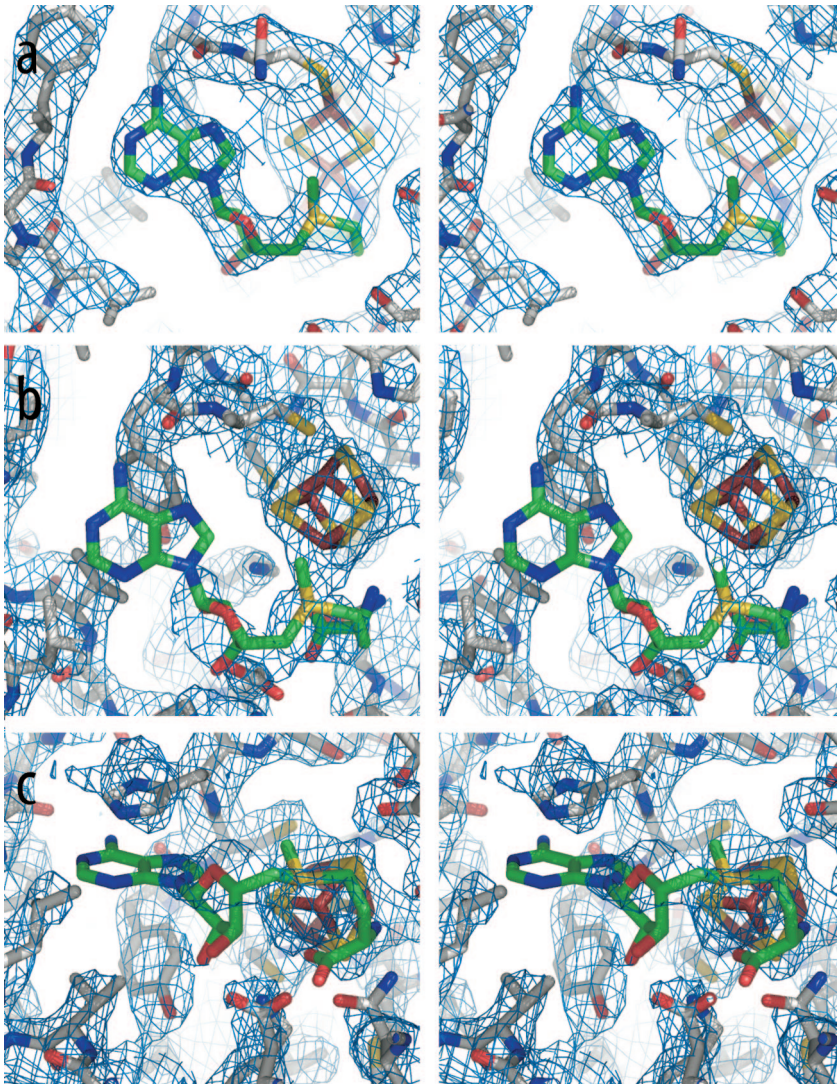

Fig. S3. Comparison of the electron density maps of the substrate-free and substrate-bound PFL-AE models at the AdoMet-binding site. Colors are as in Fig. S1. (*a*) Pept-AE model with a phase-combined composite omit electron density map calculated by CNS and contoured at 1σ. (*b*) AE model with a composite omit electron density map calculated by CNS and contoured at 1σ. AdoMet from the peptide structure is shown with green carbons. (*c*) AE model and electron density map as described in *b*, shown in an orientation rotated ≈90° about the plane of the paper from the orientation shown in *b*.

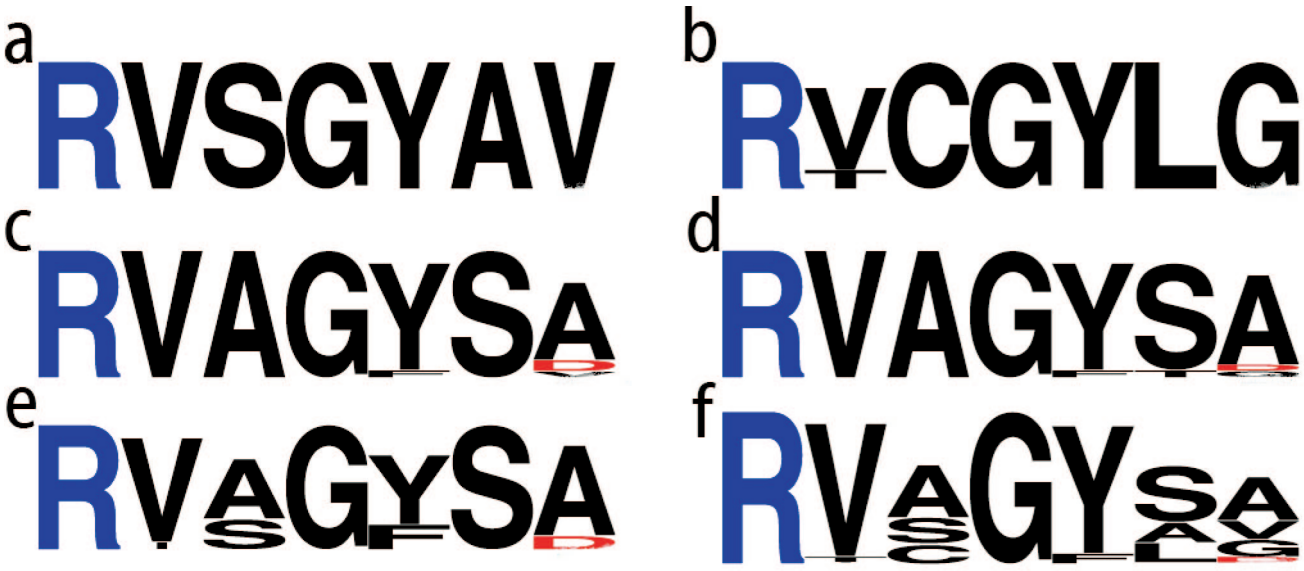

Fig. S4. Glycyl radical enzyme consensus motifs. Alignments were performed with ClustalW (17) by using 15–16 sequences for each GRE subfamily, and logos were generated by using WebLogo (18). The letter size represents the frequency at which the residue is observed at that position in the sequence. Basic residues are colored in blue, acidic residues in red, and others in black. Consensus motifs shown are: (*a*) PFL. (*b*) Class III ribonucleotide reductase. Residues at position 2 are Val and Thr. (*c*) Glycerol dehydratase. Residues at position 7 are Ala, Asp, and Val. (*d*) 4-Hydroxyphenylacetate decarboxylase. Residues at position 7 are Ala, Asp, and Gly. (*e*) Benzylsuccinate synthase. (*f*) Overall GRE. Residues at position 2 are Val and Thr; at position 7 are Ala, Val, Gly, and Asp.

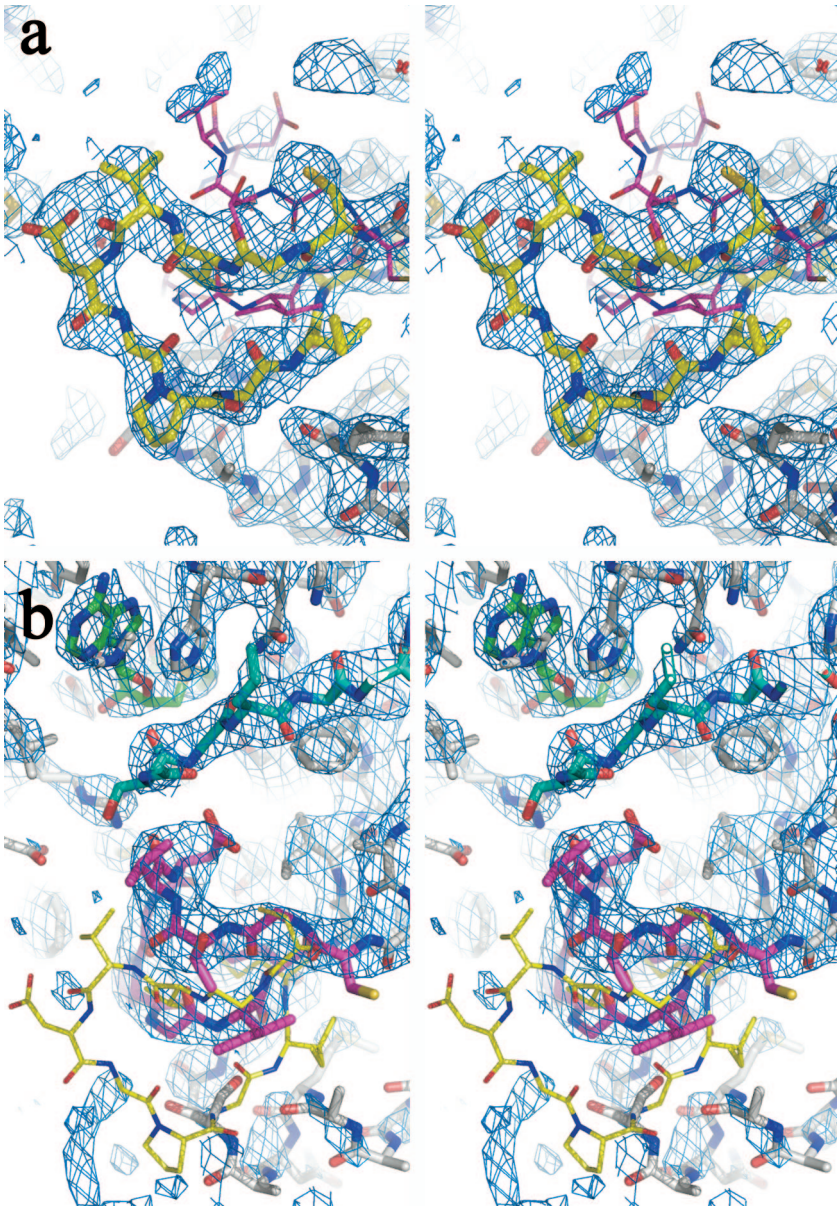

Fig. S5. Electron density for loop A. $2F_o - F_c$ composite omit map density for the (*a*) AE model and (*b*) pept-AE model, both contoured at 1σ. Loop A of the AE model is shown with yellow carbon atoms (displayed as sticks in *a* and as lines in *b*), whereas loop A from the pept-AE model is shown with carbons colored magenta (displayed as lines in *a* and as sticks in *b*).

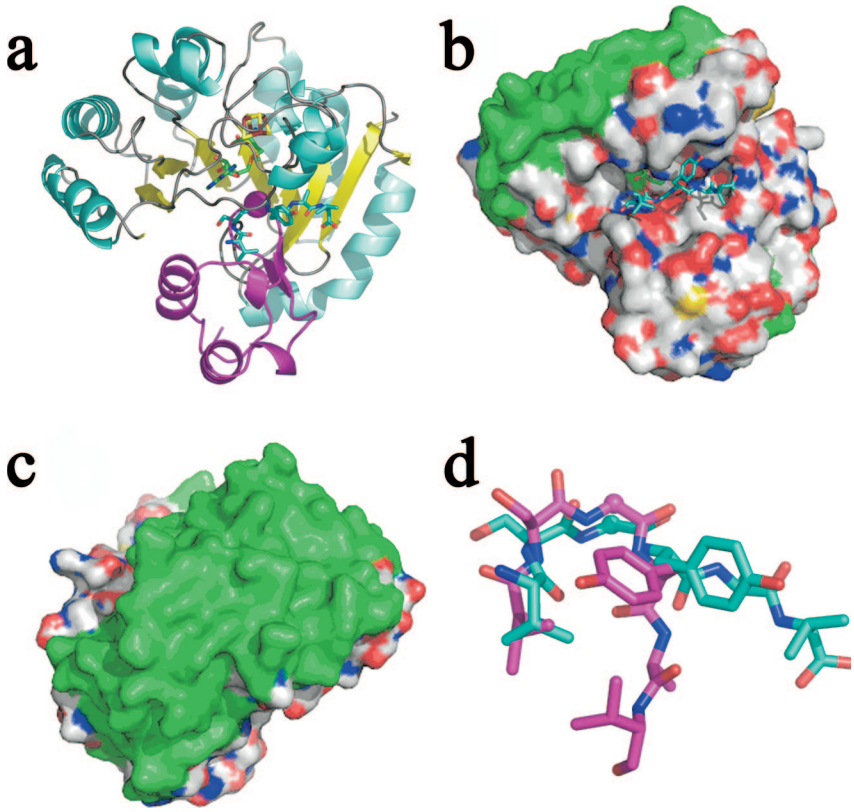

Fig. S6. Docking models. (*a*) Manual docking model. Pept-AE is colored as in Fig. 1, with the RD displayed in magenta ribbons. The peptide is shown in sticks for comparison. (*b* and *c*) Surface representation of pept-AE colored as in Fig. S2, with residues blocked from the docking calculations highlighted in green. Orientations are as in Fig. 4. (*d*) Comparison of the conformation of residues 732–737 in the pept-AE model (teal carbons) and the docked RD (magenta carbons). To make this figure, the best docking model was superimposed with the pept-AE model. The C_α of each G^{734} is shown as a small sphere.

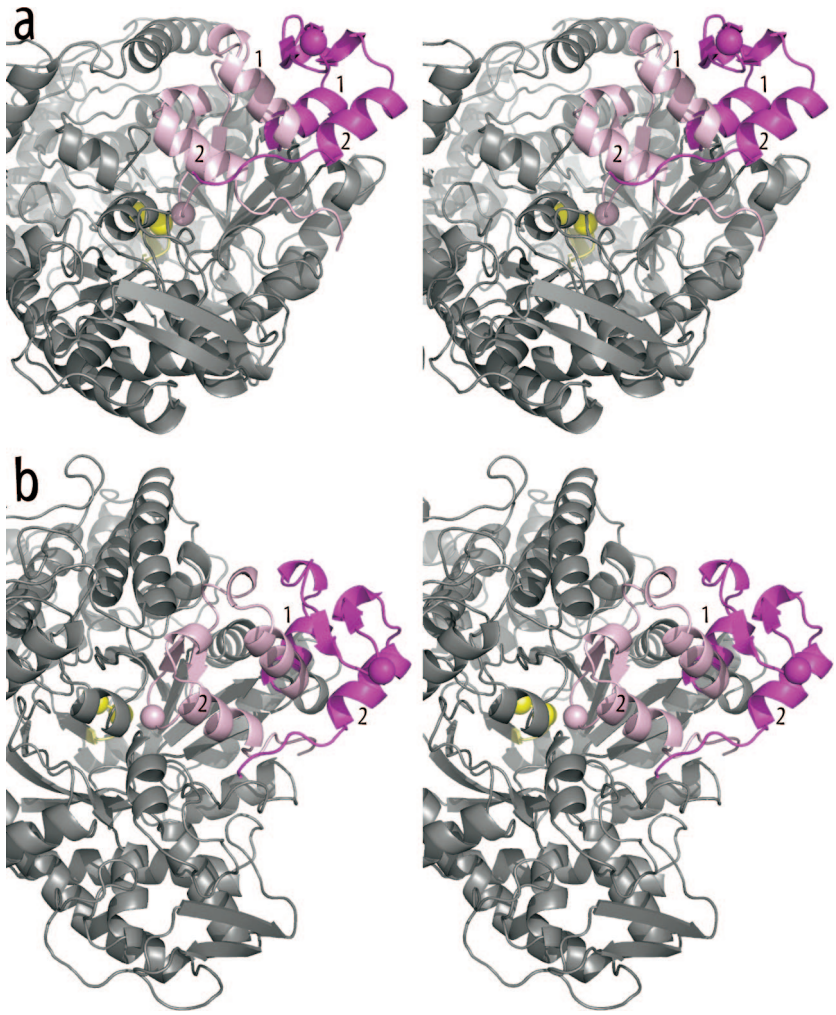

Fig. S7. Proposed PFL conformational change. Stereoview of the PFL dimer (PDB code 2PFL) showing close-up view of one monomer and focusing on the RD in two orientations. The orientation in *a* is rotated ≈60° about a horizontal axis with respect to *b*. The RD is colored light pink in its native conformation and magenta in the proposed accessible conformation. The other PFL finger loop containing the active site cysteines (C^{418} and C^{419}) is colored yellow. The C_α atoms of G^{734}, C^{418}, and C^{419} are shown in spheres.

Table S1. Data collection statistics

	AE				Pept-AE		
Dataset	Peak	Inflection	Remote	Native	Peak	Inflection	Remote
Wavelength, Å	1.73542	1.74166	1.37755	1.00000	1.73955	1.74158	1.36241
Space group	$P3_121$	$P3_121$	$P3_121$	$P3_121$	$P6_122$	$P6_122$	$P6_122$
Resolution, Å	2.87	3.20	3.70	2.25	2.90	2.80	2.77
R_{symm}, %	7.1 (30.5)	7.0 (27.8)	7.9 (26.0)	7.8 (32.2)	9.8 (42.0)	8.1 (36.9)	7.3 (35.9)
Unique Obs	9,997 (1,021)	7,363 (742)	4,680 (460)	11,183 (1,103)	12,853 (1,152)	12,677 (1,107)	14,585 (1,406)
I/σ	36.9 (7.4)	37.6 (8.3)	21.1 (7.1)	23.9 (4.0)	18.6 (2.7)	13.1 (3.1)	19.9 (3.1)
% Complete	100.0 (100.0)	99.9 (100.0)	99.8 (100.0)	99.0 (93.7)	98.5 (87.6)	88.7 (78.6)	98.6 (95.8)
Redundancy	10.8 (10.9)	10.7 (10.8)	5.0 (4.9)	9.6 (4.4)	9.8 (5.8)	5.4 (5.0)	5.5 (4.0)
a, Å	58.09	58.25	58.27	57.98	74.56	74.41	74.43
c, Å	117.27	117.64	117.85	117.37	187.98	187.46	187.61

Values in parentheses refer to the high-resolution bin. Friedel mates were not merged in the processing of peak, inflection, or remote wavelength datasets.

Table S2. Root mean square deviations between structures of AdoMet radical proteins

Protein	Pept-AE	AE
AE	0.77 (245)	
BioB	2.04 (100)	2.08 (98)
HemN	1.99 (109)	2.01 (115)
MoaA	1.56 (145)	1.47 (140)
LAM	2.05 (101)	1.97 (107)
mjTYW1*	1.71 (118)	1.99 (116)
phTYW1†	1.57 (135)	1.86 (121)

Root mean square deviations are shown in Å. The number of C_α atoms matched in the calculation is shown in parentheses.
*Enzyme from *Methanococcus jannaschii* TYW1 enzyme.
†Enzyme from *Pyrococcus horikoshii*.