X-ray structure of an AdoMet radical activase reveals an anaerobic solution for formylglycine posttranslational modification

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
X-ray structure of an AdoMet radical activase reveals an anaerobic solution for formylglycine posttranslational modification

Peter J. Goldman, Tyler L. Grove, Lauren A. Sites, Martin I. McLaughlin, Squire J. Booker, and Catherine L. Drennan

Departments of Chemistry and Biology and Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139; and Departments of Chemistry and Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802

Edited by Vern L. Schramm, Albert Einstein College of Medicine of Yeshiva University, Bronx, NY, and approved April 12, 2013 (received for review February 5, 2013)

Arylsulfatases require a maturing enzyme to perform a co- or posttranslational modification to form a catalytically essential formylglycine (FGly) residue. In organisms that live aerobically, molecular oxygen is used enzymatically to oxidize cysteine to FGly. Under anaerobic conditions, S-adenosylmethionine (AdoMet) radical chemistry is used. Here we present the structures of an anaerobic sulfatase maturing enzyme (anSME), both with and without peptidyl-substrates, at 1.6–1.8 Å resolution. We find that anSMEs differ from their aerobic counterparts in using backbone-based hydrogen-bonding patterns to interact with their peptidyl-substrates, leading to decreased sequence specificity. These anSME structures from Clostridium perfringens are also the first of an AdoMet radical enzyme that performs dehydrogenase chemistry. Together with accompanying mutagenesis data, a mechanistic proposal is put forth for how AdoMet radical chemistry is coopted to perform a dehydrogenation reaction. In the oxidation of cysteine or serine to FGly by anSME, we identify D277 and an auxiliary [4Fe-4S] cluster as the likely acceptor of the final proton and electron, respectively. D277 and both auxiliary clusters are housed in a cysteine-rich C-terminal domain, termed SPASM domain, that contains homology to ~1,400 other unique AdoMet radical enzymes proposed to use [4Fe-4S] clusters to ligate peptidyl-substrates for subsequent modification. In contrast to this proposal, we find that neither auxiliary cluster in anSME bind substrate, and both are fully ligated by cysteine residues. Instead, our structural data suggest that the placement of these auxiliary clusters creates a conduit for electrons to travel from the buried substrate to the protein surface.

Iron-sulfur cluster fold | radical SAM dehydrogenase

Posttranslational modification expands the chemical repertoire of enzymes, in some cases by generating modified amino acids that are well suited to perform specific reactions. Arylsulfatases, for example, require the co- or posttranslational formation of a catalytically essential formylglycine (FGly) moiety to perform their hydrolysis function, removing sulfate groups from a wide array of substrates (e.g., sulfated polysaccharides, sulfolipids, and steroid sulfates) (1–3). In humans, lack of sulfatase activity can lead to disease (4), while in bacteria, inhibition impairs colonization of the mucosal layer of the host gut (5). The maturation of these sulfatases involves two classes of enzymes, one that requires molecular oxygen and another that can function in its absence. FGly generating enzymes (FGEs), found in eukaryotes or aerobically living prokaryotes, generate FGly by oxidizing a cysteine residue on the target sulfatase using molecular oxygen (6, 7), whereas anaerobic sulfatase maturing enzymes (anSMEs) generate FGly from either cysteine or serine residues on their target sulfatases using S-adenosyl-l-methionine (AdoMet) radical chemistry (8–11). In addition to their importance for sulfatase chemistry, FGEs have commercial applications for generating site-specific “aldehyde tags” to use in protein-labeling technology (12). While FGEs have been characterized in terms of structure and mechanism (6, 7), far less is known about their anaerobic cousins, the anSMEs. Here we provide structural insights into these unusual AdoMet radical enzymes.

The AdoMet radical enzyme family catalyzes a diverse array of radical-based reactions, including sulfur insertions, complex chemical transformations and rearrangements, DNA and RNA modifications, and in the case of anSMEs, dehydrogenation (13) (Fig. 1). Members of this family have historically been identified by a CX3CXΦC motif (where Φ is an aromatic residue), which ligates three of the four irons of a [4Fe-4S] cluster (14, 15), leaving the fourth iron free to bind AdoMet (16). Radical chemistry is initiated by the injection of an electron via the [4Fe-4S] cluster into AdoMet, resulting in the homolytic cleavage of the molecule into methionine and a 5′-deoxyadenosyl radical (5′dA, Fig. 1). This radical species subsequently abstracts a hydrogen atom from substrate, resulting in 5′-deoxyadenosine (5′dA) and a substrate radical. Differentiation among the family members is a result of the action of this substrate radical. In anSMEs, the removal of a proton and an electron from the radical intermediate completes catalysis (10, 17) (Fig. 1).

The AdoMet radical dehydrogenase subfamily includes anSMEs and the carbohydrate natural product biosynthetic enzyme BtrN (18, 19). Interestingly, both enzymes harbor additional [4Fe-4S] clusters that are necessary for turnover (20). In the case of BtrN, one auxiliary cluster has been identified (21), while anSMEs have two (10, 17). For anSME, the sequence surrounding these two clusters, including a previously identified 7-cysteine motif (CX3_15
\[\text{G}X_{16-17}\text{C}—\text{gap—}\text{CX}_1\text{C}_1\text{C}—\text{gap—}\text{C}\] (17), places it in a ~1,400-membered AdoMet radical subfamily that was recently described by Haft and Basu through bioinformatic analysis and thought to function in the modification of ribosomally translated peptides (22). This subfamily has been designated TIGR04085 and named SPASM for its biochemically characterized founding members AlbA, PqqE, anSMEs, and MtIC, which are involved in subtilosin A, pyrroloquinoline quinone, anaerobic sulfatase, and mycofactacin maturation, respectively (22, 23). While the function of these auxiliary clusters is unknown, the 7-cysteine motif prompted speculation that members of the SPASM subfamily, including anSMEs, use an available ligation site on one of the [4Fe-4S] clusters for substrate binding (10). Direct binding of substrate

Author contributions: P.J.G., T.L.G., L.A.S., S.J.B., and C.L.D. designed research; P.J.G. and T.L.G. provided the crystallization and crystal structure determination; M.I.M. performed the crystallization and crystal structure determination; T.L.G. provided Hi66 and native protein samples; L.A.S. generated mutant constructs and performed activity assays; and P.J.G. and C.L.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4K36–4K39).

1To whom correspondence should be addressed. E-mail: cldrennan@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1302417110/-/DCSupplemental.
to an auxiliary [4Fe-4S] cluster has been observed in the molybdemen cofactor biosynthetic enzyme MoeA, another AdoMet radical enzyme (24, 25). Other possible functions specific to anSMEs include the second oxidation of substrate or substrate deprotonation (Fig. 1).

In this work, we report structures of anSMEcpe, a biochemically characterized anSME from Clostridium perfringens, which can oxidize either serines or cysteines into FGly sidechains (26, 27). We have solved the structure of a His$_6$-tagged protein construct with AdoMet bound and three structures of an untagged protein construct in the presence of AdoMet, one substrate free and two others with different peptide substrates bound. The overall structure and substrate binding characteristics of anSMEcpe differ significantly from the aerobic FGE system, which clarifies the difference in promiscuity between the two enzyme families. Surprisingly, the structures show full cysteine ligation of both auxiliary clusters, which has important implications for the anSMEcpe mechanism, as well as for SPASM family members as a whole.

**Results**

Iron anomalous signal from a dataset collected at a home Cu-K$_\alpha$ source was used to solve an initial structure of His$_6$-tagged anSMEcpe (“His$_6$, AdoMet-bound” structure; Table S1). Due to occupancy of the His$_6$ tag in the substrate binding site, additional structures of native, untagged anSMEcpe were solved with and without two substrate peptides (“AdoMet-bound,” “Kp18Cys, AdoMet,” and “Cp18Cys, AdoMet” structures; Table S1). All structures of anSMEcpe contain three [4Fe-4S] clusters including the AdoMet cluster housed in the N-terminal AdoMet radical domain and two auxiliary clusters located in the C-terminal SPASM domain. The two domains are connected by the $\alpha_6$ helix and the protein terminates with the $\alpha_6'$ helix (Fig. 2 A–C).

**Structural Features of the anSMEcpe AdoMet Radical Domain.** The N-terminal domain of anSMEcpe is a parallel ($\beta$)$_6$ partial triose phosphate isomerase (TIM) barrel, spanning residues 3–234 (Fig. 2 A and C, magenta). This fold is common to nearly all other structurally characterized members of the AdoMet radical superfamily (28, 29). We will refer to this parallel barrel as the AdoMet domain. Here, the AdoMet radical sequence motif (CX$_4$CX$_4$FC$\beta$) is found in a loop following the $\beta$ strand. C15, C19, and C22 each ligate an iron atom in one of the [4Fe-4S] clusters, referred to here as the AdoMet cluster. The fourth, so-called “unique iron,” is ligated by the amine nitrogen and carbonyl oxygen from the methionine moiety of AdoMet, as expected (16). All structures contain clear density for AdoMet bound in the active site (Fig. S14), with the exception of a substrate-bound (Cp18Cys) crocystal structure where the shape and position of density for AdoMet in chain B is minimal and inconsistent with other models.

Four previously described AdoMet binding motifs are conserved in anSMEcpe, including the “GGE” motif, the ribose motif, the “GXIXGXXE” motif, and the $\beta$$_6$ motif (Figs. S1C and S2) (28, 29). In addition, the backbone of Y21 (the hydrophobic residue in the AdoMet radical CX$_4$CX$_6$C motif) hydrogen bonds with the N6 position of adenine, and R143, just following

---

**Fig. 1.** anSME reaction. (1) Electron donation to the AdoMet radical cluster initiates homolysis of AdoMet and 5’dA$\alpha$ formation in the presence of bound substrate. (2) Substrate radical generation, (3) Deprotonation of the substrate Cys sidechain, (4) Substrate oxidation, and (5) Hydrolysis of the thioaldehyde intermediate yields the FGly moiety of the activated sulfatase.

**Fig. 2.** Structure of anSMEcpe. (A) The AdoMet domain (magenta) contains the AdoMet cluster and the (\beta)$_6$, partial TIM barrel. The SPASM domain (green) comprises most of the C-terminal segment and houses the remaining two [4Fe-4S] clusters. Two helices, $\alpha_6$ and $\alpha_6'$, are not part of either domain and are colored light blue. (B) Positions and distances between the three [4Fe-4S] clusters (stick representation with Fe in orange and S in yellow). (C) Topology of anSMEcpe. A, AdoMet cluster; I, Aux I; II, Aux II. (D) The SPASM domain, colored by the level of sequence homology between anSMEcpe and the other 280 members of TIGR04085. Conservation scores were calculated by the ConSurf server (40). Iron ligating cysteines are shown as spheres and labeled by their designation in C. Aux I and II are shown in stick representation and labeled.
α4a, stabilizes the ribosyl and carboxyl moieties of AdoMet. An arginine following α4a makes a similar interaction in the AdoMet radical proteins HemN and HydE (30, 31). On the backside of the AdoMet domain, a patch of conservation can be found following the β2 strand (Fig. S1B). This site is the proposed binding location of the physiological reductant, commonly flavodoxin (30, 32).

**SPASM Domain and Auxiliary Cluster Binding.** The α6 helix links the AdoMet domain to the C-terminal SPASM domain (Fig. 2 A–C, green), the latter containing both auxiliary clusters. According to the designation by TIGR04085, the SPASM domain initiates at C261, not C255 (the first Fe ligating cysteine in anSMEcpe). However, both cysteines are conserved in anSMEs and it is common for SPASM domain-containing proteins to have a proximal upstream cysteine (22). In the anSMEcpe structures, C255 and C261 ligate the first auxiliary cluster, Auxiliary Cluster I (Aux I), before the backbone folds into a beta hairpin (V266–Y274). Residues of this hairpin exhibit high conservation to other members of the SPASM family (Fig. 2D). Immediately following the hairpin, C276 provides the third ligand to Aux I. A variable alpha helical region follows, providing a barrier between this cluster and solvent. Aux I lies 16.9 Å from the AdoMet cluster (measured from the closest atom in each cluster).

The second portion of the SPASM domain contains the CX2CXCX2CX3C part of the 7-cysteine motif. The first three cysteines (C317, C320, and C326) provide three ligands to the second auxiliary cluster, Aux II, while the protein backbone forms small helical interactions that surround the cluster. The fourth cysteine of the motif, C330, crosses back to provide the fourth ligand to this cluster (Fig. 2F). Very high sequence conservation is found in the linear CXXXC region bridging the two clusters, where the last two residues before the final cysteine of the motif are glycines (G328 and G329 in anSMEcpe; Fig. 2D). Aux II lies 12.9 Å away from Aux I and 26.7 Å away from the AdoMet cluster (Fig. 2B). A series of loops follows CX2CXCX2CX3C before C348 (the final of the seven cysteines) occupies the final ligation site of Aux II, ending the SPASM domain and initiating the α6′ helix. This helix lies adjacent to the α6 helix and completes the barrel. The C terminus of the protein lies at the end of this helix (Fig. 2C).

**Structural Homology to MoaA.** While the AdoMet domain of anSMEcpe is very similar to other members of the AdoMet radical family, the C-terminal SPASM domain is structurally similar (rmsd 6.3 Å) to only one other AdoMet radical protein, MoaA (33). Like anSMEcpe, MoaA ligates a C-terminal auxiliary cluster that overlays well with anSMEcpe’s Aux I (Fig. S3). Following two cysteine ligands to its auxiliary cluster, MoaA has a beta hairpin that shares high sequence homology with anSMEcpe and other SPASM domain containing proteins, including G273 (G271 in anSMEcpe) in the n+3 position of the hairpin turn and Y276 (Y274 in anSMEcpe), a residue that contributes to a hydrophobic pocket adjacent to the hairpin. MoaA then has a third cysteine ligand and terminates after a helical region. It lacks both the CX2CXCX2CX3C motif and the second auxiliary cluster that are common in SPASM family members. With only these three protein ligands to its auxiliary cluster, MoaA uses the available coordination site to bind substrate (24, 25). In anSMEcpe, the final cysteine of the CX2CXCX2CX3C motif is the fourth ligand to this cluster (Fig. S3 E and F).

**Binding Specificity for Substrate Peptides.** A cavity underneath the barrel is the only access to the active site from the exterior of the protein. In all structures of anSMEcpe solved using the His8–anSMEcpe construct, multiple histidines of the C-terminal tag could be modeled into residual electron density. These residues appear to block access to the active site. However, when native protein was used for cocystalization with AdoMet and peptide substrates, peptide density reaching into the active site was present. Two substrate-bound structures were solved using peptides designed to mimic a C. perfringens (Cp18Cys) and a Klebsiella pneumoniae (Kp18Cys) sulfatase protein (see Fig. S44 for sequences). Each structure contains density for 9–11 residues of the 18mers (Fig. S54). These peptide residues stretch from the exterior of the protein to the active site, both entering and exiting via the cavity at the bottom of the barrel (Fig. 3A and B). Buried surface area of the protein–peptide interaction for the four bound peptides to anSMEcpe (two per asymmetric unit) is 786 ± 38 Å², corresponding to 60.7 ± 1.7% of the total surface area of the modeled portion of the peptides (34). This binding mode differs from the aerobic sulfatase maturing enzymes, FGEs, which use the same CXPXR motif for substrate recognition. In the aerobic system, the sulfatase maturase adopts a very different fold that lacks the internal cavity found in anSMEcpe. Instead, peptide binding and catalysis occurs on a surface-exposed region of the enzyme (Fig. 3C) (7). Further, in the aerobic system, only four of the 12 hydrogen bonds between FGE and its substrate use the substrate backbone [Protein Data Bank (PDB) ID code 2A1J]. The majority of interactions are made between peptide sidechains and the maturase, explaining the high sequence specificity in this system (7).

Other than the arginine residue of the CX(A/P)XR motif (Fig. S6 A and B), this side-chain–based sequence specificity is not seen in anSMEs. First, anSMEcpe is able to accommodate considerably more sequence variation on the substrate cysteine (Fig. S6 C and D), allowing the two peptides (Cp18Cys and Kp18Cys) to bind in almost an identical orientation (Fig. 3A). In both cases, backbone hydrogen bonds to anSMEcpe are the primary means of stabilization (Fig. S5C). Only two positions, 4 and 10, of the peptides differ, resulting in two additional hydrogen bonds at the 4 position of Kp18Cys (Figs. S4 and S5). The remaining 17 hydrogen bonding interactions are conserved between the two substrates. Of these, 12 are formed between peptide backbone and anSMEcpe and five are formed between anSMEcpe and a single peptidyl sidechain, R11 of the conserved CX(A/P)XR sulfatase motif. The extensive binding pocket created for this arginine, made up of F188, E159, and L118, uses π stacking, electrostatics/hydrogen bonding, and van der Waals interactions, respectively (Fig. S6A). These interactions appear to be the anchor for the peptide and are conserved among the other biochemically characterized anSMEs (Fig. S2).

**Identification of Catalytic Residues.** Aided by two prolines, the substrate peptide makes a tight turn in the active site, allowing the target cysteine to protrude into the deepest part of the barrel, just below AdoMet. The cysteiny1 Cβ is located 4.1 Å from the 5′ carbon of AdoMet (Fig. 4A). This distance is in agreement with previously reported distances between the 5′ position of AdoMet and the substrate hydrogen abstraction site (3.8–4.1 Å) (28). The orientation of the cysteine directs the Cβ pro-S hydrogen toward the AdoMet 5′ position, matching biochemical evidence for the enzyme’s stereoselectivity (26).

During catalysis, a general base is needed for deprotonation of the cysteine side chain to allow the formation of the thioaldehyde (Fig. 1). Analysis of the peptide-binding pocket revealed the presence of two residues with titratable side chains within 5 Å of the substrate cysteiny1 sulfur position, D277 and Y24 (Fig. 4C). Y24 is two residues downstream of the AdoMet radical domain’s CX2CXC motif, while D277 is in the SPASM domain and adjacent to C276, an Aux I ligand. To identify the catalytic residue, two mutants were generated and assayed for activity, anSMEcpe D277N and Y24F. Compared with wild type, the Y24F mutant retains 11.7% FGly production activity, while the D277N mutant only retains 0.8% activity. Along with this decreased activity, an uncoupling of the production of FGly and S-da is observed for the D277N mutant (Fig. 4F). The proximity to the substrate sulfur and a large decrease in activity imply a catalytic role for the D277 side.
chain. Two residues, Q64 and Q98, have a role in stabilizing the catalytic conformation of D277, as well as helping with stabilizing AdoMet binding. These residues have distinct conformations in the presence or absence of AdoMet and substrate peptide (SI Results and Fig. S5 D and E).

**Electron Transfer Pathway.** After hydrogen abstraction and thiol deprotonation, formation of the thioaldehyde drives oxidation of the Cβ, requiring an electron acceptor to complete catalysis. Cβ lies 8.6 and 0.9 Å from Aux I and the AdoMet cluster, respectively, indicating both clusters are within suitable distance to be electron transfer partners (35). If an electron transfer event results in the reduction of Aux I, removing the electron from the system would likely require transfer from Aux I to Aux II, as peptide binding provides a barrier between Aux I and solvent (Fig. S7). Aux I lies near the bottom opening of the anSMEcpe barrel. When no peptide is present, the cluster is 9.7 Å from bulk solvent, with the inside of the protein barrel as the closest protein–bulk solvent interface (Fig. S74). In this conformation, the cluster has a similar residue depth to both the AdoMet cluster (9.5 Å) and Aux II (8.3 Å). However, when substrate is present this avenue to solvent is cut off, and the shortest path to bulk solvent is below the barrel, 11.0 Å away (Fig. S7B). Peptide binding does not affect the residue depth of either the AdoMet cluster or Aux II.

**Discussion**

Here we present the structures of an anSME, anSMEcpe, which allow us to compare how nature evolved anaerobic as well as aerobic solutions for the same enzyme function. While both enzyme classes are designed to bind their target sulfatase, in one case, binding must involve sequestering a Cys/Ser to afford radical-based chemistry, while in the other, the target Cys must be accessible to interact with molecular oxygen (6–8, 17). We find that the aerobic FGEs and anSMEs use different protein folds, with the N-terminal domain of anSMEs sharing a classic AdoMet radical partial barrel fold (28, 29), consistent with the chemistry being performed. While the active site of FGE is on the surface of that enzyme, where it is readily accessible to molecular oxygen, the active site of anSME is buried in a cleft created between the C-terminal SPASM domain and the N-terminal AdoMet radical domain. To fit into this cavity, the target peptide adopts a relatively tight turn, perhaps explaining the preference for Ala or Pro in position of the conserved (S/C)X(A/P)XR motif. In contrast, peptides bound to FGE have no apparent conformational restraints (Fig. 3) (7).

FGEs and anSMEs also vary in their substrate selectivity. Compared with FGEs, anSMEs are able to act on a larger variety of peptide substrates. For example, anSMEcpe itself can bind and catalyze FGlY formation on a C. perfringens substrate analog as well as a K. pneumoniae substrate analog (26). Another anSME from *Bacteroides thetaiotaomicron* is responsible for activating up to 28 sulfatases under anaerobic conditions (5). From structural comparisons, we can now explain this substrate specificity variation between these enzyme classes. While the aerobic system uses a mainly side-chain–mature hydrogen bonding network for substrate stabilization, the anSME system uses a primarily backbone–mature hydrogen bonding network, with only the Arg of the (C/S)X(A/P)XR motif involved in side-chain–based hydrogen bonding. This reliance by SMEs on primarily peptide backbone-based hydrogen bonding interactions is in agreement with the much higher degree of promiscuity that exists for the anSME system in relation to the FGE system. Interestingly, pyruvate formate lyase–activating enzyme (PFL-AE), the only other structurally characterized AdoMet radical enzyme involved in protein modification, also uses primarily peptide backbone–activase interactions (32). As more structures become available, it will be interesting to see if this binding mode will be common to AdoMet radical enzymes that act on protein substrates.

These structures of an anSME also provide insight into the catalytic mechanism of this enzyme class. For catalysis, at least five steps are required (Fig. 1). For the first step of 5′dA• generation, the high degree of similarity between the structure of the anSME AdoMet radical domain and structures of other AdoMet radical enzymes suggests that all components necessary for radical generation are found in this N-terminal domain, implying...
that anSMEs share a common initiation mechanism with the rest of the AdoMet radical superfamily.

In terms of substrate radical generation, these structures show the Cys 4.1 Å from the AdoMet 5°C, in agreement with all other structures of AdoMet radical enzymes that have substrates bound (Fig. 4). Very little movement within anSMEcpe is required for substrate to bind in this catalytic position. Only two glutamines, which are well conserved in anSMEs (Fig. S2, blue), have a distinct orientation in the presence or absence of either AdoMet or peptide substrate. These residues aid both the stabilization of AdoMet and the positioning of D277 (SI Results and Fig. S5 D and E).

In contrast to radical generation, substrate deprotonation is only required by a small number of AdoMet radical enzymes. Before this work, it was not clear in anSME whether auxiliary cluster(s), enzyme residue(s), or both are involved in this reaction step. Here, the structures of anSME with peptides bound show that the substrate does not directly ligate either auxiliary cluster as previously suggested (10), and that both auxiliary clusters are far from the target Cys (8.6 Å and 20.8 Å), making it highly unlikely these clusters are involved in this deprotonation and/or inner-sphere electron transfer. Instead, the structure reveals two residues with titratable side chains that are close to the target Cys (Y24 from the AdoMet radical domain and D277 from the SPASM domain), and mutagenesis studies are consistent with D277 as the catalytic base. It is interesting that the SPASM domain, and not the AdoMet radical domain, contributes this key catalytic residue that differentiates anSME’s chemistry from that of other AdoMet radical enzymes, as this finding is consistent with previous structural studies that also showed the importance of residues outside of the partial barrel radical fold to the diversification of AdoMet radical chemistry (28, 29).

The next step, substrate oxidation, is again only required by certain subfamilies of AdoMet radical enzymes such as the dehydrogenases studied here, and others, like the heme biosynthetic enzyme HemN (30). While the Aux I and AdoMet clusters are nearly equidistant and within acceptable electron transfer distances (8.6 and 8.9 Å) from the substrate Cys (Fig. 3), we propose that Aux I is the immediate electron acceptor for this oxidation (Fig. 3, step 4). With few exceptions (36, 37), reduction of the AdoMet cluster during catalysis has only been proposed in systems that, unlike anSMEcpe, use AdoMet catalytically. Assuming that Aux I is the electron acceptor, for a subsequent turnover, it would need to be reoxidized. Here we further propose that Aux II, 12.9 Å away, performs this function. Other options for the reoxidation of Aux I are more problematic: the AdoMet cluster is too far from Aux I for direct electron transfer (16.9 Å), and the closest protein surface to which an external electron acceptor could bind appears blocked by bound substrate (Fig. S7). Thus, Aux II is the most viable candidate.

We can further consider if electrons are recycled in this reaction—that is, an electron used to homolytically cleave AdoMet in one cycle is derived from a previous cycle’s substrate oxidation. By monitoring the level of flavodoxin semiquinone depletion during anSMEcpe catalysis, Grove et al. have recently established that an electron can indeed be recycled in this fashion (26). While Aux II is also too far from the AdoMet cluster for direct electron transfer, an external electron acceptor, like flavodoxin, could accept an electron from Aux II and redeposit it into the AdoMet cluster. Electrochemical characterization of all three clusters would provide validation that (i) Aux I is the substrate radical electron acceptor during catalysis and (ii) electron transfer between Aux I and Aux II is possible. In the meantime, the structures described here reveal distances that support a role for these clusters in substrate oxidation by electron transfer, and refute other possible functions, including a role in substrate binding and deprotonation, as discussed above.

In addition to the mechanistic insight provided by these structures, visualization of the C-terminal domain of anSMEcpe clarifies the function of the recently described SPASM domain. Accession TIGR04085 designates 281 sequences as SPASM subfamily members. However, searching these 281 sequences within the Structure Function Linkage Database (SFLD) of the Radical SAM superfamily shows that these sequences are found in 153 different nodes, which contain 1,392 unique sequences (Fig. S8) (38). While this is the first structure of a SPASM domain-containing enzyme, we find that the first part of the domain, containing two of the seven cysteines of the 7-Cys motif, has been visualized before in the structure of the AdoMet radical protein MoaA. In particular, anSMEcpe shares with MoaA a conserved beta hairpin that extends the AdoMet radical beta sheet and contains cysteines on either end of the turn (Fig. 2). In both cases, these two cysteines ligate an auxiliary [4Fe-4S] cluster along with a third upstream cysteine (C255 in anSMEcpe) (Fig. S3). This beta hairpin also has high sequence homology to non-SPASM AdoMet radical dehydrogenase BtrN. Like MoaA, BtrN contains an auxiliary [4Fe-4S] cluster and, while BtrN has not been structurally characterized, the MoaA auxiliary cluster superimposes very well with Aux I in anSMEcpe (Fig. S3). Thus, this beta hairpin motif that is flanked by cysteines appears to be associated with binding an auxiliary [4Fe-4S] cluster in more than just SPASM-domain-containing proteins. Since this motif is a subdomain of SPASM, we will refer to it as a twitch subdomain.

Following the twitch subdomain, MoaA completes two helices and terminates, while anSMEcpe continues the SPASM domain with one helix leading to the next set of four cysteines arranged in a CX_CX_CX_C motif. Interestingly, while all four cysteines of this motif bind to an auxiliary cluster, they do not ligate the same cluster. The first three cysteines ligate Aux II, while the final cysteine ligates Aux I. Therefore, Aux I is coordinated by one upstream cysteine, two twitch subdomain cysteines, and one cysteine from the SPASM domain’s CX_CX_CX_C motif. MoaA shares this cluster coordination except for the last Cys; this Fe site is available for substrate binding (Fig. S3 C and F). In anSME, the bridging sequence between the final two cysteines of CX_CX_CX_C is Lys–Gly–Gly. This “XGG” sequence is highly conserved in TIGR04085 and could function to ensure a viable electron transfer pathway environment between the two auxiliary clusters. The final Cys of MoaA must be the internal “nick” or “head” cluster.

Cysteine ligation of the auxiliary clusters was accurately predicted by TIGR04085, which establishes seven cysteine ligands for the two auxiliary clusters. However, an upstream cysteine, C255 in anSMEcpe, is also involved in cluster binding, resulting in the unexpected full protein ligation of both clusters. This result refutes the idea that all SPASM domains have an available ligation site for substrate binding and indicates that at least anSMEs do not use an auxiliary cluster for this purpose. An analysis of TIGR04085 reveals that 37% of SPASM domain proteins have a cysteine that is both upstream of the SPASM domain and downstream of the AdoMet radical domain, indicating that full ligation of the auxiliary cluster may be common among members of the SPASM subfamily. Without the role of substrate ligation, these SPASM family members may use these clusters to facilitate electron flow in or out of the active site during turnover, insinuating that their mechanisms involve some kind of redox chemistry.

In the case of the SPASM family members lacking an upstream cysteine, the anSMEcpe structure indicates how substrates might coordinate Aux I. In the absence of a cysteine equivalent to C255, an available iron coordination site would be exposed to this substrate binding region (Fig. S3C). Importantly, due to the distance between the ligation site and the AdoMet binding site, the auxiliary cluster binding and hydrogen abstraction locations must be distal (in anSMEcpe, C255 is 12 Å from the substrate hydrogen abstraction site). In the case of MoaA, where substrate does directly ligate cluster, the ligation site (the N1 position of
the guanine base) is on the opposite end of the GTP substrate than the hydrogen abstraction site (the 3' hydrogen atom of the ribose; Fig. S3C, gray) (24, 25, 39). A similar mode of binding would be required for any SPASM members that use cluster ligation for substrate binding. As substrates in the SPASM family are predicted to be peptides, this would entail cluster ligation ~3 residues N-terminal of the ribose hydrogen abstraction site.

In summary, the anSMEcpe system provides a great example of the modularity of the AdoMet radical superfamily. The anSMEcpe AdoMet domain has very similar structural folds and cofactor binding motifs to the rest of the superfamily. However, the end reaction catalyzed is unique from all other structurally characterized AdoMet radical dehydrogenases, BtRn, only contains one additional [4Fe-4S] cluster. Upon structural characterization, it will be interesting to compare this enzyme to anSMEcpe and the full SPASM domain architecture.

**Materials and Methods**

His6-anSMEcpe, purified as described (26), and native anSMEcpe, constructed and purified as described in the SI Materials and Methods, were crystallized anaerothetically using the vapor-diffusion technique (SI Materials and Methods). The structure of anSMEcpe was solved using Fe anomalous data collected from a rotating copper anode source. High-resolution data were collected at beamlines 24-ID-E and 24-ID-C at the Advanced Photon Source and on X-29 at the National Synchrotron Light Source and solved by isomorphous replacement. Further data collection and model building details, as well as protocols used for creating ZYAF anSMEcpe and D277N anSMEcpe, can be found in the SI Materials and Methods. Activity assays were carried out as previously described (26), with modifications detailed in the SI Materials and Methods.

**ACKNOWLEDGMENTS.** For helpful discussions, we thank Daniel Dowling and Marco Jost. We also thank Allison Provost (Harvard University) for assistance with sequence analysis and Dennis Dean (Virginia Polytechnic Institute) for the gift of pDB1282. Data for this study were measured at Beamline X29 at the National Synchrotron Light Source. This work was supported by National Institutes of Health (NIH) Grant GM-63847 (to S.J.B.), NIH Grant GM-103268 (to S.J.B.), and National Science Foundation Grant MCB-0543833 (to C.L.D.). This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by Award RR-15301 from the National Center for Research Resources at NIH. Use of the Advanced Photon Source, an Office of Science User Facility operated for the US Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the US DOE under Contract DE-AC02-06CH11357. Financial support comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US DOE, the National Center for Research Resources (P41RR012408), and the National Institute of General Medical Sciences (P41GM103473) of NIH. C.L.D. is a Howard Hughes Medical Institution Investigator.