Structural studies of S-adenosyl-L-methionine radical enzymes involved in tRNA and natural product biosynthesis

By

Tsehai A.J. Grell
B.S., Chemistry, 2012
Morgan State University

Submitted to the Department of Chemistry
in Partial Fulfillment of the Requirements for Degree of

Doctor of Philosophy
in Biological Chemistry

at the

Massachusetts Institute of Technology

June 2018
© 2018 Massachusetts Institute of Technology.
All rights reserved.

Signature redacted

Signature redacted

Department of Chemistry
March 13th 2018

Signature redacted

Certified by

Catherine L. Drennan
Professor of Biology and Chemistry
Howard Hughes Medical Institute Professor and Investigator
Thesis Supervisor

Accepted by

Robert W. Field
Haslam and Dewey Professor of Chemistry
Chairman, Departmental Committee on Graduate Students
This doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:

Signature redacted

Professor Bradley L. Pentelute

Committee Chairman
Associate Professor of Chemistry

Signature redacted

Professor Catherine L. Drennan

Research Supervisor
Professor of Biology and Chemistry
Howard Hughes Medical Institute Professor and Investigator

Signature redacted

Professor Collin M. Stultz

Committee Member
Professor of Electrical Engineering and Computer Science,
Institute for Medical Engineering and Science
Structural studies of S-adenosyl-L-methionine radical enzymes involved in tRNA and natural product biosynthesis

By

Tsehai Ariane Julien Grell

Submitted to the Department of Chemistry on March 13th, 2018 in Partial Fulfillment of the Requirements for Degree of Doctor of Philosophy in Biological Chemistry

Abstract

Members of the S-adenosyl-L-methionine (AdoMet) radical enzyme superfamily catalyze a myriad of diverse and challenging biotransformations using a [4Fe-4S] cluster and a molecule of AdoMet to initiate radical. In this thesis, we used a combination of crystallographic and biochemical methods to identify the use of covalent catalysis and polar reactions in two AdoMet radical enzymes that catalyze the key steps in the biosynthesis of the tRNA modified bases wybutosine and queuosine. TYWI catalyzes the formation of the characteristic imidazopurine ring of wybutosine through a disputed mechanism. Here, we have garnered support for one of the proposed mechanisms, through the identification and characterization of a Schiff base between a catalytically essential lysine residue and the substrate pyruvate. The ability of TYWI to form and possibly use a Schiff base presents the first instance of a covalent catalysis in the mechanism of an AdoMet radical enzyme. In an attempt to obtain a snapshot of the active site of the queuosine biosynthetic enzyme, QueE, with AdoMet and a substrate analog, 6-carboxypterin (6-CP), we uncovered a covalent adduct between AdoMet and 6-CP. Further investigation of the mechanism by which this adduct was formed revealed a polar mechanism instead of a radical one. This result highlights the ability for AdoMet radical enzymes to use the same active site for two different reactions, polar and/or radical reactions.

The unifying characteristics of this superfamily include the canonical CX_3CXφC cluster-binding motif and a partial (β/α)_6 triose isomerase phosphate (TIM) barrel. Work in this thesis presents the structural characterization of a third QueE ortholog from Escherichia coli. Together, these three QueE orthologs revealed different variations in the core barrel architecture, which may influence binding of the biological reductant Flavodoxin. This variance in the core AdoMet radical fold emphasizes the structural diversity of this superfamily. On the other hand, we see conservation of an overall three-domain architecture for the maturation of ribosomally synthesized and post-translationally modified natural products, underlining the importance of this architecture for catalysis.

Thesis supervisor: Catherine L. Drennan

Title: Professor of Biology and Chemistry
Howard Hughes Medical Investigator and Professor
Acknowledgments

*It takes a village — African Proverb*

All of this would not be possible without the love and support of a number of persons. I would like to thank my advisor Catherine Drennan for always believing in me, supporting me, accepting me and encouraging me, for going above and beyond her role as mentor and PI and truly treating me like family. To the Drennan lab past and present, I would like to thank you for creating a wonderful environment to grow both scientifically and a personally. You truly are like family to me. I would like to thank my summer research advisor Collin Stultz, for taking a chance on me seven years ago and continuing to believe in me now. I would like to thank my collaborator Vahe and his amazing lab for being a source of wisdom and stimulating scientific conversations. I would like to thank Bradley Pentelute for wonderful conversations and directions during our annual meetings.

I would like to thank the community I have found at MIT for being a steady presence and the calm during the storm that is graduate school. Thank you. To my roommates and friends at MIT, greater Boston, from MSU and from Dominica, you have no idea what your friendship has meant to me over the years. Thank you for the love and support!

Last but definitely not least, to my family. To the Julien clan, thank you for always being there to support, to listen, to encourage and to celebrate. The little things you all do every day mean so much to me — I love you all. To my siblings, love you all so much! Thank you for your constant and never wavering love, support, encouragement, and for the much needed distractions. To my parents, thank you so much for everything and I mean EVERYTHING!! You have been my rock through this process. Thank your for teaching me from a young age that everyone fails but it is what you do after you have failed is what that counts. This one is for you!
For the Grells
# Table of Contents

Abstract 3  
Acknowledgments 4  

**Chapter I** Introduction to the S-adenosyl-L-methionine radical enzyme superfamily 11  
Chapter Summary 11  
Contributions 11  
The AdoMet Radical Core Fold 12  
Auxiliary Cluster ligation in Non-SPASM/non-Twitch AdoMet Radical Enzymes. 14  

*Roles of Auxiliary Clusters in Non-SPASM/non-Twitch AdoMet Radical Enzymes* 16  
The SPASM/Twitch Subfamily 17  

*Function of SPASM/Twitch subfamily* 18  
SPASM/Twitch domain architecture 21  

*Roles of Auxiliary Cluster in the SPASM/Twitch subclass* 23  
Conclusion 24  
Figures 25  
References 34  

**CHAPTER II** Biochemical and Structural Characterization of a Schiff base in the Radical-Mediated biosynthesis of 4-demethylwyosine 41  
Chapter Summary 41  
Contributors: 41  
Introduction 42  
Results 44  

*Overall architecture of holo-TYW1* 44  
*Visualization of a Schiff base intermediate in crystallo.* 45  
*SAM binding motifs appear to be conserved.* 46  
*A positive electrostatic surface near active site suggests location of tRNA binding site* 47  
*In vitro activity of C195S and K41A TYW1 mutants* 47  
*In vitro trapping of a Schiff base intermediate using C195S-TYW1* 48  
Discussion 50  
Conclusion 52  
Experimental procedures 53  

*Expression of tobacco etch virus (TEV) protease* 53  
*Purification of TEV protease* 53
Cloning and expression of wild type TYW1 (wtTYW1) 54
Purification of wtTYW1 54
Expression of His-tagged TYW1 55
Purification of TYW1 55
Preparation of SAM 55
Preparation of tRNA 56
Crystallization of TYW1 56
Cloning and purification of C195S and K41A TYW1 mutants 58
Trapping of Schiff base in TYW1 59
Figures 61
References 76

Chapter III 7-Carboxy-7-deazaguanine synthase – A radical S-adenosyl-L-methionine enzyme with polar tendencies 81
Chapter Summary 81
Contributions 81
Introduction 82
RESULTS 84
X-ray structure of BsQueE with a 6-CP–dAdo ester covalent adduct 84
6-CP is an alternative substrate for QueE 86
QueE catalyzes two different 6-CP products depending on the presence of reductant 87
Mg²⁺ dependence of alternative activities 89
B. multivorans QueE catalyzes the formation of both 6-dAP and 6-CP–dAdo ester 90
Discussion 90
Conclusions 93
Experimental Procedures 93
Cloning, expression, and protein purification 93
Crystallization of BsQueE 94
Crystallization of BsQueE for initial structure determination 94
Crystallization of BsQueE with AdoMet and 6-CP for determination of the adduct structure 94
Determination of the initial structure of BsQueE 95
Synthesis of substrates 96
QueE incubation with 6-CP 98
Determination of the rate of formation of 6-deoxyadenosylnpterin (6-dAP) 98
Determining the rate of formation of 6-CP-dAdo ester 99
BsQueE reaction with [U-13C7]-6-CP and [13C10]-AdoMet 99
MS/MS fragmentation 99
Mg2+ dependence for BsQueE reaction with 6-CP and AdoMet 100
ICP-MS analysis for Mg2+ content 100
Figures 101
References 114

Chapter IV Structural comparison of bacterial QueE orthologs 117
Chapter Summary 117
Contributions 117
Introduction 118
Results 120
EcQueE reveals an intermediary structure between BmQueE and BsQueE. 120
AdoMet binding motifs are conserved in the QueE homologs. 122
The active site of QueE comprises of residues from the N- and C-terminal extensions. 123
QueE homologs display differential electrostatic surfaces. 125
Modeling Flavodoxin and QueE interactions. 126
Discussion 127
Conclusion 130
Experimental procedures 130
Purification of EcQueE 130
Crystallization and Data collection of EcQueE 130
Structure determination and refinement 131
Manual docking of AdoMet and substrates 132
Homology modeling of flavodoxins from Burkholderia multivorans and Bacillus subtilis 132
Docking with FRODOCK 132
Preparation of Figures and electrostatic surfaces 133
Figures 134
References 152

Chapter V Structural characterization of an AdoMet Radical enzyme involved in sactionine bond formation 155
Chapter Summary 155
Contributions 155
Introduction

Results

Overall architecture of SkfB is modular.  
SkfB adopts a canonical AdoMet radical domain.  
The Twitch domain binds a [2Fe-2S] auxiliary cluster.  
The N-terminal peptide-binding domain contains a RiPP recognition element motif.  
Structural comparison of AdoMet radical RIPP biosynthetic enzymes.

Discussion

Conclusion

Experimental Methods

Purification of SkfB

Crystallization and data collection of SkfB

Structure Determination and refinement

Figures

References

Chapter VI S-adenosyl-L-methionine radical enzymes in perspective

Chemistry of the AdoMet radical enzyme superfamily.

SkfB is a hybrid AdoMet radical enzyme

Prediction of how peptide-binding domains may bind to AdoMet radical enzymes in two-component systems.

Figures

References
Chapter I
Introduction to the S-adenosyl-L-methionine radical enzyme superfamily

Chapter Summary

S-adenosylmethionine (AdoMet) radical enzymes use AdoMet and a [4Fe-4S] cluster to catalyze a diverse array of reactions. They adopt a partial TIM barrel fold with N- and C-terminal extensions that tailor the structure of the enzyme to its specific function. One extension, termed a SPASM domain, binds two auxiliary [4Fe-4S] clusters and is present within peptide-modifying enzymes. The first structure of a SPASM-containing enzyme, anaerobic sulfatase maturing enzyme, revealed unexpected similarities to two non-SPASM proteins, butirosin biosynthetic enzyme BtrN and molybdenum cofactor biosynthetic enzyme MoaA. The latter two enzymes bind one auxiliary cluster and exhibit a partial SPASM motif, coined a Twitch domain.

Contributions

This chapter was adapted from a MiniReview series article titled “SPASM and Twitch Domains in S-Adensoymlmethionine (SAM) Radical Enzymes” *Journal of Biological Chemistry*. 2015; 290 (7), 3964-3971
Members of the S-adenosyl-L-methionine (AdoMet) radical superfamily catalyze a wide variety of radical-mediated reactions, including complex chemical transformations and rearrangements; modifications of peptides; DNA, and RNA; dehydrogenations; and sulfur insertions\(^1\). Despite this diversity, there are unifying structural and mechanistic themes. For instance, AdoMet radical enzymes typically bind a \([4\text{Fe}-4\text{S}]\) cluster using a conserved \(\text{CX}_3\text{CX}\phi\text{C}\) motif (where \(\phi\) is an aromatic residue). This motif provides three cysteine ligands to the iron atoms of the cluster, with the fourth ligand coming from the bidentate coordination of AdoMet to the unique iron\(^2\-3\). Direct ligation of AdoMet to the cluster facilitates reductive cleavage of the C-S bond through an inner sphere electron transfer event, forming methionine and a 5'-deoxyadenosyl radical, 5'-dAdo* (Figure 1.1A)\(^4\). The abstraction of an H-atom from the substrate by 5'-dAdo*, producing a substrate radical, ends the mechanistic similarity between enzymes of this superfamily; each enzyme utilizes a different mechanism to generate product. Structures of the first seven members of the AdoMet radical superfamily were used to define a core fold for binding AdoMet and for the generation of 5'-dAdo* species (Figure 1.1). This core consists of a partial \((\beta/\alpha)_6\) triose-phosphate isomerase (TIM) barrel\(^5\) (Figure 1.2). Outside of the core fold, the structure can vary greatly, with N- and C-terminal extensions that are functionalized for binding substrates or additional cofactors. For example, the queuosine biosynthetic enzyme 7-carboxy-7-deazaguanine synthase (QueE) and pyruvate-formate lyase activating enzyme (PFL-AE) utilize residues from modest N- and C-terminal extensions to bind their respective substrates, 6-carboxy-5,6,7,8-tetrahydropterin and the glycyl loop of pyruvate-formate lyase. In addition, the N- and C-terminal of QueE forms the oligomerization surface for the physiological dimer\(^6\-7\). Several AdoMet radical enzymes have adopted N- and/or C-terminal extensions to bind different flavors of additional iron-sulfur cluster (Figure 1.4). The SPASM/Twitch and the cobalamin dependent subfamilies have adopted defined domains for the binding of one or two auxiliary clusters or a cobalamin molecule respectively\(^8\-10\) (Figure 1.4, I.6). Here we will discuss the structural variations of observed architecture used to bind additional clusters by AdoMet radical enzymes.

**The AdoMet Radical Core Fold**

AdoMet radical enzymes adopt a core fold, with exceptions described in a review by Dowling et. al\(^11\) and below. This core is responsible for AdoMet binding and radical generation, and has been defined as a partial TIM barrel \((\beta/\alpha)_6\) with six \(\alpha\)-helices making up the outside of the partial
barrel and the six parallel β-strands forming the inner face of the barrel (Figure 1.2). The active site is located within the lateral opening of the partial TIM barrel, and includes a [4Fe-4S] cluster, AdoMet and substrate binding sites. Although some AdoMet radical enzymes are composed of a complete TIM barrel (β/α)_{12-13}, most have the partial barrel architecture. Recently, the structural diversity of this enzyme superfamily was expanded when a significant deviation was reported for the queuosine biosynthetic enzyme 7-carboxy-7-deazaguanine synthase (QueE). QueE was found to have a hypermodified barrel fold in which three of the six α-helices are replaced by two loops and one 3_{10} helix.

The [4Fe-4S] cluster motif, more popularly known as the CX_{3}CXφC motif, binds the [4Fe-4S] cluster responsible for initiating radical chemistry, herein called the AdoMet radical cluster. The three cysteines of this motif reside on the loop linking β1 to α1, termed the cluster-binding loop (Figure 1.2). This motif, which is largely but not absolutely conserved, allows for binding of AdoMet to the cluster through the open coordination site on the unique iron atom of the cluster, and through hydrophobic interactions from the aromatic residue (φ). With respect to cluster binding, the largest deviations observed through crystallographic studies thus far are found in the AdoMet radical enzymes QueE, mentioned above, and ThiC, which is involved in the biosynthesis of thiamin pyrimidine moiety. In the former case, the cysteine motif has an eleven amino acid insertion between the first two cysteines (CX_{14}CXφC) and in the latter, the cluster-binding region is not part of the AdoMet radical core, but is instead found in a separate domain. For QueE, the eleven-residue insert does not alter the way in which AdoMet binds, and is instead believed to be important for binding of the physiological reductase. For ThiC, no structure is available of the fully reconstituted protein and thus the exact mode of cluster and AdoMet binding remains to be determined.

AdoMet binding is facilitated by four different motifs, some of which are more highly conserved than others. The “GGE motif”, named after sequences in pyruvate-formate lyase activating enzyme and MoaA, resides at the C-terminal end of β2 and interacts with the amino group of AdoMet, helping to orient the methionyl moiety of the AdoMet. Interactions between residues from β4, called the “ribose motif”, and the hydroxyls of the ribose moiety appear to play an important role in positioning of 5′-dAdo• with respect to substrate for H-atom abstraction.
The “GXIXGXXE motif”, which is only partially conserved in terms of sequence, is located on β5 and provides hydrophobic interactions with the adenine ring of AdoMet. Finally, the “β6 motif” is responsible for hydrogen bonding via backbone atoms to the adenine ring. Interestingly, the structures of butirosin biosynthetic enzyme BtrN from *Bacillus circulans* shows that residues at the end of β5 can replace the function of the “β6 motif.” BtrN thus contains both the “GXIXGXXE” motif and the “β6 motif” on β5. Taken together, all of these motifs, which span across the AdoMet radical core, work together to ensure that AdoMet is positioned in the correct orientation for H-atom abstraction from the substrate after homolytic cleavage.

**Auxiliary Cluster ligation in Non-SPASM/non-Twitch AdoMet Radical Enzymes.**

The AdoMet radical core acts as a building block, and many members have functionalized N- and C-terminal extensions appended to this core. A number of AdoMet radical enzymes utilize N- and C-terminal extensions to bind additional cofactors. In fact, the two largest subclasses of AdoMet radical enzymes are made up of members that bind cobalamin (*Figure I.4A*) and the SPASM/Twitch domain-containing enzymes that bind additional clusters. Outside of these two subclasses, other members of the superfamily bind additional cofactors, in particular additional clusters using less well-defined domain architecture. Here we will discuss enzymes that fall into this loose subset of non-SPASM/non-Twitch auxiliary cluster-containing enzymes: BioB, LipA, HydE, HydG, RimO, MiaB and TYW1.

The final step in the biosynthesis of the vitamin biotin is catalyzed by the AdoMet radical enzyme biotin synthase (BioB)\(^{16-17}\). BioB inserts a sulfur atom between the unactivated carbons, C\(_6\) and C\(_9\), of the substrate dethiobiotin in a radical mediated mechanism (*Figure I.3A*). A [2Fe-2S] auxiliary cluster was implicated in the mechanism of BioB\(^{12,18-19}\). The architecture of BioB revealed a full TIM barrel fold composed of the partial (β/α)\(_6\) TIM barrel of the AdoMet radical domain and two C-terminal β-strand/α-helix pairs. The auxiliary [2Fe-2S] cluster was found in the middle of the TIM barrel, ligated by three cysteines, C\(_{97}\), C\(_{128}\), C\(_{188}\), and, at the time, an unprecedented ligand R\(_{260}\), donated from the β-sheet core of the barrel\(^{12}\) (*Figure I.4B*). An N-terminal extension, comprised of two consecutive α-helices, along with an α-helix of the C-terminal extension contribute to the dimer interface of the BioB homodimer.
Similar to BioB, lipoyl synthase (LipA) catalyzes insertion of a sulfur atom at two saturated carbon sites of a protein-bound \( n \)-octanyl chain during the biosynthesis of lipoyl cofactor \(^{20}\) (Figure I.3B). The catalytic mechanism of LipA requires two molecules of AdoMet, one per sulfur insertion, and two [4Fe-4S] clusters\(^{21}\). The structure of LipA revealed that residues from the N- and C-terminal extensions to the AdoMet radical core made up the auxiliary cluster-binding site\(^{22-23}\). The auxiliary cluster of LipA was coordinated by residues from the N- and C-terminal extensions; three cysteines from a conserved N-terminal CX\(_4\)CX\(_5\)C motif, and a serine residue, S\(_{292}\), form the C-terminal extension. (Figure I.4C)

HydE and HydG are two AdoMet radical enzymes involved in the maturation of [Fe-Fe] hydrogenases (HydA) and, along with the GTPase enzyme HydF, are essential for bioassembly of the H-cluster\(^{24-25}\). The H-cluster is an unusual modified [4Fe-4S] cluster bridged to a di-iron subcluster through cysteine thiolates and the two-iron subcluster contains three carbon monoxide (CO) and two cyanides (CN\(^-\)) ligands and a dithiomethylamine (DMTA) ligand\(^{26-28}\) (Figure I.3C). HydG has been implicated in the production of the diatomic ligands, CO and CN\(^-\), from a tyrosine-derived dehydroglycine (DHG)\(^{29}\) (Figure I.3C). Structural analysis of HydG revealed that, like BioB, the protein core adopts a full TIM (\( \beta/\alpha \))\(_8\) barrel fold where two \( \beta \)-strands from the mostly \( \alpha \)-helical C-terminal domain extends the AdoMet radical domain. The \( \alpha \)-helical region of the C-terminal extension flanks the TIM barrel and ligates an auxiliary [4Fe-4S] sulfur cluster. Conversely, the N-terminal domain is also mainly \( \alpha \)-helical\(^{30-31}\) (Figure I.4E). The role of HdyE in H-cluster maturation has been more elusive than that of HydG. By process of elimination, it was believed that HydE was responsible for synthesizing the DMTA bridging ligand (Figure I.3.C)\(^{32}\). HydE folds into an AdoMet radical domain flanked by N- and C-terminal domains. The structure revealed the binding site of the auxiliary cluster is coordinated by three residues of the C-terminal extension, but further investigation of the cluster identity awaits as the crystal structure visualizes a [2Fe-2S] auxiliary which is contrary the [4Fe-S] auxiliary cluster indicated by spectroscopy\(^{33}\) (Figure I.4E).

Methylthioltransferases (MTTases), as their name suggests, transfer a methylthiol group to macromolecular substrates including the S12 ribosomal protein, catalyzed by RimO, and tRNA-
A37, catalyzed by MiaB^{34-36} (Figure I.3D,E). MiaB and RimO both contain an auxiliary cluster with an open coordination site, though only RimO has been structurally characterized. The overall architecture of RimO is modular with an N-terminal UPF0004 domain followed by an AdoMet radical core domain and a C-terminal TRAM domain. UPF0004 domain provides the three cysteines for the ligation of a [4Fe-4S] auxiliary cluster, which sits at the interface between the UPF0004 domain and the AdoMet radical domain^{37-38} (Figure I.4F).

Wye bases, such as the hypermodified base wybutosine, are a class of modified tRNA bases found adjacent to the wobble position of tRNA^{Phe} in archaea and eukaryotes^{39-40}. Wye bases contain a characteristic imidazopurine (tricyclic) ring generated by the AdoMet radical enzyme TYWI on the tRNA wobble position guanosine base using a pyruvate molecule^{41-42} (Figure I.3F). Two apo-structures of TYWI were solved and revealed a central AdoMet core, flanked by N- and C-terminal extension. Although no clusters were visualized in these structures, the three conserved N-terminal cysteines, CX_{12}CX_{12}C motif, were found in close proximity to each other and appeared poised to coordinate an auxiliary cluster^{43-44}. Based on spectroscopic analysis, this auxiliary cluster is expected to be a [4Fe-4S] cluster^{45}.

Roles of Auxiliary Clusters in Non-SPASM/non-Twitch AdoMet Radical Enzymes

A number of auxiliary cluster-containing non-SPASM/non-Twitch AdoMet radical enzymes have been identified and the roles of their auxiliary clusters are under investigation. BioB, for example, has been proposed to use its auxiliary [2Fe-2S] in a sacrificial manner, donating one of its sulfur atoms to perform a sulfur insertion reaction during biotin formation^{12, 19, 46-47}. Further evidence for cannibalism of the auxiliary cluster during sulfur insertion came from the structure of LipA. Similar to BioB, LipA contains an auxiliary [4Fe-4S] cluster that was proposed to be the source of the two sulfur atoms needed to form lipoic acid^{20, 48-49}. Through clever biochemical and structural work, a snapshot of the LipA bound to an intermediate revealed the octanoyl substrate ligated to a sulfur atom of the auxiliary cluster. The atypical auxiliary cluster ligands observed in BioB and LipA may not be unusual at all but instead help facilitate the unconventional role of these auxiliary clusters. MTTases also coordinate auxiliary clusters^{35-36}, which, unlike for BioB and LipA, are not currently believed to serve as the source of sulfur for generation of the methylthiol group^{50}. Instead, the crystal structure of RimO has led to the
proposal that its auxiliary cluster is responsible for binding a polysulfide moiety that serves as the sulfur source.\(^{38}\)

The mechanism by which TYW1 performs tricyclic ring formation is currently disputed, but the auxiliary cluster is proposed to ligate the pyruvate substrate through the unique iron, based on spectroscopic evidence.\(^{42,45,51}\) Whether pyruvate is activated through ligation of the cluster or through a covalent adduct with the catalytically essential lysine awaits further investigation. The hydrogenase cofactor maturing enzymes, HydE and HydG, both have site differentiated auxiliary cluster. No role for the auxiliary cluster of HydE has been assigned and it is debated whether the auxiliary cluster is actually necessary for function, as the cysteine ligands are not conserved amongst HydEs from different species.\(^{33}\) The auxiliary cluster of HydG, on the other hand, is crucial to the decomposition of DHG to the diatomic products, CO and CN\(^{-}\) and is proposed to be the site of [Fe(CO)\(_2\)(CN)]Cys synthon formation.\(^{29}\) Other auxiliary cluster-containing AdoMet radical enzymes include a subset of glycyl radical activating enzymes. Our understanding of this latter set of enzymes is still in its infancy and there is still a lot more to discover.

The SPASM/Twitch Subfamily

The SPASM subfamily is an example of a functionalized C-terminal extension for the binding of two auxiliary clusters. Haft and Basu recognized that enzymes with this C-terminal extension appear to be involved in the modification of ribosomally translated peptides.\(^{53-54}\) This subclass is referred to as SPASM after the biochemically characterized members, AlbA, PqqE, anSMEs, and MftC, which are involved in subtilosin A, pyrroloquinoline quinone, anaerobic sulfatase and mycofactocin maturation, respectively. The SPASM subfamily, accession TIGR04085, is composed of 281 sequences. However, recent similarity network analysis by Holliday and Babbitt using the Structure Function Linkage Database identified additionally related sequences, expanding the number to over 18,000 (Figure I.1B).

Enzymes in this subfamily were expected to use a seven cysteine motif to bind two auxiliary [4Fe-4S] clusters, leaving a unique iron for substrate binding.\(^{54,57}\) However, initial structures of a SPASM protein, the anaerobic sulfatase maturing enzyme (anSME) from Clostridium
perfringens (C. perfringens), showed full cysteine ligation to both [4Fe-4S] clusters. anSME was also the first structure of an AdoMet radical enzyme with dehydrogenase activity, and it, along with the structures of butirosin biosynthetic enzyme BtrN from Bacillus circulans, and molybdenum cofactor biosynthetic enzyme MoaA \(^*\), revealed unexpected structural homology between SPASM and non-SPASM enzymes. In particular, the comparison of anSME, BtrN and MoaA led to the identification of a truncated SPASM domain used for binding a single auxiliary [4Fe-4S] cluster, which has been termed a Twitch domain. Recently, the structural analysis of a thioether bond forming SPASM enzyme, CteB from Clostridium thermocelum ATCC 27405 (C. thermocelum) produced the first evidence for the initial hypothesis that SPASM enzymes will contain a unique iron for substrate binding.

**Function of SPASM/Twitch subfamily**

A chief function of this subfamily appears to be in the post-translational modification of peptides. Post-translational modifications expand the chemical repertoire of enzymes by generating modified amino acids that are well suited to perform specific reactions, priming peptides for cofactor biosynthesis, or rigidifying the protein or peptide (Figure 1.5).

anSME catalyzes the co- or post-translational modification of an arylsulfatase active site cysteine or serine residue to a catalytically essential formylglycine (FGly) moiety in anaerobic organisms. The active site FGly allows sulfatases to perform their hydrolysis function, removing sulfate groups from a wide array of substrates (e.g. sulfated polysaccharides, sulfolipids, and steroid sulfates). Sulfatase activity is important in humans and bacteria with a lack of activity leading to disease in humans and an inability to colonize the mucosal layer of the host’s gut upon inhibition in bacteria. anSMEs are known to use AdoMet radical chemistry, in addition to two auxiliary [4Fe-4S] clusters, to perform the dehydrogenation of a serine or cysteine residue to the FGly moiety (Figure 1.5). With the structural and mechanistic information on anSME from C. perfringens, our understanding of the anaerobic sulfatase maturating enzyme family is on the rise.

AlbA is involved in the leader peptide-dependent post-translational modification of a linear ribosomally synthesized peptide into the cyclic peptide natural product, subtilosin A. In
particular, AlbA is responsible for forming three thioether bonds between the sulfur atoms of three cysteine residues and the α carbons of two phenylalanines and one threonine, yielding a rigidified peptide (Figure 1.5)\(^{69}\). Mutation of the corresponding gene, \textit{albA}, indicates that AlbA is essential for subtilosin A biosynthesis and shows that thioether bond formation is a critical aspect of subtilosin A maturation\(^ {70}\). Peptide-derived natural products that contain thioether bonds, such as subtilosin A, are collectively known as sactipeptides and belong to a larger group of natural products which are constituted of ribosomally synthesized and post-translationally modified peptides or RiPPs. In terms of the medical relevance of this class of compounds, subtilosin A, which is produced by the soil bacterium \textit{B. subtilis}, shows antimicrobial activity against both Gram-negative and Gram-positive bacteria, as well as some human pathogens\(^ {71}\). Another sactipeptide produced by \textit{B. subtilis}, sporulation killing factor, is excreted under nutrient limitation to lyse neighboring cells in what can only be described as cannibalistic behavior\(^ {72-74}\).

A different AdoMet radical enzyme, SkfB, is responsible for the formation of the thioether bond of this sactipeptide\(^ {75}\). Despite high sequence similarity to AlbA, the C-terminal of SkfB is predicted to fold into a Twitch domain coordinating one auxiliary iron-sulfur cluster.

Bioinformatics analysis identified in clostridial genomes the co-localization of genes encoding putative AdoMet radical enzymes with genes encoding peptides\(^ {54}\). These peptides called SCIFF peptides contained variable N-terminal regions but conserved C-terminal regions containing six cysteines in forty-five-residue. A SCIFF peptide/maturase pair, Tte1186/Tte1186a, from \textit{Caldanaerobacter subterraneus} subsp. \textit{Tengcongensis} MB4 was characterized. Biochemical and spectroscopic studies demonstrated that Tte1186 was indeed an AdoMet radical enzyme and, similarly to AlbA and SkfB, was able to install a thioether linkage between a sulfur atom of a cysteine residue and the α-carbon of a residue in the C-terminal region of Tte1186a\(^ {76}\). Tte1186 contains three iron-sulfur clusters; one AdoMet cluster necessary for reductive cleavage of AdoMet, and two auxiliary clusters necessary for thioether bond formation. Recently another SCIFF maturase from \textit{C. thermocellum}, CteB, was shown to be able to install thioether bonds on its respective substrate, CteA\(^ {59}\).

PqqE is one of several enzymes that participate in the maturation of the bacterial redox cofactor pyrroloquinoline quinone (PQQ) from a peptide precursor. In particular, PQQ is formed post-
translationally from a skeleton peptide PqqA through complex rearrangements initiated by the cross-linking of glutamyl and tyrosyl sidechain\(^{77-78}\) catalyzed by PqqE (Figure 1.5). PQQ is then excised from PqqA and attached to its target enzyme. PqqE is known to reductively cleave AdoMet and contain two [4Fe-4S] clusters and a third iron-sulfur cluster, which is either a [2Fe-2S] cluster of a [4Fe-4S] cluster\(^{79}\).

StrB is involved in the maturation of a microbial cyclic peptide pheromone, streptide. Streptide is involved in bacterial communication and contains a lysine-tryptophan carbon-carbon crosslink that is post-translationally installed on the precursor peptide, StrA, by StrB\(^{80}\). Although initial characterization of StrB identified one auxiliary [4Fe-4S] cluster, recent mutagenesis and biochemical studies by Benjdia et. al show that StrB most likely contains two auxiliary clusters\(^{81}\). Further evidence for two auxiliary clusters in StrB came from the structural characterization of a homolog SuiB, which revealed two fully ligated [4Fe-4S] auxiliary clusters in addition to the AdoMet radical cluster\(^{82-83}\).

BtrN, a 2-deoxy-scyllo-inosamine (DOIA) dehydrogenase, is a member of the Twitch structural subclass of AdoMet radical enzymes, and is involved in the biosynthesis of the aminoglycoside antibiotic butirosin B. BtrN is also a member of a recently described functional subclass of the AdoMet radical enzymes, AdoMet radical dehydrogenases, with the sulfatase maturases anSME and AtsB being the only other biochemically characterized members\(^{57,63,84}\). BtrN contains a Twitch domain\(^9\), which houses one auxiliary [4Fe-4S] cluster. Using AdoMet radical chemistry, BtrN catalyzes the oxidation of the C3 hydroxyl group of DOIA by a H-atom abstraction, deprotonation and one-electron oxidation to produce the ketone group in amino-DOI (Figure 1.5)\(^{85-86}\).

The Twitch subclass enzyme MoaA, on the other hand, is involved in the biosynthesis of molybdopterin (Moco). MoaA catalyzes the first step in Moco biosynthesis, the complex rearrangement of guanosine triphosphate to (8S)-3’,8-cyclo-7,8-dihydroguanosine 5’-triphosphate (3’-8-cH\(_2\)GTP), which is then converted to Moco in subsequent steps (Figure 1.5)\(^{87-88}\). In humans, defects in molybdenum cofactor biogenesis leads to death shortly after birth, and these
patients show neurological abnormalities including untreatable seizures and attenuated brain growth.

**SPASM/Twitch domain architecture**

In addition to the core fold, AdoMet radical enzymes contain N- or C-terminal extensions. The SPASM domain at the C-terminal end of the core fold binds two auxiliary [4Fe-4S] clusters using a conserved seven-cysteine motif, CX_{9,15}GX_4CX_2CX_3CX_3C−X_{n}C. Based on the reactions performed by the biochemically characterized SPASM members and the presence of only seven cysteine residues in the conserved cysteine motif, an open coordination site on one auxiliary [4Fe-4S] cluster has been proposed to ligate the peptide substrate. Through the work of Goldman et al., we now have a snapshot of one of these enzymes, anSME, both with and without substrate bound. This structure shows that both clusters of anSME are fully ligated by cysteine residues, with auxiliary cluster I (Aux I) 16.9 Å from the AdoMet radical cluster and 12.9 Å from the second auxiliary cluster (Aux II). Aux I is ligated by four cysteine residues: one before the predicted SPASM domain seven cysteine motif, two within the motif, and one downstream cysteine. A conserved β-hairpin motif lies between the second and third coordination sites of Aux I, with an alpha helix following (α2'). This helix leads to the second auxiliary cluster (Aux II). Aux II is coordinated by four cysteine residues of the seven-cysteine motif, but not in the same order as the primary sequence. Unexpectedly, the protein chain travels back and forth between the two auxiliary clusters such that cysteines 4, 5, 6, and 8 coordinate Aux II whereas cysteine 7 is the downstream cysteine ligand to Aux I. Based on sequence homology, it was predicted that the cluster-binding architecture of anSME would be a common feature of SPASM domain-containing enzymes, with the caveat that some SPASM proteins may not display full cysteine ligation.

The protein topology around Aux I of anSME shows structural similarity to the auxiliary cluster domain of the non-SPASM AdoMet radical enzyme MoaA. MoaA also displays a ~16 Å distance between its AdoMet radical cluster and its auxiliary cluster and adopts an abridged SPASM domain architecture. This abridged SPASM fold provides three cysteine ligands to bind one [4Fe-4S] cluster, and contains the β-hairpin motif and α2', all of which are found in the first half of the SPASM domain. Based on the similarity between anSME and
MoaA, this abridged SPASM domain was subsequently coined the Twitch domain. Given that MoaA contains no cysteine ligands following α2', an open coordination site on the [4Fe-4S] cluster is used for substrate binding as was originally proposed for the SPASM enzymes.

BtrN also contains a Twitch domain that binds one [4Fe-4S] cluster similar to MoaA’s auxiliary cluster-binding domain. However, the auxiliary cluster in BtrN is fully ligated, showing the versatility of the Twitch domain. Like anSME and MoaA, the auxiliary cluster of BtrN is ~16 Å from the AdoMet radical cluster (Figure I.6C) and has a domain architecture made up of a β-hairpin followed by α2', with cysteine ligands flanking both ends of the β-hairpin. Similar to anSME, cysteine residues following α2' ligate an auxiliary cluster. Unlike anSME and MoaA, BtrN only has one cysteine ligand before the β-hairpin (Figure I.6A). Thus, both the Aux I of anSME and the Aux cluster of BtrN have full ligation by cysteine residues, although the cysteine positions in the primary sequence are different.

Overall, these structural snapshots show two different coordination modes of [4Fe-4S] clusters in Twitch domains, one in which the [4Fe-4S] cluster can bind substrate directly to an open coordination site (MoaA) and one in which it cannot (BtrN). They also reveal two different varieties of AdoMet radical dehydrogenases, one with one auxiliary [4Fe-4S] cluster (BtrN) and one with two auxiliary clusters (anSME). Together, these three structures defined the architecture of the SPASM/Twitch domain. Within the last year three new structures, SuiB, CteB and PqqE have been solved, revealing both similarities and differences to the three enzymes described above.

The structure of SuiB revealed two fully ligated [4Fe-4S] auxiliary clusters following the same cluster-binding architecture as anSME. The protein chain travels back and forth between the two clusters such that Aux I is ligated by cysteines 1,2,3, and 7 and the Aux II by cysteines 4,5,6, and 8 and the conserved hairpin β-hairpin intersperses cysteines 2 and 3. As in anSME, cysteine 1 is found before the predicted seven-cysteine cluster-binding motif but unlike anSME it is located in the linker region between the AdoMet and the SPASM domain. Aux I is found at a similar distance from the AdoMet radical cluster, 16 Å as found in anSME but the distance between Aux I and Aux II is shorter than in anSME, ~10 Å compared to 12.9 Å (Figure I.6D).
CteB also shows a similar binding architecture to anSME and SuiB but only the seven-cysteines of the conserved motif are involved in cluster-binding resulting in an open coordination site on Aux I\(^5\). In the vicinity of this unique iron, density was noted that was not accounted for by a protein ligand. The authors proposed that this density was due to the direct ligation of a cysteine residue from the N-terminal fragment of the peptide substrate CteA, which was co-crystallized with CteB (Figure I.6E). This structure presents the first look at a SPASM domain-containing enzyme with an open coordination site for substrate ligation and garners merit for the hypothesis that some member of SPASM subclass will use their auxiliary clusters, like MoaA, to bind substrate.

The structure of another founding member of the SPASM enzyme subclass, PqqE, was solved recently and revealed a number of unexpected differences\(^9\). Although PqqE fully ligates two auxiliary clusters, one of the Aux II ligands is a non-cysteine ligand, an aspartic acid from the motif CX\(_2\)CX\(_5\)DX\(_3\)C. This is in contrast to the other three characterized SPASM enzymes where this motif, CX\(_2\)CX\(_5\)CX\(_3\)C motif, has a cysteine ligand at the corresponding position. The cluster binding architecture of PqqE is similar to BtrN where cysteines 1 and 2 flank either side of the \(\beta\)-hairpin and the two final cysteines to Aux I follow \(\alpha2'\). Therefore the Aux I binding site is made up of cysteines 1,2,6 and 7 and Aux II binding site is composed of cysteine 3,4,8 and the aspartic acid ligand at position 5. The structure of PqqE also revealed a \([2Fe-2S]\) Aux I cluster bound by an unprecedented CXC motif (Figure I.6F).

**Roles of Auxiliary Cluster in the SPASM/Twitch subclass**

The structural data presented here has served to clarify the function(s) of auxiliary clusters in SPASM/Twitch enzymes\(^9,10,59,83,91\). For anSME and BtrN, crystal structures show that auxiliary clusters are fully ligated both in the presence and absence of substrates, and that substrates are bound through protein-mediated hydrogen bonding interactions at distances of 9-10 Å away from their auxiliary clusters. These long distances also suggest that auxiliary clusters do not function in substrate deprotonation. For anSME, both crystallographic and mutagenesis data support Asp277 as being responsible for the deprotonation step \(^10\). In particular, Asp277 is adjacent to the substrate and its mutation to Asn results in a protein with only 0.8% of wild-type activity and
increased uncoupling of AdoMet cleavage from product formation. For BtrN, Arg152 has been proposed to play a role in substrate deprotonation based solely on its location in the active site. Instead of binding or deprotonating substrate, the observed 9-10 Å distances between cluster and substrate are consistent with a role for the auxiliary clusters as electron acceptors during substrate oxidation. There is diversity in the number of auxiliary clusters needed for substrate oxidation. anSME has two auxiliary clusters (12.9 Å apart) to provide a route for the electron from the buried active site to the protein surface, whereas BtrN only needs one auxiliary cluster for the electron to reach the protein surface, where it would be accessible to an external electron acceptor.

As with anSME and BtrN, a protein residue, Glu319, is proposed to deprotonate the putative tryptophanyl radical intermediate in SuiB, and the auxiliary clusters instead play a role in electron transfer to the redox partners. The fully ligated auxiliary clusters PqqE may play a similar role as proposed for anSME, SuiB and BtrN. On the other hand, the site differentiated Aux I cluster of CteB appears to play a role in peptide substrate binding, as density attributed to a substrate cysteine is visualized near the cluster. It still remains to be seen if Aux I will additionally play a role in activating the donor cysteine of CteA for thioether bond formation.

**Conclusion**

After 17 years of investigation, we have not begun to scratch the surface of understanding the chemistry performed by the members of the AdoMet radical enzyme superfamily. We still have so much to learn about the mechanisms and types of reactions catalyzed by AdoMet radical enzymes, as well as the role of their additional cofactors.
Figure 1.1. The AdoMet radical superfamily sequence similarity network. (A) General mechanism of reductive cleavage of AdoMet by inner sphere electron transfer from the [4Fe-4S] cluster to the C-S bond, forming a 5'-deoxyadenosyl radical and methionine. (B) Protein sequence similarity network of the AdoMet radical enzyme superfamily with an alignment score of 16 is shown, created using the Enzyme Function Initiative (EFI) tool (http://efi.igb.illinois.edu/efi-est/). A node represents non-redundant protein with 40% or higher sequence identities, obtained from the defined AdoMet Radical Superfamily Pfam PF04055 and each edge represents sequences with values below the alignment score indicated above. Nodes shown in pink are annotated as Cbl-binding enzymes (IPR006158) whereas those in salmon are annotated as members of the SPASM/Twitch structural subclass (IPR034395). Other enzymes highlighted are colored based on their proposed functions; methylthiotransferases (RimO) in lime green, H-cluster biosynthetic enzymes (HydE and HydG) in green, glycy1-radical activating enzymes in yellow, biotin synthases (BioB) in blue, lipoic acid synthases (LipA) in teal, 4-
demethylwyosine synthases (TYW1) in cyan and 7-carbxy-7deazaguanine synthases (QueE) in maroon. Diamonds denotes structurally characterized enzymes. This figure was generated in Cytoscape$^{93}$. 
Figure 1.2. The core domain of AdoMet radical enzymes. Members of the AdoMet radical enzyme superfamily adopt a partial (β/α)$_6$ triose isomerase (TIM) barrel core fold. (A) Cartoon representation of the AdoMet radical core fold with loops omitted for clarity. The β-strands comprising the inner cavity of the barrel are colored in pink and the α-helices which flank the outside of the barrel are colored light pink. The partial (β/α)$_6$ TIM barrel architecture provides residues to ligate the co-substrate AdoMet (green) and the three cysteines of the CX$_3$CXφC cluster-binding motif are located on a loop following β1 denoted the cluster binding loop. The AdoMet radical cluster is shown in a ball and stick representation with iron atoms in orange and sulfur atoms in yellow. (B) Topology diagram of the AdoMet core with the β-strands in pink and α-helices in light pink. The cysteines of the CX$_3$CXφC motif are represented as yellow circles and the iron and sulfur atoms of AdoMet radical clusters in orange and yellow circles.
Figure 1.3. Reactions of AdoMet radical enzymes harboring an auxiliary cluster. Some members of the AdoMet radical enzyme superfamily utilize N- and C-terminal extensions to coordinate auxiliary clusters. Shown here are examples of reactions catalyzed by auxiliary cluster-containing AdoMet radical enzymes with the hydrogen atom abstracted by 5'-dAdo• indicated in blue. (A) BioB catalyzes the insertion of a sulfur atom between two unactivated carbon atoms in the biosynthesis of Biotin. (B) LipA uses two molecules of AdoMet to perform two sulfur insertion reactions. (C) The auxiliary cluster of HydG is the site of dehydroglycine decomposition and as the reaction of HydE is still unclear, the reaction of HydE, HydF and HydG is shown here. (D) RimO transfers a methylthiol group to an aspartic residue of ribosomal protein S12. (E) MiaB post-transcriptionally modifies an adenine base of several tRNAs. (F) TYW1 catalyzes the characteristic tricyclic ring formation of a class modified tRNA bases called the Wye bases.
Figure 1.4. Topology diagrams and ribbon drawings of additional cofactor-containing AdoMet radical enzymes. (A) In addition to the AdoMet radical cluster, OxsB uses a cobalamin cofactor for catalysis. The overall architecture of OxsB is shown with Cbl-binding domain of colored in pink. (B) BioB folds into a full TIM barrel completed by structural elements from the C-terminal extension. The auxiliary [2Fe-2S] cluster is ligated by residues from the core of the barrel including an unusual arginine residue indicated as a blue circle. (C) Three cysteine residues from the N-terminal extension and a serine residue (red circle) from the C-terminal extension coordinate the auxiliary [4Fe-4S] cluster of LipA. (D) In addition to completing the full TIM barrel of HydG, the C-terminal extension also ligates a [4Fe-4S] auxiliary cluster. (E) The C-terminal extension of HydE binds the [2Fe-2S] cluster at the surface of the protein. (F) The auxiliary [4Fe-4S] cluster of RimO is housed at the interface between the N-terminal domain and the AdoMet binding domain. The AdoMet radical core domain is colored in cyan, and the N- and/or C-terminal extensions are colored in yellow and grey, respectively. Structural elements that are not found in the aforementioned domains and extensions are colored white. On the topology diagrams, the cysteine ligands to the clusters are shown as yellow circles. Iron and
sulfur atoms of the clusters are represented as orange and yellow circles in the topology and as orange and yellow balls in the ball and stick representation of the ribbon diagrams.
Figure 1.5. SPASM/Twitch subfamily reactions. The involvement of SPASM/Twitch domain containing enzymes anSME, MoaA, BtrN, and AlbA/SkfB in the biosynthesis of their respective products is shown.
Figure 1.6. Topology diagrams and ribbon drawings of SPASM/Twitch subfamily members. (A). Overall architecture of one of the founding members of the SPASM subfamily, anSME, reveals a two fully ligated [4Fe-4S] auxiliary clusters coordinated by a C-terminal SPASM domain. (B) Architecture of the Twitch domain-containing enzyme MoaA. MoaA uses a similar secondary fold to anSME to ligate one [4Fe-4S] auxiliary cluster with an open coordination site for binding of substrate (yellow circle in topology diagram). (C) BtrN
coordinates one [4Fe-4S] auxiliary cluster using four cysteine ligands from the C-terminal Twitch domain. (D) SuiB ligates two [4Fe-4S] auxiliary cluster with a cysteine ligand from the linker region between AdoMet radical domain and the SPASM domain. The N-terminal binding domain was omitted for clarity. (E) CteB portrays the first SPASM domain to contain a site differentiated auxiliary cluster. The SPASM domain coordinates two [4Fe-4S] auxiliary clusters, but Aux I contains a unique iron for binding substrate (yellow circle in topology diagram) reminiscent of MoaA. The N-terminal RRE domain of CteB was eliminated for clarity. (F) PqqE fully ligates two auxiliary clusters with an unusual aspartic acid ligand (purple circle in topology diagram) to the [4Fe-4S] of Aux II cluster. In Aux I, a [2Fe-2S] cluster was observed. The AdoMet radical core domains are colored in pink, the C-terminal SPASM/Twitch domains in green, and secondary structures outside of these two domains in light blue. The iron-sulfur clusters are shown as ball and stick representations with the iron in orange and the sulfur in yellow. The positions of cysteine ligands in the topology diagrams are indicated as yellow circles. Black dashed lines indicate distances between the iron-sulfur clusters.
References


41. Young, A. P.; Bandarian, V., Pyruvate is the source of the two carbons that are required for formation of the imidazoline ring of 4-demethylwyosine. Biochemistry 2011, 50 (49), 10573-10575.


CHAPTER II
Biochemical and Structural Characterization of a Schiff base in the Radical-Mediated biosynthesis of 4-demethylwyosine

Chapter Summary
TYW1 is a radical S-adenosyl-L-methionine (AdoMet) enzyme that catalyzes the condensation of pyruvate and N-methylguanosine to form the posttranscriptional modification, 4-demethylwyosine, \textit{in situ} on transfer RNA. Two mechanisms have been proposed for this transformation, with one of the possible mechanisms invoking a Schiff base intermediate formed between a conserved lysine residue and pyruvate. Utilizing a combination of mass spectrometry and X-ray crystallography we have obtained evidence to support the formation of a Schiff base lysine adduct in TYW1. When $^{13}$C labeled pyruvate is used, the mass shift of the adduct matches that of the labeled pyruvate, indicating that pyruvate is the source of the adduct. Furthermore, a crystal structure of TYW1 provides a visualization of the Schiff base lysine-pyruvate adduct, which is positioned directly adjacent to the auxiliary [4Fe-4S] cluster. The adduct coordinates the unique iron of the auxiliary cluster through the lysine nitrogen and a carboxylate oxygen, reminiscent of how the radical SAM [4Fe-4S] cluster is coordinated by SAM. The structure provides insight into the binding site for tRNA and further suggests how AdoMet radical chemistry can be combined with Schiff base chemistry for RNA modification.

Contributors:
This chapter was written with Dr. Anthony P. Young and Professors Catherine L. Drennan and Vahe Bandarian. Dr. Young performed the cloned and purified the TYW1 constructs and performed the biochemical experiments.
Introduction

In addition to the four canonical RNA bases, transfer RNA (tRNA) molecules are adorned with a wide variety of posttranscriptionally modified bases, which range in complexity from methylations to hypermodifications and may require the actions of several enzymes, such as the 8-step transformation involved in the biosynthesis of the modified base, queuosine\(^1\)\(^8\). These modifications expand the chemical repertoire of nucleobases and have also been implicated in the stabilization of tRNA structure and the maintenance of function\(^9\)\(^10\). In fact, many hypermodified bases occur at or adjacent to the anticodon loop, where they are thought to be involved in modulating codon-anticodon interaction or stability\(^11\). The nucleoside wyosine is a hypermodified base that contains a characteristic imidazopurine (tricyclic) core (Scheme II.1)\(^12\)\(^-\)\(^15\). Wyosine derivatives are found at position 37 of the anticodon stem loop (ACSL) of tRNA\(^{p}^\)\(^{he}\) molecules from archaea and eukarya\(^16\)\(^-\)\(^18\). The presence of the large hydrophobic bulky base in position 37, which is adjacent to the wobble base in the anticodon loop, has been shown to be important in preventing frameshifts\(^19\)\(^,\)\(^20\).

The simplest wyosine derivative, 4-demethylwyosine (imG-14), is found as both a modified tRNA base in some archaeal species, and as an intermediate on the pathway to more complex wyosine derivatives\(^12\)\(^,\)\(^16\)\(^-\)\(^18\)\(^,\)\(^21\)\(^-\)\(^29\) (Scheme II.1). In vivo, imG-14 is biosynthesized from guanosine by the successive actions of two enzymes, the S-adenosyl-L-methionine (SAM) dependent methyl transferase TRM5 and the radical SAM enzyme TYW1\(^17\)\(^,\)\(^18\)\(^,\)\(^30\)\(^,\)\(^31\). TRM5 methylates the genetically encoded guanosine residue at N1 creating N-methylguanosine (m\(^1\)G)\(^30\)\(^,\)\(^31\), whereas TYW1 adds C2 and C3 of pyruvate to m\(^1\)G via a radical mechanism, forming the characteristic tricyclic ring of all wyosine derivatives\(^32\)\(^-\)\(^37\). The identity of pyruvate as the source of the carbons in the imidazolene ring was established using \(^13\)C pyruvate isotopologues in vitro, which showed that C2 and C3 of pyruvate are incorporated into the tRNA base and C1 is lost as an unidentified side product, proposed to be either carbon dioxide or formate\(^33\).

TYW1, like other enzymes in the 100,000-plus membered radical SAM superfamily\(^38\), uses a SAM-bound [4Fe-4S] cluster to reductively cleave SAM, forming methionine and a highly reactive 5’-deoxyadenosyl radical species (5’-dAdo•) that initiates radical-based chemistry\(^39\).
TYW1 also has an auxiliary iron-sulfur cluster, identified as a [4Fe-4S] cluster by an electron paramagnetic resonance (EPR) and Mössbauer study. Additional or auxiliary iron-sulfur clusters have been identified in several characterized radical SAM enzymes involved in the biosynthesis of cofactors (e.g. biotin, lipoic acid, and the molybdopterin cofactor) and antibiotics, in the posttranslational modification of proteins, and posttranscriptional modifications of RNA. A small subset of these enzymes also contain auxiliary [4Fe-4S] clusters that are ligated by three cysteine residues, leaving a site differentiated iron ion. In these cases, the enzymes are predicted to use the open coordination site to bind substrate, as seen in the structural and spectroscopic studies of molybdopterin biosynthetic enzyme MoaA. The auxiliary cluster in TYW1 is predicted to be ligated by three cysteine residues, leaving an open coordination site.

Two competing mechanisms have been proposed for how TYW1 uses its [4Fe-4S] clusters to form the tricyclic ring of wyosine derivatives. In one mechanism, pyruvate is activated by forming a Schiff base to a conserved Lys residue in the active site. Hydrogen atom abstraction from the methyl group of m'G by the 5'-dAdo radical, which has been demonstrated biochemically, results in production of 5'-deoxyadenosine (5'-dAdoH) and a methylene radical. The methylene radical in turn attacks the C2 of the immobilized pyruvate. Homolytic cleavage of the C1-C2 bond of pyruvate eliminates the equivalent of a carbon dioxide anion radical. The carbon dioxide anion radical may then be oxidized to form carbon dioxide or reduced to form formate, with the auxiliary cluster possibly playing a role. The exocyclic amino group at C2 of guanosine resolves the Schiff base in a transimination reaction that regenerates the active site Lys for a subsequent turnover cycle.

A competing mechanism has been proposed that discounts the role of the conserved Lys residue proposing, on the basis of spectroscopic studies, that the auxiliary cluster alone facilitates the activation of pyruvate. In this proposal, hydrogen atom abstraction also occurs at the methyl group of m'G, consistent with the biochemical data, and the resulting radical species forms a bond with C2 of pyruvate, prompting homolytic cleavage of the C1-C2 bond of pyruvate. In this mechanism, decarboxylation is concomitant with reduction of the auxiliary cluster to form carbon dioxide. The auxiliary cluster is proposed to catalyze the
formation of the imidazoline ring by activating the keto oxygen for elimination upon transimination\textsuperscript{35,36}. This mechanism does not provide a direct role for a catalytically essential lysine, which is proposed to play an indirect role in modulating the redox properties of the cluster\textsuperscript{36,37}

A key difference between the two proposed mechanisms is the involvement of a Schiff base in mechanism 1. Here we employ crystallography and biochemistry to investigate the formation of a Schiff base between the conserved Lys in TYW1 and pyruvate.

**Results**

**Overall architecture of holo-TYW1**

To investigate the molecular architecture of holo-TYW1, a structure of anaerobically purified and reconstituted TYW1 from *Methanocaldococcus jannaschii* was solved to 1.6-Å resolution (Figure II.1). TYW1 crystallizes as a functional monomer with the structural core adopting a partial (\(\beta/\alpha\))\textsubscript{6} TIM barrel, the radical SAM core fold found in most members of the radical SAM superfamily\textsuperscript{40,42,50,51,63-72}. This core fold includes a loop that follows \(\beta1\) and provides the three cysteines, C62, C66 and C69, of the CX\textsubscript{3}CX\(\phi\)C motif for binding the canonical radical SAM cluster (where \(\phi\) is a conserved aromatic residue)\textsuperscript{73}. In TYW1, the core fold is extended by the addition of N- and C-terminal extensions, which flank either side of the partial TIM barrel architecture (Figure II.1). The N-terminal extension begins with \(\alpha1'\), which folds into an antiparallel \(\beta\)-hairpin (\(\beta1'\) and \(\beta2'\)) that extends the inner face of the partial barrel. \(\beta2'\) is followed by two \(\alpha\)-helices, \(\alpha2'\) and \(\alpha3'\), which sit below the base of the N-terminal region of the radical SAM domain, creating a binding site for a second, or auxiliary, iron-sulfur cluster. On the other side of the radical SAM domain, the C-terminal extension folds into a nonconsecutive antiparallel \(\beta\)-sheet, \(\beta7'\) and \(\beta8'\), where \(\beta8'\) is found adjacent to \(\beta6\) (of the radical SAM domain). The C-terminal extension expands the inner face of the active site, and with \(\alpha3'\) of the N-terminal extension helping to seal off the bottom of the barrel.

In this structure, we observe one intact and one partial [4Fe-4S] cluster. The canonical [4Fe-4S] radical SAM cluster is not intact and is best accounted for by a [2Fe-3S] cluster (Figure II.1A) with a water molecule as the fourth ligand, suggesting either cluster degradation or incomplete
cluster reconstitution. In contrast, the auxiliary cluster is an intact [4Fe-4S] cluster coordinated by three conserved cysteines, C26, C39 and C52 (Figure II.1A, B, 2A).

Two apo structures of TYW1 from *Methanocaldococcus jannaschii* (apo-MjTYW1) and *Pyrococcus horikoshii* (apo-PhTYW1) have been previously reported to 2.4-Å and 2.2-Å resolution, respectively. These two structures are very similar to our MjTYW1 structure, with a rmsd of 0.56 Å for apo-MjTYW1 and 0.94 Å for apo-PhTYW1 (calculated by alignment of 222 out of 280, and 234 out of 325, alpha carbons respectively, using PyMol). A key difference between these structures and ours is the increased iron-sulfur cluster content in our structure, but the previous structures did correctly predict the locations of the cluster-binding sites on the basis of the clustering of conserved Cys residues. Interestingly, outside of the other TYW1 structures, the closest structural homologue with a Z-score of 15 and a rmsd of 3.8 Å is RlmN, a radical SAM enzyme involved in the methylation of ribosomal RNA and recently discovered to also methylate tRNA.

Additional structurally similar enzymes are pyruvate formate lyase activating enzyme (PFL-AE, Z-score of 14.5 and rmsd of 3.0 Å), viperin (Z-score 12.4 and rmsd of 3.5 Å), MoaA (Z-score 12.0 and rmsd of 3.5 Å), and QueE (Z-score of 11.8 and rmsd of 3.8 Å).

**Visualization of a Schiff base intermediate in crystalllo.**

Sequence analysis correctly predicted incomplete Cys ligation of the auxiliary cluster; only three cysteine residues are conserved (C26, C39 and C52) outside of the canonical radical SAM cluster binding CX₃CX₄C motif, and we find that only these three conserved cysteine residues coordinate this auxiliary [4Fe-4S] cluster (Figure II.2A,B). In contrast to predictions, however, there is not an open coordination site on the auxiliary cluster. Instead, we find the catalytically essential residue, K41, in close proximity (~2.8 Å) to the unique iron of the auxiliary cluster (Figure II.2A). During iterative rounds of refinement and model building, positive difference density extending from the Nz of K41 towards the unique iron of the auxiliary cluster appeared (Figure II.2A). Guided by this difference density, we modeled in and refined the Schiff base intermediate predicted to form between a lysine residue and the pyruvate substrate during catalysis. We found that a pyruvate-lysine adduct accounts well for the density (Figure II.2B). It must be noted that the density observed for adduct can also accommodate a reduced
Schiff base and therefore we cannot rule out the possibility that the species observed is an amine rather than an imine. Regardless, this covalent Schiff base adduct places the oxygen of the carboxylic acid of pyruvate and the Nz of lysine about 2.4 Å and 2.3 Å away from the unique iron of the auxiliary cluster, respectively, creating a bidentate ligation reminiscent of SAM binding to the radical SAM clusters of members of the superfamiliy \(^{40,43,59,60,67,72,75-78}\) (Figure II.2C,D).

In addition to interactions with the unique iron, the Schiff base is stabilized in the active site through water hydrogen bonds (Figure II.3A). At a distance of 2.5 and 3.0 Å, a water molecule bound by Q54 and S134 provides contacts to the carboxylic group of the pyruvate-derived portion of the Schiff base. Whereas stacking interactions from F160 help orient the C3 methyl of the pyruvate-lysine adduct in the active site, the inner face of the Schiff base binding site is made up of an intricate hydrogen bonding network between highly conserved residues, Y40, Y45, Y180 and E248, from the N-terminal extension and the radical SAM core. These interactions (Figure II.3A) serve to orient the adduct species at the base of the active site with C2 of pyruvate, the site of radical attack by the m'^G substrate radical, 12.6 Å away from the radical SAM cluster.

**SAM binding motifs appear to be conserved.**

We modeled the missing atoms of the radical SAM cluster using the positions of the cluster atoms that we could observe and modeled SAM binding to the cluster using the structure of MoaA (PDB 1TV8) as a guide (Figure II.4). In particular, SAM was positioned such that its amino group and its carboxylic group were 2.3 Å and 2.0 Å from the unique iron, respectively. SAM was also positioned in a catalytically competent orientation with the S-atom poised 3.2 Å from the unique iron. We find that when SAM is positioned in this classic orientation with respect to the radical SAM cluster, SAM can make the traditional interactions with the SAM-binding motifs \(^{75,76,79}\). For example, a carbonyl from the “GGE” motif, S\(_{136}\)G\(_{137}\)E\(_{138}\) in TWY1, is available to hydrogen bond to the amide of SAM, and D184 (ribose motif) and R221 are positioned to interact with the ribose hydroxyl groups of SAM either through a water molecule or hydrogen bond, respectively (Figure II.3B). Residues T223 (GXIXGXXE motif) and F68 (\(\phi\) in CX\(_3\)CX\(_4\)C motif) are found in close proximity to the adenine moiety of modeled SAM and
can provide hydrophobic interactions to orient the adenine ring. In addition, backbone atoms are positioned to make classic interactions, M251 (β6 motif) and F68 (CX3CXφC motif), to the N1, N6 exocyclic amine, and N7 of the adenine moiety (Figure II.3B).

A positive electrostatic surface near active site suggests location of tRNA binding site

The nucleotide modified by TYW1, m'G37, is part of the ACSL of tRNA\textsuperscript{Phe}, which when bound to its modifying enzyme, is expected to adopt a tRNA bulge conformation that has m'G37 flipped out of the loop\textsuperscript{74,80}. Due to its negatively charged phosphate backbone, the ACSL should bind a positively charged site on the protein. Electrostatic calculations (Figure II.5) show a large region of positive electrostatic charge that leads from the active site cavity to the tip of the N-terminal extension near α3' and α1a (Figure II.5). Modeling suggests that this patch of positive charge is appropriately the correct size for a 20 base pair ACSL (Figure II.5). In this model, the m'G37 sits on top of the active site with the rest of the ACSL extended along the positively charged protein surface (Figure II.5). The active site of TYW1 is solvent exposed in the absence of nucleotide substrate, but the binding of a tRNA\textsuperscript{Phe} molecule to TYW1 would be expected to close off the active site, making it less accessible to solvent.

There is a space in between SAM and the Schiff base in the active site cavity that is a suitable size for a m'G binding site (Figure II.3C,D). Modeling m'G into this putative binding site, positions the N1-methyl group of m'G (the site of hydrogen atom abstraction by 5'-dAdo•), 3.2 Å away from the C-5' of SAM (Figure II.3D). The N1-methyl group is also 4.9 Å from the C2 of the pyruvate-derived portion of the Schiff base intermediate (Figure II.3D). In this location, the base of m'G would be capable of making π-π stacking interactions with the adenine ring of SAM, and forming hydrogen bonds (~2.4 Å) between O6 and R221, and (~2.9 Å) between N7 and K250 (Figure II.3C). The phosphate group could be stabilized by hydrogen bonding interactions with K24, distance modeled at 2.6 Å (Figure II.3C).

\textit{In vitro activity of C195S and K41A TYW1 mutants}

K41 has been shown previously to be essential \textit{in vivo}\textsuperscript{50} and here we wanted to verify that a K41A variant protein is unable to produce imG-14 \textit{in vitro}. As a control, we also prepared a C195S TYW1 variant. C195 is adjacent to the radical SAM cluster but is not a direct cluster
ligand. The extracted ion chromatogram at \( m/z \) 322 (imG-14 \([M+H^+]) obtained when C195S TYW1 (red), wild type TYW1 (black), and K41A TYW1 (blue) are incubated in the presence of tRNA, reductant, SAM, and pyruvate is shown in Figure II.6. There is a peak present in both the wild type and C195S samples that is not present in the negative control (no enzyme) or in the K41A sample that corresponds to imG-14. Therefore, at least under the conditions of the assay, C195S variant is catalytically active and K41A is not active. This finding supports the importance of K41 in catalysis. The C195S variant is used in the subsequent Schiff base trapping experiments.

**In vitro trapping of a Schiff base intermediate using C195S-TYW1**

To unambiguously demonstrate formation of a Schiff base between TYW1 and pyruvate, C195S-TYW1 was incubated with NaCNBH₃ and NaCNBH₃ in the presence of SAM, reductant, and substrate. Reactions were treated with iodoacetamide prior to tryptic digestion to ensure that oxidation of cysteines did not occur. The shortest tryptic fragment of TYW1 containing a pyruvate-modified K41 would be NCYK, where \( K \) denotes the modified K41 and \( C \) denotes a carbamidomethylated cysteine residue. Since a modified \( K \) at the cleavage site for trypsin may lead to a missed cleavage, we also considered that the shortest observable fragment would be NCYKSK. The fragment NCYK would have an \( m/z \) 584.2497 \([M+H^+]\), while modification of the lysine, NCYK, would shift the mass of the peptide to \( m/z \) 656.2708 \([M+H^+]\). The fragment NCYKSK peptide should have an \( m/z \) of 799.3767 \([M+H^+]\), whereas a pyruvate-modified lysine, NCYKSK, would shift the mass to \( m/z \) 871.3978 \([M+H^+]\).

The ability of TYW1 to form a Schiff base with pyruvate was investigated by incubating the enzyme under anaerobic conditions in a mixture that contained all components of the assay, including pyruvate, yeast tRNA, methyl viologen, dithionite, and SAM. In addition, NaCNBH₃ was also added in the incubations to stabilize the Schiff base that forms by reduction. The samples were digested with trypsin, the cysteine residues alkylated, and the resulting fragments subjected to LC-MS analysis, following the incubations. The extracted ion chromatograms (Figure II.7) reveal a peak at \( m/z \) 871, corresponding to NCYKSK, as well as a peak at \( m/z \) 584.2497 corresponding to NCYK. These results were encouraging as they suggested that the K
could be modified; however, we could not eliminate the possibility that the modified K was present in the purified enzyme.

To determine the source of the pyruvate appended to K41, identical incubations were carried out with 1,2,3-$^{13}$C$_3$pyruvate. The resulting extracted ion chromatograms (Figure II.8) reveal peaks at m/z 874, which would be expected for a NCYKSK fragment where the 1,2,3-$^{13}$C$_3$-pyruvate is present. Interestingly, we also observe a significant peak at m/z of 871, which corresponds to unlabeled pyruvate appending K41. This peak likely results from pyruvate that is adventitiously bound to the enzyme through K41. In fact, we point out that the structure of the pyruvate-modified K41 reported above was solved with protein that had never been incubated with pyruvate in vitro, and was likely purified with the Schiff base that had formed during protein production in vivo, or during the purification and reconstitution of the protein under reducing conditions.

To further confirm that the peak at m/z 874 corresponds to NCYKSK, which is obtained by pyruvate-modification and subsequent reduction with NaCNBH$_3$, we repeated the experiment with NaCNB$_2$H$_3$. In this experiment we would expect that the peak at m/z 874 that is observed with 1,2,3-$^{13}$C$_3$-pyruvate would shift to 875 in the presence of NaCNB$_2$H$_3$. Indeed, as predicted, when TYWI was incubated in the presence of cyanoborohydride the m/z peak for NCYKSK shifts from 874.4085 to 875.4156, which is within 1 ppm of the theoretical mass of the peptide containing three $^{13}$C and one $^2$H atoms (Figure II.9A and B). These data unambiguously show that pyruvate can modify K41 to a Schiff base.

In the experiments with the NaCNB$_2$H$_3$ and 1,2,3-$^{13}$C$_3$-pyruvate, as with the experiments with unlabeled reductant, we also observe a peak at 871.399 corresponding to NCYKSK that does not undergo reduction. This result suggests that the pool of pyruvate-modified K41 that we observe is already reduced in the protein. To probe this further we repeated the experiment in the absence of added pyruvate (Figure II.9C and D). Indeed, the peak at 871.399 does not shift when NaCNB$_2$H$_3$ is used, confirming the presence of modified K41 in the sample. We cannot eliminate the possibility that the reduction of this pool occurs in the course of the incubation, as the structural data do not clearly distinguish between an imine and an amine.
To determine which reaction components were required for the Schiff base to form, a series of control reactions were setup with TYW1 in the presence of 1,2,3-$^{13}$C$_3$-pyruvate and either NaCNBH$_3$ or NaCNB$_2$H$_3$ (Figure II.10). The controls show that SAM, dithionite, or tRNA are not required, as the peak at $m/z$ 874.4085 ± 1 ppm (with NaCNBH$_3$) or $m/z$ 875.4148 ± 1 ppm (with NaCNB$_2$H$_3$) is present in all the samples. Therefore, the formation of the Schiff base is not dependent upon the presence of the tRNA or the oxidation state of the cluster, which is in agreement with previously published spectroscopic studies that show changes in the environment of the auxiliary cluster in the presence of pyruvate in the absence of SAM or reductants$^{35,52}$.

Discussion

The mechanism by which TYW1 catalyzes the addition of two carbon atoms from pyruvate to extend the ring of m$^1$G37, forming the characteristic imidazopurine of the wyosine bases found in tRNA$^{\text{Phe}}$, has been disputed$^{32-36,52}$. Although both mechanistic paradigms suggest a role for the auxiliary cluster in activating pyruvate, one suggests a more intimate interaction between TYW1 and pyruvate, which is mediated by a Schiff base linkage. Herein, we present biochemical and structural evidence to support a role for a Schiff base in the reaction catalyzed by TYW1.

Our structure presents the first snapshot of the active site of TYW1 with iron sulfur clusters bound. This structure reveals electron density for an auxiliary [4Fe-4S] cluster coordinated by the three conserved cysteines, confirming the results of a previous EPR and Mossbauer study. As predicted, the conserved and catalytically important K41 is found in close proximity to the unique iron of the auxiliary cluster. K41 has been shown previously to be essential for in vivo activity$^{50}$; the recombinant K41A variant is shown here to be catalytically inactive, further emphasizing its essential role. We observe electron density consistent with pyruvate forming a Schiff base intermediate with K41, and find that the adduct is within hydrogen bonding distance of the cluster, consistent with earlier EPR and Mossbauer studies$^{35,52}$. In addition, the chemical in vitro trapping of a pyruvate-lysine adduct using pyruvate, 1,2,3-$^{13}$C$_3$-pyruvate and either
NaCNBH$_3$ or NaCNB$_2$H$_3$ further confirms that TYW1 is able to form a Schiff base intermediate between pyruvate and K41, which is inline with mechanism 1 in Scheme II.2.

A mechanistic paradigm for TYW1 that is consistent with all structural, biochemical, and spectroscopic evidence to date is shown in Scheme II.3. A Schiff base between K41 and pyruvate is proposed to precede the reductive cleavage of SAM and the subsequent hydrogen abstraction from m$^1$G. The resulting m$^1$G methylene radical can attack the C2 of the cluster-bound Schiff base, forming a new C-C bond and leading to a subsequent decarboxylation. Although the fate of C1 of pyruvate remains to be established, the ligation of the Schiff base to the cluster sets up the adduct for one electron oxidation by the auxiliary cluster, which would lead immediately to decarboxylation to carbon dioxide. Reduction of the cluster and concomitant formation of carbon dioxide would place the cluster in the +1 oxidation state, which is in agreement with previous Mossbauer and EPR studies$^{35,52}$. For subsequent turnover cycles, flavodoxin could reoxidize the auxiliary cluster and potentially deliver the electron back to the radical SAM cluster, preparing it for the next round of catalysis as has been observed in AtsB and anSMEcpe$^{81}$.

Visualization of a Schiff base intermediate raises the question why TYW1 needs both a Schiff base intermediate and an iron sulfur cluster in wyosine base formation, when either in theory is sufficient to perform the requisite chemistry; both can act as an electron sink during catalysis and both can bind and appropriately orient the substrate. Hints may be found in revisiting the roles of Schiff base intermediates in biochemistry and the structure of TYW1. Schiff bases are used enzymatically, producing a highly reactive imine intermediate. As in other radical SAM enzymes, the core of TYW1’s structural architecture comprises of a partial ($\beta/\alpha$)$_6$ TIM barrel which buries the active site within the inner face of the barrel. This enclosure of the active site serves to sequester the reactive radical species produced during catalysis and decrease off pathway interactions: a strategy used by both radical SAM enzymes and cobalamin dependent enzymes. In TYW1 the active site cavity is largely solvent exposed, but would be sealed off with binding of the bulky tRNA substrate. The active site cavity as well as the surface leading from the active site cavity is largely positively charged, and tRNA is predicted to bind with the anticodon stem loop making contact with this positive patch of protein surface, orienting the
anticodon loop to sit on top of the active site. Binding of the tRNA will not only effectively seclude the active site from solvent, but also impede the diffusion of substrates into the active site. No other channel for substrate entry into TYW1 has been found and it is possible that all three substrates, SAM, pyruvate and tRNA$^\text{Phe}$, access the active site through the opening observed in the structure. This active site architecture imposes an order of binding on the substrates as the negatively charged pyruvate molecule and SAM molecule will not be able to enter the activity site once tRNA is bound. Taking this idea into consideration, as well as that TYW1 is purified with some reduced Schiff base, the Schiff base in TYW1, in addition to potentially orienting the pyruvate for chemistry, could play a role in tethering pyruvate to the open active site that exists in the absence of the tRNA, priming it for catalysis once the other substrates bind. Therefore, the Schiff base may be playing a complex role by directing binding, orientation, and catalysis in TYW1.

The work presented here strongly supports a mechanistic paradigm where a Schiff base plays a central role in activating the pyruvate for catalysis (Scheme II.3). It remains to be seen if the cluster plays a role in resolving the radical intermediate hypothesized to be formed and exist on C1 during this transformation. Identification of the side product that results from C1 (carbon dioxide or formate) will be highly informative in this regard.

Conclusion

Radical SAM enzymes are a treasure trove of interesting chemical reactivity and can proceed through a wide variety of mechanisms to facilitate challenging transformations. Our understanding of the roles of additional cofactors in radical SAM chemistry has been hazy at best, with most additional cofactors having no concrete roles assigned to them. For auxiliary [4Fe-4S] clusters, the proposed roles include substrate binding and electron transport. In the biochemical and structural work presented here, we provide support for a Schiff base intermediate in catalytic mechanism of TYW1. TYW1 is therefore emerging as a member of a subset of radical SAM enzymes that in addition to employing radical mechanisms to activate their substrates, they also utilize covalent ones.
Experimental procedures

Expression of tobacco etch virus (TEV) protease

Escherichia coli SG1300009 pLacIRARE cell line was transformed with TEV (both generous gifts from Hazel Holden at University of Wisconsin-Madison)\(^{53}\), and plated on an lysogeny broth-Lennox (LB) agar plate containing 34 μg/mL chloramphenicol and 100 μg/mL ampicillin. A single colony from this plate was used to inoculate a 150 mL culture of LB containing 34 μg/mL chloramphenicol and 100 μg/mL ampicillin that was grown overnight at 37°C with shaking at 225 RPM.

The overnight culture was used to inoculate twelve 2.8 L Fernbach flasks, each containing 1 L of LB containing 34 μg/mL chloramphenicol and 100 μg/mL ampicillin. They were grown at 37°C with shaking at 200 RPM until the OD\(_{600}\) ~ 0.6. The temperature was set to 16°C and protein expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside to a final concentration of 0.1 mM. After eighteen hrs, cells were harvested by centrifugation and flash frozen with liquid N\(_2\).

Purification of TEV protease

TEV protease was purified at 4°C. Approximately 30 g of cells were resuspended in 50 mM potassium phosphate (KPi) (pH 7.4) buffer containing 0.5 M potassium chloride (KCl), 20 mM imidazole, 20% (v/v) glycerol, and 1 mM phenylmethane sulfonyl fluoride (PMSF). Cells were lysed using a Branson sonificator at 50% amplitude with 15 sec on and 59.9 sec off, for a total of 12 minutes of lysis time, while stirring on ice. Cell lysate was clarified via centrifugation at 18,500 xg at 4°C for 30 min.

The clarified lysate was loaded onto a 5 mL HiTrap chelating HP column charged with nickel sulfate and preequilibrated in 50 mM KPi (pH 7.4) buffer containing 0.5 M KCl, 20 mM imidazole, and 20% (v/v) glycerol. When all the lysate was loaded, the column was washed with 50 mL of the above buffer. Protein was eluted with a gradient to 100% 50 mM KPi (pH 7.4) buffer containing 0.5 M KCl, 0.5 M imidazole, and 20% (v/v) glycerol over 40 mL. Fractions containing protein were identified via SDS-PAGE and the concentrated fractions were pooled and then dialyzed against 4 L of 50 mM KPi (pH 7.4) buffer containing 0.5 M KCl, 20
mM imidazole, and 20% (v/v) glycerol at 4°C with two changes. Glycerol was added to the protein to a final concentration of 50% (v/v), and the protein was aliquoted and stored at -20°C \(^{53}\).

**Cloning and expression of wild type TYW1 (wtTYW1)**

To create a wtTYW1 (pAY613) construct containing a TEV protease site, the codon optimized *Methanocaldococcus jannaschii* TYW1 gene was cut from pET28a (pAY429) \(^{33}\) using the *HindIII* and *NdeI* cut sites and inserted into pET28J vector digested with *HindIII* and *NdeI*. Expression of wtTYW1 was as previously described \(^{33}\).

**Purification of wtTYW1**

wtTYW1 was purified in a Coy anaerobic chamber (approximately 97% N\(_2\), 3% H\(_2\)) at room temperature. Cell paste was resuspended in 20 mM Tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) (pH 8) and lysed by sonicating with a Branson sonicator at 50% amplitude for 15 min lysis time with 15 sec on and 45 sec off. Cell lysate was clarified by centrifugation at 18,500 xg at 4°C for 30 min.

Clarified lysate was transferred to a sealed bottle and heated for 30 min in an 80°C water bath outside the anaerobic chamber and was then cooled for 15 min in an ice bath. The lysate was again clarified by centrifugation at 18,500 xg at 4°C for 30 min.

Solid ammonium sulfate was added to the clarified lysate to a final concentration of 1 M, which was loaded onto a 40-mL butyl-sepharose FF column equilibrated in 20 mM Tris-HCl (pH 8) buffer containing 1 M ammonium sulfate. Following loading of the lysate, the column was washed with equilibration buffer until the absorbance returned to baseline. Adsorbed protein was eluted with a step gradient to 20 mM Tris-HCl (pH 8) and brown fractions were pooled.

Pooled fractions were loaded onto two 5 mL HiTrap chelating HP columns connected in series charged with nickel sulfate and equilibrated in 50 mM KPi (pH 7.4) buffer containing 0.5 M KCl, and 50 mM imidazole. The column was washed with equilibration buffer and when absorbance had returned to baseline, the adsorbed protein was eluted by a step gradient to 50 mM KPi (pH 7.4) buffer containing 0.5 M KCl, and 0.5 M imidazole. Brown fractions were
pooled and desalted into 50 mM piperazine-N,N’-bis(2-ethanesulfonic acid)-sodium hydroxide (PIPES-NaOH) (pH 7.4) buffer containing 2 mM dithiothreitol (DTT).

The His-tag was removed by incubation overnight with TEV protease at room temperature. The solution was then loaded onto two 5 mL HiTrap chelating HP columns connected in series, charged with nickel sulfate and equilibrated in 50 mM KPi (pH 7.4) buffer containing 0.5 M KCl, and 50 mM imidazole. The protein that flowed through the column was pooled and desalted into 50 mM PIPES-NaOH (pH 7.4) buffer containing 2 mM DTT.

Protein concentrations were determined by the Bradford method and a correction factor of 0.32 was applied. Reconstitution of iron sulfur clusters was performed with desalted TYW1 by stirring at room temperature in the presence of 10 molar equivalents of iron(III) chloride and sodium sulfide for 4 hrs. Precipitated protein was then removed by centrifugation and the protein was desalted into 50 mM PIPES-NaOH (pH 7.4) buffer containing 150 mM KCl, and 2 mM DTT.

Desalted protein was loaded onto a Sephacryl 16/60 S-200 column equilibrated in 50 mM PIPES-NaOH (pH 7.4) buffer containing 150 mM KCl, and 2 mM DTT. The peak corresponding to monomeric protein was pooled, concentrated, and flash frozen with liquid N2.

*Expression of His-tagged TYW1*

TYW1 with an N-terminal His-tag was expressed using pAY429 as previously described.

*Purification of TYW1*

TYW1 with an N-terminal His-tag was purified as described above for wild-type TYW1 with the following changes. The clarified lysate was heated to 80°C in a mineral oil bath within an anaerobic chamber in order to minimize exposure to oxygen. The TEV cleavage step was also omitted.

*Preparation of SAM*

SAM was prepared as described previously.
Preparation of tRNA

Soluble RNA was extracted from a YPL207W knockout strain of Saccharomyces cerevisiae as previously described\(^3\).

Crystallization of TYW1

His-tagged TYW1 was crystallized using the sitting-drop vapor diffusion method within an anaerobic chamber (MBraun) under a nitrogen atmosphere (O\(_2\) < 0.1-1.2 ppm) at 21°C. All sparse matrix screening was performed using a Mosquito pipetting robot (TTP LabTech). In order to obtain data quality crystals, a microseed matrix screening technique was utilized. A microcrystal seed stock was prepared by mixing 1 \(\mu\)L of protein solution (containing 8 mg/mL of TYW1, 50 mM PIPES pH 7.4, 150 mM KCl, 2 mM DTT and 5 mM SAM) with 1 \(\mu\)L of reservoir solution (0.2 M K/Na tartrate tetrahydrate, 25% (w/v) PEG 3350). The very-thin plate-like crystals, which appeared after 24 hrs, were harvested by pipetting the entire drop (2 \(\mu\)L in total) into a seed bead tube (Hampton Research) containing 50 \(\mu\)L of stabilization solution (0.2 M Potassium Sodium tartrate tetrahydrate, 30% (w/v) PEG 3350). After vortexing for 2 min, 450 \(\mu\)L of stabilization buffer was added to the tube to make the seed stock and serial dilutions were performed to prepare a 100-fold and a 1000-fold dilution for the microseed matrix screening experiments.

Data quality crystals were obtained in the screening tray by mixing 0.25 \(\mu\)L of protein solution (containing 9.7 mg/mL of TYW1, 50 mM PIPES pH 7.4, 150 mM KCl, 2 mM DTT and 5 mM SAM) with 0.2 \(\mu\)L reservoir solution (100 mM KH\(_2\)PO\(_4\)/Na\(_2\)HPO\(_4\) pH 6.2, 10% (w/v) PEG 3000) and 0.05 \(\mu\)L of the 1000-fold diluted seed stock, and incubating the mixture over a 70 \(\mu\)L reservoir. Dark-brown rod-like (50 \(\mu\)m x 100 \(\mu\)m x 40 \(\mu\)m) crystals appeared after two days. The tray was transferred to a Coy laboratory products’ antechamber at 24°C and under a 95% argon, 5% hydrogen atmosphere, for harvesting. Crystals were cryoprotected by transferring them in four steps of increasing glycerol concentration into a cryogenic solution containing 100 mM KH\(_2\)PO\(_4\)/Na\(_2\)HPO\(_4\) pH 6.2, 15% (w/v) PEG 3000 and 20% (v/v) glycerol before flash freezing in liquid nitrogen.
Structure Determination of TYW1

A native dataset of TYW1 was collected at the Advanced Photon Source (Argonne, IL) on beamline 24-ID-C using the Pilatus-6M pixel array detector at a wavelength 0.9792 Å (12662 eV) and a temperature of 100 K. During data collection, the crystal was continuously translated along its major macroscopic axis, using the continuous vector strategy. Data were indexed, integrated and scaled in HKL2000 \(^5^4\) in the spacegroup P2\(_1\)2\(_1\)2\(_1\) to 1.64 Å resolution.

An initial model of TYW1 containing one molecule in the asymmetric unit was solved by molecular replacement in Phaser \(^5^5\) (LLG and TFZ scores of 1763.8 and 41.8 respectively) using the previously published structure of TYW1 from Methanocaldococcus jannaschii (PDB 2Z2U) as an all atom search model and the full resolution of the processed dataset. After one round of rigid body refinement, ill-fitting sections of the initial model were deleted and the resulting model was subjected to ten rounds of simulated annealing to minimize existing model bias. Iterative rounds of model building into unambiguous electron density and refinement were performed in Coot \(^5^6\) and Phenix \(^5^7\), respectively. In the initial stages of refinement, Coot was used to place the iron sulfur clusters, whereas water molecules were manually added during advanced stages of refinement. Positive F\(_o\)-F\(_c\) difference density was used to guide the modeling of the covalent adduct formed at K41. A new residue was created by merging a pyruvate molecule and lysine molecule to produce a Schiff base ligand, KAC, in Jligand \(^5^8\). Coot was used to dock the ligand into the structure. In addition to preparing the residue, respective link records and CIF files used during docking and refinement were prepared in Jligand.

Verification of the final structure was guided by simulated annealing composite omit maps and analysis of Ramachandran and rotamer outliers. Side chains without visible density were stubbed at the last visible atom and residues without unambiguous electron density were not included in the model. This refinement procedure resulted in a final structure of TYW1 that includes residues 0-72, 82-255, 260-311 (of 331 residues including the purification tag and linker), a [2Fe-3S] radical SAM cluster, and a [4Fe-4S] auxiliary cluster. It contained 96.9% of its residues in the most favored region of the Ramachandran plot, 3.1% in the additionally favored region and 0.0% in the outlier region. R\(_{\text{work}}\) and R\(_{\text{free}}\) are 0.1760/0.2024, respectively. All refinement statistics are summarized in Table1. The radical SAM and the auxiliary clusters were refined.
with occupancies of 0.84 and 0.58 respectively, indicating that about 80% of molecules in the crystal contained a [2Fe-3S] radical SAM cluster and about 60% have a [4Fe-4S] auxiliary cluster bound. We observed electron density that was best fit by a double conformation of the pyruvate-lysine adduct species (Figures II.11A, B). In the major conformation (refined occupancy of 0.57), the Schiff base is ligated to the auxiliary cluster. However, in the minor conformation (refined occupancy of 0.43) the pyruvate moiety is rotated such that the carboxylate group is facing the Schiff base binding pocket and methyl group facing the auxiliary cluster (Figure II.11C). We posit that the minor conformation of the Schiff base exists in the 40% of the molecules that also lack an auxiliary cluster, and is stabilized through side chain interactions (Figure II.11D). The remainder of this manuscript will focus on the dominant conformation of the Schiff base species.

Visualization of a [2Fe-3S] radical SAM cluster was unexpected given that the protocols used for both anaerobic cluster reconstitution and anaerobic crystallization are the same as have been used successfully previously and given that the auxiliary cluster of TYW1 is intact. Thus the source of the difficulty in observing an intact radical SAM cluster is not obvious. Density for SAM was not apparent in the structure of TYW1, although SAM was present in the crystallization mixture. The absence of SAM in the structure is likely due to cluster disorder. Many unsuccessful attempts were made to obtain a structure of holo-TYW1 with SAM bound (data not shown). Therefore, substrates bound models of TYW1 were achieved by manual docking in Coot.

To model the radical SAM cluster and the bound SAM, the corresponding regions of the MoaA structure (PDB 1TV8) were used as a guide. Protein figures were generated in PyMol, and electrostatic calculations were performed with APBS plugin within PyMol.

**Cloning and purification of C195S and K41A TYW1 mutants**

C195S and K41A TYW1 mutants were created following the Stratagene QuikChange site-directed mutagenesis protocol and pAY613 as template. To create the K41A mutant the following primers were used, 5’-CACAAAACCCTTACGCATCAAMAATTCTAC-3’ and 5’-GTGAATTTTGTAGAATTTTTGA-3’ with the mutated bases shown in bold. To
create the C195S mutant the following primers were used, 5’-CCTACCGTGGATTTCCGGGTTAAAAAAGAATAC-3’ and 5’-GTATTCTTTTTTACCGCCGGAAATGCGACCGGTAGG-3’ with the mutated bases shown in bold. The mutant proteins were expressed and purified as described above for wtTYW1.

Activity assays of C195S and K41A variants - The assays were performed in in 0.1 M Tris-HCl (pH 8), 0.1 M KCl, 4 mM DTT, 2 mM SAM, 2 mM pyruvate, 10 mM sodium dithionite, 1.5 mM methyl viologen, 200 µg yeast tRNA, and 100 µM protein. The reactions were incubated at 50°C overnight. The RNA was extracted and digested to the nucleoside level as previously described. The resulting nucleotide mixture (20 µL) was injected onto a Thermo Vanquish UHPLC interfaced with a Thermo LTQ OrbiTrap XL. The analytes were separated on a Thermo hypersil gold C18 column (150 x 2.1 mm) equilibrated in 50 mM ammonium acetate (pH 6) (solution A). Solution B consisted of 40% acetonitrile (Fisher Optima LC/MS grade). The separation program was as follows with a flow rate of 200 µL/min: 0-3 min, 0% B; 3-3.25 min, 0-0.2% B; 3.25-3.5 min, 0.2-0.8% B; 3.5-3.75 min, 0.8-3.2% B; 3.75-4 min, 3.2-5% B; 4-7 min, 5-25% B; 7-10 min, 25-50%; 10-12 min, 50-75% B; 12-12.1 min, 75-100% B; 12.1-15 min, 100% B; 15-15.1 min, 100-0% B; 15.1-18 min, 0% B. The LTQ OrbiTrap XL was operated in positive ion mode with the FT analyzer set to a resolution of 100,000.

Trapping of Schiff base in TYW1
The assays to trap the Schiff base intermediate were performed in 0.1 M Tris-HCl (pH 8.0) buffer containing 0.1 M KCl, 4 mM DTT, 2 mM SAM, 10 mM pyruvate (either 1,2,3-13C3-pyruvate or 1-13C1-pyruvate), 10 mM sodium dithionite, 1.5 mM methyl viologen, 200 µg yeast tRNA, 10 mM sodium cyanoborohydride (NaCNBH3) or sodium cyanoborodeuteride (NaCNBH3) (freshly prepared), and 100 µM C195S-TYW1. The reactions were incubated for 7 hrs in the glovebox at room temperature. Following the incubation, 20 units of trypsin (bovine pancreas) dissolved in 100 mM Tris-HCl (pH 8.0) was added to each assay, and assays were incubated at room temperature overnight. The following day, DTT was added to a final concentration of 9.5 mM and assays were incubated for 45 min at 56 °C. The reactions were cooled to room temperature and 2-iodoacetamide, dissolved in 100 mM ammonium bicarbonate, was added to a final concentration of 23 mM and assays were incubated in the dark at room
temperature for 30 min. Trypsin and other large molecules were removed by filtration through a PES 10K centrifugal filter and 80 μL of the filtrate was injected onto a Thermo Vanquish UHPLC interfaced with a Thermo LTQ OrbiTrap XL. The analytes were separated on a Thermo hypersil gold C18 column (150 x 2.1 mm) equilibrated in 0.1% TFA (Fisher Optima LC/MS grade) (solution A). Solution B consisted of acetonitrile: 0.1% TFA (Fisher Optima LC/MS grade). The separation program was as follows with a flow rate of 200 μL/min: 0-1 min, 0% B; 1-6.5 min, 0-30% B; 6.5-6.6 min, 30-100% B; 6.6-9.6 min, 100% B; 9.6-9.7 min, 100-0% B; 9.7-12.7 min, 0% B. The LTQ OrbiTrap XL was operated in positive ion mode with the FT analyzer set to a resolution of 100,000.

Controls to determine which components were required for the trapping of a Schiff base were carried out as described above in the presence of 1,2,3-13C3-pyruvate and either NaCNBH3 or NaCNBH2H3 in the absence of one of the following components, SAM, sodium dithionite, or tRNA. The resulting protein was analyzed as described above.

The Schiff base trapping assay was repeated with wtTYW1 to ensure that the wild type protein formed the same adduct as the C195S protein. The assay was repeated as described above using wtTYW1 instead of C195S-TYW1.
Scheme II.1. The biosynthetic pathway for \textit{imG-14}
Scheme II.2. The two hypothesized mechanisms for TYW1
Figure II.1. Overall architecture of TYW1. A. TYW1 adopts a partial (β/α)_6 TIM barrel fold (light pink), which is expanded at both the N-terminus (light blue) and C-terminus (light orange). The radical SAM cluster and auxiliary [4Fe-4S] cluster are shown as spheres (iron in orange and sulfur in yellow) surrounded by 2Fo-Fc electron density contoured at +1 σ (blue). When the radical SAM cluster is refined as a [4Fe-4S] cluster, negative difference density (-3 σ, red) is visible, suggesting either incomplete assembly of the cluster or partial cluster degradation. The location of K41 is also shown. B. Topology diagram of the overall structure of TYW1, colored as in A. The cysteine ligands to the clusters are shown as yellow spheres. Position of K41 is represented as a blue sphere.
Figure II.2. A Schiff base intermediate in TYW1 is visualized. A. An intact [4Fe-4S] auxiliary cluster is shown surrounded by 2Fo-Fc electron density, contoured at 1 σ (blue). Three cysteine ligands, C26, C39 and C52, bind the auxiliary cluster, producing a site differentiated iron ion, which is in close proximity to K41. Positive difference density (green) contoured at +3 σ is shown extending from Nz of K41 to the unique iron. B. To account for difference electron density, a Schiff base between K41 and the substrate pyruvate (KAC41) was modeled and refined. A 2Fo-Fc omit density map (cyan) contoured at 1 σ is shown for the auxiliary cluster, cysteine ligands and the covalent adduct. C. The Schiff base intermediate, KAC41, ligates the unique iron of auxiliary cluster of TYW1 in a bidentate fashion through interactions with the Nz atom (blue) and a carboxylic oxygen (red). D. SAM ligating the unique iron of the SAM cluster in MoaA (PDB 1TV8), utilizing the amino nitrogen (blue) and a carboxylic oxygen (red). All residues are shown as sticks.
Figure II.3. TYW1 active site structure and models. A. Binding pocket of KAC41 is shown, with interactions represented as dashed lines. B. SAM molecule (slate) modeled in the active site with dashed lines indicating potential hydrogen bonds. Positions of classic SAM radical motifs are labeled. C. m¹G (teal) modeled into the active site with potential interacting residues shown as sticks. D. The m¹G (teal) modeled structure from panel C is shown in a different orientation with the dashed lines highlighting the distances between the sites of hydrogen atom abstraction.
Figure II.4. SAM binding motifs of MoaA and TYW1. **A.** MoaA (PDB 1TV8). **B.** Model of SAM binding to TYW1 using MoaA as a guide. Residues putatively involved in positioning of SAM are represented in sticks and colored based on their role/motif. The β6 motif (green), the GXIXGXXE motif (maroon) and the CX3CXΦC motif (light pink) have been previously shown to position the adenine moiety of SAM, whereas the ribose motif (tan) orients the ribose rings. The GGE motif (teal) and residues indicated in purple orient the amino group and the carboxylic group of SAM respectively, to ensure proper ligation to the radical SAM cluster, shown as balls-and-sticks (iron atoms in orange and sulfur atoms in yellow).
Figure 11.5. Electrostatic surface of TYW1. The surface of TYW1 is colored based on its electrostatic properties with positively charged areas in blue and negatively charged areas in red. A path of positive electrostatic surface leads from the active site to a putative binding site for tRNA\textsuperscript{Phe}. SAM (C in slate) and KAC41 (C in light blue) are shown in sticks. A model of a 20 base pair tRNA anticodon stem loop, (PDB 1EHZ in orange) is shown against this surface. Site of tRNA base modification, m'G37 (C in teal), is shown as sticks in the active site.
Figure II.6. Activity of TYW1 with RNA. The extracted ion chromatogram from m/z 322.1 to 322.12 showing that imG-14 forms in the presence of wildtype TYW1 and the C195S variant, but not with K41A or when TYW1 is not included in the assay.
Figure 11.7. Extracted ion chromatograms of all observable tryptic peptides bearing Lys41.

The extracted ion chromatograms at m/z 584.25 (NCYK), 656.27 (NCKK), 799.37 NCKSK), and 871.39 (NCKKS) ± 0.01.
Figure 11.8. Extracted ion chromatogram of the Schiff base intermediate. The extracted ion chromatogram at m/z 874 and 871 showing the tryptic digest fragment containing the trapped Schiff base-lysine adduct when labelled and unlabelled pyruvate are used. The peak corresponding to isotopically enriched pyruvate adduct of the peptide at 7 mins is only observed in the presence of labelled substrate.
Figure II.9. The mass spectrum of a tryptic fragment of TYW1. TYW1 incubated with; A. 1,2,3-$^{13}$C$_3$ pyruvate in the presence of NaCNBH$_3$, B. 1,2,3-$^{13}$C$_3$ pyruvate in the presence of NaCNBH$_3$, C. no pyruvate in the presence of NaCNBH$_3$, and D. no pyruvate in the presence of NaCNBH$_3$. 
Figure II.10. Mass spectrum of the Schiff base intermediate. The mass spectrum of a tryptic fragment of TYW1 incubated with 1,2,3-$^{13}$C$_3$ pyruvate in the absence of SAM, dithionite, or tRNA in the presence of either NaCNBH$_3$ or NaCNB$_2^3$H$_3$. 

NaCNBH$_3$

871.3987

875.4145

871.3981

871.3984

NaCNB$_2^3$H$_3$

874.4088

875.4148

871.3977 871.3984

874.4093

m/z

m/z

874.4093

875.4148

871.3984

875.4139

871.3984

875.4139

871.3977

875.4139

871.3975

m/z

m/z

m/z

m/z
Scheme II.3. Mechanism 1 revised to highlight structural features of TYW1 active site.
Figure II.11. Double conformations of pyruvate-lysine adduct in TYW1. A. Two conformations of the pyruvate-lysine adduct, KAC41, are shown modeled into 2Fo-Fc electron density (blue) contoured at 1 σ. B. The same two conformations of KAC41 as in A are shown here modeled into 2Fo-Fc omit map density (cyan) contoured at 1 σ. The major conformation of the Schiff base is planar (gray), is contained in ~60% molecules in the crystal, and is observed to ligate the auxiliary cluster in a bidentate fashion. The minor conformation is not planar (salmon), is found in the remaining ~40% of the molecules, and is observed with its carboxyl group rotated away from the cluster. C. Close-up of the two conformations of KAC41, showing that the carboxylate of the minor KAC41 conformation (salmon) is rotated 90° from the position of the carboxylate of the major KAC41 conformation (gray). D. Slight rearrangements of side chains are observed in response to the multiple conformations of KAC41. For example, Arg221, which is thought to be involved in binding of both m'G37 (gray) and SAM (not shown), flips from its major conformation (gray) to an orientation that provides hydrogen bonds (~2.4 and 2.5 Å) to an oxygen of KAC41 in its minor conformation (salmon). Radical addition to C2 of the pyruvate moiety by m'G37 is expected when the adduct is in the dominant planar conformation. The hydrogen bonds and ligation of Schiff base to the cluster is shown in black dashed lines. The purple dashed lines represent the hydrogen bonding-network formed by residues in the binding pocket.
<table>
<thead>
<tr>
<th>Table II.1. TYW1 Data and Refinement Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TYW1</strong></td>
</tr>
<tr>
<td>Beamline</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Cell dimensions (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>Unique reflections</td>
</tr>
<tr>
<td>Completeness</td>
</tr>
<tr>
<td>Redundancy</td>
</tr>
<tr>
<td>(R_{sym})</td>
</tr>
<tr>
<td>(CC_{1/2})</td>
</tr>
<tr>
<td>(l/\sigma(l))</td>
</tr>
</tbody>
</table>

**Model Refinement**

| Resolution limits (Å)                         | 48.3-1.64 (1.67-1.64)                    |
| \(R_{work}/R_{free}\)                        | 0.1760/0.2024                            |
| Reflections                                   | 47,297                                   |
| No molecules in asu                           | 1                                        |
| No atoms                                      |                                           |
| Protein                                       | 2586                                     |
| Iron sulfur clusters                          | 13                                       |
| Adduct                                        | 28                                       |
| Water                                         | 210                                      |
| \(B\)-factors (Å\(^2\))                      |                                          |
| Protein                                       | 30.4                                     |
| Iron sulfur clusters                          | 31.4                                     |
| Adduct                                        | 39.8                                     |
| Water                                         | 47.9                                     |
| R.M.S. deviations                             |                                          |
| Bond Lengths, (Å)                             | 0.007                                    |
| Bond Angles (°)                               | 0.889                                    |
| Rotamer outliers (%)                          | 1.45                                     |
| Ramachandran Plot (%)                         |                                          |
| Most Favored                                  | 96.9                                     |
| Additionally allowed                          | 3.1                                      |
| Disallowed                                    | 0.0                                      |

Highest-resolution shell is shown in parentheses

\(^1\)Advanced Photon Source, Argonne National Laboratory set.

\(^5\) 5% reflections used for test set
References


(9) Motorin, Y.; Helm, M. Biochemistry 2010, 49 (24), 4934.

(10) Lorenz, C.; Lünse, C.; Mörì, M. Biomolecules 2017, 7 (2), 35.


(32) Young, A. P.; Bandarian, V. Biochemistry 2015, 54 (23), 3569.
(33) Young, A. P.; Bandarian, V. Biochemistry 2011, 50 (49), 10573.
(67) Dowling, D. P.; Bruender, N. A.; Young, A. P.; McCarty, R. M.; Bandarian, V.; Drennan, C. L. *Nature Chemical Biology* 2013, 10 (2), 106.
(69) The PyMol Molecular Graphics System, Version 2.0 Schrödinger, LLC.
(84) Grell, T. A. J.; Goldman, P. J.; Drennan, C. L. J. Biol. Chem. 2015, 290 (7), 3964.
Chapter III

7-Carboxy-7-deazaguanine synthase – A radical S-adenosyl-L-methionine enzyme with polar tendencies

Chapter Summary

Radical S-adenosyl-L-methionine (AdoMet) enzymes are widely distributed and catalyze diverse reactions. AdoMet binds to the unique iron atom of a site-differentiated [4Fe-4S] cluster and is reductively cleaved to generate a 5'-deoxyadenosyl radical, which initiates turnover. 7-Carboxy-7-deazaguanine (CDG) synthase catalyzes a key step in the biosynthesis of 7-deazapurine containing natural products. 6-Carboxypterin (6-CP), an oxidized analog of the natural substrate 6-carboxy-5,6,7,8-tetrahydropterin (CPH₄), is shown to be an alternate substrate for CDG synthase. Under reducing conditions that would promote the reductive cleavage of AdoMet, 6-CP is turned over to 6-deoxyadenosylpterin, presumably by radical addition of the 5'-deoxyadenosine followed by oxidative decarboxylation to the product. By contrast, in the absence of the strong reductant, dithionite, the carboxylate of 6-CP is esterified to generate 6-carboxypterin-5'-deoxyadenosyl ester. Structural studies with 6-CP and AdoMet also reveal electron density consistent with the ester product being formed in crystallo. The differential reactivity of 6-CP under reducing and non-reducing conditions highlights the ability of AdoMet radical enzymes to carry out both polar and radical transformations in the same active site.

Contributions

Nathan Bruender purified BsQueE and performed the biochemical analysis. Daniel Dowling crystallized BsQueE. Daniel Dowling and Tsehai Grell solved the structure of BsQueE. The manuscript was written with Nathan Bruender, Daniel Dowling, Vahe Bandarian and Catherine L. Drennan.
Introduction

The radical S-adenosyl-L-methionine (AdoMet) superfamily is a group of enzymes that harness the reductive cleavage of AdoMet to carry out complex radical-mediated transformations. The superfamily was initially identified on the basis of a conserved CxxxCxxC motif, which binds a site-differentiated [4Fe-4S] cluster whereby the three cysteine-thiolates coordinate the cluster. The fourth iron interacts with the α-amino and α-carboxylate of AdoMet. To date, with only a few notable exceptions, the mechanisms of action for all AdoMet radical enzymes that have been proposed involve radical-mediated transformations that are initiated by the 5′-deoxyadenosyl radical (5′-dAdo•), which is generated from the reductive cleavage of AdoMet (Figure III.1A).

The exception to this is the enzyme Dph2 involved in diphthamide biosynthesis, which generates a 3-amino-3-carboxypropyl radical. The RNA methylases RlmN and Cfr consume two equivalents of AdoMet catalyzing both polar and radical-mediated group transfer reactions in the same catalytic cycle. The first equivalent of AdoMet methylates an active site Cys releasing S-adenosylhomocysteine. Binding and reductive cleavage of the second equivalent of AdoMet initiates the radical-mediated transfer of the methyl group from the active site methyl-Cys residue to either the C2- or C8-position of A2503 in 23S rRNA. Finally, the cobalamin-dependent AdoMet radical enzyme TsrM catalyzes the methylation at C8 of the indole on tryptophan using an AdoMet-derived methyl group. The mechanism of this reaction remains to be established, but formation of 5′-deoxyadenosine (5′-dAdoH) is not observed in this transformation.

Recent bioinformatic analysis of the AdoMet radical superfamily has revealed >113,000 homologs in genome sequences, nearly all of which are enzymes that have been placed in this superfamily on the basis of the presence of the conserved CxxxCxxC sequence. However, the ability of RlmN, Cfr, or TsrM to carry out both polar and/or radical transformations suggests that the conserved sequence motif may not ideally describe the expected range of reactivity. Perhaps some of the putative AdoMet radical proteins could utilize the activated sulfonium in AdoMet to catalyze polar group transfer chemistry. If true, it is likely that non-reductive group transfer chemistry may be an as yet uncharacterized promiscuous activity in AdoMet radical enzymes.

7-Carboxy-7-deazaguanine (CDG) synthase (QueE) is a member of the AdoMet radical superfamily that catalyzes the radical-mediated ring rearrangement required to convert 6-carboxy-5,6,7,8-tetrahydropterin (CPH₄) into CDG, which is the precursor to all
pyrrolopyrimidine metabolites observed in nature. Prior investigations have shown that QueE contains one [4Fe-4S] cluster that mediates the reductive cleavage of AdoMet (Figure III.1A). Reducing equivalents can be supplied in vitro from NADPH via ferredoxin(flavodoxin):NADP+ reductase and flavodoxin or via the chemical reductant dithionite. Once formed, the 5'-dAdo* abstracts the C6 hydrogen atom from the substrate to initiate the ring rearrangement and subsequent elimination of N5 from the pterin.

Recent X-ray crystal structures of the QueE homolog from Burkholderia multivorans (BmQueE) in complex with AdoMet and substrate (CPH₄), product (CDG), or the substrate analog 6-carboxypterin (6-CP), have provided invaluable snapshots of the active site of the protein (Figure. III.1B, Figure III.2). The overall structure of the BmQueE is similar to other AdoMet radical enzymes, except for an 11-residue insertion in the AdoMet radical cysteine motif that forms a 3₁₀-helix at the surface of the protein. The CPH₄-bound structure of QueE shows that the 5'-position of AdoMet is within 3.9 Å of C-6 of the substrate, which biochemical studies have shown to be the site of the H-atom abstracted by the 5'-dAdo* to initiate catalysis (Figure III.1B). These observations are complemented by spectroscopic studies in other systems that suggest the role of AdoMet radical enzymes is to generate and shield/protect radical intermediates from off-pathway reactions. The enzymes act as a scaffold to protect the generated radical species by shielding it from solvent and by providing a framework to prevent the radical intermediate from moving great distances by either tightly binding the intermediate or positioning the necessary reacting partners within van der Waals (VDW) distances of each other, favoring on-pathway reactions.

6-CP is an oxidized analog of the QueE substrate. The structure of 6-CP complexed with QueE revealed that the 5'-carbon of the deoxyadenosine moiety of AdoMet is 3.2 Å away from one of the carboxylate oxygen atoms and 4.9 Å away from the sp² hybridized C-6 of 6-CP (Figure III.1C). This close proximity between the substrate analog and cofactor suggests that the enzyme may instead catalyze a group transfer reaction in cases where an H-atom is not present.

In our continuing structure/function studies of QueE, we made a serendipitous discovery in the structure of the Bacillus subtilis QueE (BsQueE) homolog crystallized in the presence of
AdoMet and 6-CP. The electron density in the active site did not resemble that of 6-CP, but rather, the substrate analog appeared to have undergone a modification that involved attachment to the 5'-dAdo of AdoMet. Therefore, we initiated studies to determine if 6-CP is an alternative substrate. Herein, we show that 6-CP is indeed a substrate for the QueE homologs from both *B. subtilis* and *B. multivorans*. However, unlike the H-atom abstraction catalyzed with CPH₄, 6-CP undergoes two distinct catalytic outcomes. One of these leverages the ability of the enzyme to bind and reductively cleave AdoMet, but instead of H-atom abstraction, leading to a radical addition. In the other, by contrast, QueE binds AdoMet and facilitates group transfer in the absence of dithionite to form the 6-carboxypterin-5'-deoxyadenosyl (6-CP-dAdo) ester covalent adduct observed in the crystal structure. These findings have implications in the functional role of uncharacterized enzymes in the AdoMet radical superfamily.

**RESULTS**

*X-ray structure of BsQueE with a 6-CP–dAdo ester covalent adduct*

An initial X-ray crystal structure of dimeric His₆ tagged *BsQueE* was determined to 2.55 Å resolution by Multiwavelength Anomalous Dispersion (MAD) phasing, which was then used to phase a 2.4 Å resolution structure of dimeric *BsQueE* that contains a 6-CP–dAdo ester covalent adduct (*Figure III.3, Table III.1*). Similar to the *BmQueE* structure, the monomeric subunit of *BsQueE* is composed of a partial (β/α)₆ TIM barrel fold, which is characteristic of AdoMet radical enzymes.¹⁷ ¹⁸ The active site is located within the lateral opening of the partial barrel (*Figure III.3A*), and is flanked by N- and C- terminal regions, the latter of which protects the active site from solvent.

Although we crystallized *BsQueE* with AdoMet and 6-CP, the electron density was not consistent with intact versions of these two molecules being bound in the active site. When an intact AdoMet was refined, negative difference density appeared for the C5'-S bond of AdoMet, suggesting that AdoMet has undergone cleavage to form methionine and 5'-dAdo (*Figure III.4A*). Additionally, positive difference density appeared between the C5' position of AdoMet and a carboxylate oxygen of 6-CP, indicating that a new ester bond may have been formed between these two molecules, requiring a rotation of about 30° around the ribose to bring the 5'-
dAdo moiety toward 6-CP (Figure III.4B). The resulting adduct can be described as a 5'-deoxyadenosyl ester of 6-CP (6-CP-dAdo) and is an excellent fit to omit map density (Figure III.3B). Methionine, the other cleavage product of AdoMet, is also observed bound as expected for a AdoMet radical protein, i.e. coordinating the unique iron of the [4Fe-4S] cluster through its α-carboxyl and α-amino groups (Figure III.3B).

The newly formed 6-CP-dAdo ester adduct spans the AdoMet and 6-CP binding sites on BsQueE with the adenine and pterin rings residing in nearly perpendicular planes relative to one another (Figure III.3B). The enzyme employs known AdoMet binding motifs for interactions with the 5'-dAdo moiety of the newly formed adduct (Figure III.5)21, and despite the 30° rotation of the 5'-dAdo about the ribose required to form the adduct, the interactions between the ribose and the protein are maintained (Fig. Figure III.5). The planar pterin ring of 6-CP moiety of the adduct is also bound similarly to how 6-CP was bound in the BmQueE•6-CP•AdoMet structure.15 It is positioned in the active site by a number of π-π and electrostatic interactions from residues located in the protein core and the N- and C-terminal extensions (Fig. Figure III.3C). The pterin ring moiety of the 6-CP–dAdo ester adduct stacks with the Phe28 and His233 residues from the N- and C-terminal extension and is further held in place by various hydrogen bonds, including hydrogen bonds to the N3 and exocyclic amine by the carboxyl group of the C-terminal residue, Val243. One new hydrogen bond between the protein residue Gln16 and the pterin ring, positioning N8 of the pterin, was identified in this structure containing the covalent adduct (Figure III.3C). Although this glutamine residue is conserved in BmQueE, it is not in the proper orientation or within sufficient distance (5.9 Å from N8 of the pterin) to interact with 6-CP in that structure.15 Instead, a water molecule is observed to bridge an interaction between this glutamine and the pterin N8 atom.

The mechanism by which 6-CP–dAdo ester forms in the active site of QueE cannot be gleaned from the structures alone. Although it has not been demonstrated that cluster reduction by the X-ray beam can lead to the reductive cleavage of AdoMet, it has not been ruled out either. In this scenario, formation of the 6-CP adduct could occur by a radical addition route. However, it is also possible that the close juxtaposition of AdoMet to 6-CP when both are bound to the enzyme
could allow for the carboxylate oxygen of 6-CP to attack the 5’-position of the deoxyadenosine, displacing methionine to form the 6-CP–dAdo ester adduct.

**6-CP is an alternative substrate for QueE**

To probe whether the 6-CP–dAdo ester observed in the crystal structure could have arisen from a reaction related to the radical-mediated ring contraction observed with CPH₄, we set up assays with 6-CP in the presence of AdoMet and BsQueE under reducing conditions using dithionite. The reactions were quenched with acid and analyzed via LC-MS. A new product was observed to elute at ca. 39 min when BsQueE, 6-CP, and AdoMet were present in the reaction. The appearance of this new product was dependent on the presence of enzyme, 6-CP, and AdoMet (Figure III.6A). The rate of formation of this species, measured in a separate experiment, is ~0.0013± 0.0002 min⁻¹, which is 150-fold slower than that for formation of CDG from CPH₄.

Inspection of the UV-visible-spectrum of the new product eluting at 39 min revealed features at 350 nm reminiscent of 6-CP, as well as a substantial peak at 260 nm (Figure III.6B). The simplest explanation for this observation is the presence of adenine originating from AdoMet, since the formation of this product is AdoMet dependent (Figure III.6Ac). Therefore, we examined if the new product represents a covalent adduct between 5’-dAdo and 6-CP by mass spectrometry. The theoretical m/z of such a product is expected to be 457.1333 ([M+H]⁺). However, the mass spectrum of the new peak eluting at 39 min exhibits a m/z of 413.1424, which is 43.9909 amu lighter than a simple adduct between 6-CP and 5’-dAdo (Figure III.6Ca). The loss of CO₂ (43.9898 amu) from a 5’-dAdo-6-CP covalent adduct would yield a m/z of 413.1435, which is within 2.7 ppm of the observed mass.

The identity of the new product was explored using isotopically enriched AdoMet and 6-CP. AdoMet was synthesized enzymatically from unlabeled methionine and [U-¹³C₁₀] ATP to label the 10 carbons in 5’-dAdo moiety of AdoMet. [U-¹³C₇] 6-CP was synthesized by permanganate oxidation of [U-¹³C₇] CPH₄, which itself was obtained by the successive actions of GTP cyclohydrolase I and CPH₄ synthase on [U-¹³C₁₀] GTP. Mass spectrometry revealed the [¹³C₁₀] AdoMet and [U-¹³C₇] 6-CP to be >87% and >95% enriched, respectively (Figure III.7 and III.8). Isotopically enriched AdoMet and 6-CP analogs were incubated with QueE and the mass spectra of the resulting products were compared to reactions containing natural abundance 6-CP.
and natural abundance AdoMet (Figure III.6C). When 6-CP is substituted with [U-\(^{13}\)C\(_7\)] 6-CP, the m/z for the resulting product shifts from 413.1424 to 419.1625, consistent with the retention of six of the seven possible \(^{13}\)C enriched carbons from 6-CP (theoretical m/z 419.1636, 2.6 ppm error) (Figure III.6Cb). By contrast, with \([^{13}\text{C}\_10\text{-dAdo}]\) AdoMet the m/z shifts from 413.1424 to 423.1760 (Figure III.6Cc). The observed 10 amu shift (theoretical m/z 423.1770, 2.4 ppm error) indicates that 5'-dAdo from AdoMet is incorporated in the new product. This finding is consistent with the increase at 260 nm for the product (Figure III.6B). When both \([^{13}\text{C}\_10\text{-dAdo}]\) AdoMet and [U-\(^{13}\)C\(_7\)]-6-CP are incubated with QueE, the mass shifts by 16 amu from m/z 413.1424 to 429.1952 (theoretical m/z of 429.1971, 4.4 ppm error) indicating that the new species was derived from both 6-CP (minus one carbon) and the 5'-dAdo moiety of AdoMet (Figure III.6Cd).

Mass spectral fragmentation of the products from the reaction mixtures with natural abundance and isotopically enriched substrates were carried out to further probe the structure of the adduct. Three fragments are observed by MS/MS of the product ([M+H]^+ m/z of 413.1424) in reactions containing natural abundance AdoMet and 6-CP, which are consistent with adenine (m/z: observed, 136.0616; theoretical, 136.0623), dehydrated decarboxylated pterin-ribose (m/z: observed, 260.0776; theoretical, 260.0784), and decarboxylated pterin-ribose (m/z: observed, 278.0881; theoretical, 278.0890) (see Figure III.6Da, E). MS/MS analysis of the corresponding [M+H]^+ peaks in reactions containing isotopically enriched substrates exhibited the expected isotopic shifts and were readily assigned to the fragments observed with the natural abundance substrate, with exception of expected isotopic enrichment due to incorporation of \(^{13}\)C from 6-CP (Figure III.6Db), AdoMet (Figure III.6Dc), or both (Figure III.6Dd). The fragmentation patterns observed with the isotopically enriched product are readily mapped onto that obtained with the natural abundance product. These data are consistent with the new product being 6-deoxyadenosylpterin (6-dAP) (Figure III.6E). However, this product is not consistent with the adduct observed in the structure of BsQueE solved in the presence of AdoMet and 6-CP.

QueE catalyzes two different 6-CP products depending on the presence of reductant

Formation of 6-dAP requires incubating BsQueE, AdoMet, and 6-CP in the presence of dithionite. As a control, to determine if the product with m/z of 413.1424 was dependent on the
presence of reductant, BsQueE was incubated with 6-CP in the absence of dithionite. Surprisingly, a new peak appeared at ca. 42 min in the chromatogram of the reaction where only dithionite was omitted (Figure III.9b), which is distinct from the peak at 39 min observed in the presence of dithionite (Figure III.9d). As with the species at 39 min, the appearance of the 42 min species requires the presence of BsQueE (Figure III.9a,c). The formation of this new species requires BsQueE, AdoMet, and 6-CP (Figure III.10Aa), and it is not observed when any of these components are omitted (Figure III.10Ac-e). As with the 6-dAP product, this new product exhibits UV-visible spectral features consistent with both 6-CP and adenine, but its retention time of 42 min is significantly different suggesting that it is not identical to 6-dAP (Figure III.10B). The rate of formation of this product ($5.6 \pm 0.9 \times 10^{-5} \text{ min}^{-1}$) is 23- and 3600-fold slower than formation of 6-dAP and CDG, respectively. Although the strong reductant dithionite was omitted from these reactions, they were carried out under reducing conditions in the presence of 10 mM DTT in an anaerobic chamber. Control experiments, however, show that the new product at 42 min is observed even in the absence of DTT (Figure III.10Ab).

We examined the identity of the new product via mass spectrometry. The species that elutes at ca. 42 min exhibits a $m/z$ of 457.1320 ([M+H]$^+$) (Figure III.10Ca). This new product is most consistent with the 6-CP–dAdo ester adduct (theoretical $m/z = 457.1333$) observed in the crystal structure. The identity of the product was further probed with natural abundance and $^{13}$C enriched AdoMet and 6-CP. In contrast to 6-dAP, when the reactions were carried out in the absence of dithionite with [U-$^{13}$C$_7$]-6-CP the mass of the 42 min peak shifts from 457.1320 to 464.1560 (theoretical $m/z$ 464.1567, 1.5 ppm error) (Figure III.10Cb), which is consistent with retention of all the carbons of 6-CP. Replacing natural abundance AdoMet with $[^{13}$C$_{10}$-dAdo] AdoMet results in a 10 amu shift in the $m/z$ for [M+H]$^+$ from 457.1320 to 467.1654 (theoretical $m/z$ 467.1668, 3.0 ppm error) (Figure III.10Cc). Finally, when both isotopically enriched substrates are incubated with BsQueE the [M+H]$^+$ mass shifts by 17 amu from $m/z$ 457.1320 to 474.1906 (theoretical $m/z$ 474.1902, 0.8 ppm error) (Figure III.10Cd).

To probe the structure of the new product, the [M+H]$^+$ peak from each reaction mixture was subjected to MS/MS fragmentation. The fragmentation spectrum obtained from the natural abundance product is shown in (Figure III.10Da). The observed fragments are consistent with
adenine (m/z: observed, 136.0614; theoretical, 136.0623), 6-CP (m/z: observed, 208.0461; theoretical, 208.0471), dehydrated 6-CP-ribose (m/z: observed, 304.0669; theoretical, 304.0682), 6-CP-ribose (m/z: observed, 322.0774; theoretical, 322.0788), 6-CP-deoxyadenosine that has lost the equivalent of two water molecules (m/z: observed, 421.1106; theoretical, 421.1122), and dehydrated 6-CP-deoxyadenosine (m/z: observed, 439.1211; theoretical, 439.1227) (Figure III.10Da,7E). MS/MS spectra of the corresponding [M+H]⁺ peaks in reactions containing isotopically enriched substrates exhibit the expected isotopic shifts and are readily assigned to the fragments observed with the natural abundance substrate, with exception of expected isotopic enrichment due to incorporation of ¹³C from 6-CP (Figure III.10Db), AdoMet (Figure III.10Dc), or both (Figure III.10Dd). The fragmentation patterns observed with the isotopically enriched product are readily mapped onto that obtained with the natural abundance product, further confirming the structural assignment of 6-CP–dAdo ester (Figure III.10E). Therefore, the product observed in the aforementioned BsQueE crystal structure forms in the absence of reductant by taking advantage of the inherent reactivity of AdoMet towards nucleophilic attack.

**Mg²⁺ dependence of alternative activities**

Previous structural and functional investigations of BsQueE and BmQueE revealed an unexpected requirement for Mg²⁺ in the conversion of CPH₄ to CDG. Structures of BmQueE complexed with substrate (CPH₄) and product (CDG) show that the Mg²⁺ provides key contacts with the substrate and product, thus anchoring the molecules in the active site. Interestingly, Mg²⁺ is not observed in the BsQueE structure of the 6-CP–dAdo ester adduct described above and the BmQueE complexed with 6-CP.¹⁵ Therefore we sought to determine if the divalent cation was required for the formation of either 6-dAP or 6-CP–dAdo ester. We observed no significant change in the rates of formation for 6-dAP or 6-CP–dAdo ester when Mg²⁺ was omitted from the reaction or when it was added to a final concentration of 2 mM. However, it is possible that Mg²⁺ was present in the reaction mixture as a contaminant from the purified protein, AdoMet, or 6-CP. Therefore, we subjected an aliquot of each reaction mixture to ICP-MS to analyze for the presence of Mg²⁺. The concentration of Mg²⁺ in each reaction mixture was not detectable above the ~1μM background in the trace-metals grade nitric acid used to prepare the sample. This shows that unlike the wild-type function of QueE, the formation of either 6–CP product is not enhanced by the presence of Mg²⁺.
B. multivorans QueE catalyzes the formation of both 6-dAP and 6-CP–dAdo ester

The structural investigation of B. multivorans QueE revealed that this homolog was capable of binding 6-CP in the active site in a similar fashion to that of the B. subtilis QueE. The two homologs share 21% sequence identity and 31% sequence similarity. Therefore, we sought to determine if 6-CP was an alternative substrate for the QueE homolog from B. multivorans to determine if the divergent outcomes of turnover with 6-CP in the presence or absence of reductant is an intrinsic property of the enzyme and not specific to the B. subtilis protein. Indeed, when BmQueE was incubated with 6-CP in the presence of reductant, a new product was observed in the extracted ion chromatogram (Figure III.11A). Based on both the retention time (ca. 37 min) and MS data (m/z of 413.1425) the species is 6-dAP (Fig 8B).

In parallel to the BsQueE, in the absence of reductant, we observe a peak in the extracted ion chromatogram of the reaction at 42 min (Figure III.11Ca), which requires the enzyme Figure III.11Cb). The m/z of the species eluting at 42 min, 457.1322, is identical to 6-CP-dAdo ester (Figure III.11D). Therefore, both BsQueE and BmQueE catalyze identical polar and radical mediated transformations utilizing 6-CP as an alternate substrate.

Discussion

Since the unification of the AdoMet radical superfamily in 2001, the number of members has expanded from ~600 to >113,000. All members of the AdoMet radical superfamily characterized to date have been shown to utilize a [4Fe-4S] cluster, which is typically coordinated by the CxxxCxxC motif, to mediate the reductive cleavage of AdoMet commonly affording the dAdo• (Figure III.1A). The dAdo• abstracts a H-atom from substrate to initiate a myriad of radical-mediated transformations.

While under some conditions, alternate reactivity with substrate analogs have been noted leading to novel products, to date, with two notable exceptions, AdoMet participates by radical mechanism. Two members of the superfamily (RlmN and Cfr) are unique in that they utilized AdoMet for both polar and radical transformations. The homologous enzymes RlmN and Cfr catalyze the methylation of C-2 or C-8 of the adenine ring of A2503 in 23S rRNA. Similarly,
TsrM catalyzes the methylation of tryptophan in an as yet undetermined mechanism that does not appear to entail reductive cleavage of AdoMet.\textsuperscript{8-10} The structure of RlmN reveals that as with other AdoMet radical enzymes it adopts the partial TIM barrel (\(\beta/\alpha\)\textsubscript{6}) fold and the three cysteine thiolates of the CxxxCxxC motif coordinate three iron atoms of a catalytically essential [4Fe-4S] cluster.\textsuperscript{26} The fourth iron of the cluster is coordinated by the \(\alpha\)-amino and \(\alpha\)-carboxylate moieties of AdoMet. Functionally, both RlmN and Cfr reductively cleave AdoMet, as expected for a member of the AdoMet radical superfamily. However, what sets these two enzymes apart from the rest of the superfamily is that they catalyze methyl transfer from AdoMet coordinated to the [4Fe-4S] cluster to an active site Cys residue using a polar substitution mechanism, in addition to reductively cleaving a second equivalent of AdoMet at the [4Fe-4S] cluster to mediate the methyl transfer from the methyl-Cys to the \(A_{2503}\).\textsuperscript{6-7} RlmN and Cfr are examples of AdoMet radical enzymes that use AdoMet in both a polar and radical capacity in the same catalytic cycle. The involvement of AdoMet in polar group transfer is a common reaction that occurs in a variety of biological processes including gene regulation and metabolite biosynthesis. However, enzymes that utilize AdoMet exclusively as a methyl donor are structurally distinct from the AdoMet radical superfamily, adopting an \(\alpha\beta\alpha\)-sandwich fold that is reminiscent of the Rossmann fold.\textsuperscript{27}

The physiological role of QueE is to harness 5'-dAdo• from reductively cleaved AdoMet to catalyze the radical mediated ring rearrangement in the conversion of CPH\textsubscript{4} to CDG.\textsuperscript{14} This enzyme has been extensively characterized both structurally and functionally.\textsuperscript{14-16} Structures of \textit{Bm}QueE show that the enzyme is able to bind an oxidized analog of the substrate, 6-CP, in a manner similar to that of the substrate (CPH\textsubscript{4}) and product (CDG) (\textbf{Figure III.2}).\textsuperscript{15} We initiated the biochemical studies in this paper when we observed an unusual product, 6-CP–dAdo ester, in the X-ray crystal structure of \textit{Bs}QueE•6-CP• AdoMet complex. We are able to produce the 6-CP–dAdo ester \textit{in vitro} when QueE is incubated with 6-CP and AdoMet in the absence of dithionite. Under reducing conditions, an alternate product, 6-dAP was observed and characterized. This product is analogous to that observed for the AdoMet radical enzymes, MqnE, in the futalosine biosynthetic pathway, and HydE, in the assembly the [FeFe]-hydrogenase active site.\textsuperscript{28-29} Additionally, Knappe and coworkers observed that pyruvate
formate-lyase-activating enzymes catalyzed a similar addition of 5'-dAdo• to the olefinic β carbon of a dehydroalanine residue in a dehydroalanine-containing octapeptide.\textsuperscript{30} In the presence of the strong reductant dithionite, QueE generates the 5'-dAdo•, which subsequently reacts with 6-CP via a radical addition reaction leading to the decarboxylation of 6-CP and the formation of 6-dAP (Scheme III.1, right). Although 6-CP is not the natural substrate, this reaction is not surprising as the current understanding of AdoMet radical enzymes is that they generate radical intermediates and stabilize them by providing scaffolds that minimize off-pathway reactions.\textsuperscript{19-20} One mechanism to favor on-pathway reactions is to properly orient the substrate within VDW distances of the 5'-dAdo•.

6-CP was modified by AdoMet in the absence of reductant, but instead of a radical addition, we observed formation of an ester linkage between 6-CP and the 5'-deoxyadenosine moiety of AdoMet. To explain this, we propose a polar substitution mechanism (Scheme III.1, left) analogous to that observed in methyltransferases and the AdoMet radical enzymes RlmN and Cfr.\textsuperscript{6-7} Close inspection of the BmQueE structure with 6-CP bound shows that one of the carboxylate oxygen atoms is only 3.2 Å from the C5' of AdoMet, where it could participate in nucleophilic substitution (Figure III.1C). The BsQueE structure further shows that a carboxylate oxygen is within 2.9 Å of R30, a conserved residue capable of stabilizing the deprotonated form of the 6-CP carboxylate and potentially activating it for nucleophilic attack on C5' of AdoMet (Figure III.3C). The resulting adduct characterized in this study provides evidence of now a third AdoMet radical enzyme that is capable of using AdoMet coordinated to a [4Fe-4S] cluster for polar group transfer. However, what sets this apart from RlmN and Cfr is the fact that 6-CP is not a natural substrate for QueE, therefore the formation of 6-CP–dAdo ester and 6-dAP are promiscuous activities for this AdoMet radical enzyme. Unlike the radical mediated ring contraction reaction of CPH\textsubscript{4} to form CDG, the polar and radical additions of 5-dAdo to 6-CP do not appear to require magnesium. This finding may suggest that magnesium is involved in minimizing off-pathway reactivity with the substrate. Certainly, the binding of the carboxylate of CPH\textsubscript{4} to the magnesium may block formation of the ester adduct. Additional studies with alternate substrate analogs should aid in delineating the role(s) of the active site residues and the magnesium ion in directing the promiscuous activities of QueE.
The current work in the AdoMet radical superfamily has provided overwhelming evidence that these enzymes use the AdoMet bound [4Fe-4S] cluster to reductively cleave SAM to initiate radical-mediated reactions. However, the recent observations of RlmN, Cfr, and now QueE utilizing the same fold to catalyze polar group transfer from AdoMet questions the paradigm that all proteins containing the CxxxCxxC motif reductively cleave AdoMet. The sulfonium of AdoMet activates it to transfer 5'-dAdo, methyl, or 3-amino-3-carboxypropyl moieties to any properly positioned acceptor molecule by nucleophilic displacement. Booker and colleagues have provided a beautiful example of polar methyl group transfer by a cluster-bound AdoMet.

The studies presented here are the first to demonstrate the polar transfer of 5'-dAdo. Our prediction is that as more presumed AdoMet radical enzymes are studied, that additional polar group transfer from AdoMet will be discovered.

Conclusions
The original report by Sofia and coworkers nearly 20 years ago where the AdoMet radical superfamily was identified by the presence of the CxxxCxxC motif occurred before the explosion of genome sequences. Moreover, in the intervening decades, studies from several laboratories have uncovered a surprising range of reactivity. The studies on polar methyl group transfer in RlmN and Cfr, the methylation catalyzed by TsrM, and this manuscript show that, at least in principle, moieties attached to the sulfonium of AdoMet can be transferred by a polar route in AdoMet radical enzymes. To be sure, the formation of 6-CP-dAdo ester is a promiscuous activity in a protein that is designed to do an entirely different type of transformation. However, it is now generally accepted that new activities emerge in enzymes by optimization of low-level reactions. We propose that with >113,000 annotated AdoMet radical enzymes that it is only a matter of time before enzymes whose sole function is not methyl group transfer or radical chemistry will emerge.

Experimental Procedures

Cloning, expression, and protein purification
The genes encoding for His6-tagged BsQueE and BmQueE were cloned and expressed and the proteins used in this investigation were purified as previously described.
Crystallization of BsQueE

All data quality crystals of BsQueE were obtained with a His6-tagged construct through sparse matrix screening using a Mosquito pipetting robot (TPP LabTech) housed within an anaerobic chamber (MBraun) under nitrogen atmosphere ($O_2 < 0.1$ ppm) at 21°C. Enzymatically synthesized and purified AdoMet\(^{32}\) used in crystallography was provided by the Broderick lab at Montana State University and the oxidized substrate 6-CP was obtained from Sigma Aldrich (>98% purity). For harvesting, crystals of BsQueE were transferred to another anaerobic chamber (COY Laboratory products) under 95% argon, 5% hydrogen at 25°C, cryoprotected and then cryo-cooled in liquid nitrogen within this chamber.

Crystallization of BsQueE for initial structure determination

Crystals were grown by sitting drop vapor diffusion using a mixture 0.150 µL of protein solution [5 mg/mL BsQueE, 5 mM AdoMet, 0.05 M PIPES•NaOH (pH 7.4) and 10 mM DTT] and 0.150 µL of reservoir [0.1 M Tris•HCl (pH 8.5), 20% (w/v) PEG 8000 and 0.2 M MgCl\(_2\)] over a 70 µL reservoir. Diffraction quality crystals (Table III.1), with dimensions of 50 µm × 100 µm × 30 µm, grew within a month. Crystals were harvested in the reservoir solution, cryoprotected in 0.05 M Tris•HCl (pH 8.5), 0.1 M MgCl\(_2\), 22% (w/v) PEG 8000, 20% (v/v) glycerol and immediately plunged into liquid nitrogen.

Crystallization of BsQueE with AdoMet and 6-CP for determination of the adduct structure

Sitting drops with crystals contained 0.150 µL protein solution [5 mg/mL BsQueE, 5 mM AdoMet, 5 mM 6-CP, 0.05 M PIPES•NaOH (pH 7.4) and 10 mM DTT] mixed with 0.150 µL of reservoir [0.1 M trisodium citrate (pH 4.2), 1.5 M (NH\(_4\))\(_2\)SO\(_4\)] over a 70 µL reservoir. Prior to setting up the sparse matrix screen, 6-CP dissolved in DMSO was added to the protein, yielding a final concentration of 5% (v/v) DMSO in the protein solution. The mixture was centrifuged to separate any insoluble 6-CP from the protein. Crystals of BsQueE with AdoMet and 6-CP (Table S1) grew at room temperature within a month to approximately 150 µm × 150 µm × 150 µm. Crystals were harvested in reservoir solution, cryoprotected with 0.1 M trisodium citrate (pH 4.2), 1.5 M (NH\(_4\))\(_2\)SO\(_4\) and 20% (v/v) glycerol and immediately flash cooled in liquid nitrogen.
Determination of the initial structure of BsQueE

The initial structure of BsQueE was solved by iron Multiwavelength Anomalous Dispersion (MAD) using data collected at a remote wavelength (0.9999 Å) and the iron peak (1.7389 Å) wavelength to 2.55 and 3.0 Å resolution respectively, on beamline 8.2.1 at the Advanced Light Source (Berkeley, CA). The Fe peak data set was collected using the inverse beam data collection method in 20° wedges and 1° oscillation steps. Both the remote and peak data sets were processed in HKL2000 (Table III.1) and used to find heavy-atom sites in the program SOLVE in phenix AutoSol. Two heavy atom sites were identified with occupancies greater than 1.0, corresponding to two [4Fe-4S] clusters per BsQueE dimer in the asymmetric unit. Heavy atom sites were refined and experimental electron density maps were generated using the program SHARP/autoSHARP. Secondary structure elements were traced into the figure of merit-weighted F₀ maps (FOM of 0.24 to 3.1 Å resolution) in the program COOT using the structure of the radical AdoMet enzyme, pyruvate formate-lyase activating enzyme (PFL-AE [PDB accession code 3CB8]), as a guide. The resulting polyalanine model was used to generate a new mask for solvent flattening in SHARP and the improved 3.1 Å resolution electron density maps allowed for the assignment of ninety percent of the side chains to the initial structure, using the CX₃CX₂C motif and the [4Fe-4S] cluster as initial anchors for the protein sequence.

For crystallographic refinement, data collected at the remote wavelength were converted from intensities to structure factors using the French and Wilson method in the ccp4 suite of programs. Rigid body refinement was followed by deformable elastic network assisted simulated annealing refinement in CNS 1.3 against the remote data set without a sigma cutoff, and using non-crystallographic symmetry (NCS) restraints for all atoms. Iterative rounds of refinement and model building in COOT were performed, followed by gradual extension of the phases to the full 2.55 Å resolution of the remote data set with DM and solvent flattening in SHARP. The [4Fe-4S] cluster and ligating cysteine ligands were refined using restraints from the 0.93 Å published X-ray crystal structure of a high-potential iron protein from Chromatium vinosum (PDB accession code 1B0Y). The amino acid backbone of AdoMet (N, Ca, Cβ, C, and O atoms) has been built and refined in monomer A only. NCS restraints were taken off all and atoms and water molecules were added to the structure. Final rounds of refinement were done in phenix, including TLS parametrization. To verify the structure, composite omit maps were...
generated. The final model contained residues G4-D193 and S201-W237 (of 243), an [4Fe-4S] cluster and AdoMet in monomer A and residues A2-N239 (of 243) and an [4Fe-4S] cluster in monomer B. For both monomers, the N-terminal His₆-tag was disordered and was not observed in the final structure (Table III.1).

_Determination of the BsQueE adduct structure_

A data set of BsQueE with AdoMet and 6-CP was collected at the Advanced Photon Source (Argonne, IL) on beamline 24-IDE at a wavelength of 0.9792 Å, and was indexed, integrated and scaled in HKL200 to 2.4 Å resolution (Table III.1). Phases were determined using the program PHASER with the initial BsQueE structure without ligand, [4Fe-4S] cluster, or water molecules as a search model. A solution was found that contained four QueE molecules in the asymmetric unit, with LLG and TFZ scores of 3599 and 49.1, respectively. The structure was refined using NCS restraints, first in CNS 1.3 against all data (no sigma cutoff) and later in phenix. NCS restraints were loosened over the course of the refinement using $R_{\text{free}}$ as a guide. The refinement protocol included simulated annealing, energy minimization, and B-factor refinement. Parameter files for the [4Fe-4S] cluster, AdoMet, 6-CP, and 6-CP-dAdo ester were generated using the program eLBOW in phenix. Iterative rounds of refinement and model building in COOT were carried out until $R_{\text{factors}}$ converged. Composite omit maps were calculated to verify the final structure. The final model includes residues K3-V243 of 243 residues of the protein chain; the N-terminal expression tag, Met1 and Ala2 have been omitted due to lack of discernable density. Each monomer also contained a [4Fe-4S] cluster, methionine, and the 6-CP-dAdo ester adduct. See (Table III.1) for full data processing, refinement and validation statistics.

_Synthesis of substrates_

AdoMet and $[^{13}C_{10}]$ AdoMet were synthesized enzymatically following the same protocol previously published for AdoMet except 3.5 mg of $[^{U-13}C_{10}]$ ATP (Santa Cruz Biotechnology) was used in place of ATP and the final volume of the reaction was 0.52 mL for the synthesis of $[^{13}C_{10}]-$AdoMet only. After 5 h incubation at room temperature the reaction was quenched by addition of HCl to adjust the pH to 5.0. The reaction was incubated on ice for 30 min then
centrifuged to remove the precipitate. The 0.52 mL reaction was then diluted to 20 mL with 1 mM sodium acetate (pH 5.0) buffer and loaded onto a 1.6 cm × 26.5 cm SP sepharose column equilibrated in 1 mM sodium acetate (pH 5.0) buffer. After loading, the column was washed with 0.1 L of 1 mM sodium acetate (pH 5.0) and [13C10]-AdoMet was eluted with a linear gradient from 1 mM sodium acetate (pH 5.0) to 1 M HCl over 0.2 L. Fractions that contained the [13C10]-AdoMet were identified via HPLC analysis, pooled and lyophilized to dryness. Dried AdoMet was dissolved in a minimal volume of water and the concentration was determined by measuring the UV-visible spectrum of AdoMet on an Agilent 8453 diode array spectrophotometer using the extinction coefficient of adenine at 259 nm (ε259nm = 15,400 M⁻¹cm⁻¹). [13C10] AdoMet was >95 % pure and 87% enriched with 13C based on LC-MS analysis using a LTQ Orbitrap XL (Thermo Fisher) mass spectrometer connected to an Agilent 1100 HPLC equipped with a diode-array detector.

An aliquot of the pure, concentrated sample (10 µL) was injected onto an 4.6 × 250 mm Eclipse XDB-C18 reverse phase column (Agilent) equilibrated in 0.05 M ammonium acetate (pH 6.0) (solution A) operated at 0.3 mL/min. Solution B was 40% acetonitrile/60% H2O. The separation program was as follows: 0–3 min, 0% B; 3–4.4 min, 0–0.2% B; 4.4–5.8 min, 0.2–0.8% B; 5.8–7.2 min, 0.8–1.8% B; 7.2–8.6 min, 1.8–3.2% B; 8.6–10 min, 3.2–5% B; 10–25 min, 5–25% B; 25–40 min, 25–50% B; 40–49 min, 50–75% B; 49–52 min, 75% B. After each run the C18 column was washed with 11 mL of 100% acetonitrile and equilibrated with 10 mL of 0.05 M ammonium acetate (pH 6.0). The LTQ-Orbitrap XL was operated in positive ion mode detecting with FT analyzer set to a resolution of 100,000, 1 microscan, and 200 ms maximum injection time.

Natural abundance 6-CP was obtained from Sigma-Aldrich, dissolved in 20 mM Tris•HCl (pH 8.0) and quantified by UV-visible spectroscopy (ε363nm = 9,330 M⁻¹cm⁻¹, pH 8.0). [U-13C7]-6-CP was generated by permanganate oxidation of [U-13C] CPH4, which was synthesized enzymatically (10 mL reaction) from [U-13C] GTP (Sigma-Aldrich) as previously described.2 [U-13C7]-CPH4 was oxidized to [U-13C7]-6-CP with potassium permanganate at room temperature; the reaction progress was monitored by LC-MS analysis. Potassium permanganate was added to the 10 mL reaction to a final concentration of 7 mM. The reaction was stirred at room
temperature for 60 min. After stirring, 10 mL of ethanol was added to the reaction and subsequently diluted to 40 mL in 0.01 M ammonium bicarbonate (pH 7.8). The diluted reaction was then loaded onto a 20 mL DEAE FF column (GE life sciences) equilibrated in 0.01 M ammonium bicarbonate (pH 7.8). After loading, the column was washed with 80 mL of 0.01 M ammonium bicarbonate (pH 7.8). [U-\textsuperscript{13}C\textsubscript{7}]-6-CP was eluted using a linear gradient from 0.01 – 0.5 M ammonium bicarbonate (pH 7.8) over 300 mL collecting 5 mL fraction. Fractions that contained the [U-\textsuperscript{13}C\textsubscript{7}]-6-CP were identified via LC-MS analysis, pooled, and lyophilized to dryness. [U-\textsuperscript{13}C\textsubscript{7}]-6-CP was dissolved in a minimal volume of 20 mM Tris-HCl (pH 8.0) and the concentration was determined by as described above. The purity and isotopic enrichment of [U-\textsuperscript{13}C\textsubscript{7}]-6-CP was assessed by LC-MS analysis using the same method described above for [\textsuperscript{13}C\textsubscript{10}]-AdoMet.

**QueE incubation with 6-CP**

All experiments described below were setup under anaerobic conditions (95% N\textsubscript{2}, 5% H\textsubscript{2} atmosphere in a Coy anaerobic chamber). Each reaction contained 0.05 M PIPES•NaOH (pH 7.4), 10 mM DTT, 2 mM MgSO\textsubscript{4}, 0.5 mM AdoMet, 0.1 mM BsQueE or BmQueE, 0.5 mM 6-CP, and 0 or 10 mM dithionite in a final volume of 0.1 mL. All components except 6-CP were mixed and allowed to incubate at room temperature for 10 min. Turnover was initiated by the addition of 6-CP. Control reactions where AdoMet, BsQueE or BmQueE, or 6-CP was omitted were setup in the same manner for both the plus and minus dithionite condition. Each reaction was quenched by adding 10 μL of 30% (w/v) TCA after 15 h. The samples were centrifuged at 14,000×g for 5 min to remove the precipitate. Each reaction was analyzed for a new product by injecting a 100 μL aliquot from each reaction mixture onto a 4.6 × 250 mm Eclipse XDB-C18 reverse phase column (Agilent) using the method described above.

**Determination of the rate of formation of 6-deoxyadenosylpterin (6-dAP)**

Each reaction contained 0.05 M PIPES•NaOH (pH 7.4), 10 mM DTT, 2 mM MgSO\textsubscript{4}, 2 mM AdoMet, 0.01 mM BsQueE, and 10 mM dithionite. The concentration of 6-CP was varied at 0.1 mM, 0.3 mM, or 1 mM in a final volume of 0.1 mL. All components except 6-CP were mixed and allowed to incubate at room temperature for 10 min. Turnover was initiated by the addition of 6-CP. Each reaction was quenched with 10 μL of 30% (w/v) TCA after 5 min. The samples
were centrifuged at 14,000\,\times\,g for 5 min to remove the precipitate. The concentration of 6-dAP was determined by injecting 60 \,\mu\,L of the reaction mixture onto a 4.6 \times 250 \,mm Eclipse XDB-C18 reverse phase column (Agilent) using the method described above and integrating the area under the 6-dAP peak when monitoring at 350 nm. The area was converted to concentration using a calibration curve constructed by injecting 60 \,\mu\,L aliquots of a range of 6-CP standards of known concentration running the same method.

**Determination of the rate of formation of 6-CP–dAdo ester**

Each reaction contained 0.05 M PIPES\,NaOH (pH 7.4), 10 mM DTT, 2 mM MgSO$_4$, 2 mM AdoMet, and 0.05 mM BsQueE in a final volume of 0.3 mL. The concentration of 6-CP was either at 0.25 mM or 0.5 mM. Turnover was initiated by the addition of 6-CP. Aliquots (70 \,\mu\,L) were quenched by adding 7 \,\mu\,L of 30\% (w/v) TCA after 30, 90, and 180 min. The samples were centrifuged at 14,000\,\times\,g for 5 min to remove the precipitate. The concentration of 6-CP–dAdo ester was determined by injecting 60 \,\mu\,L of the reaction mixture onto a 4.6 \times 250 \,mm Eclipse XDB-C18 reverse phase column (Agilent) using the method described above and integrating the area under the 6-CP–dAdo ester peak when monitoring at 350 nm. The area was converted to concentration using a calibration curve constructed by injecting 60 \,\mu\,L aliquots of a range of 6-CP standards of known concentration running the same method.

**BsQueE reaction with [U-\textsuperscript{13}C\textsubscript{7}]-6-CP and [\textsuperscript{13}C\textsubscript{10}]-AdoMet**

Three reactions were setup with and without dithionite and analyzed in the same manner as described above except one reaction contained [\textsuperscript{13}C\textsubscript{10}]-AdoMet with natural abundance 6-CP, a second reaction contained [U-\textsuperscript{13}C\textsubscript{7}]-6-CP with natural abundance AdoMet, and a third reaction contained both [\textsuperscript{13}C\textsubscript{10}]-AdoMet and [U-\textsuperscript{13}C\textsubscript{7}]-6-CP.

**MS/MS fragmentation**

Each new product was selected for MS/MS analysis. The appropriate peak from each reaction was selected for both HCD and CID fragmentation in separate runs. All HCD fragmentation experiments had an isolation width of 2.0 \,m/z, 35\% collision energy, and 0.1 ms activation time. All CID fragmentation experiments had an isolation window of 2.0 \,m/z, 35\% collision energy, Q 0.25, and 30 ms activation time. All MS/MS experiments were detected using the FT analyzer set to a resolution of 100,000, 1 microscan, and 200 ms maximum injection time.
**Mg\textsuperscript{2+} dependence for BsQueE reaction with 6-CP and AdoMet**

Each reaction contained 0.05 M PIPES-NaOH (pH 7.4), 10 mM DTT, 0 or 10 mM DT, 0 or 2 mM MgSO\textsubscript{4}, 2 mM AdoMet (obtained from Sigma-Aldrich), and 0.3 mM 6-CP in a final volume of 0.1 mL. BsQueE was added to each reaction to the final concentration of 0.01 mM for the reactions containing dithionite and 0.05 mM for the reactions where dithionite was omitted. All components except for 6-CP were mixed and allowed to incubate at room temperature for 10 min. The reactions were initiated upon addition of 6-CP. Control reactions were setup and run as described above except BsQueE was omitted from the reaction. The reactions were quenched by the addition of 10 μL of 30% (w/v) TCA after incubating for 40 min, for the reactions containing dithionite, or 180 min, for the reactions were dithionite were omitted. The concentrations of 6-dAP and 6-CP-dAdo ester were determined by injecting 100 μL of the reaction mixture onto a 4.6 × 250 mm Eclipse XDB-C18 reverse phase column (Agilent) using the method described above and integrating the area under the 6-dAP or 6-CP-dAdo ester peak when monitoring at 350 nm. The areas were converted to concentration using a calibration curve constructed by injecting 100 μL aliquots of a range of 6-CP standards of known concentration running the same method.

**ICP-MS analysis for Mg\textsuperscript{2+} content**

The Mg\textsuperscript{2+} ion content of the QueE/6–CP reactions in the presence or absence of dithionite were determined by ICP-MS at the Center for Water, Ecosystems, and Climate Science in the Department of Geology and Geophysics at the University of Utah. A 0.5 mL reaction was setup as described in the previous section and diluted to 5 mL in 1% (v/v) trace-metals grade nitric acid prior to analysis.
Figure II.1. Schematic of radical AdoMet chemistry and active site view for the ligand complexes of *Burkholderia multivorans* QueE. (A) The [4Fe-4S] cluster of a radical AdoMet enzyme is reduced from the +2 to the +1 oxidation state by an electron from an external source. In vitro, electrons are commonly supplied from NADPH via the biological reducing system Fpr/Fld or from chemical reductants such as dithionite. AdoMet is reductively cleaved to form 5'-dAdo* and 1-methionine upon inner sphere electron transfer from the [4Fe-4S] cluster to the sulfonium of AdoMet. 5'-dAdo* abstracts a H-atom from the substrate to initiate the catalytic cycle. Some AdoMet radical enzymes reform the cofactor at the end of the catalytic cycle. (B) For BmQueE, the 5'-carbon of the deoxyadenosine moiety of AdoMet (maroon) is 3.9 Å from the C-6 carbon of the substrate CPH4 (salmon) (PDB ID 4NJI). The [4Fe-4S] cluster is displayed as yellow and orange spheres. Nitrogen atoms are in blue, oxygen are in red, and sulfur are in yellow. (C) The 5'-carbon of the deoxyadenosine moiety of AdoMet (maroon) is 3.2 Å and 4.9 Å from a carboxylate oxygen and the C-6 of 6-CP (grey), respectively (PDB ID 4NJG).
**Figure III.2. Structures of BmQueE** Overlay of the structures of BmQueE complexed with CPH₄ (salmon) (PDB ID 4NJI), CDG (green) (PDB ID 4NJK), and 6-CP (slate). The [4Fe-4S] cluster (yellow and orange spheres) and AdoMet (maroon) are observed in nearly the same position in each of the structures. The ligands are bound with minimal differences as necessary to accommodate the small structural changes.
Table III.1. Data collection and refinement statistics for BsQueE structures\(^\text{a}\).

<table>
<thead>
<tr>
<th></th>
<th>BsQueE Fe Peak(^\text{a})</th>
<th>BsQueE Initial structure</th>
<th>BsQueE Adduct</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Beamline</strong></td>
<td>ALS 8.3.1</td>
<td>ALS 8.3.1</td>
<td>APS 24-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ID-E</td>
</tr>
<tr>
<td><strong>Wavelength (Å)</strong></td>
<td>1.7389</td>
<td>0.9999</td>
<td>0.9792</td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>(P_{21}2_{1}2_{1})</td>
<td>(P_{21}2_{1}2_{1})</td>
<td>(P_{21}2_{1}2_{1})</td>
</tr>
<tr>
<td><strong>Cell dimensions (Å)</strong></td>
<td>54.3, 79.4, 122.5</td>
<td>54.2, 79.3, 122.2</td>
<td>53.9, 136.8, 141.2</td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>50.0-3.00 (3.11-3.00)</td>
<td>50.0-2.55 (2.64-2.55)</td>
<td>50.0-2.41 (2.46-2.41)</td>
</tr>
<tr>
<td><strong>Completeness</strong></td>
<td>99.5 (98.9)</td>
<td>99.3 (98.0)</td>
<td>98.8 (93.8)</td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td>7.7 (7.5)</td>
<td>3.9 (3.6)</td>
<td>3.2 (2.7)</td>
</tr>
<tr>
<td><strong>(R_{sym})</strong></td>
<td>0.083 (0.494)</td>
<td>0.07 (0.46)</td>
<td>0.099 (0.461)</td>
</tr>
<tr>
<td><strong>(I/\sigma(I))</strong></td>
<td>16.8 (4.2)</td>
<td>17.7 (2.3)</td>
<td>10.7 (2.1)</td>
</tr>
</tbody>
</table>

**Model Refinement**

<table>
<thead>
<tr>
<th>Resolution limits (Å)</th>
<th>50.0-2.55</th>
<th>50.0-2.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R_{work}/R_{free})</td>
<td>21.6/24.8</td>
<td>21.7/24.3</td>
</tr>
<tr>
<td><strong>Reflections</strong></td>
<td>17,694</td>
<td>40,805</td>
</tr>
<tr>
<td><strong>No molecules in asu</strong></td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td><strong>No atoms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>3627</td>
<td>7599</td>
</tr>
<tr>
<td>[Fe-4S]</td>
<td>16</td>
<td>32</td>
</tr>
<tr>
<td>L-Met</td>
<td>9</td>
<td>36</td>
</tr>
<tr>
<td>Adduct</td>
<td>-</td>
<td>132</td>
</tr>
<tr>
<td>Water</td>
<td>85</td>
<td>188</td>
</tr>
<tr>
<td><strong>B-factors (Å(^2))</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>60.3</td>
<td>35.2</td>
</tr>
<tr>
<td>[4Fe-4S]</td>
<td>67.1</td>
<td>30.9</td>
</tr>
<tr>
<td>L-Met</td>
<td>80.7</td>
<td>31.0</td>
</tr>
<tr>
<td>Adduct</td>
<td>-</td>
<td>32.3</td>
</tr>
<tr>
<td>Water</td>
<td>49.9</td>
<td>33.3</td>
</tr>
<tr>
<td><strong>R.M.S. deviations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond Lengths, (Å)</td>
<td>0.076</td>
<td>0.008</td>
</tr>
<tr>
<td>Bond Angles (°)</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>Rotamer outliers (%)</td>
<td>1.5</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>Ramachandran Plot (%)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most Favored</td>
<td>95.1</td>
<td>95.8</td>
</tr>
<tr>
<td>Additionally allowed</td>
<td>4.9</td>
<td>4.2</td>
</tr>
<tr>
<td>Disallowed</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

\(^a\)Coordinates have been deposited to the Protein Data Bank with PDB accession codes: 5TH5 and 5TGS

Highest-resolution shell is shown in parentheses

& Bijvoet pairs were scaled separately in this data set.

\(^\text{A}\) 5% reflections used for test set
Figure II.3. Adduct bound structure of BsQueE. (A) The monomeric subunit of dimeric BsQueE is composed of a central partial TIM domain (green) with N-terminal (light blue) and C-terminal (pink) extensions. The adduct and l-methionine carbons are colored teal and purple, respectively. The iron and sulfur atoms of the cluster are colored orange and yellow respectively. (B) Simulated annealing omit density maps for the AdoMet radical [4Fe4S] cluster, l-methionine (colored purple) and adduct refined as 6-CP-dAdo ester (teal), contoured at 3σ. (C) Interactions between the 6-CP moiety of 6-CP-dAdo ester (teal) and protein residues. Protein is colored according to regions with the AdoMet radical core in green N-terminal extension in light blue and C-terminal in pink extension.
Figure II.4. Formation of the BsQueE adduct species. (A) Difference electron density maps obtained when refining data with intact AdoMet (maroon) and 6-CP (slate). Positive and negative difference densities are displayed as green and red meshes, respectively, contoured at 3σ. The iron-sulfur cluster is in orange and yellow. (B) In order to form the adduct (teal), the ribose ring of 5'-dAdo must rotate about 30° from initial position in AdoMet.
Figure II.5. Walleyed stereoview of the binding interactions of BsQueE with the 5'-dAdo moiety of the adduct species. The 5'-dAdo moiety binds to BsQueE using previously characterized AdoMet binding motifs; residues of the ribose motif (slate) interact with the hydroxyl groups of the ribose ring through the hydroxyl and amino groups of Ser127 and Lys129 respectively, the β5 or GXIXGXXE motif (green) provides hydrophobic interactions to the adenine ring, and the β6 motif (salmon) provides interaction to the nitrogen atoms of the adenine ring positioning it in the active site via hydrogen bonds from the amide of Gln188 and the backbone amide of Asn191. Two residues from the cluster-binding loop (tan) provide both π-π interactions as well as hydrogen bonds to assist in the orientation of the adenine ring in the active site.
Figure II.6. QueE catalyzes the conversion of 6-CP to a new pterin-containing species in the presence of dithionite and AdoMet. (A) HPLC chromatogram of reactions monitored at 350 nm showing that QueE can turnover 6-CP in the presence of AdoMet to a new product (*) with retention time of 39 min (a). This peak was not observed in the control reactions where either BsQueE (b), AdoMet (c), or 6-CP (d) were omitted. 6-CP elutes at 15 min under these conditions. (B) The UV-visible spectra of 6-CP and the new product eluting at ca. 15 and 39 min, respectively. (C) Mass spectra of product eluting at 39 min isolated from reaction of BsQueE under reducing conditions with natural abundance 6-CP and natural abundance AdoMet (black) (a); [U-13C7] 6-CP and natural abundance AdoMet (blue) (b); natural abundance 6-CP and [13C10-dAdo]- AdoMet (red) (c); or [U-13C7]-6-CP and [13C10-dAdo]-AdoMet (purple) (d). The species that is +22 amu relative to the [M+H]+ corresponds to [M+Na]+. (D) HCD fragmentation analysis of the 6-CP adduct formed under reducing conditions with natural abundance 6-CP and natural abundance AdoMet (black) (a); [U-13C7] 6-CP and natural abundance AdoMet (blue) (b); natural abundance 6-CP and [13C10-dAdo]- AdoMet (red) (c); or [U-13C7]-6-CP and [13C10-dAdo]- AdoMet (purple) (d). (E) The fragmentation patterns in (D) allow assignment of the new species at 39 min to 6-dAP.
Figure 11.7. Analysis of isotopically-enriched AdoMet. (A) HPLC chromatogram of $[^{13}\text{C}_{10}]$-AdoMet monitoring at 260 nm and (B) the mass spectrum of the peak that elutes at 22 min. (C) Mass spectrum focusing on the region from 395 to 415 m/z. AdoMet enriched with ten $^{13}$C atoms has a m/z of 409.1779. The peaks from 404.1616 to 408.1747 represent AdoMet enriched with five to nine $^{13}$C atoms. Natural abundance AdoMet has a [M]+ of 399.1445, which is a result of natural abundance ATP contaminant from the cell lysate with overexpressed AdoMet synthetase used to synthesize the isotopically enriched AdoMet. Based on the HPLC chromatogram and mass spectrum, $[^{13}\text{C}_{10}]$-AdoMet was $>95\%$ pure and 87 % enriched with $^{13}$C at all ten carbons of the deoxyadenosine moiety.
Figure III.8. Analysis of isotopically enriched 6-CP. (A) HPLC chromatogram of [U-$^{13}$C$_7$]-6-carboxypterin monitoring at 350 nm and (B) the mass spectrum of the peak that elutes at ca. 17 min. (C) Mass spectrum focusing on the region from 205 to 225 m/z. 6-CP enriched with seven $^{13}$C atoms has a m/z of 215.0703. Based on the HPLC chromatogram and mass spectrum, [U-$^{13}$C$_7$]-6-carboxypterin was >99% pure and 95 % enriched with $^{13}$C at all seven carbons.
Figure II.9. HPLC chromatogram of BsQueE reactions containing AdoMet and 6-CP monitored at 350 nm. In the absence of dithionite, a new peak was observed that elutes at 42 min (b) instead of the peak that elutes at 39 min (d) that was observed in the presence of dithionite. Both the 39 and 42 min peaks require BsQueE (c and a, respectively).
Figure II.10. Conversion of 6-CP to 6-CP-dAdo ester under non-reducing conditions. (A) HPLC chromatogram monitored at 350 nm of reactions showing that BsQueE can turnover 6-CP in the presence of AdoMet to a new product denoted by * (a). To rule out that DTT, which was present in the reaction was not responsible, the reactions were repeated in the absence of DTT and the same result was obtained (b). However, this species was not observed in the control reactions where either BsQueE (c), AdoMet (d), or 6-CP (e) were omitted. (B) The UV-visible spectra of 6-CP and the new product, eluting at 15 and 42 min, respectively, show that the new species has spectral features of 6-CP and an additional feature ~260 nm. (C) Mass spectra of product (eluting at ca. 42 min) isolated from reactions of BsQueE in the absence of dithionite with natural abundance 6-CP and natural abundance AdoMet (black) (a); [U-13C7]-6-CP and natural abundance AdoMet (blue) (b); natural abundance 6-CP and [13C10-dAdo]-AdoMet (red) (c); or [U-13C7]-6-CP and [13C10-dAdo]-AdoMet (purple) (d). The species at +22 amu relative to the [M+H]+ corresponds to [M+Na]+. (D) CID fragmentation analysis of the new reaction product obtained with natural abundance 6-CP and natural abundance AdoMet (black) (a); [U-13C7] 6-CP and natural abundance AdoMet (blue) (b); natural abundance 6-CP and [13C10-dAdo]-AdoMet (red) (c); or [U-13C7]-6-CP and [13C10-dAdo]-AdoMet (purple) (d). (E) The fragmentation patterns in (D) allow assignment of the new species at 42 min to 6-CP-dAdo ester.
Figure III.11. *B. multivorans* QueE catalyzes the formation of 6-dAP and 6-CP–dAdo ester in the presence or absence of reductant, respectively. (A) LC-MS extracted ion chromatograms monitoring at m/z 412.5-413.5 of reactions containing natural abundance 6-CP and natural abundance AdoMet incubated with (a) or without (b) BmQueE in the presence of reductant. (B) Mass spectra of product eluting at ca. 37 min isolated from reaction of BmQueE under reducing conditions with natural abundance 6-CP and natural abundance AdoMet. (C) LC-MS extracted ion chromatograms monitoring at m/z 456.5-457.5 of reactions containing natural abundance 6-CP and natural abundance AdoMet incubated with (a) or without (b) BmQueE in the absence of reductant. (D) Mass spectra of product eluting at ca. 42 min isolated from reaction of BmQueE with natural abundance 6-CP and natural abundance AdoMet in the absence of reductant.
Scheme II.1 Proposed mechanisms of the polar (left) and radical addition (right) reactions to form 6-CP–dAdo ester and 6-dAP, respectively.
References

Chapter IV
Structural comparison of bacterial QueE orthologs

Chapter Summary
7-carboxy-7-deazaguanine synthase, QueE, catalyzes the radical mediated ring contraction of 6-carboxy-5,6,7,8-tetrahydropterin, forming the characteristic pyrrolopyrimidine core of all 7-deazaaguanine products. QueE is member of the S-adenosyl-L-methionine (AdoMet) radical enzyme superfamily, which harnesses the reactivity of radical intermediates to perform challenging chemical reactions. Members of AdoMet radical enzyme superfamily utilize a canonical binding motif, a CX₃CXφC motif, for binding a [4Fe-4S] cluster and a partial (β/α)₆ TIM barrel fold for the arrangement of AdoMet and substrates for catalysis. Although variations to both the cluster-binding motif and the core fold have been observed, visualization of drastic variations in the structure of QueE from Burkholderia multivorans called into question whether a re-haul of the defining characteristics of this superfamily was in order. Surprisingly, the structure of QueE from Bacillus subtilis revealed an architecture more reminiscent of the classical AdoMet radical enzyme. With these two QueE structures revealing varying degrees of alterations to the classical AdoMet fold, a new question arises: what is the purpose of these alterations? In this chapter, we present the structure of a third QueE enzyme from Escherichia coli, which establishes the middle range of the spectrum of variation observed in these orthologs. With these three homologs we compare and contrast the structural architecture and make hypotheses about the role of structural variations in binding and recognizing the biological reductant, flavodoxin.

Contributions
Nathan Bruender carried out the construct design and protein purifications. Daniel Dowling and Benjamin Bell performed the crystallography and Chi Ngyuen and Tsehai Grell solved the structure. We would like to thank Nathan Bruender and Vahe Bandarian for stimulating discussion.
Introduction

7-Deazapurines or pyrrolopyrimidine-containing compounds are a structurally diverse class of nucleoside analogs found in secondary metabolites, natural products and the hypermodified tRNA bases queuosine and archaeosine. Queuosine is a universal tRNA modification occupying the wobble position of His, Tyr, Asp and Asn tRNAs in all kingdoms of life, whereas, archaeosine is found within the dihydrouridine loop of most archaeal tRNAs. The 7-deazaupurine core found in >30 compounds is produced from guanosine triphosphate (GTP) in three steps (Figure IV.1). The final step converts 6-carboxy-5,6,7,8-tetrahydropterin (CPH₄) to 7-carboxy-7-deazaguanine (CDG) in an unprecedented rearrangement catalyzed by the Mg²⁺- and S-adenosyl-L-methionine (AdoMet) dependent enzyme CDG synthase (QueE).

QueE is member of the AdoMet radical enzyme superfamily, a rapidly growing family of enzymes, which utilize the reductive cleavage of a molecule AdoMet ligated to a [4Fe-4S] cluster to initiate radical chemistry. The identification of a conserved motif, CX₃CXₚC motif, which allowed for binding of the essential [4Fe-4S] cluster with a site differentiated iron for AdoMet ligation unified the AdoMet radical enzyme superfamily. Direct ligation of AdoMet to the unique iron of the cluster positions the C5'-S bond of AdoMet to be cleaved homolytically through the inner sphere transfer of an electron form the cluster. The highly reactive intermediate, 5'-deoxyadenosyl radical or 5'-dAdo•, generated can abstract a hydrogen-atom (H-atom) from diverse substrates, initiating a variety of chemically challenging and complex reactions.

AdoMet radical enzymes adopt a full or partial (β/α)₆ triose-phosphate isomerase (TIM) barrel fold, where the inner cavity comprises of a parallel β-sheet flanked by α-helices. The AdoMet radical cluster sits at the head of this barrel with the unique iron facing the cavity of the barrel. Through coordination of the unique iron of the AdoMet radical cluster and interactions with conserved binding motifs provided by the AdoMet radical core fold AdoMet is oriented within the barrel. In addition, the core fold provides a solvent protected cavity for the production of the radical intermediates necessary for the catalytic mechanisms undertaken by members of the superfamily. Variations to the AdoMet core fold have been observed. For example, BtrN, a butirosin biosynthetic enzyme and OxsB, an enzyme involved in the
biosynthesis of oxetanocin A both lack β6 of the partial TIM barrel substituting a loop in its place. Surprisingly, AdoMet binding was not perturbed and residues from the loop supplemented the binding interactions normally provided by the β6 residues.

The structure of QueE from *Burkholderia multivorans* (*Bm*QueE) revealed drastic deviations to two pillars of the AdoMet radical enzyme structure, the AdoMet radical core fold and the cluster-binding motif: *Bm*QueE folded into a parsed-down partial (β6/α3) TIM barrel, in comparison to a partial (β/α)6 TIM barrel, and contained a modified cluster-binding motif, a CX14CXϕC motif. As noted previously, these variations did not alter cluster binding or AdoMet ligation in *Bm*QueE. Furthermore, they are not necessary for CDG formation as other QueE enzymes utilize the canonical cluster-binding motif (Figure IV.2) and QueE from *Bacillus subtilis* (*Bs*QueE) adopts a partial (β6/α5) TIM barrel fold with minimal variations. It has been hypothesized that the alterations to the *Bm*QueE scaffold are involved in binding of the physiological reductant, flavodoxin. Flavodoxins reduce the AdoMet radical cluster from the 2+ state to the 1+ state necessary for catalysis. The reductant-binding site, which places the flavin mononucleotide (FMN) cofactor of flavodoxin within electron transfer distance from the AdoMet radical cluster, is proposed to comprise of the cluster binding loop as well as residues from the loops connecting β2 and β4 to α2 and α4. The flavodoxin—flavodoxin reductase system from *Escherichia coli* (*E. coli*) has been successfully used to deliver reducing equivalents to AdoMet radical enzymes of different species, in fact, *Bs*QueE shows increased production of CDG when incubated with the *E. coli* biological reducing system in comparison to dithionite. However, maximal production of CDG is observed for *Bs*QueE in the presence of the one of the native flavodoxins, *BsFld*. Surprisingly, the converse is true for *Bm*QueE: more CDG is observed in the presence of dithionite than when *Bm*QueE is incubated with the *E. coli* flavodoxin—flavodoxin reductase system.

Over the last 17 years, our understanding of the AdoMet radical family has increased. Through combination of biochemical and structural analysis, we have begun to uncover the mechanisms by which members of the AdoMet radical superfamily tailor their structures and active sites to the type of chemistry they perform, and to some extent, the substrates they catalyze. Little work, however, has focused of understanding the reasons for structural differences between AdoMet
radical enzyme orthologs. Here we compare and contrast the structures of three QueE homologs, from *B. multivorans*, *B. subtilis* and *E.coli* (Figure IV.2, IV.3) and use these structures, along with the structures of the cognate flavodoxins to begin to tease out the structural determinants for differential reductant binding.

**Results**

*EcQueE reveals an intermediary structure between BmQueE and BsQueE.*

The crystal structure of QueE from *Escherichia coli* (*EcQueE*) was determined to 2.1-Å resolution by multi-wavelength anomalous dispersion (MAD) phasing and *R*\textsubscript{work} and *R*\textsubscript{free} of 0.205 and 0.238, respectively (Table IV.1). In the final structure, electron density was observed for most of the crystallization construct with the exception of the first 9 residues of the N-terminal hexahistidine tag (His\textsubscript{6}tag), residues 192-196 and the final 10 residues. *EcQueE* folds into a structural and functional head-to-tail homodimer, reminiscent of the published QueE structures from *Burkholderia multivorans* (*BmQueE*) and *Bacillus subtilis* (*BsQueE*)\textsuperscript{15-16} (Figure IV.4A). The overall structure of *EcQueE* is similar to that of *BmQueE* (rmsd 1.8 Å) and *BsQueE* (rmsd 2.9 Å), with variations in the structure and orientation of the loops and α-helices (Figure IV.4B).

All three QueE homologs fold into variants of AdoMet radical core domain with extensions at the N- and C-termini. The N-terminal extensions of QueE structures comprise of a single anti-parallel β-strand, β1’ (Figure IV.5A-B), which is found adjacent to β1 of the AdoMet radical core. In *EcQueE*, the linker and the first residue of the His\textsubscript{6}tag is visible, forming an additional α-helix, α1’, at the N-terminal of the enzyme (Figure IV.5A). The C-terminal of *BmQueE* and *BsQueE* folds into a β-strand/α-helix (β7’/α7’) pair, where β7’ is found adjacent to β6 (Figure IV.5B,C). In the *EcQueE* structure, α7’ of the β7’/α7’ pair is not visible due to the final 10 amino acids of the structure being disordered (Figure IV.5A). The N- and C-terminal extensions were shown to be important for both substrate binding and dimerization in *BmQueE* and *BsQueE* and it is expected to serve the same function in *EcQueE*. Mutual interactions between the β1’-loop-β1 and β7’ of the adjacent QueE monomers creates a dimeric interface such that the β-strands of the N- and C-terminal extensions, not only extend the monomeric inner face (made up
of β-sheets) of the QueE barrel, but form an inter-monomer-ten-stranded β-sheet which is thought to resemble a crown (Figure IV.5A-C).

The core of the QueE homolog structures adopt three unique partial TIM barrels folds, where each variant differs in the number and type of α-helices flanking the conserved parallel β-sheet. The previously published structure of BmQueE shows the greatest variation of the three homologs and of the whole AdoMet radical superfamily characterized to date. BmQueE sports a vastly parsed down AdoMet radical fold, a β6/α3, where short loops, L3 and L4, replaced α3 and α4 and a short 3_{10}-helix, 3_{10}H5, replaced α5. (Figure IV5.B). The variations in the AdoMet radical core of BsQueE are the most conservative of the three homologs. BsQueE folds into a partial β6/α5 TIM barrel, which contains a non-traditional short 3_{10}-helix, 3_{10}H3 in place of α3 (Figure IV.5C). Similarly, the AdoMet radical domain of EcQueE folds into a partial β6/α5 TIM barrel with a variation at the α3 position, but this change is not as conservative as that seen in BsQueE (Figure IV.5A). In EcQueE, α3 is replaced by a short loop (L3), reminiscent of the α3 alternative in BmQueE, and has a long loop, L4, connecting β4 to a very short α4.

The three QueE homologs, to date, are the smallest structurally characterized AdoMet radical enzymes, with BmQueE spanning only 210 amino acid residues, EcQueE, 223 amino acid residues and BsQueE 243 amino acid residues. The second smallest non-QueE AdoMet radical enzyme structurally characterized, PFL-AE (246 amino acid residues), shows surprising structural similarities to the QueE homologs, in particular BsQueE^{27}. PFL-AE adopts a normal AdoMet radical core, a (β/α)6 TIM barrel, and contains N- and C-terminal extensions, β1' and β7', which closely resemble those found in QueE (Figure IV.5D). Unlike QueE, PFL-AE is a monomer, thus these terminal extensions do not play a role oligomerization. However, similar to QueE, the N-terminal extension is involved in substrate binding.

In all three structures, electron density was present for a [4Fe-4S] cluster bound by three cysteine ligands, leaving a site-differentiated iron. Sequence analysis revealed an 11 amino acid insertion in the cluster-binding loop of BmQueE, resulting in a CX_{14}CXφC sequence instead of the canonical CX₃CXφC cluster-binding motif. Surprisingly, the insertion did not affect cysteine
positioning and cluster binding and the cysteine ligands from cluster binding loop superimposed well with other AdoMet radical enzymes (Figure IV.4B, IV.5B). Instead, the insertion folds into a short $3_{10}$-helix, $3_{10}H1$, found on top of the AdoMet radical cluster and further sequesters the cluster from solvent as well as increases the negative charge in that area. Following the cluster-binding motif, the loop folds into a short β-strand, $\beta 2'$, before transitioning into $\alpha 1$, another structural addition outside of the AdoMet radical core of BmQueE. Sequence and structural analysis of BsQueE and EcQueE revealed a canonical cluster binding loop motif, CX3CXW'C, which positions the cluster at the top of the AdoMet radical barrel. Akin to BmQueE, the transition to the $\alpha 1$ from the cluster-binding motif proceeds through $\beta 2'$ in BsQueE. In EcQueE, on the other hand, we see further additional structural elements; following the cluster binding motif, the loop folds first into a short $\alpha$-helix, $\alpha 2'$, which proceeds into $\beta 2'$ and subsequently into $\alpha 1$ of the AdoMet radical core.

AdoMet binding motifs are conserved in the QueE homologs.

The structures of QueE homologs revealed various substrate and product bound states but only in BmQueE have we visualized intact AdoMet and substrates (Table IV.2). Even with the variant AdoMet core folds, the previously described AdoMet binding motifs are conserved$^{11-12.15}$ (Figure IV.2). This conservation of motifs were particular surprising in the case of the extensively modified AdoMet radical fold of BmQueE (Figure IV.5B, IV.6A). To compare the AdoMet binding sites of the three homologues, AdoMet was manually docked into EcQueE and BsQueE. In the case of BsQueE, modeling of intact AdoMet was guided by the position of the adenosyl moiety of the 6-carboxypterin-5'-deoxyadenosyl (6-CP—dAdo) ester, and AdoMet bound BmQueE (PDB ID 4NJI) (Figure IV.6B). The adenine ring of 6-CP—dAdo previously revealed the adenine binding AdoMet motifs in BsQueE, therefore AdoMet was anchored by positioning the carboxylic group, amino group and sulfonium atom, ~2.0, ~2.3 and ~3.2 Å away from the unique iron, respectively, and overlaying the adenine ring of AdoMet with that of 6-CP—dAdo. The structure of BmQueE (PDB ID 4NJI) was used to manually dock AdoMet into the active site of EcQueE. AdoMet was positioned in a catalytic orientation with its carboxylic group, amino group and sulfonium atom, ~2.0, ~2.3 and ~3.2 Å away, respectively, from the unique iron of the AdoMet radical cluster. Although residues corresponding to the AdoMet binding motifs are present in EcQueE, positioning of AdoMet in the catalytic orientation created
clashes with nearby residues. This clashing indicates that some rearrangement of the EcQueE active site will have to occur upon substrate or AdoMet binding in order to accommodate the AdoMet molecule (Figure IV.6C).

In both BmQueE and BsQueE the α-amino and carboxyl groups of the methionyl moiety are positioned to bind the unique iron of the AdoMet radical cluster through interaction from the so-called GGE motif (G91GE93 in BmQueE and G82GD84 in BsQueE) and a basic residue (K135 BmQueE and K129 BsQueE), respectively (Figure IV.6A, IV.6B). Likewise the GGE and K138 appears to be conserved in EcQueE; residues G95GE95 are in the correct orientation to H-bond with the AdoMet α-amino group, but K138, would need to rotate towards the active site to interact with the carboxyl group of AdoMet (Figure IV.6C). In its current position, K138, clashes with a hydroxyl group of the ribose ring and thus would be expected to move upon AdoMet binding. In BmQueE and BsQueE, S133 and K135, and S127 and K129, respectively, make up the ribose motif and provide hydrogen bonds to the hydroxyl groups of the ribose ring of AdoMet. In EcQueE, S136 and K138 could play an equivalent role. Hydrophobic interactions from the GXIXGXXE or β5 motif, V151 (BmQueE), V163 (BsQueE) and interestingly P161 (EcQueE), do or should orient the adenine ring in the active site. Although residue V162, found adjacent to P161 in EcQueE (Figure IV.2), may upon AdoMet and substrate binding rearrange to provide hydrophobic interactions to AdoMet, both sequence and structural analysis indicate P161 will provide the β5 motif interaction. Additional interactions to the adenine ring are provided by F48, D40, D176 and Q173 (BmQueE), and W40, D52, D191 and Q188 (BsQueE). To allow for AdoMet binding to EcQueE, W37 will have rotate to populate a planar rotamer, which could stack with the adenine ring, and D39 and S192 will have to move towards the active site, bringing the side chain carboxyl group and backbone carbonyl, respectively, within hydrogen bonding distance of the adenine ring. Residue Q189 overlays well with the corresponding residue in BmQueE, Q173, and appears to be in a conformation sufficient to interact with the adenine ring (Figure IV.6C).

The active site of QueE comprises of residues from the N- and C-terminal extensions.

To predict substrate binding modes in BsQueE and EcQueE, CPH_{4} and 6-CP bound structures of BmQueE (PDB ID 4NJI and PDB ID 4NJG, respectively) were used to model substrate binding to BsQueE (Figure IV.7B,E), and to EcQueE (Figure IV.7C,F)^{15-16}. 6-CP moiety of 6-CP—
dAdo adduct bound to BsQueE, was used as a guide to model 6-CP and CPH4. In EcQueE the substrates were modeled by aligning the active site of EcQueE to that of BmQueE. The QueE active site is found in the lateral opening of the partial TIM barrel and comprises of residues from the AdoMet radical core and N- and C-terminal extensions (Figures IV. 7). The pterin ring of 6-CP is oriented in the active site through several interactions, including π-π stacking with a His and Phe residue in both BmQueE and BsQueE. In EcQueE, the residue predicted to be involved in π-π stacking interactions to the substrate, H217, is the last ordered residue observed in the structure and is turned away from the active site. Instead, residue T217 appears to be in an orientation to interact with the pterin ring. In addition, the disordered C-terminal does not allow for visualization the C-terminal plug in EcQueE, which is provided by the carboxylate moiety of final residue of the protein, P210 in BmQueE or V245 in BsQueE. These C-terminal residues provide interactions to the N2 exocyclic amino group, N3 and the C4 carbonyl group, whereas residues R27 and T90 in BmQueE, R30 and S81 in BsQueE, and possibly R27 and T92 in EcQueE position the C6 carboxyl group. Hydrogen bonds from the backbone of N-terminal residues, G14 and L12 in BmQueE and EcQueE and G17 and Ile15 in BsQueE, further positions 6-CP in the active site.

Oxidized 6-CP was shown to bind similarly to substrate, CPH4, in BmQueE (Figure IV.7D). Q13 rotates slightly from its position in the 6-CP bound structure of BmQueE to provide hydrogen bonds to the newly protonated N8 of CPH4 (Figure IV.7D). In both BsQueE and EcQueE, the corresponding residue is already observed in a competent orientation to provide hydrogen bonds to the N8 of CPH4 (Figure IV.7E,F). In addition to positioning the carboxylic moiety closer to R27 and T90 in BmQueE, and corresponding residues in the two other QueE homologs, the ring puckering of CPH4 creates a binding site for the catalytic Mg2+ (Figure IV.8A-C). Only one residue, T51 (BmQueE) directly ligates Mg2+ and residues D50 and H204 (BmQueE) indirectly ligate Mg2+ through the ligation of metal bound water molecules (Figure IV.8A). Corresponding residues are found in BsQueE, S43, D42 and H223 and in fact we see water molecules in the vicinity of the modeled Mg, bound by these residues, which are reminiscent of what is seen in BmQueE (Figure IV.8B). In EcQueE, with rearrangement, T40 could coordinate Mg2+ and residues D39 and H217 could position water molecules to interact with Mg2+.
QueE homologs display differential electrostatic surfaces.

Generally two flavodoxins are found per organism. In the case of E. coli, flavodoxin denoted here as EcFldA has been shown to reduce a number of AdoMet radical enzymes and has been structurally characterized. Although both flavodoxins from B. subtilis reduce BsQueE, YkuN herein BsFld, showed more robust turnover and will be used in the analysis below. It is proposed that B. multivorans has three flavodoxins and we are not sure which of the three if not all are able to reduce BmQueE. In this analysis the most distinct flavodoxin in terms of predicted structure and charge is discussed, herein called BmFld. To understand the structural determinants of protein-protein interactions between QueE and flavodoxin, the structures of the cognate flavodoxins for BsQueE BsFld (Figure IV.10B) and BmQueE BmFld (Figure IV.10C) were predicted by homology modeling. In each case, 6 homology models were obtained. The overall architecture of flavodoxins (Figure IV.10A) shows five parallel \( \beta \)-strands flanked on either side by five helices, \( \alpha/\beta \) fold. The final strand, \( \beta_5 \), of EcFldA (PDB ID 1AHN) is split in two due to a \( \sim 20 \) amino acids insertion, indicative of a long chain flavodoxin (Figure IV.9, IV.10A). BsFld, on the other hand, appears to be a short chain flavodoxin (Figure IV.9, IV.10.B) and does not contain any alterations to the \( \alpha/\beta \) fold. Although BmFld is the right length to be considered a long chain flavodoxin, 170-180 residues, sequence and homology structure analysis reveals that it lacks the of the typical long chain insertion. Instead, BmFld has an insertion of \( \sim 20 \) amino acid after \( \beta_2 \) (Figure IV.9, IV.10C) that is not seen in either the short chain fold of the BsFld or the long-chain fold of EcFldA. This insertion extends the \( \beta_2-\alpha_2 \) and \( \alpha_2-\beta_3 \) loops, and based on homology is expected to place \( \alpha_2 \) adjacent to the FMN cofactor instead of it flanking the \( \beta \)-sheet. Furthermore, \( \alpha_2 \) appears to split in two with a loop insertion between the two segments that extends below the plane of FMN (Figure IV.10C). Flavodoxins bind their cofactor, flavin mononucleotide (FMN), at the C-terminal end of the \( \beta \)-strands with the loops playing significant roles in binding the phosphate group and isoalloxazine ring. This binding site is also predicted in the homology models of BsFld and BmFld generated (Figure IV.10D).

The electrostatic surfaces of QueE homologs and the corresponding flavodoxins were estimated in order to investigate whether charge-charge complementarity could play a role in QueE-flavodoxin interactions, (Figure VI.11). The electrostatic charge surfaces of the BsQueE (Figure VI.11B) and EcQueE (Figure IV.11C) contains large positive patches, on either side of the
cluster binding loop, and a large patch on the on the back of the proteins corresponding a number
loops/α-helices that flank the outside of the barrel. In particular, the N-terminal of BsQueE, the
N- and C-terminal ends of the cluster binding loop, β2', L2, the N-terminal of α2, L3, L4 and
the N-terminal of α4 constitutes a large positive patch surrounding the [4Fe-4S] cluster. In
EcQueE, the positive patches corresponded to the N-terminal of the protein, the cluster-binding
loop, α1', L2, and the N-terminal of L4. BmQueE shows a drastically decreased positive surface
with only a large patch of positive charge situated near the N-terminal of the protein consisting
with β1', N- and C-terminal of the cluster-binding loop and β2' (Figure IV.11A). In fitting with
the decreased positive surface charge observed in BmQueE, the cognate flavodoxin, BmFld,
(Figure IV.11D) is expected to have a highly positive surface, especially in the vicinity of the
FMN cofactor in comparison to the almost completely negative surfaces of BsFld (Figure
IV.11E) and EcFldA (Figure IV.11F).

Modeling Flavodoxin and QueE interactions.
The QueE-flavodoxin protein-protein interaction was predicted using the fast rotational docking
method for each QueE-flavodoxin cognate pair (Figures IV.12, IV.13, IV.14) and two of the top
predicted orientations of flavodoxin are shown in reference to QueE (Figure IV.12A, IV.13A
and IV.14A)37. Flavodoxins are predicted to bind their physiological partners using the surface
surrounding the cofactor, specifically the dimethyl end of FMN38-41. In EcQueE, the predicted
binding site of the negatively charged EcFldA appears to be consistent with the positive patches
observed on EcQueE surface (Figure IV.12B) and these interactions bring the FMN cofactor in
close proximity to the [4Fe-4S] AdoMet radical cluster, approximately 12 and 14 Å in the
predicted protein-protein interactions (Figure IV.12C,D). To orient the FMN cofactor in close
proximity of the [4Fe-4S] cluster without blocking the active site, EcFldA binds to the top of
EcQueE. In one of the docking results, EcFldA binds towards the C-terminal of EcQueE,
contacting the cluster binding loop, the loop following β4 (L4), L5 and L6, which is currently
disordered in our structure (Figure IV.12C). In the second predicted interaction EcFldA sits on
top of the cluster, positioning FMN ~14 Å away from the cluster by interacting with the cluster
binding loop and α1' and the N-terminal ends of α5 and α6 (Figure IV.12D). Similarly, in
BsQueE, predicted binding interactions with BsFld correspond to the large positive patch of
BsQueE, which is comprised of the loops and N-terminal ends of α2-α4 (Figure IV.13B). In
particular we observe interactions between $B$s$F$ld and $\beta1'$, the cluster binding loop, $\beta2'$, $\alpha2$, L3, $\alpha3$, L4 and $\alpha4$ of $B$s$Q$ueE (*Figure IV.13C,D*). The binding of flavodoxin to $B$s$Q$ueE in the orientation described brings the FMN cofactor within $\sim$17 Å of the AdoMet radical cluster in each of the modeled interactions shown. Although $B$m$Q$ueE-flavodoxin interactions are predicted to occur in some areas of positively charged surface on $B$m$Q$ueE, more interactions are observed with the negatively charged surfaces, than was seen in the two other homologs (*Figure IV.14B*). As in Ec$Q$ueE, the $B$m$F$ld predicted binding surfaces are comprised of secondary elements from both the N- and C-terminal, with the cluster binding loop and L4 being the common interactions between the two flavodoxin binding orientations (*Figure IV.14C,D*). In the predicted protein-protein interactions shown, the closest methyl group from the dimethyl end of FMN is approximately 9 Å (*Figure IV.14C*) and $\sim$12 Å (*Figure IV.14D*).

**Discussion**

Here we present the structure of Ec$Q$ueE in the absence of substrate. The previous QueE structures revealed non-traditional AdoMet core folds, and Ec$Q$ueE is no different. With different alterations in each enzyme, these QueE structures are a wealth of structural diversity allowing us to start teasing out the conformational changes, which occur upon CPH$_4$/6-CP and AdoMet binding, and explore the reasons for structural differences amongst enzymes catalyzing the same reaction in different organisms.

The $B$m$Q$ueE structure portrayed a new minimal architecture for an AdoMet radical enzyme that still performed catalysis. This new architecture revealed a highly altered AdoMet radical domain, a partial ($\beta6/\alpha3$) TIM barrel where the $\alpha$-helices are replaced by short loops, flexible enough to provide conformational flexibility to the rigid barrel. In addition to the substitution of $\alpha3$, $\alpha4$ and $\alpha5$ with L3, L4 and $3_{10}$H5, $B$m$Q$ueE has an altered cluster-binding motif, CX$_{14}$CX$\phi$C. The 11 amino acid insertion between the fist two cluster-binding cysteines is reconciled by formation of a $3_{10}$ helix, $3_{10}$H1’, which sits above the barrel and the [4Fe-4S] cluster, further secluding the AdoMet radical cluster from solvent. As described below, $B$s$Q$ueE and Ec$Q$ueE have canonical cluster-binding motifs. $B$m$Q$ueE further deviates from other AdoMet radical enzymes in its decreased positive surface charge in the vicinity of the cluster when compared to $B$s$Q$ueE and Ec$Q$ueE and the majority of other AdoMet radical enzymes.
The variations to the cluster binding motif and the AdoMet radical core do not appear to be necessary for the ring contraction of CPH₄ as both sequence analysis and structural analysis revealed that homologs such as BsQueE and EcQueE have canonical CX₃CXφC cluster-binding motifs. Although each of the QueE homolog sports a variant AdoMet radical cluster domain, these variations are not conserved and vary in extremity. As mentioned above, the variant AdoMet radical core of BmQueE was drastically pared down. In contrast, the BsQueE structure displays a moderate replacement of α₃ to a short 3₁₀ helix, 3₁₀H₃, and most resembles the canonical AdoMet radical core fold with a β₆/α₅ fold. The EcQueE structure presented another variation to the AdoMet radical core fold. In the EcQueE structure, once again we see replacement of α₃ to a loop and although α₄ is present, it is a very short helix connected to β₄ by a long loop, L₄, more reminiscent of L₄ found in BmQueE than the α₄ found in BmQueE. Therefore, EcQueE represents a structural intermediate between the drastic deviations found in BmQueE and the conservative alterations of BsQueE.

Even with these differential variations to the AdoMet core fold, AdoMet and substrate binding are conserved in each of the homologs. The QueE homologs employ all the AdoMet binding motifs previously described to bind AdoMet in AdoMet radical enzymes. In addition, the binding sites for CPH₄/6-CP overlay well in BmQueE and BsQueE showing that substrate binding is highly conserved amongst the homologs and are not affected by the variations to the AdoMet radical core fold that are observed. The EcQueE structure gives a view of QueE active site void of substrate. This active site shows the potential to bind substrate and taken with the substrate bound structures of BmQueE and BsQueE, underscores the conformational changes that QueE undergoes upon substrate binding. For example, in order for QueE to bind substrate, there needs to be a general clamping down of the active site where binding residues move towards the active site.

In contrast to the BmQueE electrostatic surface, which showed reduced positive patches in the vicinity of the AdoMet radical cluster, the surface behind the AdoMet radical cluster corresponding to the loops and N-terminal ends of α₁-α₄, of BsQueE is positive. The electrostatic surface of EcQueE shows a number of positive patches corresponding to the cluster-binding loop and the N-terminal region of the protein. It is hypothesized that the alterations to
the to the cluster-binding loop in *Bm*QueE and the AdoMet core in the three QueE homologues, as well as the differences in electrostatic surface potentials, may play a part in specific binding of the biological reductant, flavodoxins. In fact, the deviations occur in and around the secondary structures previously predicted to form the flavodoxin binding site and observed in our binding site predictions; the cluster-binding loop, as well as L2 and L4 (loops following β2 and β4). Consistent with this hypothesis, increased production of CDG in *Bm*QueE is observed in the presence of the chemical dithionite than in the presence of the biological reductant system, *E. coli* flavodoxin—flavodoxin reductase. This is in direct contrast to what has been detected with other AdoMet radical enzymes, where the *E. coli* reductant system promotes robust turnover in non-cognate systems *in vitro*. *Bs*QueE shows increased activity in the presence of *E. coli* flavodoxin-flavodoxin reductase than in the presence of dithionite alone, but the most robust activity is displayed in the presence of its cognate flavodoxin—flavodoxin reductase system, *Bs*Fld—flavodoxin reductase system.

The increase in negative electrostatic charge surface of *Bm*QueE makes it compatible with its highly positive cognate flavodoxins, *Bm*Fld, and less so with the highly negative *Ec*FldA. Moreover, the structural variations to *Bm*QueE, in particular, within the cluster-binding loop, could further occlude binding of *Ec*FldA and act to fine-tune the interactions between *Bm*QueE and *Bm*Fld. It is possible that the insertion after β2 that perturbs the overall architecture of *Bm*Fld creates a binding site that accommodates the cluster binding loop insertion of *Bm*QueE. These corresponding changes in surface electrostatics and structure of *Bm*QueE and *Bm*Fld could solidify a lock and key type mechanism for interaction, setting up a strong dependence on both charge and surface complementarity for binding. These structural variations could explain the lack of activation observed when *Ec*FldA was used with *Bm*QueE. Even with the long chain insert, the structure of *Ec*FldA and *Bs*Fld should be very similar, as well as their electrostatic surface charge potentials. Based on charge-charge complementarity, both *Bs*QueE and *Ec*QueE appear compatible with *Ec*FldA, consistent with the biochemical data. The unique structural feature of each QueE or AdoMet radical enzyme may provide further fine-tuning of the protein-protein interaction, thus allowing the cognate flavodoxin to activate the enzyme more robustly than the non-cognate flavodoxins. Although *Ec*FldA can bind and activate *Bs*QueE, it is not quite
a correct fit, allowing for suboptimal activation. This could explain why the cognate \( Bs\text{Fld-BsQueE} \) interaction promotes the highest activity.

**Conclusion**

Although variations to the \( \text{AdoMet} \) radical core has been observed before outside of the QueE system, these changes have been attributed to tailoring of the enzyme to the chemistry performed and/or substrate binding. Structural analysis of three QueE orthologs, which perform identical chemistry on the same substrate, revealed both structural and electrostatic differences. We believe these variations serve to dictate binding to their cognate biological reductant. Charge-charge complementarity could serve as hard discriminant, preventing flavodoxins with incompatible charges surfaces from binding to QueE. Surface complementarity (dictated by the structure) can further fine-tune these interactions, allowing for activation of the enzyme. It is only when there is both charge and surface complementarity does full activation of the enzyme occur. Thus we expect some sort of co-evolution of flavodoxins-ligand pairs, to allow for the complementarity needed for optimal activation.

**Experimental procedures**

*Purification of EcQueE*

\( Ec\text{QueE} \) was purified as previously described for \( Bs\text{QueE} \).^42

*Crystallization and Data collection of EcQueE*

Crystallization conditions for His\text{tagged} \( Ec\text{QueE} \), were initially identified by sparse matrix screening within a room temperature MBraun anaerobic chamber using a TTP Mosquito pipetting robot and optimized by sitting drop vapor diffusion within a Coy scientific anaerobic chamber. Data quality crystals were obtained by equilibrating drops containing 1.5 \( \mu \text{L} \) of protein (10 mg/mL in 50 mM Tris pH 8.0 and 10 mM dithiothreitol) and 0.5 \( \mu \text{L} \) of reservoir (175-200 mM magnesium chloride, 25-30% PEG 400 and 100 mM TRIS pH 8.5) over a reservoir of 500 \( \mu \text{L} \). Brown 200-300 \( \mu \text{m} \times 30 \mu \text{m} \) rod like-crystals were obtained after 24 hours. Crystals were harvested from the mother liquor with no further cryoprotecting and cryo-cooled in liquid nitrogen within the Coy anaerobic chamber.
Diffraction data was collected at the Advanced Photon Source (Argonne, IL) at beamline 24-ID-C, using a Pilatus 6M pixel detector at 100 K. Data was collected on the same crystal at two different wavelengths. An Fe-peak data set was collected in six 35° wedges using an inverse beam method (Friedel mates were measured consecutively, rotating the crystal 180° every 120 frames with 0.3° oscillation steps and an exposure time of 0.3 s) at a wavelength of 1.7384 Å to 2.6-Å resolution. The remote data set was collected at a wavelength of 0.9792 Å to 2.1-Å resolution, using the continuous vector scan method (the crystal was continuously translated along it major crystallographic axis during data collection). All data was processed in the HKL2000 in the space group P2₁2₁2₁.

Structure determination and refinement
The structure of EcQueE was solved using Fe multi-wavelength anomalous dispersion (MAD) technique and contains two molecules per asymmetric unit. Two Fe sites were identified using the remote and peak data sets trimmed to 4-Å resolution in ShelxD/E⁴³ in HKL2MAPS with occupancies above 0.9. Heavy atom site refinement, experimental map generation, automated model building and density modification were performed in SOLVE and RESOLVE in Phenix AutoSol⁴⁴. The figure of merit-weighted electron density map (FOM=0.64 to 4-Å resolution) obtained was sufficient for tracing protein secondary structure elements manually in coot. The automated model was extensively rebuilt to produce a model for one monomer in the asymmetric unit. The second monomer was placed in the asymmetric unit using RESOLVE in Phenix AutoBuild⁴⁵ and the resulting model was subjected to iterative rounds of refinement, and density modification using xtriage, phenix.refine⁴⁶ and the resolution was extended to the full-length of the data, 2.1-Å resolution. The resulting R-factors were 25.6% and 29.9% working and free R-factors respectively.

Iterative rounds of model building in coot and refinement in Phenix⁴⁶ using atomic coordinates, atomic displacement parameters (B-factors) and non-crystallographic symmetry (NCS) restraints, without sigma cutoffs, completed the model. In advanced stages of refinement, water molecules were manually added in coot⁴⁷ and in final stages, NCS restraints were released and refinement included translation, libation, screw (TLS) parameterization with one TLS group per monomer. The model was validated using simulated annealing composite omit maps calculated in Phenix.
Analysis of geometry using MolProbity\textsuperscript{48} indicated that 96.45\%, 3.55\%, and 0.0\% in the favored, allowed and disallowed regions of the Ramachandran plot respectively. The final structure of \textit{EcQueE} contained 224 residues (out of 243) and a [4Fe-4S] cluster in chain A and 229 residues (out of 243) and a [4Fe-4S] cluster in chain B. In both chains, the His\textsubscript{6} tag linker region containing the Tobacco Etch Virus (TEV) protease cleavage site is visible as well as a His residue from the His\textsubscript{6} tag tag. Crystallography software packages were compiled by SBGrid\textsuperscript{49}.

\textit{Manual docking of AdoMet and substrates}

Docking of AdoMet, 6-CP and CPH\textsubscript{4} molecules into \textit{EcQueE} was performed manually in coot using \textit{BmQueE} (PDB ID 4NJI, 4NJG) as a guide. In \textit{BsQueE} (PDB ID 5TH5) the 6-CP—dAdo binding foretold the binding interactions of AdoMet and substrate, therefore the adduct was used, in addition to \textit{BmQueE} (PDB IDs 4NJI, 4NJG), to configure intact AdoMet, 6-CP and CPH\textsubscript{4} in the active site of \textit{BmQueE}.

\textit{Homology modeling of flavodoxins from \textit{Burkholderia multivorans} and \textit{Bacillus subtilis}}

Homology modelling of flavodoxins, \textit{BsFld} and \textit{BmFld}, were calculated using Modeller software\textsuperscript{52}. For \textit{BsFld}, a flavodoxin from \textit{Streptococcus pneumoniae} (PDB ID 5LJL and sequence identity 42\%) was used as template for generating the homology model, where as the best template determined by Modeller for the construction of \textit{BmFld} was a flavodoxin from \textit{Aquifex ecolicus} (PDB ID 2ARL and sequence identity of 29\%).

\textit{Docking with FRODOCK}

Protein-protein interactions between QueE and flavodoxin were predicted using the fast rotational docking method in FRODOCK\textsuperscript{37}, which takes into account shape complementarity, van der Waals potential, electrostatic potentials, desolvation potentials and coarse-grained knowledge-based protein-docking potentials. Monomeric models of \textit{BmQueE}, \textit{BsQueE} and \textit{EcQueE} were generated by deleting chain B of each structure. The homology models of \textit{BsFld} and \textit{BmFld} generated in Modeller and \textit{EcFldA 1AHN} were used in the docking studies.

The top 40 predicted complexes for each QueE-flavodoxin cognate pair, as scored by FRODOCK, were evaluated based on the distance between the closest methyl group of the
dimethyl group end of the FMN cofactor of flavodoxin and the closest atom of the [4Fe-4S] cluster of QueE. Predictions which placed FMN further than 20 Å away from the AdoMet radical cluster were discarded and the reasonable models were further evaluated based on positioning of the Flavodoxin molecule. Predicted models in which flavodoxin occluded the binding site of QueE were also discarded. Out of the resulting models for each QueE-flavodoxin, the two models with the shortest distance between cofactors were used as the models for discussing QueE-flavodoxin interactions in this chapter.

Preparation of Figures and electrostatic surfaces
All crystallographic figures were created with PyMOL Software and electrostatic surface potentials were calculated using the Adaptive Poisson-Boltzmann Solver plugin implemented in PyMOL, using default parameters\textsuperscript{50-52}.
Figure IV.1. Biosynthetic pathway of preQ₀, the common precursor of 7-deazapurines. 7-cyano-7-deazapurine, preQ₀ is biosynthesized from guanosine triphosphate (GTP) through the actions of four enzymes. In the first step, GTP cyclohydrolase I (GCH I) converts GTP to 7,8-dihydroneopterin triphosphate (H₂NTP). In the second step, QueD converts H₂NTP into 6-carboxy-5,6,7,8-tetrahydropterin (CPH₄), which then undergoes a ring contraction to form 7-carboxy-7-deaguanine (CDG). This quintessential ring contraction is performed by a metal-dependent, AdoMet radial enzyme, CDG synthase, QueE. QueC converts CDG to preQ₀, which is further modified to obtain 7-deazapurines such as the tRNA modifications, queuosines and archaeosines, as well as natural products such as toyocamycin.
**Figure IV.2. QueE sequence alignment.** Sequences include QueE from *Burkholderia multivorans* ATCC 17616 (A0A0H3KB22), *Rhodoth errum marinus* ATCC 43812 (D0MDN0), *Clostridium acetobutylicum* ATCC 824 (Q97D55), *Candidatus Methanoperedens nitroreducens* (A0A062UW24), *Bacillus subtilis* str 168 (O31677), *Agrobacterium fabrum* str C58 (A9CF16), *Rhodopirellula baltica* DSM 10527 (Q7UVG8), *Methanospirillum hungatei* JF-1 ATCC 27890 (Q2FS67), *Nitrososphaera gargensis* (K01746), *Escherichia coli* str K12 (P64554), *Buchnera aphidicola* subsp. *Schizaphis graminum* str Sg (Q8K9D9), *Methanothrix soehngenii* ATCC 5969 (F4BU01). UniprotKB identification numbers are shown in brackets for the corresponding sequences. The cluster binding motif residues are boxed in yellow, whereas residues involved in AdoMet binding are indicated by a blue circle (single residue) and blue boxes (sequential residues) and the AdoMet binding motifs are labeled. Residues involved in CPH4/6-CP ligation are denoted by purple rectangles (single residue) and purple boxes (sequential residues). The protein residues involved in binding the catalytic magnesium ion are indicated by a green circle.
Figure IV.3. Sequence similarity network of the AdoMet radical enzyme subfamily, 7-carboxy-7-deazaguanine synthases (QueE). The protein sequence similarity network\textsuperscript{53} for the QueE AdoMet radical subfamily, obtained from the Structure Function Linkage Database (http://sfld.rbvi.ucsf.edu/django), is visualized in Cytoscape\textsuperscript{54}. Each node represents sequences that share 50% identity or higher and node connections are filtered at a Blast Probability of 10\textsuperscript{-25}. Nodes representing \textit{B. multivorans}, \textit{B. subtilis} and \textit{E. coli} QueE are shown as coral diamonds and sequences used in the sequence alignment (Figure IV.2) are shown as lilac circles.
Table IV.1. Data collection and refinement statistics for EcQueE structures.

<table>
<thead>
<tr>
<th></th>
<th>EcQueE Fe Peak¹</th>
<th>EcQueE Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beamline</td>
<td>APS 24-ID-C</td>
<td>APS 24-ID-C</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.7384</td>
<td>0.9792</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Cell dimensions (Å)</td>
<td>70.3, 74.5, 103.7</td>
<td>70.3, 74.5, 103.7</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.0-2.60 (2.69-2.60)</td>
<td>50.0-2.10 (2.18-2.10)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>32,108</td>
<td>31,076</td>
</tr>
<tr>
<td>Completeness</td>
<td>99.4 (98.4)</td>
<td>96.0 (76.0)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.8 (2.8)</td>
<td>8.6 (5.1)</td>
</tr>
<tr>
<td>R₁ (0.132)</td>
<td>0.132 (0.411)</td>
<td>0.153 (0.489)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>10.4 (3.1)</td>
<td>17.7 (2.3)</td>
</tr>
<tr>
<td>CC1/2</td>
<td>(0.806)</td>
<td>(0.879)</td>
</tr>
</tbody>
</table>

**Model Refinement**

<table>
<thead>
<tr>
<th>Resolution limits (Å)</th>
<th>50.0-2.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rwork/Rfree ³</td>
<td>20.5/23.8</td>
</tr>
<tr>
<td>No. molecules in asu</td>
<td>2</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>3516</td>
</tr>
<tr>
<td>[Fe-4S]</td>
<td>16</td>
</tr>
<tr>
<td>Water</td>
<td>220</td>
</tr>
<tr>
<td>B-factors (Å²)</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>27.7</td>
</tr>
<tr>
<td>[4Fe-4S]</td>
<td>26.4</td>
</tr>
<tr>
<td>Water</td>
<td>27.8</td>
</tr>
<tr>
<td>R.M.S. deviations</td>
<td></td>
</tr>
<tr>
<td>Bond Lengths, (Å)</td>
<td>0.004</td>
</tr>
<tr>
<td>Bond Angles (°)</td>
<td>0.770</td>
</tr>
<tr>
<td>Rotamer outliers (%)</td>
<td>0.54</td>
</tr>
<tr>
<td>Ramachandran Plot (%)</td>
<td></td>
</tr>
<tr>
<td>Most Favored</td>
<td>96.45</td>
</tr>
<tr>
<td>Additionally allowed</td>
<td>3.55</td>
</tr>
<tr>
<td>Disallowed</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Highest-resolution shell is shown in parentheses

¹ Bijvoet pairs were scaled separately in this data set.

² $R_{sym} = \frac{\sum_{hkl} \sum_{i=1} |I_{hkl}^i - \overline{I}_{hkl}|}{\sum_{hkl} \sum_{i=1} |I_i|}$

³ R-factor = $\sum(|F_{obs}|-k|F_{calc}|)/\sum(|F_{obs}|)$ and R-free of the R value for a test set of reflections consisting of 5% of the diffraction data not used in refinement.
<table>
<thead>
<tr>
<th>QueE ortholog</th>
<th>PDB ID</th>
<th>Ligands bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>BmQueE</td>
<td>4NJG</td>
<td>AdoMet and 6-CP</td>
</tr>
<tr>
<td></td>
<td>4NJH</td>
<td>AdoMet, CPH₄, and Na⁺</td>
</tr>
<tr>
<td></td>
<td>4NJI</td>
<td>AdoMet, CPH₄, and Mg²⁺</td>
</tr>
<tr>
<td></td>
<td>4NJJ</td>
<td>AdoMet, CPH₄, and Mn²⁺</td>
</tr>
<tr>
<td></td>
<td>4NJK</td>
<td>AdoMet, CDG, and Mg²⁺</td>
</tr>
<tr>
<td>BsQueE</td>
<td>5TGS</td>
<td>Methionine</td>
</tr>
<tr>
<td></td>
<td>5TH5</td>
<td>6-CP—dAdo ester adduct</td>
</tr>
<tr>
<td>EcQueE</td>
<td>none</td>
<td>none</td>
</tr>
</tbody>
</table>

Table IV.2. Available X-ray crystal structures of QueE.
Figure IV.4. Structure of QueE from *Escherichia coli*. (A) Structure of EcQueE, shown in cartoon, folds into a head-to-tail functional dimer with the dimer interface composed of interactions between the N-terminal (light pink) and C-terminal (grey) extensions. The modified AdoMet core, a partial (β6/α5) TIM barrel, is shown in blue. (B) EcQueE (blue) monomer overlays well with the monomers of BsQueE, PDB ID: 5TH5 (translucent light green) and BmQueE, (PDB ID 4NJI), (translucent yellow). In both panels, [4Fe-4S] clusters are shown in a ball and stick representation, where iron is colored orange and sulfur is colored yellow.
Figure IV.5. The overall structure of QueE homologs and PFL-AE. Topology diagrams of structurally characterized QueE enzymes, (A) EcQueE, (B) BmQueE and (C) BsQueE are shown. The core AdoMet domains are colored blue for EcQueE, yellow for BmQueE and green for BsQueE whereas the N- and C-terminal extensions are colored light pink and grey (respectively) for all three structures. The differences between the QueE homologs structure are shown in bold and the corresponding secondary structure element denoted in magenta and the dashed line delineates the dimer interface. (D) The topology diagram of PFL-AE is shown with the N- and C-terminal extensions colored pink and slate respectively and AdoMet domain colored in coral. The iron atoms of the [4Fe-4S] clusters are colored orange and sulfur atoms are colored yellow. Cysteine ligands to the [4Fe-4S] cluster are shown as yellow circles. Structural elements outside the AdoMet radical core fold are labeled with a prime.
Figure IV.6. AdoMet binding pocket in QueE homologs. AdoMet binding within the AdoMet core (translucent cartoons) is shown for (A) BmQueE (PDB ID 4NJI), (B) BsQueE (PDB ID 5TH5) and (C) EcQueE. The binding pockets are composed of residues (sticks), which can provide hydrogen bonds (red) to AdoMet (white). The irons (orange) and the sulfurs (yellow) of the [4Fe-4S] AdoMet radical cluster are shown as spheres. In (B), the intact AdoMet molecule is modeled using the adenosyl moiety of the 6-carboxypterin-5'-deoxyadenosyl ester adduct (PDB ID 5TH5) and an intact AdoMet molecule (PDB ID 4NJI) as a guide. The AdoMet binding pocket of EcQueE (blue) (C) is shown overlaid with the binding pocket from BmQueE (white) to highlight the changes that need to be made (red arrows) to allow binding of the modeled AdoMet (white) molecule.
Figure IV.7. Substrate binding pocket. Residues (sticks) that comprise the substrate-binding pocket are shown for each of the QueE orthologs. (A) 6-CP binds to BmQueE (PDB ID 4NJG) through interactions from the AdoMet radical core (yellow) and the N- and C-terminal extensions (light pink and grey, respectively). (B) 6-CP and AdoMet were modeled into the active site of BsQueE (PDB ID 5TH5) using the 6-CP—dAdo ester adduct as a guide. The AdoMet radical core of BsQueE is colored in green and the N- and C-terminal extensions are colored in pink and grey, respectively. (C) The active site of EcQueE is shown overlaid with the
active site of 6-CP bound BmQueE (PDB ID 4NJG) (white). EcQueE is colored with the AdoMet radical domain in blue, the N-terminal extension in pink and the C-terminal extension in grey. (D) CPH₄ is bound to the active site by residues from the N-terminal extension (pink), the AdoMet radical core fold (yellow) and the C-terminal extension (grey) of BmQueE (PDB ID 4NJ1). (E) In the modeled orientation, CPH₄ appears to interact with the C-terminal domain (green) of BsQueE in addition to the N- and C-terminal extensions, pink and grey, respectively. (F) The active site of EcQueE, AdoMet radical domain in blue and N- and C-terminal extensions in pink and grey, respectively, is shown overlaid with the active site of BmQueE (white). The substrates, 6-CP and CPH₄, are shown in lilac, the irons (orange) and the sulfurs (yellow) of the [4Fe-4S] AdoMet radical cluster are shown as spheres, AdoMet is shown in light blue and the hydrogen bonds are shown as red dashes.
Figure IV.8. Magnesium binding in QueE homologs. (A) CPH$_4$ binding in BmQueE (PDB ID 4NJI) (yellow) creates a magnesium-binding site. (B) CPH$_4$ binding in BsQueE (PDB ID 5TH5) (green) is expected to create a magnesium-binding site similar to that seen in BmQueE. (C) The magnesium-binding site of EcQueE (blue) is shown overlaid with the CPH$_4$ bound BmQueE (PDB ID 4NJI) (white). The catalytically essential magnesium is represented as a green sphere, AdoMet is shown in white and CPH$_4$, in lilac. Water molecules (red spheres) necessary for magnesium binding are shown and hydrogen bonds are shown as red dashes.
Figure IV.9. Flavodoxins sequence alignment. Sequences include flavodoxins from Helicobacter pylori, Escherichia coli, Anacystis nidulans, Aquifex aeolicus, Desulfovibrio gigas, Clostridium beijerinckii, Streptococcus pneumonia TIGR4, Bacillus subtilis (YkuN), Bacteroides fragilis NCTC 9343, and Burkholderia multivorans. The sequence alignment is colored according to secondary structure, blue for β-strands and α-helices and the insertion for long chain flavodoxins and the chain insertion in flavodoxins from B. multivorans and B. fragilis are denoted with a box.
Figure IV.10. Overall structure of Flavodoxins. The overall structure of (A) EcFldA (1AHN), (B) BsFld (homology model) and (C) BmFld (homology model) are shown in cartoon. The flavin mononucleotide (FMN) cofactor is colored in yellow. (D) A closer look at the FMN binding site of the overlaid flavodoxin models, EcFldA (green), BsFld (cyan) and BmFld (magenta) is shown. The FMN molecules are colored corresponding to the enzyme.
Figure IV.11. Electrostatic surface charge for QueEs and the cognate Flds. The solvent accessible electrostatic surface representations of QueEs, (A) BmQueE, (B) BsQueE and (C) EcQueE are displayed with the cluster binding loop pointing out of the page. (D) BmFld homology model with FMN colored pink, (E) BsFld homology model with FMN colored cyan and (F) EcFldA structure (PDB ID 1AHN) with FMN colored green are also displayed as electrostatic surface representations of the cognate flavodoxins. Electrostatic potentials are depicted on a colorimetric scale from red to blue for -1 to +1 kTe⁻¹.
Figure IV.12. Docking models of EcFldA X-ray structure to EcQueE X-ray structure. (A) Two of the top docking results of EcFldA (α trace), lilac and magenta, to EcQueE (cartoon) are shown. EcQueE is colored with the AdoMet radical core domain in blue and N-terminal extension in light pink and C-terminal extension in grey. (B) The top predicted interactions between EcFldA (α trace), lilac and magenta to EcQueE, shown as an electrostatic surface charge representation with electrostatic potentials colored from red to blue for $-1$ to $+1 \text{ kT}^{-1}$. (C) A zoomed in view of the predicted interface of the protein-protein interaction between EcFldA (lilac), and EcQueE (blue). (D) The predicted interface between EcFldA (magenta), and EcQueE colored with the AdoMet radical core domain in blue and the N- and C-terminals in light pink and grey, respectively. The secondary elements of EcQueE are labeled with the cluster-binding loop denoted as CBL. FMN cofactor (sticks) is colored based on the corresponding EcFldA model (lilac and magenta) and the black dashed lines indicate the distance from the FMN in EcFldA to the [4Fe-4S] in EcQueE. The [4Fe-4S] cluster is shown as a ball and stick representation with iron in orange and sulfur in yellow.
Figure IV.13. Docking models of BsFld homology model to BsQueE X-ray structure. (A) Two of the top predicted protein-protein interactions between BsFld (cα trace), orange and yellow, and BsQueE (cartoon) colored with the AdoMet radical domain in green, N-terminal extension in light pink and C-terminal extension in grey. In (B) the top two docking models of BsFld (cα trace), orange and yellow, are shown relative to the surface of BsQueE. The electrostatic surface charge representation of BsQueE is shown with the electrostatic potentials colored from red to blue for -1 to +1 kT e\(^{-1}\). (C) A zoomed in view of an interface between one of the BsFld docking results (yellow), and BsQueE is shown with the AdoMet radical core domain and the N-terminal extension of BsQueE colored in green and light pink, respectively. (D) A zoomed in view of the second interface between the BsFld docking results (orange), and BsQueE is shown. The AdoMet radical core domain and the N-terminal extension of BsQueE are colored in green and light pink, respectively. The secondary elements of BsQueE are labeled with the cluster-binding loop denoted as CBL. FMN cofactor (sticks) is colored based on the corresponding BsFld model (orange and yellow) and the distance from FMN to the [4Fe-4S] in BsQueE is displayed as black dashes.
Figure IV.14. Docking models of BmFld homology model to BmQueE X-ray crystal structure. (A) Two of the top predicted docking results of BmFld (ca trace), teal and slate, to BmQueE (cartoons) are shown. BmQue. (B) The top two docking models of BmFld (ca trace), teal and slate, are shown relative to the surface of BmQueE. BmQueE is shown as an electrostatic surface charge representation, with the electrostatic potentials colored from red to blue for -1 to +1 kTe$^{-1}$. (C) A closer look at the binding interface of one of the predicted interactions between BmFld (teal) and BmQueE. (D) The second predicted protein-protein interface between BmFld (slate) and BmQueE is shown. BmQueE is colored the AdoMet radical domain in yellow, N-terminal extension in light pink and C-terminal extension in grey. The secondary elements of BmQueE are labeled with the cluster-binding loop denoted as CBL. FMN cofactor (sticks) is colored based on the corresponding BmFld model (slate and cyan) and the distance (black dashes) from the FMN in BmFld to the [4Fe-4S] in BmQueE is shown.
References


Chapter V
Structural characterization of an AdoMet Radical enzyme involved in sactionine bond formation

Chapter Summary
Sactipeptides are a subclass of ribosomally synthesized and post-translationally modified peptides, which contain a unique thioether bond, called a sactionine linkage, between the sulfur atom of a cysteine residue and the alpha-carbon of an acceptor residue. Sactionine linkages are imperative for the spermicidal, antifungal and antibacterial properties of sactipeptides and are installed by members of the S-adenosyl-L-methionine radical enzyme family. In particular, the enzymes involved in sactionine linkage synthesis are predicted to be members of the SPASM and Twitch subfamily of AdoMet radical enzymes and contain multiple iron-sulfur clusters necessary for catalysis. One hypothesis is that the additional cluster is used to ligate the precursor peptide. Here we present the structure of an AdoMet radical enzyme involved sactionine linkage formation in the maturation of the sactipeptide sporulation killing factor. SkfB contains a C-terminal Twitch domain, which binds an auxiliary cluster, and a N-terminal domain, which is reminiscent of the RiPP recognition elements (RRE) involved in peptide binding.

Contributions
Nathan Bruender carried out construct preparation and protein purifications. We would like to thank Nathan Bruender and Vahe Bandarian for stimulating discussion feedback and troubleshooting of the data in this chapter.
Introduction

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a large class of structurally and functionally diverse natural products produced by all kingdoms of life, and can be categorized into subclasses based on the structural modifications they contain\textsuperscript{1-3}. In general, RiPP biosynthetic pathways encode a precursor peptide(s), and enzymes involved in modifying/tailoring the precursor peptide, exporting of the mature RiPP and in some cases regulation\textsuperscript{4-6}. The biosynthesis of a RiPP natural product commences with the post-translational modification of the precursor peptide core, which is guided by interactions between the leader peptide sequence of the precursor peptide and the modifying enzymes. The modifying enzymes recognize the leader peptide sequence by interactions with a conserved structural motif called RiPP recognition element (RRE)\textsuperscript{7}. The leader peptide sequence has been shown to bind to this motif by extending the anti-parallel β-sheet or wing of the wing helix turn helix (wHTH) RRE motif and is thought to guide the core peptide into the active site of these enzymes for modification\textsuperscript{8-9}.

A number of S-adenosyl-L-methionine (AdoMet) radical enzymes have been identified in RiPP biosynthetic pathways and are implicated in performing challenging and fascinating rearrangements in the maturation of RiPP natural products, such as thioether bond formation, epimerization and carbon-carbon bond formation\textsuperscript{10-11}. Members of the AdoMet radical superfamily proceed through a radical mediated mechanism, initiated by the reductive cleavage of an AdoMet molecule directly coordinated to an essential [4Fe-4S] cluster (the AdoMet radical cluster)\textsuperscript{12-14}. The direct ligation of AdoMet to the cluster allows for reductive cleavage of the C5'-S bond AdoMet to produce the highly reactive 5'-deoxyadenosyl radical (5'-dAdo\textsuperscript{•}) intermediate. 5'-dAdo\textsuperscript{•} subsequently abstracts a hydrogen atom from substrate forming 5'-deoxyadenosine and a substrate radical, which undergoes further transformation to form the desired product. A number of features for the binding of the invariable [4Fe-4S] cluster and AdoMet molecule have been identified and are conserved amongst members of the superfamily\textsuperscript{15-17}. In most AdoMet radical enzymes, the AdoMet radical cluster is coordinated by three cysteines, arranged in a CX\textsubscript{3}CXφC motif\textsuperscript{18}. AdoMet provides the fourth ligand to the cluster through a bidentate coordination using the amino and carboxylic groups of the amino moiety. Coordination to the cluster situates AdoMet in the cavity of the AdoMet radical.
domain. The core of AdoMet radical enzymes adopts a partial (β/α)_6 Triose isomerase phosphate (TIM) barrel fold which positions a number of conserved AdoMet radical motifs to orient AdoMet within the barrel\textsuperscript{16-17}. Sulfur to alpha-carbon cross-linked peptides (sactipeptides) is a subclass of RiPP natural products characterized by the presence of a thioether bond between the sulfur atom of a cysteine residue and the α-carbon of an acceptor residue, a so-called sactionine linkage (Figure V.1B). The sactionine linkages is installed by a co-localized AdoMet radical enzymes, herein called sactisynthases\textsuperscript{19} (Figure V.1A)\textsuperscript{19-24}, which have been predicted to, in addition to an AdoMet radical core, contain an N-terminal peptide binding domain with an RRE motif, and a C-terminal SPASM/Twitch domain (Figure V.2)\textsuperscript{25-26}. The SPASM subclass was named after the four biochemically characterized enzymes AlbA, PqqE, anSME and MfrC, involved in the maturation of peptide based products, subtilosin A, pyrroloquinoline quinone, anaerobic sulfatase maturating-enzyme and mycofactocin, respectively (Figure V.3)\textsuperscript{4-5,25}. These enzymes contain a C-terminal domain with a conserved seven-cysteine motif (CX_{9-15}GX_{4}C_{gap}CX_{2}CX_{5}CX_{3}C_{gap}C) predicted to coordinate two additional or auxiliary clusters, resulting in an open coordination site on one of the auxiliary cluster for binding substrate. The Twitch domain is a truncated version of the SPASM domain that binds only one auxiliary cluster\textsuperscript{27}. SPASM/Twitch domains, initially visualized in the structures of anSME (SPASM), butirosin biosynthetic enzyme, BtrN (Twitch) and molybdenum cofactor biosynthetic enzyme MoaA (Twitch) (Figure V.3), utilize a β-hairpin followed by an α-helix motif to ligate two or one auxiliary cluster, respectively\textsuperscript{27-29}. In anSME and BtrN, the auxiliary clusters were fully ligated, whereas an open conformation site for substrate-binding was found in MoaA\textsuperscript{30}. The fully ligated auxiliary one (Aux 1) cluster of anSME and BtrN was proposed to oxidize the product radical to generate the product during the final steps of catalysis. Therefore, two roles for auxiliary clusters are possible: to bind substrate and/or to oxidize the product radical\textsuperscript{28}. The difference in the number of auxiliary clusters between anSME and BtrN, two versus one, was proposed to allow electrons to reach the protein surface from the buried active site of anSME\textsuperscript{25,27-28}. Hence, it is not clear why two clusters are needed if Aux I coordinates the substrate.
Sactisynthases were proposed to use Aux I to ligate and possibly activate the cysteine donor residue of the precursor peptide during catalysis due to the fact that AlbA only contained seven cysteine residues to ligate two clusters. Recently, a structure of a sactisynthase, CteB, involved in sactinione formation in the SCIFF peptide system revealed an open coordination site on Aux I\textsuperscript{11}. In the crystal structure, density was observed near the unique iron for a cysteine residue from the N-terminal fragment of the precursor peptide, CteA, used in crystallography. This structure provided the first look at a SPASM enzyme with a site differentiated auxiliary cluster bound to substrate and sheds light on the cluster coordination environment of sactisynthases.

In this work we present the structure of a Twitch containing sactisynthase SkfB. SkfB creates a sactinione linkage between the cysteine residue, C\textsubscript{4}, and the \(\alpha\)-carbon of M\textsubscript{12} of the precursor peptide, SkfA (Figure V.1B), during the maturation of the sactipeptide sporulation killing factor in \textit{Bacillus subtilis}. The structure of SkfB with AdoMet bound closely resembles the structures of a RiPP biosynthetic enzyme involved in streptide biosynthesis, SuiB (Figure V.3), as well as the sactisynthase CteB. The structure reveals an unexpected auxiliary cluster bound to the Twitch domain with an open coordination site.

**Results**

\textit{Overall architecture of SkfB is modular.}

The structure of SkfB from \textit{Bacillus subtilis} in complex with AdoMet was solved to 1.29-Å resolution using iron single-wavelength anomalous dispersion technique (Table V.1). The final structure contained one molecule in the asymmetric unit with residues 12-321 and 330-401, a molecule of AdoMet bound to a [4Fe-4S] AdoMet radical cluster and [2Fe-2S] cluster. The SkfB architecture is modular, folding into three distinct domains, an N-terminal peptide binding domain, an AdoMet radical domain and a C-terminal Twitch domain, where the C-terminal domain wraps back around to the N-terminal end of AdoMet radical domain interacting with both the AdoMet radical domain and the N-terminal domain (Figure V.4).

\textit{SkfB adopts a canonical AdoMet radical domain.}

The central domain of SkfB (residues 106-321) folds into a partial (\(\beta/\alpha\))\textsubscript{6} triose phosphate isomerase (TIM) barrel, with 6 parallel \(\beta\)-strands comprising the inner face of the barrel and 6 \(\alpha\)-
helices flanking the outside of the barrel, providing flexibility to the rigid barrel architecture (Figure V.4A-C). This architecture is characteristic of AdoMet radical enzymes and provides a binding site for cofactors and substrates in a protected cavity and facilitates the corresponding radical mechanisms of members of the superfamily. Herein, we will refer to the partial (β/α)₆ TIM barrel architecture as the AdoMet radical domain. Residues from the canonical CX₃CXϕC motif, C₁₁₇, C₁₂₁ and C₁₂₄, are located on the loop following β₁ and ligate the essential [4Fe-4S] cluster, the AdoMet radical cluster, leaving a unique iron for direct coordination of AdoMet (Figures V.4A-C, V.5A).

In the structure of SkfB, the α-amino and α-carboxyl groups of AdoMet are found 2.3 Å and 2.2 Å, respectively, away from the site differentiated iron the AdoMet radical cluster. Conserved interactions from the AdoMet radical domain helps position AdoMet in a catalytically competent orientation in the active site. Hydrogen bonds from the conserved GGE motif, G₁₆₀, G₁₆₁, and G₁₆₂, and R₂₂₃ position the α-amino and α-carboxyl moieties of AdoMet. The ribose motif, S₂₁₁, and residue R₂₂₃, located adjacent to helix α₄a, position the hydroxyls of the ribose ring in the active site and is thought to guide the 5’-deoxyadenosyl radical upon homolytic cleavage of AdoMet. The adenine moiety of AdoMet is oriented in the active site through hydrophobic interactions from the GXIXGXXE or β₅-motif, T₂₅₁, and from the CX₃CXϕC motif, F₁₂₃. Polar backbone interaction from the β₆-motif, L₂₅₈, and from the hydrophobic residues of the CX₃CXϕC motif, F₁₂₃, and Y₁₂₅ (Figure V.6) further help to orient AdoMet in the active site.

The Twitch domain binds a [2Fe-2S] auxiliary cluster.

Following helix α₆, the SkfB structure wraps back to the N-terminal of the AdoMet radical domain and places β₁’ of the C-terminal domain adjacent to β₁, extending the inner parallel β-sheet by one strand (Figure V.4B). Most of the linker between the AdoMet radical domain and the C-terminal domain is disordered in SkfB and the first visible residue is 330. The C-terminal folds into a 3-stranded anti-parallel β-sheet, β₁’-β₃’, followed by two α-helices, a short α₁’ and a longer α₂’. (Figures V.4A,B,D). This domain ligates a [2Fe-2S] auxiliary cluster, Aux I, using three cysteine ligands (Figure V.5B), one from the loop following β₂’, C₃₅₁, and two from the loop following α₂’, C₃₈₅ and C₃₈₇ (Figure V.4A,B,D), reminiscent of the SPASM/Twitch
This C-terminal cluster binding architecture is expected, as SkfB is a member of the SPASM/Twitch subclass of AdoMet radical enzymes (Figure V.2, V.7A). The identity of the auxiliary cluster as a [2Fe-2S] cluster, however, is a surprise and it is unclear whether the presence of this cluster type is an artifact or a physiologically relevant cluster.

Members of the SPASM/Twitch subfamily use a C-terminal domain to bind either one auxiliary cluster, such as in the founding Twitch enzymes, BtrN (Figure V.7 B) and MoaA (Figure V.7C), or two auxiliary clusters as in anSME (Figure V.7D), PqqE (Figure V.7E), CteB (Figure V.7F) and the SuiB (Figure V.7G) 7-2, 1-33. The SPASM/Twitch domain architecture consists of a β-hairpin, β1’ and β2’, followed by α2, which intersperses the cysteine ligands to the clusters (Figure V.7); ligands to both auxiliary clusters are found just before and just after the β-hairpin and on the C-terminal loop following α2. Although they use a very similar architecture, members of the SPASM/Twitch superfamily differ in ligation environment of the auxiliary clusters and the identity of the clusters.

The SPASM enzymes anSME, SuiB and PqqE, and the Twitch enzyme BtrN have fully ligated auxiliary clusters, whereas the twitch enzyme MoaA, and the SPASM enzyme CteB have an open coordination site on Aux I, which ligates the corresponding substrate. The current structure of SkfB reveals an open coordination on Aux I, which was surprising as the sequence of SkfB predicts five cysteines in the C-terminal region, four of which are conserved cluster ligands (Figure V.7H). The fourth conserved cysteine, C333, was expected to provide the ligand to the cluster before the β-hairpin in SkfB and is visible about 12 Å away from the cluster on the linker region that connects the AdoMet domain to the Twitch domain (Figure V.8).

The structures of SkfB and PqqE both revealed a [2Fe-2S] Aux I cluster bound by a Twitch and a SPASM domain, respectively, (Figure V.4D, V.9A-D). The [2Fe-2S] cluster observed at the Aux I cluster-binding site of PqqE was unexpected and as with SkfB, we are not sure what it means. Both enzymes utilize a CXC sequence, C385 and C387 in SkfB and C323 and C325 in PqqE, following helix α2 to ligate the visualized [2Fe-2S] cluster, which provides ligands to each of the irons of the cluster and has not been seen in the other SPASM/Twitch enzymes structurally characterized (Figure V.9A,C). When the Aux I sites of SkfB and PqqE are overlaid with a
Twitch and SPASM enzyme, respectively, we find that the observed [2Fe-2S] cluster is superimposed with the Aux I [4Fe-4S] cluster and we note that these cluster-binding sites could accommodate [4Fe-4S] clusters with minor perturbations (Figure V.9).

The N-terminal peptide-binding domain contains a RiPP recognition element motif.

The N-terminal domain of SkfB (residues 12-105) sits below the N-terminal side of the AdoMet domain and the C-terminal Twitch domain (Figure V.4A,B). It folds first into a three-stranded anti-parallel β-sheet, β1''-β3'' followed by helical bundle comprised of four consecutive α-helices, α1''-α4'', which stacks against the lower face of the anti-parallel β-sheet (Figure V.4E). Following α4'', SkfB folds into a fourth β-strand, β4'', before connecting to the AdoMet radical domain. It is the upper face of the anti-parallel β-sheet that forms the interface with the AdoMet radical domain and Twitch domain.

The N-terminal domain architecture of SkfB (Figure V.4E, V.10A) is reminiscent of a peptide binding protein that is involved in the biosynthesis of pyrroloquinoline quinone PqqD (Figure V.10B) and the peptide binding domains of a nisin biosynthetic enzyme, NisB (Figure V.10C) and a cyanobactin biosynthetic enzyme, LynD (Figure V.10D), an architecture termed the RIPP recognition element (RRE)8-9, 34-35. The RRE is responsible for binding the substrate precursor peptide in RIPP biosynthetic enzymes and folds into a winged helix-turn-helix (wHTH) motif composed of a three-stranded anti-parallel β-sheet and three α-helices. In SkfB, the anti-parallel β-sheet, β1''-β3'', and α1''-α3'' constitute the RRE motif. As visualized in NisB and LynD, the leader sequence of the precursor peptide binds to the RRE motif by forming a β-strand, extending the anti-parallel β-sheet or the “wing” (Figure V.10C,D). This interaction positions the leader peptide into the active site of RIPP modification enzymes.

Structural comparison of AdoMet radical RIPP biosynthetic enzymes.

In addition to SkfB, the structure of two other AdoMet radical enzymes involved in the biosynthesis of RiPPs have been determined, CteB (Figure V.11A) and SuiB (Figure V.11B). Similar to SkfB, both CteB and SuiB are members of the SPASM/Twitch subfamily and adopt a modular overall architecture comprised of three domains; an N-terminal peptide binding domain, a central AdoMet radical domain and a C-terminal SPASM domain, in the place of the C-
terminal Twitch domain visualized in SkfB. CteB and SuiB structures contained three [4Fe-4S] clusters including the AdoMet radical cluster, bound by the canonical CX₃CXφC motif, and two auxiliary clusters, Aux I and Aux II, bound by the SPASM domain. In addition, the linker between the AdoMet radical domain and the SPASM domains were visualized in these structures. In SuiB, this linker region donates the first cysteine ligand to Aux I, resulting in two fully ligated auxiliary clusters. In CteB, Aux I has an open coordination site, which is ligated by a cysteine from the N-terminus of the peptide substrate CteA. Although having slightly different C-terminal domains, the AdoMet radical domain and SPASM domains of CteB and SuiB overlay well with the AdoMet radical and Twitch domains of SkfB structure (Figure V.11C,D). The main difference between SkfB and CteB or SuiB is the ligation of Aux I and also the location of the N-terminal peptide-binding domain.

Both N-terminal domains adopt a typical RRE fold. As aforementioned for NisB and LynD, the leader peptide of CteA adopts a β-strand conformation and interacts with β3” of CteB, extending the antiparallel β-sheet. The N-terminal domain of CteB is found on the opposite side of the AdoMet radical domain than in SkfB, interacting with the C-terminus of the AdoMet radical and SPASM domains instead (Figure V.11C). A long flexible linker between β1 of AdoMet radical domain and α3” of the N-terminal domain traces the outer surface CteB’s SPASM domain. Surprisingly the leader peptide, SuiA, binds to SuiB without interacting with the RRE domain. Instead SuiA interacts with the AdoMet radical domain, the C-terminal domain and the linker between these two domains. The N-terminal of SuiB is located below the N-terminus of the AdoMet radical and SPASM domain, similar to SkfB. Shorter linkers between the peptide-binding domain and the AdoMet radical domain are observed in SuiB (5 residues) and SkfB (9 residues) than is seen in CteA (22 residues).

**Discussion**

Here we present the structure of an AdoMet radical Twitch enzyme, SkfB, involved in maturation of a sactipeptide natural product. The structure of SkfB allows us to compare and contrast the structural differences between the previously solved SPASM domain containing RiPP biosynthetic enzymes, SuiB and CteB, and also presents another data point for teasing out
information on the mechanisms of precursor peptide binding and catalysis in sactionine linkage synthesis.

In terms of the overall structure, SkfB resembles SuiB and CteB and adopts a modular architecture. The Twitch domain of SkfB overlays well with the SPASM domains of CteB and SuiB, but ligates one auxiliary cluster, Aux I, instead of two, a noted difference between SPASM and Twitch enzymes. Surprisingly, the SkfB structure revealed a [2Fe-2S] Aux I cluster ligated by a CX\textsubscript{33}CXC sequence. This cluster identity was unexpected as the initial biochemical and spectroscopic characterization proposed that SkfB contained a [4Fe-4S] auxiliary cluster ligated by a CX\textsubscript{4}CXC sequence. Moreover, most of the SPASM/Twitch enzymes characterized to date contain [4Fe-4S] auxiliary clusters. However, the recent structure of the SPASM enzyme PqqE revealed a [2Fe-2S] cluster in the Aux I site, which has been reported to be able to bind both [4Fe-4S] and [2Fe-2S] clusters. Interestingly, like SkfB, the two-cysteine ligands to Aux I after helix \( \alpha 2' \) are provided by a CXC sequence, a motif found primarily in the coordination of [2Fe-2S] clusters\textsuperscript{36-37}. Because of the proximity of the cysteine ligands in CXC sequences, both cysteines are expected to ligate the same iron of the [2Fe-2S] clusters. This is not the case in SkfB or PqqE; the CXC sequence ligates a different iron atom of the [2Fe-2S] cluster. This coordination geometry could allow for [4Fe-4S] cluster binding in SkfB and PqqE with, minimal perturbations of the Aux I site and in both cases, the site appears to be amenable to rearrangement. The malleability of this site could explain why different cluster identities have been observed in the biochemical and structural data. It is possible that the CXC sequence might be indicative of a physiologically relevant [2Fe-2S] cluster or it could make a [4Fe-4S] cluster more susceptible to degradation. To understand which species is catalytically relevant, enzyme activity needs to be correlated to cluster content.

The coordination state of the auxiliary clusters of members of the SPASM and Twitch family has long been of debate. Although both anSME and BtrN structures revealed fully ligated clusters, an insufficient number of cysteines in the AlbA C-terminal sequence (seven cysteines to ligate two auxiliary clusters) led to proposition that AlbA, and other sactisynthases, may utilize an open coordination site for substrate ligation. It was also hypothesized that these enzymes could utilize a non-cysteine ligand as has been seen in LipA (Ser)\textsuperscript{38-39}, BioB (Arg)\textsuperscript{40} and recently PqqE.
(Asp)$^{32}$, as the final ligand to the cluster$^{41}$. On the other hand, the C-terminal Twitch domain of SkfB contains five conserved cysteines, four of which correspond with conserved auxiliary cluster ligands in both Twitch and SPASM domains. Therefore, the structures of both SPASM and Twitch sactisynthases were highly anticipated in hope they would shed light on the cluster coordination conundrum. The first structure of a sactisynthase, CteB, revealed that in fact there is an open coordination site on Aux I of the SPASM domain whereas Aux II is fully ligated. Interestingly, extra density inconsistent with a protein residue was observed near the site-differentiated iron of Aux I and was attributed to a cysteine residue (cysteine 3 of CteA in Figure V.1B) from the N-terminal fragment of CteA co-crystallized with CteB. The N-terminal fragment used in co-crystallization encompassed the leader peptide sequence (residues -18 to -1) and the first three residues of the core peptide (residues 1-3) of CteA (Figure V.1B, Table V.2) but only residues -18 to -10 and 2 to 3 were ordered in the crystal. This cysteine ligation to Aux I could foreshadow binding of the reacting cysteine (cysteine 14) in the presence of full length CteA (Figure V.1B). Electrochemical analysis of BtrN revealed a very low potential -765 mV for the fully ligated Aux I cluster$^{42}$. It is possible that the difference in coordination environments of the Aux I cluster, fully ligated versus site differentiated, could modulate the redox potentials of Aux I, tuning it to the needs of the enzyme.

The structure of SkfB showed only three ligands to Aux I resulting in an open coordination site although all four SPASM/Twitch cluster ligands to the Aux I site were conserved. In fact, the position of the open coordination site in SkfB corresponds to the cluster ligand before the β-hairpins (denoted as C1), which has been observed in all SPASM and Twitch enzymes structurally characterized thus far (Figure V.7). The equivalent cysteine in the SkfB, C$_{333}$, is the third visible residue on the C-terminal side of the disordered linker region between the AdoMet radical domain and C-terminal domain and is ~12 Å away from the cluster. Therefore, to say with any certainty where the location of C$_{333}$ is, a structure of SkfB with an intact linker is needed (Figure V.8).

The three RiPP biosynthetic AdoMet radical enzymes structurally characterized thus far present three different orientation of the RRE containing peptide-binding domain. In the structures of SkfB and SuiB, the peptide-binding domain sits below the N-terminal of the AdoMet radical
core and SPASM/Twitch domain whereas the peptide domain of CteB interacts with helix α6 of CteB on the C-terminal side of the AdoMet radical core. This difference in peptide binding orientation between the three could be substrate and structure specific and the mobility of the peptide-binding domain might be important for positioning substrate. It is likely that SkfA will bind to SkfB using the leader peptide sequence to extend the anti-parallel β-sheet of the RRE, as seen with CteB. Whether a conformational change of the RRE domain occurs upon SkfA binding remains to be seen.

**Conclusion**

The structure of SkfB revealed a Twitch domain-containing sactisynthase, which catalyzes thioether bond formation, similar to SPASM domain-containing sactisynthases CteB and AlbA. As with anSME and BtrN, the CteB and SkfB structures highlight that both the SPASM and Twitch fold can be used to catalyze similar chemistry. It awaits to be seen what the physiological auxiliary cluster identity is and its coordination environment.

**Experimental Methods**

*Purification of SkfB*

SkfB was purified as previously described.\(^1\)

*Crystallization and data collection of SkfB*

SkfB was crystallized within a MBraun anaerobic chamber at 21°C and under a N\(_2\) atmosphere using the sitting drop vapor diffusion method. Initial crystals of SkfB were obtained by sparse matrix screening using a TTP (LabTech) mosquito pipetting robot. To obtain data quality crystals 1.5 μL of SkfB [10 mg/mL, 50 mM PIPES pH 7.4, 150 mM KCl, 10 mM DTT and 5 mM AdoMet] was mixed with 0.5 μL of precipitant [0.25 M magnesium formate and 15% PEG 3350] and equilibrated over a 500-μL reservoir. Brown crystals grew within 24 hours and were transferred to a Coy scientific anaerobic chamber (95% Ar and 5% H\(_2\) atmosphere) for harvesting. To cryoprotect SkfB crystals, the crystals were transferred in 4 steps into a cryoprotectant solution containing 0.2 M magnesium formate, 20% PEG 3350 and glycerol. Upon each transfer, the glycerol concentration was increased to a final concentration of 20%. Crystals were then flash frozen in liquid nitrogen within the Coy chamber.
An anomalous data set and a native data set of SkfB were collected on the same crystal. The anomalous data set was collected to a resolution of 1.9 Å using an in-house Cu-Kα rotating anode source (Rigaku) with an imaging plate (RAXIS IV, Rigaku). In order to obtain phasing information, the dataset was collected using pseudo inverse beam in four 90°-wedges with 1° increments. Following the home data collection, the crystal was shipped to the Advanced Photon Source (Argonne, IL) where a native data set was collected to a 1.29-Å resolution at beamline 24-IDC using a Pilatus 6M pixel detector at a temperature of 100 K and a wavelength of 0.9792 Å. All data were processed in HKL 2000. A native data set was collected to a 1.29-Å resolution at beamline 24-IDC at the Advanced Photon Source (Argonne, IL) using a Pilatus 6M pixel detector at a temperature of 100 K and a wavelength of 0.9792 Å. All data were processed in HKL 2000 in C2 symmetry spacegroup.

Structure Determination and refinement
The structure of SkfB with AdoMet was solved by iron single-wavelength anomalous dispersion (SAD) method in a C2 spacegroup that contained one molecule in the asymmetric unit. Using anomalous data trimmed to 2.5-Å resolution, 6 Fe sites (corresponding to one [4Fe-4S] cluster and one [2Fe-2S] cluster) were identified in ShelxD\textsuperscript{44} in HKL2MAP and refined in SHARP/autosharp\textsuperscript{45-46}. The output experimental maps, with an initial Figure of Merit (FOM) of 0.545 to 2.5 Å resolution, were of sufficient quality to trace the secondary structure of the AdoMet radical partial (β/α)\textsubscript{6} TIM barrel domain and the C-terminal Twitch domain, using BtrN (PDB ID 4M7T) and MoaA (PDB ID 1TV8) as guides. This initial model was used to define the solvent boundary during a subsequent round of solvent flattening in SOLOMON\textsuperscript{47} and phase extension to 2.0-Å resolution. Loop regions and side chains with visible density were built into the resulting electron density.

When a near-completed model of SkfB was obtained, containing ~350 residues out of 410, an AdoMet molecule and the AdoMet radical cluster, the structure was refined against the full range of the native data set (50.0-1.29 Å resolution) using rigid body, atomic coordinates and atomic displacement (B-factor) refinements in Phenix\textsuperscript{48}. The resulting $R_{\text{work}}$ and $R_{\text{free}}$ were 24.4% and 29.5% respectively. The model was completed by iterative rounds of model building in coot\textsuperscript{49}.

166
and refinement in Phenix\textsuperscript{48}. Waters were added manually in coot during final rounds of refinement and disordered side chains were truncated to the last atom with discernable density. The final model contained residues 12-321 and 330-401, one molecule of AdoMet, a [4Fe-4S] AdoMet radical cluster, and a [2Fe-2S] auxiliary cluster.

Composite omit maps calculated in Phenix\textsuperscript{48} were used to validate the final structure and MolProbity\textsuperscript{50} was used to analyze the model geometry. MolProbity indicated that 97.92\% of residues were found in favored region, 2.08\% in the allowed region and 0.0\% in the disallowed region of the Ramachandran plot and \textasciitilde97\% of residues have favorable rotamers. Figures were generated in PyMol\textsuperscript{51-52}. Crystallography software packages were compiled by SBGrid\textsuperscript{53}.
**Figure V.1. Sactipeptide biosynthesis.** (A) Several RiPP biosynthetic gene clusters involved in the biosynthesis of sactipeptides have been identified. These biosynthetic clusters encode precursor peptides (teal), AdoMet radical enzymes (red) responsible for installing the unique sactionine linkages characteristic of the sactipeptide natural products, peptidases (blue) responsible for cleaving the leader peptide and ATP–binding cassette transporters involved in transporting the sactipeptide out of the producing organism. (B) The precursor peptides for each biosynthetic gene cluster in (A) are shown with the leader peptide, which binds to the biosynthetic enzymes, in grey and the mature peptide, which undergoes post-translational modifications in teal. The sactionine linkages formed between the sulfur atom of a cysteine residue and the α-carbon of an acceptor residue is shown in red. Final maturation steps of SboA and SkfA involve a N- to C- terminal macrocyclization (dashed blue lines) step concurrent with a leader peptide cleavage step. Maturation of SkfA also includes disulfide bond formation (purple).
Figure V.2. SPASM/Twitch subclass sequence similarity network. The protein similarity network for the SPASM/Twitch subclass is visualized in Cystoscape at a blast probability of $10^{-20}$. Sequences were obtained from the Structure Function Linkage Database (http://sfld.rbvi.ucsf.edu/django) and each node represents sequences that share 50% identity. Nodes corresponding to the previously solved members of the SPASM/Twitch subclass are represented as blue diamonds. SkfB, which is structurally characterized in this chapter, is shown as a purple diamond and AlbA is identified as a green triangle.
Figure V.3. Reactions of SPASM/Twitch subclass members. The sactisynthases AlbA and SkfB catalyze thioether bond (sactione linkage) formation on their precursor peptides SboA and SkfA during the maturation of the sactipeptides subtilosin A and sporulation killing factor respectively. Whereas, anSME is involved in installing the catalytically essential formylglycine in the active site of sulfatases. MoaA catalyzes the first step of molybdopterin biosynthesis, a complex rearrangement of guanosine triphosphate (GTP) and BtrN catalyzes the dehydrogenation of 2-deoxy-scyllo-inosamine (DOIA) during the biosynthesis of butirosin. In the first step of pyrroloquinoline quinone (PQQ) maturation, PqqE, in the presence of the peptide binding protein PqqD, forms a new carbon bond between glutamate and tyrosine residues of PqqA. Similarly, SuiB installs a lysine to tryptophan cross-link on SuiA in the biosynthesis of the RiPP natural product streptide. AlbA, anSME, PqqE and SuiB are SPASM enzymes whereas MoaA, BtrN and SkfB are Twitch enzymes.
Table V.1. Data collection and refinement statistics for SkfB structure

<table>
<thead>
<tr>
<th></th>
<th>SkfB</th>
<th>SkfB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fe Peak</td>
<td>Remote</td>
</tr>
<tr>
<td>Beamline</td>
<td>Saturn 944</td>
<td>APS 24-ID-C</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.5418</td>
<td>0.9792</td>
</tr>
<tr>
<td>Space group</td>
<td>C2</td>
<td>C2</td>
</tr>
<tr>
<td>Cell dimensions (Å)</td>
<td>76.5, 83.1, 61.3</td>
<td>76.5, 83.1, 61.3</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.0-1.97 (2.00-1.97)</td>
<td>50.0-1.29 (1.34-1.29)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>52,216</td>
<td>92,230</td>
</tr>
<tr>
<td>Completeness</td>
<td>97.8 (69.0)</td>
<td>96.5 (76.2)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.6 (2.4)</td>
<td>5.6 (2.6)</td>
</tr>
<tr>
<td>( R_{\text{sym}} )²</td>
<td>0.043 (0.119)</td>
<td>0.037 (0.499)</td>
</tr>
<tr>
<td>( I/\sigma(I) )</td>
<td>29.5 (7.7)</td>
<td>43.6 (1.8)</td>
</tr>
<tr>
<td>CC1/2</td>
<td></td>
<td>(0.729)</td>
</tr>
</tbody>
</table>

**Model Refinement**

- Resolution limits (Å): 50.0-1.29
- \( R_{\text{work}}/R_{\text{free}} \)³: 15.91/18.09
- No. molecules in asu: 1
- No. atoms:
  - Protein: 2,932
  - [Fe-S]: 12
  - AdoMet: 27
  - Water: 465
- \( B \)-factors (Å²):
  - Protein: 21.3
  - [Fe-S]: 18.2
  - AdoMet: 14.0
  - Water: 33.7
- R.M.S. deviations:
  - Bond Lengths (Å): 0.014
  - Bond Angles (°): 2.789
  - Rotamer outliers (%): 2.58
  - Ramachandran Plot (%):
    - Most Favored: 97.92
    - Additionally allowed: 2.08
    - Disallowed: 0.0

---

1 Bijvoet pairs were scaled separately in this data set.

2 \( R_{\text{sym}} = \frac{\sum_{hkl} \sum_{i=1} \left| I_{i}^{hkl} - \overline{I}_{hkl} \right| / \sum_{hkl} \sum_{i=1} | I_{i} |}{2} \)

3 \( R\)-factor = \( \frac{\sum_{i} (|F_{\text{obs}}| - |k|F_{\text{calc}}|)/\sum_{i} |F_{\text{obs}}|} \) and \( R\)-free of the \( R \) value for a test set of reflections consisting of 5% of the diffraction data not used in refinement.
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Description</th>
<th>Substrates bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>anSME</td>
<td>SPASM sulfatase-maturating enzyme</td>
<td>AdoMet, residues 9-11 of a (18mer) sulfatase peptide mimic</td>
</tr>
<tr>
<td>MoaA</td>
<td>Twitch molybdenum cofactor biosynthetic enzyme</td>
<td>AdoMet, GTP</td>
</tr>
<tr>
<td>BtrN</td>
<td>Twitch butirosin biosynthetic enzyme</td>
<td>AdoMet, 2-deoxy-scyllo-inosamine (DOIA)</td>
</tr>
<tr>
<td>CteB</td>
<td>SPASM sactipeptide maturating enzyme</td>
<td>AdoMet, residues -18 to -10 and 2 to 3 of a (18mer) N-terminal CteA peptide none</td>
</tr>
<tr>
<td>PqqE</td>
<td>SPASM pyrroloquinoline quinone biosynthetic enzyme</td>
<td>AdoMet, residues -13 to -1 of the precursor peptide of SuiA</td>
</tr>
<tr>
<td>SuiB</td>
<td>SPASM Streptide biosynthetic enzyme</td>
<td>AdoMet</td>
</tr>
<tr>
<td>SkfB</td>
<td>Twitch sactipeptide maturating enzyme</td>
<td>AdoMet</td>
</tr>
</tbody>
</table>

Table V.2. Structurally characterized members of the SPASM/Twitch AdoMet radical subclass.
Figure V.4. Overall architecture of SkfB. (A) The partial \( (\beta/\alpha)_6 \) TIM barrel of the AdoMet radical domain (magenta) houses the [4Fe-4S] AdoMet radical cluster and an AdoMet molecule (lilac). The C-terminal adopts a Twitch fold architecture (green) providing three visible ligands to a [2Fe-2S] auxiliary cluster, Aux I. The N-terminal peptide-binding domain (yellow) displays the architecture of the RiPP recognition element (RRE) motif identified in peptide-binding domains of RiPP biosynthetic enzymes. (B) Topology diagram of SkfB colored by domain; the AdoMet radical domain in magenta, the Twitch domain in green and the RRE domain in yellow. Cysteine residues are indicated by yellow circles and the iron and sulfur atoms of the clusters are shown as orange and yellow circles, respectively. The clusters are shown as ball and stick representations with iron atoms colored orange and sulfur atoms colored yellow. (C-E) show a closer look at the individual domains of SkfB. (C) The active site of SkfB is located within the inner cavity of the AdoMet radical domain (magenta) and is comprised of six parallel \( \beta \)-strands.
The binding sites of the AdoMet radical cluster and AdoMet (lilac) are located at the top of the barrel. (D) The C-terminal Twitch domain (green) contains the canonical elements of a SPASM/Twitch fold, a β-hairpin, β1' and β2', followed by α2'. In SkfB a short β-strand, β3', and α-helix, α1' are found in between β2' and α2'. (E) The N-terminal peptide-binding domain of SkfB (yellow) folds into a three-stranded antiparallel β-sheet (β1''-β3'') and a consecutive helical bundle (α1''-α4''), reminiscent of the winged helix turn helix (wHTH) motif that comprises the RiPP recognition element. The N-terminal domain ends with a parallel β-strand, β4'' adjacent to β1''. The clusters shown are shown as ball and stick representations with irons atoms colored orange and sulfur atoms colored yellow.
Figure V.5. Iron sulfur clusters and AdoMet density. (A) The characteristic CX$_3$CXφC motif, composed of C$_{117}$, C$_{121}$ and C$_{124}$ in SkfB, coordinates the AdoMet radical cluster. An AdoMet molecule ligates the site-differentiated iron through a bidentate interaction with the nitrogen of the α-amino moiety and oxygen from the α-carboxyl moiety. (B) A [2Fe-2S] cluster is observed in the Aux I site of the Twitch domain, bound by a CX$_3$CXC sequence, C$_{351}$, C$_{385}$ and C$_{387}$. 2Fo−Fc composite omit density is shown in blue and contoured at 1σ.
**Figure V.6. AdoMet binding motifs are conserved in SkfB.** The AdoMet binding pocket is located within the partial TIM barrel (translucent pink) and includes interactions from the GGE motif, G_{160}, G_{161}, E_{162} (cyan), the ribose motif, S_{211} and R_{223} (tan), GXIXGXXE or β5 motif, T_{251} (green) and β6 motif (yellow), L_{281} R_{223} positions both the carboxyl group and the ribose ring. Hydrophobic interactions from the CX{sub}3CXΦC motif, F_{123} and the adjacent residue, Y_{125} (pink), orient the adenine ring of AdoMet (lilac) in the active site. The AdoMet radical cluster is shown in ball and stick representation with irons atoms colored orange and sulfur atoms colored yellow.
Figure V.7. Comparison of auxiliary cluster binding by SPASM/Twitch domains. The canonical structural elements of the Twitch (A-C) and SPASM (D-F) domain from are shown here for each enzyme. (A) SkfB (green) binds a [2Fe-2S] cluster using one cysteine following the β-hairpin and the two cysteines following α2', similar to BtrN. SkfB appears to have an open coordination site; although the cysteine corresponding to the cluster ligand before the β-hairpin
motif is present, this cysteine is ~12 Å away from the cluster. (B) The Twitch domain of BtrN (orange) provides four cysteine residues to fully ligate a [4Fe-4S] auxiliary cluster. (C) MoaA (teal) uses a Twitch domain to bind a [4Fe-4S] cluster with an open coordination site. The unique iron is ligated by substrate, GTP, (teal circle). (D) anSME (purple) utilizes a SPASM motif to bind two fully ligated [4Fe-4S] clusters, Aux I and Aux II. (E) PqqE (blue) bind two fully ligated clusters, a [2Fe-2S] Aux I cluster and a [4Fe-4S] Aux II cluster using a unique aspartic acid (red circle) ligand. (F) CteB (pink) also binds two [4Fe-4S] auxiliary clusters with an open coordination site on Aux I to which a cysteine residue from the substrate, CteA (pink sphere), binds in the structure. (G) The SPASM domain of SuiB ligates two auxiliary clusters. The first cysteine ligand to Aux I is provided by the linker region connecting the SPASM domain to the AdoMet radical domain. The cluster binding cysteines are shown as yellow circles and the cluster iron and sulfur atoms are represented as orange and yellow circles, respectively. The linker regions are denoted in grey. (H) Sequence alignment for the C-terminal of SPASM/Twitch enzymes, with the sequence numbers indicated for SkfB. Cluster coordinating cysteines are denoted in red and cysteines corresponding to the seven-cysteine motif (C_{X_9}G_{X_4}C_{gap}C_{X_2}C_{X_3}C_{X_3}C_{gap}C) are indicated by yellow circles and the conserved glycine residue of this motif is denoted by a purple rectangle. Green boxes denote the CXC of SkfB and PqqE whereas the blue box indicates the cysteine previously proposed to bind the Aux 1 cluster of SkfB (C_{X_4}C_{X_3})^{20}. 

178
Figure V.8. Location of the four expected cysteine cluster ligands to Aux I. Based on sequence analysis, four conserved cysteines are expected to coordinate Aux I of SkfB. Instead, the Aux I cluster-binding site was composed of three cysteine, C_{351}, C_{385}, and C_{387}. The fourth cysteine, C_{333}, is located \(-12\) Å away from the cluster and is one of the first ordered residues on the C-terminal end of the disordered AdoMet radical core domain to Twitch domain linker (residues 322-329 are disordered).
Figure V.9. \([2\text{Fe-2S}]\) cluster binding at the Aux I site of PqqE and SkfB. (A) The Aux I site of the PqqE SPASM domain (green) has been shown biochemically to bind either a \([4\text{Fe-4S}]\) cluster or a \([2\text{Fe-2S}]\) cluster. In the recent structure of PqqE the Aux I cluster-binding site is populated by a \([2\text{Fe-2S}]\) cluster with two of the ligand provided by a CXC motif, C_{323} and C_{32}. (B) anSME (white) contains a \([4\text{Fe-4S}]\) cluster at the Aux I cluster-binding site ligated by C_{255}, C_{261}, C_{276}, and C_{330}. (C) When overlaid with anSME (white), the cysteine positions of the PqqE (green) active site appear to be amenable to binding a \([4\text{Fe-4S}]\) cluster. (D) The Twitch Aux I site of SkfB (green) can bind a \([2\text{Fe-2S}]\) cluster with cysteines from the CXC motif, C_{385} and C_{387}, binding different irons of the cluster. (D) The Twitch Aux I site of Brn (white) fully ligates a \([4\text{Fe-4S}]\) auxiliary cluster. (E) The Aux I site of SkfB (green) overlays well with the Aux I site Brn (white). To allow for \([4\text{Fe-4S}]\) cluster binding, the loop containing residues of the CXC motif of SkfB would have to move slightly. This rearrangement should be possible as this loop corresponds to the C-terminal end of the structure and is highly flexible.
Figure V.10. RRE domains bind the leader peptide of RiPP precursor peptides. The canonical RiPP recognition element (RRE) is comprised of a three strand anti-parallel β-sheet (purple) and three consecutive α-helices (cyan). (A) The N-terminal domain of SkfB folds into a RRE motif comprising of β1''-β3'' and α1''-α3''. Following the RRE, the N-terminal domain folds into an additional α-helix, α4'', and β-strand, β4'', (yellow). (B) The small peptide-binding protein PqqD (PDB ID 5SXY) is a standalone RRE domain. (C) The RRE domain of NisB (PDB ID 4WD9) binds the leader peptide sequence of the peptide substrate NisA (blue) by extending the anti-parallel β-sheet, or the wing. (D) The leader peptide of PatE, (blue) also binds to the RRE domain of LynD (PDB ID 4V1T) through interactions with the wing.
Figure V.11. Structures of other AdoMet radical enzymes involved in RIPP biosynthesis. (A) CteB exhibits a tri-modular fold composed of peptide-binding or RRE domain (yellow) followed by an AdoMet binding domain (magenta), which binds the AdoMet radical cluster and AdoMet (lilac). The C-terminal end of CteB binds two clusters, Aux I and Aux II, using a SPASM domain architecture (green). The leader sequence of CteA (blue) binds to the RRE domain by extending the anti-parallel β-sheet. (B) SuiB demonstrates a similar modular fold to CteB (A). Instead of using the peptide-binding domain (yellow), SuiB binds the leader sequence of SuiA (blue) substrate using the insertion (grey) between the AdoMet radical domain (magenta) and the SPASM domain (green). (C) The AdoMet radical and twitch domains of SkfB (white) overlay well with CteB but the peptide-binding domains are located on opposite sides of the AdoMet radical domain. (D) SkfB (white) overlays well with SuiB and the positions of the
N-terminal domains show modest differences. A view of the overlaid N-terminal domains SkfB and SuiB, rotated 180°, is shown in the inset.
References


Chapter VI

*S*-adenosyl-L-methionine radical enzymes in perspective

Understanding the chemistry performed by *S*-adenosyl-L-methionine (AdoMet) radical enzymes has been an area of active study for more than two decades. Although they appeared initially to be a simpler version of adenosylcobalamin-dependent enzymes, garnering the description ‘a poor man’s adenosylcobalamin’, over the years the AdoMet radical enzymes family has continued to surprise us with it’s wealth of chemistry and structural architecture. With over 113,000 members identified to date, many more surprises lay ahead.

Chemistry of the AdoMet radical enzyme superfamily.

Members of the AdoMet radical superfamily catalyze a wide range of chemistry from complex purine ring rearrangements to methylations and sulfur insertions at saturated carbons using common chemical and structural features. The hallmark of these enzymes is their ability to harness radical chemistry through the reductive cleavage of a molecule of AdoMet directly ligated to a [4Fe-4S] cluster. The canonical CX₃CXDC motif binds the essential [4Fe-4S] cluster with an open coordination site for AdoMet, allowing for the characteristic homolytic cleavage of AdoMet to generate 5’-deoxyadenosyl radical (5’-dAdo•), which initiates catalysis.

Structural characterization of the first members of the superfamily identified a core TIM barrel fold for binding the AdoMet radical cluster, AdoMet and substrates. With these canonical features identified, it may have appeared that the superfamily was somewhat figured out – that was not the case.

Further structural investigations uncovered variations to both the AdoMet radical core fold and the cluster binding motifs. Modified folds were found in a number of AdoMet radical enzymes including the three homologs of the queuosine biosynthetic enzyme, QueE, discussed in Chapters III and IV as well as the Twitch domain-containing butirosin biosynthetic enzyme, BtrN. Most striking was the structure of ThiC, an enzyme involved in thiamin biosynthesis. The ThiC structure portrayed a separate C-terminal domain for cluster-binding, although the protein core adopts a full TIM barrel fold. QueE from *Burkholderia multivorans* also revealed the largest deviation from the cluster binding motif noted to date, a CX₁₄CXΦC motif. However,
characteristic features, such as AdoMet binding and cluster reduction, necessary for the radical-mediated mechanisms of AdoMet radical enzymes, did not appear to be disturbed in these enzymes\textsuperscript{12-15}. On the other hand, although the cobalamin-dependent AdoMet radical enzyme involved in the biosynthesis of thiostrepton A, TsrM, contains a [4Fe-4S] cluster bound by the canonical cluster-binding motif and is able to bind a molecule of AdoMet, it does not appear to perform radical chemistry\textsuperscript{16}. Furthermore, the QueE homolog from \textit{Bacillus subtilis} discussed in chapter III is capable of using the same core fold and active site to catalyze two different reactions, a radical reaction and a polar reaction\textsuperscript{14}. This duality of the AdoMet radical enzyme binding site is also seen in the methyltransferases RlmN and Cfr, which use two molecules of AdoMet per round of catalysis. One molecule is used to methylate an active site cysteine and the second molecule is reductively cleaved to generate a 5'-dAdo\textsuperscript{•}, which abstracts a hydrogen atom for the methylated cysteine to initiate catalysis\textsuperscript{17-18}. Therefore, the ability of AdoMet radical enzymes to ligate an AdoMet radical cluster or adopt a canonical AdoMet radical core fold is not always indicative of a radical mechanism.

AdoMet radical enzymes have been shown to use hybrid mechanisms to generate their products. QueE\textsuperscript{12, 19} and the pyruvate formate lyase activating enzyme (PFL-AE)\textsuperscript{20} both catalyze metal-dependent product formation. In Chapter II, we highlighted the possibility that an AdoMet radical enzyme involved in wybutosine biosynthesis, TYWI, utilizes covalent catalysis in addition to radical chemistry. This is not the only case where another type of catalysis is used in concert with radical chemistry. It is possible that we will uncover more AdoMet radical enzymes that use either metal or covalent catalysis as we continue to explore this superfamily.

**SkfB is a hybrid AdoMet radical enzyme**

In Chapter V, the structure of an enzyme, SkfB, implicated in the formation of an unusual sulfur to \(\alpha\)-carbon thioether bond in the maturation of sporulation killing factor was presented. The overall architecture of SkfB adopts three distinct domains and SkfB appears to be a hybrid AdoMet radical proteins. SkfB has a classic core fold most similar to that of SuiB (rmsd 2.3 Å, PDB ID 5V1T)\textsuperscript{21}, a streptide biosynthetic enzyme, whereas, the C-terminal Twitch domain aligns best with the SPASM domain of a sactisynthase involved in SCIFF peptide maturation,
CteB (rmsd 4.1 Å, PDB ID 5WGG)\textsuperscript{22}. It’s N-terminal RRE domain is reminiscent of the RRE domain of LynD involved in cyanobactin biosynthesis (rmsd 4.0 Å, PDB ID 4V1T)\textsuperscript{23} and the Aux 1 cluster observed is like that of PqqE, a SPASM enzyme involved in the biosynthesis of pyrroloquinoline quinone (PQQ), but has no RRE domain\textsuperscript{24}. One might have expected SkfB to share the closest similarities with other Twitch domain proteins, BtrN\textsuperscript{13} and MoaA\textsuperscript{25}, but this isn’t the case. Also, it might be expected that SkfB would show the highest similarities to other sactisynthases and though there are shared features, the other sactisynthases are SPASM containing proteins and have different Aux 1 coordination environments. Thus, the SkfB structure emphasizes that AdoMet radical enzymes are modular.

The modular nature of AdoMet enzymes allows for customization of domains depending on the reaction catalyzed and the substrate. The classical AdoMet core fold could be swapped for a modified one and the most variability occurs with the N- and C-terminal extensions/domains. The N-terminal region could fold into an extension or a domain, such as an RRE containing domains for a peptide substrate. A C-terminal SPASM domain could be swapped out for a Twitch domain, based on the size of your substrate and/or how many clusters would be need to get electrons to/from the protein surface. Furthermore, the cluster environment could be customized – a fully ligated cluster of a site differentiated one based on the redox potential needed or whether a substrate coordination site is needed. This customizability lends itself to the diverse nature of the AdoMet radical superfamily as a few core building blocks can be combined in different ways to achieve vastly different outcome.

**Prediction of how peptide-binding domains may bind to AdoMet radical enzymes in two-component systems.**

Including SkfB, there are now four structures of AdoMet radical enzymes that are involved in RiPP biosynthesis and are predicted to use an RRE motif to bind their peptide substrate\textsuperscript{21-22, 24}. SkfB, CteB and SuiB all represent a fused or one component system where an N-terminal RRE domain is fused to a SPASM/Twitch AdoMet radical enzyme. The structure of PqqE on the other hand, does not have an N-terminal peptide-binding domain. Instead, a small peptide-binding protein, PqqD, encoded by the *pqq* operon, interacts with PqqE and is believed to bind the
peptide substrate PqqA\textsuperscript{26}. PqqE/PqqD represents a two-component system where the RRE motif is found on a stand-alone protein.

With the one-component enzyme family members, interactions between the peptide-binding domain and SPASM/Twitch-containing AdoMet radical domains have been visualized. These structures could act as a template for predicting binding of the peptide-binding domain to the AdoMet radical biosynthetic enzymes. The different positions of the N-terminal RRE domains in SuiB, SkfB and CteB, gives us a view of a range of interactions that are possible in the two-component systems, like PqqE and PqqD. Using SkfB (rmsd 2.7 Å to PqqE) and CteB (rmsd 10.1 Å to PqqE) as bookends for the positions of the N-terminal peptide domain, two models for PqqD and PqqE interactions were obtained (Figure VI.1). In both models, PqqD appears to able to interact with PqqE with no significant rearrangements of either protein, indicating that either could represent the binding mode of PqqD to PqqE. It is also possible that the interactions between the two-component system will present new binding interactions. Therefore, the structure of a complex between SPASM/Twitch AdoMet radical enzyme and the peptide binding protein is highly anticipated.

In summary, although there is still a lot to be learnt from the AdoMet radical enzyme superfamily we have made some discoveries, which are presented in this thesis and may aid in tackling the structure and chemistry of the AdoMet radical enzyme superfamily.
Figures

Figure V1.1. Modeled interaction of the peptide-binding protein, PqqD, and the AdoMet radical RiPP biosynthetic enzyme, PqqE. (A) Modular architecture of SkfB with the core AdoMet radical domain colored in magenta, the C-terminal Twitch domain in green and the N-terminal binding domain in yellow. (B) The interactions between PqqE and PqqD were modeled based on SkfB (A). PqqE (PDB ID 6C8V) is shown with AdoMet radical core domain in magenta and the SPASM domain in green. The peptide-binding domain PqqD (PDB ID 5SXY) is shown in yellow. (C) Overall architecture of CteB (PDB ID 5WGG) with the core AdoMet radical domain colored in magenta, the C-terminal SPASM domain in green and the N-terminal binding domain in yellow. (D) The interactions between PqqE and PqqD were modeled based on CteB (C). PqqE is shown with AdoMet radical core domain in magenta and the SPASM domain in green. PqqD is shown in yellow.
References

Tsehai A.J. Grell

Department of Chemistry • Massachusetts Institute of Technology, Building 68-688 • 77 Massachusetts Avenue • Cambridge, MA 02139 • (404) 539 4369 • tsehaig@mit.edu

EDUCATION

Massachusetts Institute of Technology, MIT, Cambridge, MA
Ph.D. Candidate, Biological Chemistry
(scheduled) June 2018

Morgan State University, MSU, Baltimore, MD
Bachelor of Science, Chemistry, Summa cum laude
May 2012

Dominica State College, Commonwealth of Dominica
Associate of Arts and Sciences, Biology, Chemistry, Physics
May 2008

PUBLICATIONS


ORAL PRESENTATIONS

MIT-PIER Hamburg Workshop. “Beautiful but delicate: Dissecting the complexities of metalloenzymes.” Cambridge, MA, January 2017


MIT Biology Department Data and Drinks. “S-Adenosylmethionine Radical Superfamily: Understanding function one domain at a time.” Cambridge, MA, December 2014

MSU Undergraduate and Graduate Student Research Symposium. “Metal Assisted and Microwave Accelerated Evaporative Crystallization: 2. The effect of Surface Functionality and Solvent Volume.” Baltimore MD, October 2010

POSTER PRESENTATIONS

MIT Center for Environmental Health Sciences Annual Poster Session. “7-Carboxy-7-deazaguanine synthase: A radical S-Adenosyl-L-methionine enzyme with polar tendencies.” January 2017, Cambridge MA


MIT Department of Biology, Building 68 Annual Retreat. “SPASM and Twitch domains in AdoMet Radical Enzymes.” North Falmouth, MA, June 2015

MIT Summer Research Program Poster Session. “Homology Modeling of Nickel Bound Helicobacter pylori NIKR.” Cambridge, MA, August

NOBCCHE 38th Annual Meeting. “Metal Assisted and Microwave Accelerated Evaporative Crystallization: Applications to Amino Acids.” Houston, TX, April 2011


13th Annual Undergraduate Research Symposium in the Chemical and Biological Sciences- UMBC, “Metal Assisted and Microwave Accelerated Evaporative Crystallization: Applications to Amino Acids.” Baltimore, MD, October 2010


Morgan State University CURE Symposium. “Metal Assisted and Microwave Accelerated Evaporative Crystallization.” Baltimore, MD, July 2010

RESEARCH EXPERIENCE

Graduate Research Assistant, November 2012-Present
Department of Chemistry, MIT
Advisor: Catherine L. Drennan
Thesis: “Structural studies of S-adenosyl-L-methionine radical enzymes involved in tRNA and natural product biosynthesis”

Undergraduate Research Intern, June 2011-August 2011
MIT Summer Research Program (MSRP), MIT
Advisors: Dr. Collin Stultz and Dr. Catherine L. Drennan

Undergraduate Research Assistant, June 2010-May 2012
Department of Chemistry, Morgan State University
Advisor: Kadir Aslan

HONORS, ACTIVITIES AND AWARDS

2018 MIT Martin Luther King Jr. Leadership Award
2017 Dr. Robbin Chapman Excellence Through Adversity Award
MIT Graduate Woman of Excellence
ACS Travel Award Poster Prize, Enzymes, Coenzymes and Metabolic Pathways GRC
2012 Outstanding Teaching Assistant, MIT Chemistry Department
National Science Foundation Fellowship
MIT Office of the Dean of Graduate Education Diversity fellowship
Morgan’s President’s Award for Exceptional Creativity
George H. Spalding Award
Maryland Section of the American Chemical Society
Dean’s Award for Outstanding Student Research, Creativity and Scholarly Achievements

2011
The American Institute for Chemists Award, Morgan State University Department of Chemistry
Undergraduate Ambassador, 38th NOBCCHE Conference
3rd Place winner in the Division of Colloids and Surface Materials ACS Poster Competition, 241st National Meeting.

2010
Most Outstanding Researcher Award, Continuing Undergraduate Research Experience (CURE) Morgan State University
Member of the Golden Key Honor Society

TEACHING AND MENTORING EXPERIENCE

- Stereotype threat training facilitator, Chemistry TA training, Fall 2016 and 2017
- Guest Lecturer for Physical Chemistry, Wellesley College MA, Spring semester 2015, 2016 and 2017
- Undergraduate Student Mentor, MIT, Spring 2015 and Summer 2015
- Program Assistant for the MIT Summer Research Program (MSRP), Summer 2014
- TA for Biophysical Chemistry Techniques, MIT, Spring Semester 2014
- TA for Principles of Chemical Science, MIT, Fall 2012 and Spring 2013
- MSU Chemistry Club Outreach Program, Baltimore, 2010-2012
- High school teacher of Chemistry, Biology, Mathematics and General Science, Dominica, 2008-2009

LEADERSHIP, OUTREACH AND SERVICE EXPERIENCES

- Graduate Student of Color Coalition (GSCC) Co-President, MIT, May 2017- Present
- Academy of Courageous Minority Engineer (ACME), Treasurer and Co-President, MIT, 2016-2016 and 2016-2017
- Black Graduate Student Association (BGSA) Treasurer, MIT, 2014-2016
- i-Trek (non-profit) Fundraising and Trek committee member and Advisory board, MIT, 2013-present
- Chemistry Graduate Student Council (CGSC) Executive board member, MIT, 2012-present
- Morgan State University Chemistry Club Vice-President, MSU, 2011-2012
- La Falaise House Medical Laboratory Medical Lab Assistant, Commonwealth of Dominica, 2009
REFERENCES

Professor Catherine L. Drennan
Departments of Biology and Chemistry
Howard Hughes Medical Institute
Massachusetts Institute of Technology
77 Massachusetts Ave., Rm. 68-680
Cambridge, MA 02139
Phone: (617) 253-5622
E-mail: cdrennan@mit.edu

Professor Vahe Bandarian
Department of Chemistry
University of Utah
315 S 1400 E, RM 3627
Salt lake City, UT 84112
Phone: (801) 581-6366
E-mail: vahe@chem.utah.edu

Professor Collin M. Stultz
Departments of Electrical Engineering and Computer science.
Institute for Medical Engineering and Science
Massachusetts Institute of Technology
77 Massachusetts Ave., Rm. 36-796
Cambridge, MA 02139
Phone: (617) 253-4961
E-mail: cmstultz@mit.edu

Bradley L. Pentelute
Department of Chemistry
Massachusetts Institute of Technology
77 Massachusetts Ave., Rm. 18-596
Cambridge, MA 02139
Phone: (617) 324-0180
E-mail: blp@mit.edu