HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2014 January 18.

Published in final edited form as:

Nature. 2013 June 6; 498(7452): 123–126. doi:10.1038/nature12180.

Structure-guided discovery of carboxy-SAM as a novel metabolite modulating tRNA function

Jungwook Kim[1,*], **Hui Xiao**[2], **Jeffrey B. Bonanno**[1], **Chakrapani Kalyanaraman**[3], **Shoshana Brown**[4], **Xiangying Tang**[1], **Nawar F. Al-Obaidi**[1], **Yury Patskovsky**[1], **Patricia C. Babbitt**[4], **Matthew P. Jacobson**[3], **Young-Sam Lee**[5], and **Steven C. Almo**[1,6,*]

[1]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY 10461, USA.

[2]Department of Pathology, Albert Einstein College of Medicine, Bronx, NY 10461, USA.

[3]Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA 94158, USA.

[4]Department of Bioengineering and Therapeutic Sciences, University of California at San Francisco, San Francisco, CA 94158, USA.

[5]Department of Biology, Johns Hopkins University, Baltimore, MD 21218, USA.

[6]Department of Physiology & Biophysics, Albert Einstein College of Medicine, Bronx, NY 10461, USA.

Abstract

Identifying novel metabolites and characterizing their biological functions are major challenges of the post-genomic era. X-ray crystallography can reveal unanticipated ligands which persist through purification and crystallization. These adventitious protein:ligand complexes provide insights into new activities, pathways and regulatory mechanisms. We describe a new metabolite, carboxy-S-adenosylmethionine (Cx-SAM), its biosynthetic pathway and its role in tRNA modification. The structure of CmoA, a member of the SAM-dependent methyltransferase superfamily, revealed a ligand in the catalytic site consistent with Cx-SAM. Mechanistic analyses demonstrated an unprecedented role for prephenate as the carboxyl donor and the involvement of a unique ylide intermediate as the carboxyl acceptor in the CmoA-mediated conversion of SAM to Cx-SAM. A second member of the SAM-dependent methyltransferase superfamily, CmoB,

Correspondence and requests for materials should be addressed to J.K. (jungwook.kim@einstein.yu.edu) or S.C.A. (steve.almo@einstein.yu.edu).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author Contributions J.K. performed cloning, protein purification, crystallography, and functional assay. H.X. performed MS analysis of the *in vitro* assay. Y.S.L. performed LC-MS analysis of the CmoA-bound ligand and chemical synthesis of Cx-SAM. XT performed the NMR experiments. N.F.A.-O. performed thermal denaturation studies. C.K. and M.P.J. performed computational modeling. S.B. and P.C.B. performed the bioinformatics analysis. J.B.B. and Y.P. assisted in crystallographic validation and analyzed crystallographic ligand binding results, respectively. J.K. and S.C.A. designed the study, analyzed the data and wrote the manuscript.

Atomic coordinates and structure factors for the reported crystal structure are deposited in the Protein Data Bank under the accession code 4GEK.

The authors declare no competing financial interests.

recognizes Cx-SAM and acts as a carboxymethyltransferase to convert 5-hydroxyuridine (ho5U) into 5-oxyacetyl uridine (cmo5U) at the wobble position of multiple tRNAs in Gram negative bacteria[1], resulting in expanded codon-recognition properties[2,3]. CmoA and CmoB represent the first documented synthase and transferase for Cx-SAM. These findings reveal new functional diversity in the SAM-dependent methyltransferase superfamily and expand the metabolic and biological contributions of SAM-based biochemistry. These discoveries highlight the value of structural genomics approaches for identifying ligands in the context of their physiologically relevant macromolecular binding partners and for aiding in functional assignment.

tRNAs contain numerous post-transcriptional modifications, with nearly 100 distinct modifications reported[1]. Nucleotides at the wobble position are the most frequent targets for such modifications, as they confer efficient and accurate pairing between anticodons and cognate codon sequences. For example, wobble uridines in Gram negative bacteria are often modified at C5 to 5-oxyacetyl uridine (cmo5U) (Fig. 1a), which allows multiple tRNAs to decode four of their respective degenerate codons[3]. This expanded recognition results from structural and tautomeric constraints imposed by the 5-oxyacetyl modification[4].

Mutagenesis studies established that genes responsible for chorismate biosynthesis are necessary for cmo5U formation and it was demonstrated that one carbon atom of the oxyacetyl group originates from SAM[5,6]. Gene disruption studies established that two members of the SAM-dependent methyltransferase superfamily (SDMT), CmoA and CmoB, were required for cmo5U formation[2]. Inactivation of *cmoA* resulted in formation of incompletely modified tRNAs, with hydroxy uridine (ho5U) and methoxy uridine (mo5U) observed instead of cmo5U. In *cmoB*-defective mutants, only ho5U was detected. Despite these observations, the roles of CmoA and CmoB in the transformation of ho5U to cmo5U remain unclear. Wobble uridine hydroxylation is dependent on an unidentified enzyme (Fig. 1a).

The New York Structural Genomics Research Consortium (NYSGRC) determined the structure of *Escherichia coli* CmoA, which revealed unexpected electron density features at the predicted SAM binding site. When SAM was modeled, residual electron density suggestive of a carboxylate group was observed adjacent to the S-methyl group. Refinement of this structure at 1.50Å supports a covalent linkage between the S-methyl group and the putative carboxylate, consistent with the previously unknown metabolite carboxy-SAM (Cx-SAM; (2S)-4-[{[(2S,3S,4R,5R)-5-(6-amino-9H-purin-9-yl)-3,4-dihydroxy-tetrahydrofuran-2-yl]methyl}(carboxylatomethyl)sulfonio]-2-ammoniobutanoate) (Fig. 2 and Supplementary Fig. 1). Liquid chromatography–mass spectrometry (LC-MS) analysis of purified CmoA, unambiguously confirmed Cx-SAM as the CmoA-bound ligand which persisted through purification and crystallization (Fig. 3 and Supplementary Fig. 2).

The overall structure of *E. coli* CmoA is similar to that reported for the *Haemophilus influenzae* ortholog (67% sequence identity; PDB entry 1IM8; RMSD of 0.51Å for 222 equivalent C_α atoms)[7]. Retrospective refinement of the *H. influenza* CmoA structure revealed electron density features consistent with Cx-SAM and contacts similar to those observed in the *E. coli* enzyme. These observations support the existence of conserved pathways involving Cx-SAM in the Gram negative bacteria.

The S-carboxymethyl group of Cx-SAM in the *E. coli* CmoA catalytic site forms a bidentate polar interaction with the side-chain guanidinium of Arg-199, which is invariant among CmoA orthologs. The 2'- and 3'-hydroxyl groups of Cx-SAM form hydrogen bonds with the side-chain of Asp-89; the equivalent residue in all other SAM-dependent methyltransferases is aspartate or glutamate. Other highly conserved residues among CmoA orthologs are contributed from helices α1, α6 and α7 (Supplementary Fig. 3), which appear crucial for substrate binding and are not present in other members of the SDMT superfamily. The Cx-SAM binding pocket is largely hydrophobic, with no functionality capable of deprotonating the S-methyl group of SAM closer than 4.6Å. Adjacent to the Cx-SAM binding site is a partially hydrophobic cavity which is the likely binding site for an additional ligand/substrate (**see below**).

Given data implicating chorismate, or a related metabolite, in wobble uridine oxyacetylation in Gram negative bacteria[5,6], chorismate was examined as the potential carboxyl donor. Chorismate supported the CmoA-mediated formation of Cx-SAM as evidenced by LC-MS (Fig. 3b) and MS/MS (Supplementary Fig. 4). In addition to Cx-SAM, phenylpyruvate was produced with similar kinetics (Fig. 3c and Supplementary Fig. 5). The lack of a facile pathway for the direct conversion of chorismate to phenylpyruvate suggested that chorismate undergoes a rearrangement prior to CmoA-catalyzed Cx-SAM formation. Among the several biologically characterized products of chorismate, only prephenate possesses a scaffold consistent with the observed LC-MS and MS/MS fragmentation data.

When tested for *in vitro* production of Cx-SAM, prephenate was a more efficient substrate than chorismate (Fig. 3d) and the lag in Cx-SAM production observed with chorismate was absent. This behavior is consistent with the slow non-enzymatic conversion of chorismate to prephenate, which is utilized by CmoA for Cx-SAM formation (Supplementary Fig. 6). Furthermore, production of phenylpyruvate from prephenate was unambiguously confirmed by NMR and LC-MS of the *in vitro* assay solution (Supplementary Figs 7 and 8). These activities are CmoA-specific as CmoB does not show any Cx-SAM synthase activity (data not shown). We propose that prephenate is the biologically relevant substrate and the source of the carboxylate in the CmoA-catalyzed reaction.

These results are consistent with the CmoA-catalyzed decarboxylation and concomitant loss of hydroxide from prephenate to yield phenylpyruvate, water and carbon dioxide, which is the source of the carboxylate functionality in Cx-SAM (Fig. 1b). In support of this hypothesis, uniformly ^{13}C-labeled chorismate ([U-^{13}C] chorismate) transferred ^{13}C-carbon dioxide to the product, Cx-SAM (Supplementary Fig. 9), unambiguously demonstrating that prephenate is the carboxyl donor in the CmoA-catalyzed formation of Cx-SAM and validating the overall reaction proposed for CmoA. To our knowledge, this is the first example of prephenate serving as the carboxyl group donor in a carboxytransfer reaction.

There is precedence for decarboxylation of prephenate, as prephenate dehydratase (PDT) catalyzes the decarboxylation of prephenate with concomitant loss of hydroxide to generate phenylpyruvate in a fashion similar to that suggested for CmoA[8]. In the PDT-catalyzed reaction, elimination of the hydroxyl group as water is facilitated by the participation of Thr172 as a general acid to protonate the leaving hydroxide group. The threonine side-chain

is not sufficiently acidic to directly protonate the prephenate hydroxyl (i.e., the conjugate acid of the prephenate hydroxyl is expected to behave similarly to an alcohol, with a pK_a of ~-2, while the pK_a of threonine is ~16 in water); thus, other mechanisms must be operative as the major driving force for this reaction. Hilvert and Cleland proposed that geometric distortion drives decarboxylation and the favorable energetics associated with aromatization reduces the bond order of the hydroxyl, shifting its reactivity towards that of hydroxide and allowing for efficient general acid catalysis[8]. Similar considerations are relevant to the CmoA-catalyzed reaction.

Computational docking of prephenate in the CmoA catalytic site suggests a pose in which the carboxylate and hydroxyl groups of prephenate sandwich the S-methyl group of SAM, consistent with the observed transcarboxylation reaction (Fig. 2 and Supplementary Fig. 10). Trievel recently highlighted the importance of carbon-oxygen (CH-O) hydrogen bonds in the recognition and presentation of the S-methyl group in canonical S_N2-based SAM-dependent methyltransfer reactions[9]. In the prephenate-bound model of CmoA (Fig. 2 and Supplementary Fig. 10), the hydroxyl oxygen of prephenate and the S-methyl carbon of SAM form a potential 3.4Å CH···O hydrogen bond. Notably, the prephenate hydroxyl is poorly positioned for an in-line S_N2 attack on the S-methyl group, consistent with the lack of prephenate methylation. Instead, this arrangement suggests a substrate-assisted mechanism in which the departing prephenate OH group abstracts a proton from the S-methyl group of SAM, generating water and the nucleophilic ylide (stabilized carbanion), which is carboxylated by the carbon dioxide (Fig. 1b). Importantly, the pK_a of the trimethylsulfonium cation (a model of the S-methyl group in SAM) has been reported as 18.9 and 18.2 in water and DMSO[10,11], respectively. These values are similar to that of the general acid (Thr172) in the PDT-catalyzed reaction and are consistent with elimination of hydroxide from prephenate as water and formation of the reactive ylide intermediate.

Because of its importance in the mechanism, we sought direct evidence for the formation of the sulfonium ylide. Presentation of CmoA with triply-deuterated SAM and prephenate results in formation of doubly-deuterated SAM, consistent with formation of the ylide intermediate by deuteron abstraction and regeneration of SAM by protonation (Fig. 3f). In the absence of prephenate, only triply-deuterated SAM was observed. In the absence of CmoA, no exchange was observed in mixtures of [$^2H_3$-methyl]-SAM and prephenate. Partitioning of the ylide between product formation (Cx-SAM) and return to substrate (SAM) was quantified by determining the ratio between solvent-exchanged SAM and Cx-SAM: 92.0:8.0 and 97.3:2.7 at pH 7.3 and 8.5, respectively, providing compelling evidence that the postulated ylide intermediate is on the reaction coordinate for Cx-SAM formation.

The most common biological fate of the SAM S-methyl group is intermolecular transfer catalyzed by methyltransferases, including epigenetic marking of DNA and histone targets[12,13], and a wide range of small molecule transformations[14,15]. By analogy, we hypothesized that Cx-SAM is utilized in a CmoB-catalyzed transcarboxymethylation reaction in the biosynthesis of cmo5U. Total RNA[16] and purified tRNAs[2] from cmoB-deficient cells (*i.e.*, ho5U-containing RNA) were used as substrates in *in vitro* assays with prephenate, SAM and CmoA, with or without recombinant CmoB. After the transfer reaction, RNAs were treated with P1 nuclease and the resulting 5'-nucleotide

monophosphates analyzed by MS (Fig. 4). CmoB catalyzed carboxymethyl transfer from *in situ*-generated Cx-SAM to ho5U-containing RNAs, as a species with mass corresponding to oxyacetyl-5-uridine-5'-monophosphate (cmo5UMP) was clearly detected. CmoA alone does not exhibit carboxymethyltransferase activity. Therefore, we conclude that Cx-SAM is the substrate for the CmoB-dependent transcarboxymethylation of ho5U-containing tRNAs (Fig. 1c).

In most Gram-negative species, *cmoA* and *cmoB* are co-conserved and immediately adjacent to each other in the genome, supporting the demonstrated functional relationship. Most importantly, the unique *in vitro* activities assigned to CmoA and CmoB are fully consistent with all reported genetics findings relevant to cmo5U modification[2,5,6]. Our own genetics and mutagenesis studies add further support, as plasmid-based expression of wild-type CmoA restores production of cmo5U in CmoA-deficient (Δ*cmoA*) *E. coli*, while the D89L mutant lacking *in vitro* biochemical activity failed to complement *in vivo* (Supplementary Fig. 11). Finally, *in vivo*- and *in vitro*-generated Cx-SAM-bound CmoA exhibited comparable behavior in the carboxymethylation of ho5U-containing RNAs, supporting the relevance of our *in vitro* studies (see Supplementary Fig. 12).

Elaboration of the S-methyl group of SAM with a variety of functional groups has been pursued by chemical biologists to support studies including the generation of modified nucleic acid[17,18] and protein[19] substrates, as well as genome-wide assignments of protein methyltransferase targets[19]. It is notable that Cx-SAM represents the first SAM derivative demonstrated to have been shaped by evolution for a biologically meaningful function. The electrophilic properties of the SAM sulfonium center have been exploited to realize expanded functional diversity, as exemplified by Cx-SAM, via two unique activities within the SDMT superfamily (Supplementary Fig. 13a). Sequence analysis supports the existence of CmoA orthologs throughout the Gram negative proteobacteria, as well as in the *Verrucomicrobia* and some *Cyanobacteria* (Supplementary Fig. 13b). It remains to be discovered how widespread these mechanisms are and whether there exist additional biologically relevant SAM-analogues.

The fortuitous identification of a bound ligand in a crystal structure (such as Cx-SAM in CmoA) is by no means unusual. We estimate that ~3–5% of all structures determined by the NYSGRC contain organic ligands derived from the expression host (typically *E. coli*), which persisted through purification and crystallization. Frequently this can be anticipated; *e.g.*, finding NAD/NADH bound to targets annotated as oxidoreductases, or PLP in annotated aminotransferases. However, unanticipated ligands are also identified, including nucleotides, amino acids, carbohydrates and lipids bound to proteins from a range of bacterial species, providing clues regarding catalytic activity and biological function (see examples in Supplementary Fig. 14).

In summary, direct structural observation identified the novel metabolite Cx-SAM, leading to the discovery of unique Cx-SAM synthase and carboxymethyltransferase activities involved in tRNA wobble base modification. These findings reveal new functional diversity in the SDMT superfamily, expand the metabolic and biological contributions for SAM-based biochemistry and presage the discovery of new metabolites and biological processes.

This work highlights the power of structural genomics approaches for the discovery of new metabolites, pathways and biology.

Full Methods

Cloning and Protein Purification

The *cmoA* gene was amplified from genomic DNA of *E.coli* BL21 by PCR, cloned into LIC-pET46a (Novagen) and verified by the DNA sequence analysis (Genewiz). *E. coli* BL21 (DE3) cells (Invitrogen) were transformed with vectors harboring the *cmoA* gene, grown in LB containing 100 mg/mL ampicillin at 37°C and induced with 0.5 mM IPTG at an OD_{600} of ~1. Cells were incubated overnight at 25°C and harvested by centrifugation. Cell pellets were resuspended with Bugbuster (Novagen) at room temperature for 30 min, the lysates centrifuged at 41,400*g* for 30 min and the supernatants applied to Ni-agarose (Qiagene) columns pre-equilibrated with buffer A (20 mM HEPES, pH 7.5 and 150 mM KCl). The recombinant protein was eluted with 150 mM imidazole in buffer A and further purified by size exclusion chromatography on a HiLoad Superdex 200 column (GE) equilibrated with buffer A. Final purity was over 95% as verified by SDS-PAGE analysis. For the measurement of OD_{280} of the purified CmoA, the enzyme was denatured to separate SAM/Cx-SAM from the protein by multiple rounds of mixing with 8M guanidine chloride solution followed by spin-filtration. An extinction coefficient of ε_{280} = 18.7 $cm^{-1}mM^{-1}$ was used to calculate the yield of the nucleoside-free CmoA as calculated from the amino acid sequence. The *cmoB* gene was amplified from genomic DNA of *E. coli* BL21 by PCR and cloned into LIC-pET30a (Novagen). The purification of *E. coli* CmoB was identical to that of CmoA, except for the use of kanamycin as the selectable marker. In addition, the affinity tag was removed by thrombin (Novagen) cleavage after elution of the recombinant CmoB from the Ni-agarose resin. The yield was quantitated using an extinction coefficient of ε_{280} = 72.5$cm^{-1}mM^{-1}$, as calculated from the amino acid sequence.

The D89L mutant of CmoA was generated by QuickChange (Stratagene) with primers 5'-TTGCAAAATTATTGCCATCCTCAACTCCCCGGCGATGATT-3' and 5'-AATCATCGCCGGGGAGTTGAGGATGGCAATAATTTTGCAA-3', and the plasmid of the wild-type CmoA as the template for PCR. Purification of the D89L mutant was similar to that of the wild type described above; some purifications included the addition of anion exchange and gel filtration chromatography steps. For anion exchange separation, a MonoQ column (GE) was equilibrated with Buffer A (20 mM Tris-HCl, pH 8.5 and 150 mM KCl) and 1 mL sample was loaded. A linear gradient of Buffer B (20 mM Tris-HCl, pH=8.5 and 1 M NaCl) was used to elute bound protein. Wild-type CmoA eluted as a single peak on MonoQ, while the D89L mutant exhibited two peaks. Eluted proteins were analyzed by SDS-PAGE, pooled, concentrated and loaded on a Superdex75 column (GE), equilibrated with Buffer A.

Crystallization and Structural Determination of CmoA

Purified CmoA was crystallized by sitting drop vapor diffusion at 21°C by mixing 1 μL of the protein with 1 μL of reservoir solution (0.2 M Li_2SO_4, 0.1 M Bis-Tris:HCl, pH 5.5, and 25% PEG3350) and equilibrating over 0.1 mL of reservoir solution. X-ray data were

collected on an ADSC QUANTUM 315 CCD detector at the NSLS beam line X29A and processed with HKL3000[20]. Diffraction data from CmoA crystals were collected at 100K, and at a wavelength λ = 0.9790Å, which were consistent with space group, $P2_12_12$ (a = 65.32, b = 78.68, c = 92.37Å), with two molecules per asymmetric unit. Molecular replacement was performed using the structure of *H. influenzae* YecO (pdb code 1IM8) as the search model with MOLREP[21]. Subsequent model building and refinement were performed with Coot[22] and REFMAC5[23]. The final model was refined to 1.50Å with R_{work} = 0.17 and R_{free} = 0.20 (Supplementary Table I). All residues are in allowed regions and no outliers are found in a Ramachandran plot.

Time Course Assay of Cx-SAM Production

The time-dependent formation of carboxyl-SAM was monitored using [^{14}C-methyl]-SAM with either prephenate or chorismate. The assay mixture contains 20 mM sodium phosphate, pH 6.8, 0.2 mM [^{14}C-methyl]-SAM (Perkin Elmer), and either 0.2 mM chorismate or prephenate. The assay was initiated by adding purified 2 μM CmoA to the assay mixture of 20 μL total volume of. A 2 μL aliquot was periodically withdrawn and mixed with an equal volume of 0.1 M HCl to quench the reaction, then 1 μL was spotted on TLC plate. The TLC was developed with buffer composed of ammonium sulfate/sodium acetate, pH 6.0/iso-propanol (79/19/2). The plate was air-dried and exposed to a phosphor screen imager (GE) for 2 days. The image plate was scanned with Molecular Dynamics Storm 860 PhosphorImager System with ImageQuant software.

Time Course Assay of Phenylpyruvate Production

Phenylpyruvate formation was monitored at λ = 320 nm as described previously[24]. The assay mixture contained 20 mM sodium phosphate, pH 6.8 and 0.2 mM prephenate, with or without 0.2 mM SAM in a total volume of 0.5 mL. The reaction was initiated by adding 2 μM CmoA to the assay solution. A 70 μL aliquot of the reaction mixture was periodically withdrawn and added to 30 μL of 5 M NaOH. The absorbance at λ = 320 nm was measured and the net OD_{320} was recorded by subtracting residual absorbance arising from contaminating phenylpyruvate. The net absorbance was converted to the concentration of phenylpyruvate using a standard curve prepared with commercially obtained phenylpyruvate (Sigma-Aldrich). Non-enzymatic turn-over of prephenate to phenylpyruvate was measured with a sample composed of 20 mM sodium phosphate, pH 6.8, 0.2 mM prephenate, and 0.2 mM SAM in 0.5 mL, without the addition of the enzyme.

Verification of Cx-SAM Co-purified with Recombinant CmoA by LC-MS

A 10 μL recombinant protein solution was diluted with 10 μL water and then with 190 μL of methanol. The mixture was centrifuged at room temperature for 10 min (16,000g), and the supernatant was used for the analysis. For each injection, an 80 μL aliquot was subjected to liquid chromatography–mass spectrometry analysis (Agilent 1200 HPLC coupled with Agilent 6210 AccurateMass electronspray mass spectrometer; ESI positive ion mode detection, 4 GHz, m/z range from 50–1200; Phenomenex Luna NH2 column 5 μm bead size, 100 Å pore size, 150×2 mm) using a gradient system described in literature[25]. Data were analyzed using Agilent Mass Hunter software package.

MS Analysis of CmoA Assay

10 mM prephenate (or 10 mM chorismate initially) was incubated with 10 mM SAM and 10 μM CmoA in a total of 0.5 mL solution at room temperature overnight. An aliquot of 10 μL of the reaction mixture was mixed with 100 μL methanol, which was then infused into a 12T Agilent IonSpec FT-ICR-MS (Agilent Technologies, Inc., CA). Cx-SAM (*m/z* = 443.1373) was monitored in positive mode, and phenylpyruvate (*m/z* = 163.0404) was monitored in negative mode. The Agilent 12T QFT-ICR routinely provides better than 5 ppm mass accuracy with external calibration.

Solvent Isotope Exchange of Deuterated SAM

The assay solution was prepared by mixing 10 mM Tris, pH 8.0, 0.5 mM [$^2H_3$-methyl] SAM (CDNisotope, Quebec), 0.5 mM prephenate and 10 μM CmoA in 0.5 mL solution. The reaction was incubated for 4 hr at room temperature and was quenched by filtering the enzyme with a spin column (MWCO = 10 kD). The sample was then analyzed by mass spectrometry as described above. To examine whether solvent proton exchange at S-methyl of SAM is prephenate-dependent manner, a sample without prephenate was prepared and analyzed in an identical fashion.

Assay of CmoB Reaction

Carboxymethyltransfer activity of CmoB was examined with a solution containing 50 mM Tris, pH 8.0, 4 mM $MgCl_2$, 1 mM prephenate, 1 mM SAM, 1 μM CmoA and total RNA extracted from *cmoB*-mutant *E. coli* cells as described before[16], or purified tRNAs[2], where the total volume of the assay is 50 μL. The reaction was initiated by adding 6 μM CmoB and incubated at the room temperature for 2 hr. One unit of P1 nuclease (US Biological) was added to the assay solution and incubated at 65°C for 1 hr to convert polynucleotides into 5'-nucleotide monophosphates. The P1 nuclease-treated sample was mixed with 100 μL methanol and vortexed before centrifugation at 13,200 RPM for 2 min. An aliquot of supernatant was injected to a 12T Agilent IonSpec FT-ICR-MS and analyzed in negative mode.

Computational ligand docking

To create a model of the substrate prephenate bound to CmoA, we first removed the carboxylate group from the product Cx-SAM to create SAM in the CmoA catalytic site. The resulting complex was subjected to a protein preparation protocol, during which hydrogen atoms were added; protonation states of His residues were examined and adjusted if necessary; side-chains of Thr, Tyr and Asn residues were optimized for hydrogen bonding interactions; and the entire structure was finally energy minimized such that heavy atom positions remained within 0.3Å of the starting coordinates.

Initial attempts to dock the substrate prephenate to this resulting model failed due to inadequate space for the ligand. We hypothesized that charged residues in the binding site required conformational changes to accommodate the ligand. Specifically, in the Cx-SAM-bound structure, Arg199 formed a salt-bridge interaction with the carboxylate group of Cx-SAM (Fig. 2c), and other charged residues in the active site such as Lys165 and Glu203

either pointed toward the solvent or blocked portions of the active site. Therefore, we used an induced fit docking procedure, in which side-chains of residues that are within 5Å of the docked prephenate pose were treated as conformationally flexible[26]. Induced fit docking uses a combination of a molecular-mechanics energy function and an empirical scoring function-based energy to rank the ligand poses. We re-ranked the induced fit docking poses using a molecular-mechanics based energy function that has been used successfully in many applications of metabolite docking. The lowest energy binding pose thus identified is shown in Fig. 2d.

Thermal stability of wild-type and D89L mutant CmoA

The fluorescence monitored thermal denaturation of wild-type and mutant CmoA was performed using a 7900HT RT-PCR system (Applied Biosystems). Briefly, 20 μL of each protein at 10 μM concentration was mixed with 0.5 μL of 200× Sypro orange solution and pipetted into separate wells of a 384-well PCR plate. After centrifugation to remove air bubbles, the plate is loaded into the PCR machine and the temperature ramped from 20 to 99°C, in 1°C increments with a dwell time of 6 s. The negative first derivative of the fluorescence change (−dRFU/dT) for each protein is plotted against temperature, and the melting temperature is defined as the minimum in the −dRFU/dT curve. The wild type from MonoQ exhibited a melting temperature (T_m) of 55.6 ± 0.1°C, while the two D89L mutant fractions displayed T_ms that were 2–3°C lower. The behavior the wild-type and mutant species show that they are full folded under the conditions (i.e., temperature) used *in vitro* activity assays; the lack of activity exhibited by the D89L mutant is thus the consequence of a catalytic defect and not due to issues related to thermodynamic stability.

Chemical Synthesis of Cx-SAM

3.0 mg of S-adenosyl-L-homocystein (SAH) was dissolved in 0.5 mL 150 mM ammonium bicarbonate. To this solution, 2-iodoacetatic acid (100 mg) was added. The mixture was incubated at 37°C for 12 hr with constant agitation. The progress of the reaction was monitored by thin layer chromatography (SiO2) using a solvent system composed of methanol:aqueous 1.5 N ammonium bicarbonate (10:1 vol./vol.). R_f= 0.6 and 0.3 for SAH and Cx-SAM, respectively. After the reaction was completed, 12 mL methanol was added and the mixture incubated at 4°C overnight. Precipitates were collected by centrifugation at 4°C (2,000g for 30 min), washed twice with ice-cold methanol and dissolved in 0.10 mL deionized water. The product was purified using an HILIC as described above. Concentration of Cx-SAM was determined spectroscopically, assuming an extinction coefficient of SAM (ε_{260} = 154 $cm^{-1}mM^{-1}$).

Assay of Non-Enzymatic Formation of Prephenate from Chorismate

The rate of conversion from chorismate to prephenate in the absence of CmoA was measured *in vitro*[24]. The assay solution contained 10 mM sodium phosphate (pH 6.8), 0.2 mM SAM, and 0.2 mM chorismate in 0.5 mL. An aliquot of 80 μL was withdrawn periodically and added to 5 μL of 4.5 M HCl. The mixture was then incubated at 37°C for 15 min and combined with 15 μL of 12 M NaOH before the absorbance at 320 nm was measured to determine phenylpyruvate.

NMR Analysis of the CmoA Assay Mixture

To define the nature of prephenate-derived product subsequent to donation of the carboxylate, the *in vitro* assay was scaled up using 0.5 mM prephenate, 0.5 mM SAM, and 10 μM CmoA in 10 mL total volume, so as to maximize the yield of products for NMR analysis. The reaction was incubated at room temperature for 8hrs, and the enzyme filtered using a spin column (MWCO 10 kD) before lyophilization. The lyophilized sample was dissolved in 0.6 mL D_2O (Cambridge Isotope Laboratory, MA) and ^{1}H-resonance data was collected with a Bruker DRX-300 NMR spectrometer.

Partitioning of the Ylide Intermediate

The pH dependent ylide partitioning assay was performed with 0.2 mM prephenate, 0.05 mM [$^2H_3$-methyl] SAM in either 10 mM ammonium acetate, pH 7.3, or 10 mM ammonium bicarbonate, pH 8.5. The assay was initiated by adding 2 μM CmoA, with incubation at room temperature for 2 hr before being analyzed by mass spectrometry as described below. Total amount [$^2H_3$-methyl]- and [$^2H_2{}^1$H-methyl]-SAM remaining after the reaction was determined by adding a known amount of unlabeled SAM to the assay mixture as an internal standard. The concentration of Cx-SAM was determined by subtracting the remaining SAM after the reaction from the initial quantity added. The amount of solvent exchanged ([$^2H_2{}^1$H-methyl]) and non-exchanged ([$^2H_3$-methyl]) SAM were calculated from the relative amplitude of corresponding MS peaks and the total SAM concentration. Based on the absolute concentrations of [$^2H_2{}^1$H-methyl]-SAM and Cx-SAM the partitioning of the ylide intermediate back to reactant and forward to product was calculated (e.g., partitioning back to SAM is calculated as ([$^2H_2{}^1$H-methyl]-SAM)/([$^2H_2{}^1$H-methyl]-SAM + Cx-SAM)).

Verification of Enzymatic Formation of Phenylpyruvate Using LC-MS

A Shimadzu HPLC, with two LC-20AD pumps, was used to generate a gradient with 50 μL/min flow rate. Solvent A was 5% acetonitrile in H_2O and 0.1% formic acid while solvent B consisted of 95% acetonitrile in H_2O and 0.1% formic acid. 50 μL of assay sample, which was used for NMR analysis above, was loaded onto a 1.0 × 50 mm C18 column (Phenomenex, CA). After desalting with 5 solvent B for 5 min, bound phenylpyruvic acid was eluted with a 30 min gradient composed of 5% to 95% B. The effluent was directly delivered into the 12T QFT-ICR-MS (Agilent Technologies, Inc., CA) for mass analysis.

Preparation of [U-^{13}C] chorismate

Production and purification of chorismate using *Aerobacter aerogenes 62-1* followed the methods developed previously[27,28]. *A. aerogenes 62-1* was generously provided by Drs. Jared Parker and Christopher T. Walsh at Harvard Medical School. Overnight culture (1 mL) of *A. aerogenes 62-1* was added to 50 mL Medium A, and incubated at 30°C until OD_{600} reached ~1. Cells were pelleted at 3,000g and washed with 25 mL Medium B which did not contain glucose. After pelleting once again, cells were resuspended in 25 mL Medium B with 0.2 g of [U-^{13}C] glucose (purchased from Cambridge Isotope Laboratory). Cells were grown at 30°C for 15 hr for the production of labeled chorismate, harvested by centrifugation and discarded. The supernatant was filtered and loaded on Hypercarb HPLC column (10×100 mm) equilibrated in 10 mM ammonium acetate, pH 9.9. Chorismate was

eluted with a linear gradient of acetonitrile and identified by monitoring OD_{275} of each fraction. ^{13}C-chorismate was confirmed by mass spectrometry ($^{13}C_{10}H_{10}O_6Na$, observed m/z = 259.0707, calculated m/z = 259.0711). Pooled chorismate was lyophilized, resuspended in water and quantitated by UV absorption at 275 nm (ε275 nm = 2630 M^{-1} cm^{-1})[28].

Assay of CmoA with [U-^{13}C] chorismate

The assay mixture (0.1 mL) contained 20 mM sodium phosphate, pH 6.8, 0.2 mM SAM, 0.2 mM [U-^{13}C] chorismate, and 2 μM CmoA. The reaction was incubated overnight at the room temperature before analysis by mass spectrometry.

cmoA Complementation

For the complementation assay, the *cmoA* gene was inserted between *Kpn*I and *Hind*III sites in the pQE30a (Qiagen) expression vector. *cmoA*-deficient *E. coli* cells (from KEIO collection) were transformed with either empty vector (pQE30a), plasmid bearing wild-type *cmoA* or plasmid bearing biochemically inactive D89L *cmoA*. Transformed cells were typically grown in 50 mL LB media at 37°C and induced with 0.5 mM IPTG at an OD_{600} of ~1. Cells were incubated overnight at 25°C and harvested by centrifugation. Total RNA was extracted and treated with one unit of P1 nuclease at 65°C for 1 hr. The hydrolyzed nucleotide samples were analyzed using LC/MS (Waters Symmetry C18 Column, 100Å, 3.5 μm, 2.1 mm × 150 mm coupled to 12T Agilent IonSpec FT-ICR-MS) in negative mode. A linear gradient of 20% to 95% acetonitrile and 0.1% formic acid was used over 15 min. The identification of cmo5UMP demonstrates the *in vivo* formation of cmo5U at the wobble position.

Cmo5UMP assay using *in vivo*- and *in vitro*-generated Cx-SAM

^{13}C-CmoA was purified from cells grown in M9 minimal media containing ^{13}C-glucose as the sole carbon source. The media contained 1× M9 salts (Sigma), 2 mM $MgSO_4$, 0.2 mM $CaCl_2$, and 0.4% [U-^{13}C]-glucose (Cambridge Isotope Laboratory) in 0.5 L. These growth conditions yield ^{13}C-Cx-SAM bound CmoA, which was purified as described above. Notably, the occupancy of ^{13}C-Cx-SAM within ^{13}C-CmoA was nearly 100% as determined by MS analysis, which is higher than that typically observed in samples prepared from cells grown in LB media (see Supplementary Fig. 2). 10 μM ^{13}C-Cx-SAM-CmoA complex was added to an assay mixture containing 20 mM ammonium acetate, pH 7.3, 4 mM $MgCl_2$, 0.2 mM ^{12}C-prephenate, 0.2 mM ^{12}C-SAM, 20 μM CmoB, and total RNA isolated from CmoB-deficient cells. The assay solution was incubated at room temperature for 30 min, quenched with 2 units of P1 nuclease and incubated at 65°C for 1 hr. Cmo5UMP formation was analyzed using LC/MS as described above (see Supplementary Fig. 12).

Network Analysis

BLAST[29] e-values for sequences in the Pfam[30] Methyltransf_18 family were obtained from the Structure Function Linkage Database (SFLD)[31]. SFLD BLAST searches are performed by comparing each sequence in a superfamily against each other. For efficiency, searches are performed by BLASTing bundles of 100 query sequences against all other superfamily

sequences. Results are post processed to obtain the equivalent blast2seq e-value (independent of database size) based on bit score. Cytoscape[32] networks were created from these BLAST results at several different e-value cutoffs, and using either the full sequence set or the subset of sequences most closely related to *E. coli* CmoA (based on BLAST e-value). Tools used for visualization of protein networks were created by the UCSF Resource for Biocomputing, Visualization, and Informatics and are available from the Resource (http://www.rbvi.ucsf.edu). Each node in the network represents a single sequence in the Pfam Methyltransf_18 family (or a subset thereof) and each edge represents the pairwise connection with the most significant BLAST E-value (better than the cut-off) connecting the two sequences. Connections between nodes are only shown if the e-value of the best Blast hit between two sequences is at least as good as the specified e-value cutoff. Lengths of edges are not meaningful except that sequences in tightly clustered groups are relatively more similar to each other than sequences with few connections. The nodes were arranged using the yFiles organic layout provided with Cytoscape version 2.8. Annotation information retrieved from Swissprot[33] (functional annotation) and NCBI[34,35] (phylum), and calculated via a MUSCLE[36] multiple sequence alignment (conservation of R199 in *E. coli* CmoA) was associated with each node as applicable.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank J. Parker and C.T. Walsh for providing the *Aerobacter aerogenes 62-1*strain. We are indebted to V. Schramm and J. Gerlt for critical discussions and reading of manuscript. This work was supported by NIH grants GM094662 (to S.C.A.), GM093342 (to S.C.A., M.P.J. and P.C.B.) and the Albert Einstein Cancer Center. This publication was made possible by the Center for Synchrotron Biosciences grant, P30-EB-009998, from the National Institute of Biomedical Imaging and Bioengineering (NIBIB). Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-98CH10886.

References

1. Czerwoniec A, et al. MODOMICS: a database of RNA modification pathways. 2008 update. Nucleic Acids Res. 2009; 37:D118–D121. [PubMed: 18854352]
2. Nasvall SJ, Chen P, Bjork GR. The modified wobble nucleoside uridine-5-oxyacetic acid in tRNAPro(cmo5UGG) promotes reading of all four proline codons in vivo. Rna. 2004; 10:1662–1673. [PubMed: 15383682]
3. Nasvall SJ, Chen P, Bjork GR. The wobble hypothesis revisited: uridine-5-oxyacetic acid is critical for reading of G-ending codons. Rna. 2007; 13:2151–2164. [PubMed: 17942742]
4. Weixlbaumer A, et al. Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. Nat Struct Mol Biol. 2007; 14:498–502. [PubMed: 17496902]
5. Bjork GR. A novel link between the biosynthesis of aromatic amino acids and transfer RNA modification in Escherichia coli. J Mol Biol. 1980; 140:391–410. [PubMed: 6160251]
6. Hagervall TG, Jonsson YH, Edmonds CG, McCloskey JA, Bjork GR. Chorismic acid, a key metabolite in modification of tRNA. J Bacteriol. 1990; 172:252–259. [PubMed: 2104604]
7. Lim K, et al. Crystal structure of YecO from Haemophilus influenzae (HI0319) reveals a methyltransferase fold and a bound S-adenosylhomocysteine. Proteins. 2001; 45:397–407. [PubMed: 11746687]

8. Van Vleet J, Kleeb A, Kast P, Hilvert D, Cleland WW. 13C isotope effect on the reaction catalyzed by prephenate dehydratase. Biochim Biophys Acta. 2010; 1804:752–754. [PubMed: 19948253]
9. Horowitz S, Yesselman JD, Al-Hashimi HM, Trievel RC. Direct evidence for methyl group coordination by carbon-oxygen hydrogen bonds in the lysine methyltransferase SET7/9. J Biol Chem. 2011; 286:18658–18663. [PubMed: 21454678]
10. Crosby J, Stirling CJM. Elimination and addition reactions. Part XIX. Elimination of phenoxide from [small beta]-substituted ethyl phenyl ethers: the nature of activation in 1,2-elimination. Journal of the Chemical Society B: Physical Organic. 1970:671–679.
11. Bordwell FG. Equilibrium acidities in dimethyl sulfoxide solution. Accounts Chem Res. 1988; 21:456–463.
12. Arrowsmith CH, Bountra C, Fish PV, Lee K, Schapira M. Epigenetic protein families: a new frontier for drug discovery. Nat Rev Drug Discov. 2012; 11:384–400. [PubMed: 22498752]
13. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet. 2009; 10:295–304. [PubMed: 19308066]
14. Luka Z, Mudd SH, Wagner C. Glycine N-methyltransferase and regulation of S-adenosylmethionine levels. J Biol Chem. 2009; 284:22507–22511. [PubMed: 19483083]
15. Vevodova J, et al. Structure/function studies on a S-adenosyl-L-methionine-dependent uroporphyrinogen III C methyltransferase (SUMT), a key regulatory enzyme of tetrapyrrole biosynthesis. J Mol Biol. 2004; 344:419–433. [PubMed: 15522295]
16. Kowtoniuk WE, Shen Y, Heemstra JM, Agarwal I, Liu DR. A chemical screen for biological small molecule-RNA conjugates reveals CoA-linked RNA. Proc Natl Acad Sci U S A. 2009; 106:7768–7773. [PubMed: 19416889]
17. Dalhoff C, Lukinavicius G, Klimasauskas S, Weinhold E. Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases. Nat Chem Biol. 2006; 2:31–32. [PubMed: 16408089]
18. Dalhoff C, Lukinavicius G, Klimasauskas S, Weinhold E. Synthesis of S-adenosyl-L-methionine analogs and their use for sequence-specific transalkylation of DNA by methyltransferases. Nat Protoc. 2006; 1:1879–1886. [PubMed: 17487172]
19. Binda O, et al. A chemical method for labeling lysine methyltransferase substrates. Chembiochem. 2011; 12:330–334. [PubMed: 21243721]

Methods References

20. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. Acta Crystallogr D Biol Crystallogr. 2006; 62:859–866. [PubMed: 16855301]
21. Lebedev AA, Vagin AA, Murshudov GN. Model preparation in MOLREP and examples of model improvement using X-ray data. Acta Crystallogr D Biol Crystallogr. 2008; 64:33–39. [PubMed: 18094465]
22. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–2132. [PubMed: 15572765]
23. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1997; 53:240–255. [PubMed: 15299926]
24. Dopheide TA, Crewther P, Davidson BE. Chorismate mutase-prephenate dehydratase from Escherichia coli K-12. II. Kinetic properties. J Biol Chem. 1972; 247:4447–4452. [PubMed: 4261395]
25. Lorenz MA, Burant CF, Kennedy RT. Reducing time and increasing sensitivity in sample preparation for adherent mammalian cell metabolomics. Anal Chem. 2011; 83:3406–3414. [PubMed: 21456517]
26. Kalyanaraman C, Bernacki K, Jacobson MP. Virtual screening against highly charged active sites: identifying substrates of alpha-beta barrel enzymes. Biochemistry. 2005; 44:2059–2071. [PubMed: 15697231]

27. Gibson F. Chorismic acid: purification and some chemical and physical studies. Biochem J. 1964; 90:256–261. [PubMed: 5834235]

28. Parker JB, Walsh CT. Olefin isomerization regiochemistries during tandem action of BacA and BacB on prephenate in bacilysin biosynthesis. Biochemistry. 2012; 51:3241–3251. [PubMed: 22483065]

29. Altschul SF, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25:3389–3402. [PubMed: 9254694]

30. Punta M, et al. The Pfam protein families database. Nucleic Acids Res. 40:D290–D301. 10.1093/nar/gkr1065. [PubMed: 22127870]

31. Pegg SC, et al. Leveraging enzyme structure-function relationships for functional inference and experimental design: the structure-function linkage database. Biochemistry. 2006; 45:2545–2555. [PubMed: 16489747]

32. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics. 27:431–432. 10.1093/bioinformatics/btq675. [PubMed: 21149340]

33. Consortium U. Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Res. 2012; 40:D71–D75. [PubMed: 22102590]

34. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2009; 37:D26–D31. 10.1093/nar/gkn723. [PubMed: 18940867]

35. Sayers EW, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2009; 37:D5–D15. 10.1093/nar/gkn741. [PubMed: 18940862]

36. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004; 5:113. [PubMed: 15318951]

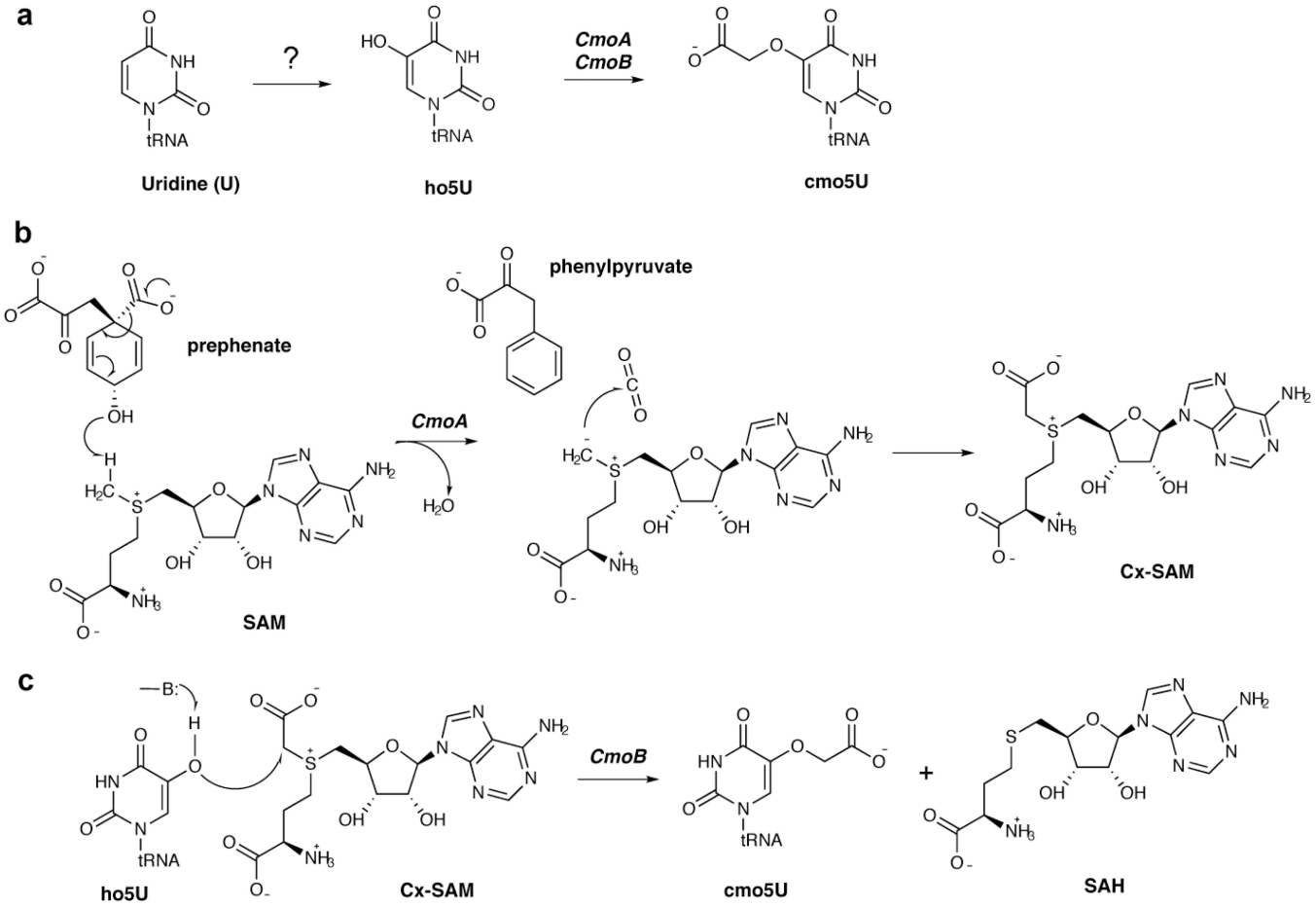

Figure 1. Proposed chemical mechanism for the biosynthesis of cmo5U

a, Previously identified biosynthetic pathway for cmo5U at wobble uridines. First, the wobble uridine is converted to ho5U by an unknown mechanism, followed by the action of CmoA and CmoB. **b,** Mechanism for CmoA-catalyzed Cx-SAM formation from SAM and prephenate. **c,** Mechanism for CmoB-catalyzed formation of cmo5U from ho5U and Cx-SAM.

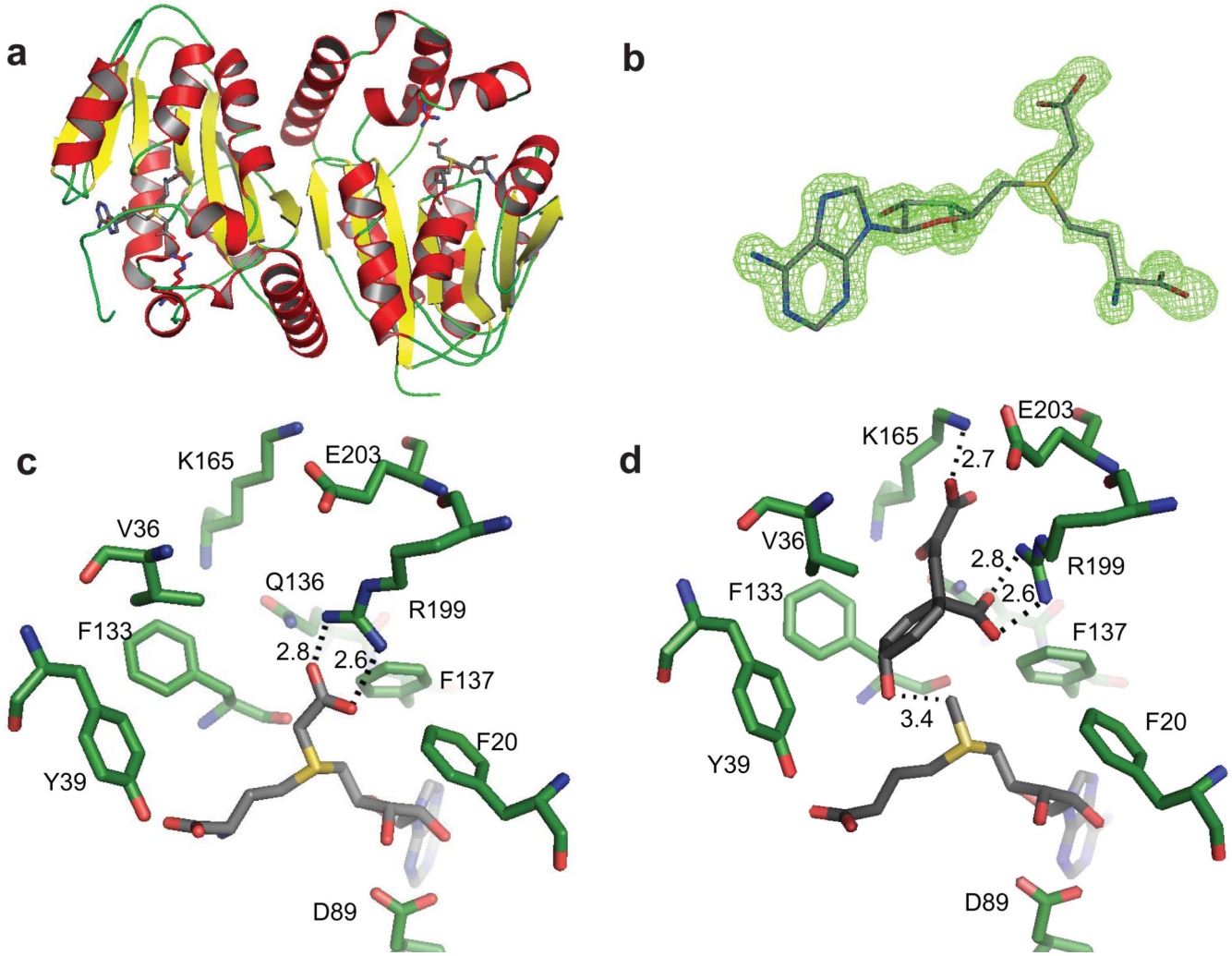

Figure 2. Structure of the CmoA:Cx-SAM complex

a, Overall dimeric structure of *E. coli* CmoA, with α-helices, β-sheets and loops colored red, yellow and green, respectively. Cx-SAM and Arg-199 are represented as sticks. **b,** *Fo-Fc* difference Fourier synthesis, calculated at 1.5 Å resolution with the ligand omitted, contoured at 5σ around the modeled Cx-SAM ligand. **c,** Catalytic site of CmoA. Protein carbon atoms are colored green and Cx-SAM carbon atoms gray. Oxygen and nitrogen atoms are colored red and blue, respectively. Ionic interactions between Cx-SAM and the side-chain of Arg-199 are depicted as dashed lines (distances in Ångstroms). **d,** Computationally-predicted pose of prephenate in the CmoA catalytic site.

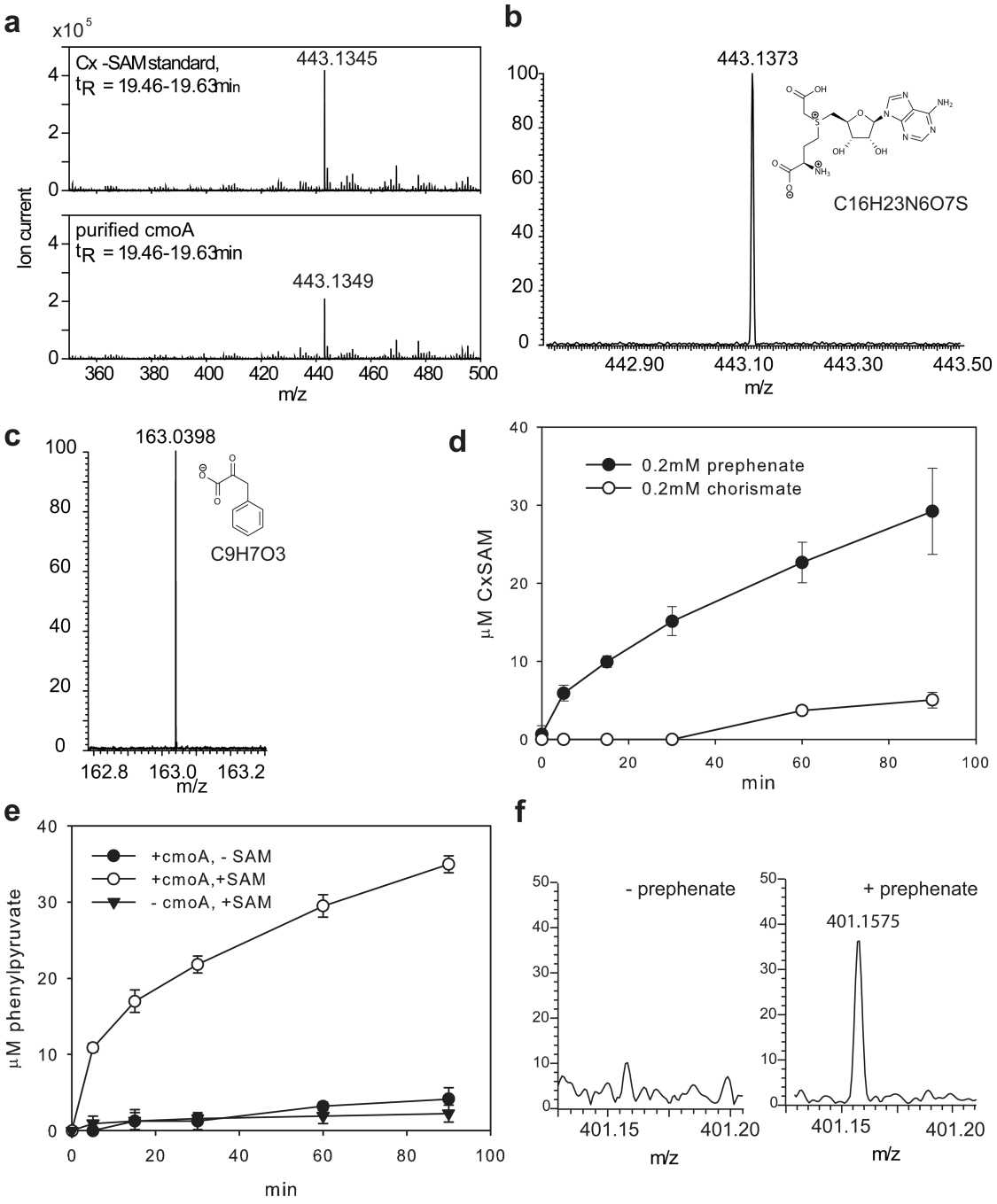

Figure 3. Identification of low molecular weight compounds associated with CmoA-mediated Cx-SAM production

a, ESI-TOF mass spectra of a Cx-SAM standard (top) and the low molecular weight compound copurifying with CmoA (bottom). Peak m/z: 443.1345 and 443.1359 (errors: −0.9 and +2.3 ppm) for Cx-SAM standard and the compound that copurified with recombinant CmoA, respectively. **b,** Detection of Cx-SAM by in an *in vitro* assay containing SAM, chorismate and CmoA. **c,** Detection of phenylpyruvate ($C_9H_7O_3$) formation in the assay mixture by MS in negative mode (*m/z* = 163.0402 observed;

163.0395 calculated). **d,** Time course of Cx-SAM production in an *in vitro* assay of CmoA. The assay solution contained 20 mM sodium phosphate pH 6.8, 0.2 mM [^{14}C-methyl]-SAM, 0.2 mM prephenate or chorismate, and 2 μM CmoA. Error bars represent the standard deviation of three data sets. **e,** Time course of the phenylpyruvate formation from prephenate. The assay mixture contained 20 mM sodium phosphate, pH 6.8, 0.2 mM prephenate, 0.2 mM SAM (open circles and inverted triangles), and 2 μM CmoA (open circles and filled circles). Error bars represent the standard deviation of three data sets. **f,** Solvent proton exchange of [$^{2}H_3$-methyl]-SAM catalyzed by CmoA. The sample contained 10 mM Tris, pH 8.0, 0.5 mM [$^{2}H_3$-methyl]-SAM and 10 μM CmoA, with or without 0.5 mM prephenate. The reaction was performed at room temperature for 4 hr. In the presence of prephenate, doubly deuterated SAM (calculated m/z = 401.1576) was observed.

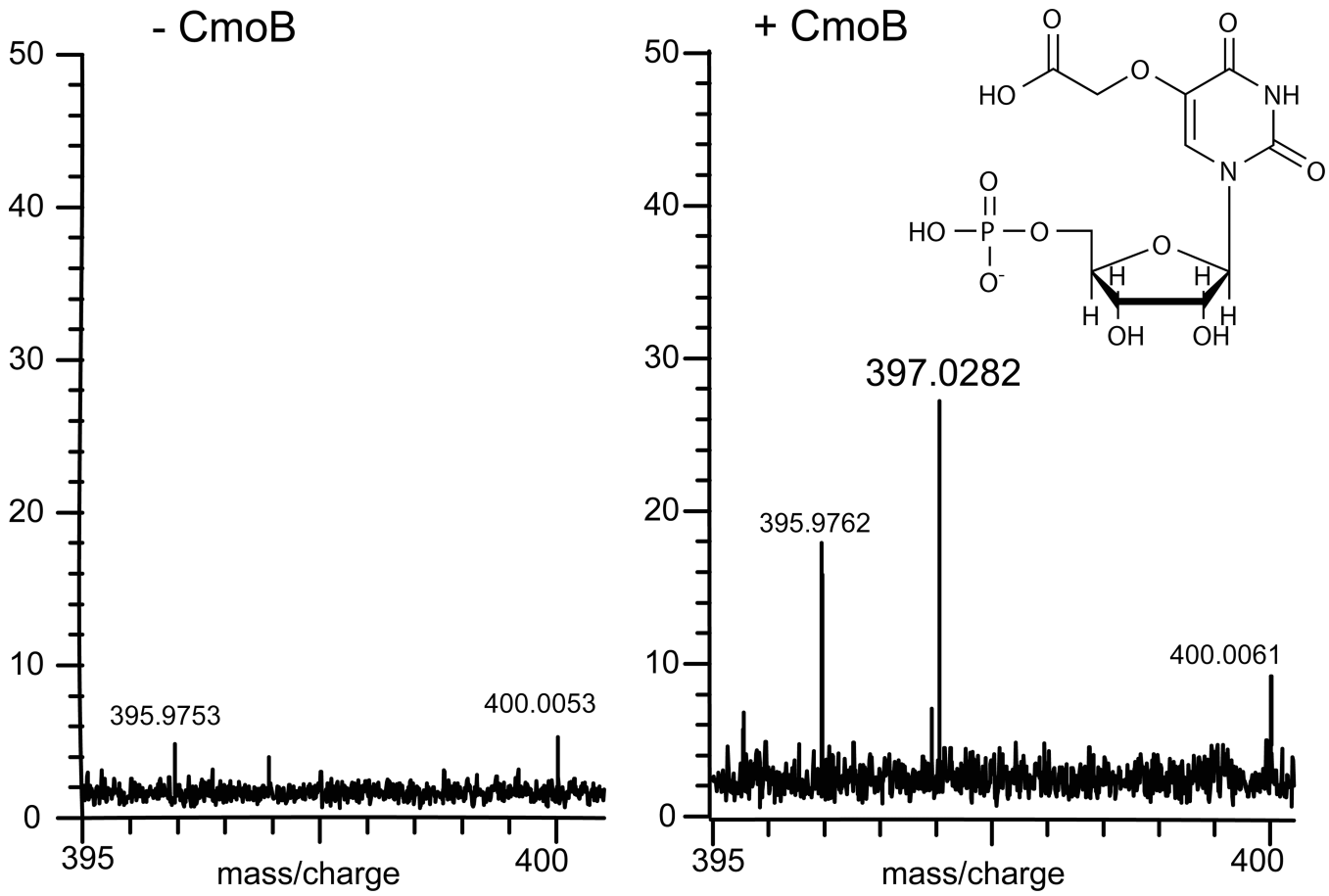

Figure 4. ***In vitro*** **assay of CmoB-catalyzed carboxymethyltransfer activity**

Total RNA extracted from CmoB mutant cells was used as a substrate, and Cx-SAM was generated *in situ* by the action of CmoA on prephenate and SAM. RNA was digested with P1 nuclease to 5'- nucleotide monophosphates followed by MS analysis. *Left*) No CmoB was added. *Right*) Addition of purified *E. coli* CmoB resulted in the detection of cmoUMP ($C_{11}H_{14}O_{12}N_2P$, m/z observed =397.0282, calculated = 397.0284) in negative ion mode.