Caffeine Junkie: an Unprecedented Glutathione S-Transferase-Dependent Oxygenase Required for Caffeine Degradation by Pseudomonas putida CBB5

Ryan M. Summers, Jennifer L. Seffernick, Erik M. Quandt, Chi Li Yu, Jeffrey E. Barrick, Mani V. Subramanian

Center for Biocatalysis and Bioprocessing and Department of Chemical and Biochemical Engineering, The University of Iowa, Coralville, Iowa, USA; Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, St. Paul, Minnesota, USA; Center for Systems and Synthetic Biology, Institute for Cellular and Molecular Biology, Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas, USA

Caffeine and other N-methylated xanthines are natural products found in many foods, beverages, and pharmaceuticals. Therefore, it is not surprising that bacteria have evolved to live on caffeine as a sole carbon and nitrogen source. The caffeine degradation pathway of Pseudomonas putida CBB5 utilizes an unprecedented glutathione-S-transferase-dependent Rieske oxygenase for demethylation of 7-methylxanthine to xanthine, the final step in caffeine N-demethylation. The gene coding this function is unusual, in that the iron-sulfur and non-heme iron domains that compose the normally functional Rieske oxygenase (RO) are encoded by separate proteins. The non-heme iron domain is located in the monooxygenase, ndmC, while the Rieske [2Fe-2S] domain is fused to the RO reductase gene, ndmD. This fusion, however, does not interfere with the interaction of the reductase with N7- and N3-demethylase RO oxygenases, which are involved in the initial reactions of caffeine degradation. We demonstrate that the N7-demethylation reaction absolutely requires a unique, tightly bound protein complex composed of NdmC, NdmD, and NdmE, a novel glutathione-S-transferase (GST). NdmE is proposed to function as a noncatalytic subunit that serves a structural role in the complexation of the oxygenase (NdmC) and Rieske domains (NdmD). Genome analyses found this gene organization of a split RO and GST gene cluster to occur more broadly, implying a larger function for RO-GST protein partners.

Received 17 May 2013 Accepted 20 June 2013
Published ahead of print 28 June 2013

Address correspondence to Mani Subramanian, mani-subramanian@uiowa.edu.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/JB.00585-13.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.
MATERIALS AND METHODS

Chemicals. Caffeine, 7-methylxanthine, xanthine, ammonium acetate, acetic acid, and 2,4-pentanedione were purchased from Sigma-Aldrich (St. Louis, MO). Trypsitone, yeast extract, Soytone, and agar were purchased from Becton, Dickinson and Company (Sparks, MD). NADH and isopropyl-β-D-thiogalactopyranoside (IPTG) were obtained from RPI Corp. (Mt. Prospect, IL). Restriction enzymes, Taq DNA polymerase, and Phusion high-fidelity DNA polymerase (the default polymerase used in all PCRs unless otherwise noted) were purchased from New England BioLabs (Ipswich, MA). High-pressure liquid chromatography (HPLC)-grade methanol (J. T. Baker, Phillipsburg, NJ) was used in all chromatographic studies.

Molecular biology procedures. The full ndmE and orf9 sequences were determined by sequencing two PCR-amplified DNA fragments obtained from an EcoRI genomic DNA library (6) as described in the supplemental material. CBB5 strains harboring a single gene knockout of ndmC, ndmD, or ndmE were created by the method of Link et al. (7).

All cloning steps are described in full in the supplemental material. ndmC was cloned into the pET28a(+) plasmid at the NdeI and XhoI sites with a 5′-terminal hexahistidine (His6) tag to form plasmid p32CHis. ndmE was cloned into the pET28a(+) plasmid at the NdeI and EagI sites with a fused N-terminal His6 tag, resulting in plasmid p28HisE.

Similarly, the ndmCDE genes, including native CBB5 intergenic DNA sequence surrounding ndmD, were cloned into the pET28a(+) plasmid at the NdeI and Eagl sites with a fused N-terminal His tag, resulting in plasmid p28HisE. NdeI and Eagl sites with a His tag fused only to the N terminus of ndmC, creating plasmid p28HisCDE. ndmA and ndmB were cloned into the pACYCDuet-1 vector at the NcoI and BamHI sites (for ndmA) and the NdeI and KpnI sites (for ndmB), resulting in plasmid p32CHis.

Heterologous expression of ndm genes. Plasmids p32CHis, p28HisE, and p28HisCDE were transformed into E. coli BL21(DE3) for overproduction of recombinant proteins. Plasmid dAB was transformed into E. coli BL21(DE3) already containing p28HisCDE. Expression of all genes was carried out in the same manner unless otherwise noted. The cells were grown in LB broth with appropriate antibiotic at 37°C and 225 rpm. When cell density reached an optical density at 600 nm (OD600) of 0.5, sterile iron(III) chloride was added to the culture at a final concentration of 10 μM, and the culture was shifted to 18°C and 250 rpm for incubation to increase the solubility of recombinant protein produced. Riboflavin was also added to cultures expressing ndmD to a final concentration of 2.5 μM. IPTG at a final concentration of 0.1 mM was added to induce gene expression when the OD600 reached 0.8 to 1.0. Induced cells were incubated at 18°C for 18 to 20 h and harvested by centrifugation at 4,000 × g for 10 min at 4°C. Cells were washed twice in 50 mM potassium phosphate (KPi) buffer (pH 7.5). Cells were lysed by passing twice through a chilled French press at 138 MPa. The lysates were centrifuged at 30,000 × g for 20 min, and supernatants were saved as cell extracts. Following cell lysis and centrifugation, the cell extract was loaded onto a 20-ml (bed volume) nickel-nitrilotriacetic acid (Ni-NTA) column (GE Healthcare) preequilibrated in Buffer A at a flow rate of 2 ml · min⁻¹. The column was washed with 80 ml Buffer A to remove unbound protein. Bound protein was then eluted with 40 ml Buffer B (50 mM KPi with 300 mM NaCl and 125 mM imidazole) and concentrated using Amicon ultrafiltration units (molecular weight cutoff [MWCO], 10,000). The concentrated enzyme solution was dialyzed at 4°C in 50 mM KPi buffer containing 5% glycerol and 1 mM dithiothreitol (DTT) (KPGD buffer). All purified enzymes were stored short term on ice and at ~−80°C for long-term storage.

Resting cell assays. CBB5 and its derivative strains containing single gene deletions were inoculated from frozen glycerol stocks and grown overnight in a modified M9 mineral salts medium containing 4 g · liter⁻¹ Soytone and 2.5 g · liter⁻¹ caffeine (M9CS medium) (4) at 30°C with 225 rpm rotary shaking. The cells were then inoculated into fresh M9CS medium (inoculum volume was 1% of the final volume) and again grown overnight. Upon reaching an OD600 of 1.5 to 2, cells were harvested by centrifugation at 4,000 × g for 10 min at 4°C. Cells were washed twice in 50 mM potassium phosphate (KPi) buffer (pH 7.5). Resting cell assays were conducted by resuspending cells in 2 ml 50 mM KPi buffer containing 300 μM caffeine or 7-methylxanthine and incubated at 30°C with shaking at 400 rpm in an incubated microplate shaker (VWR, Radnor,

FIG 1 (A) Genetic organization of the 14.8-kb genomic DNA fragment containing the Alx operon for N-demethylation of caffeine to xanthine in P. putida CBB5. Arrows indicate the position and orientation of each orf. Red arrows represent genes required for N-demethylation, while blue arrows represent genes predicted to metabolize formaldehyde. (B) Gene organization for homologous ndmA-CDE gene clusters. One or two different ndmC homologs may appear in the gene cluster. Conserved domains from the Conserved Domains Database (CDD) are depicted below the genes. The SRPBCD superfamily contains the C-terminal catalytic domain of alpha oxygenases (pink with red border), HcaE members are ring-hydroxylating dioxygenases (blue-gray), the Rieske superfamily contains the N-terminal Rieske domain of alpha oxygenases (purple), the FNR-like superfamily is the ferredoxin reductase (FNR) domain that binds the NADH/FADH cofactor (red), HMP is the ferredoxin reductase family 1 (gray), fer2 is the 2Fe-2S cluster/ferredoxin domain (green), and Gst is the glutathione S-transferase (light blue). (C) SDS-PAGE of Sephacryl S200-purified His₆-NdmCDE (right lane). Sizes of markers in the molecular mass ladder (left lane) are given on the left.

Summers et al.
RESULTS

Determination of domain architecture within NdmCDE. The Alx operon encodes caffeine degradation in CBB5, including the ndmA and ndmB genes. NdmC (orf7) is a 32.5-kDa protein homologous to the C terminus of many ROs. The Conserved Domains Database (CDD) (11) identified an SRPBC (START/RHO_alpha_C/PITP/Bet_v1/CoxG/CalC) ligand-binding superfamily domain within this region (Fig. 1B). This superfamily domain usually has a deep, hydrophobic ligand binding pocket containing non-heme Fe(II) and is also present in NdmA and NdmB (6). The N-terminal Rieske center domain that generally accompanies such proteins is lacking in NdmC. Curiously, this domain is attached to the N terminus of NdmD. With the exception of this extra domain, NdmD has the typical RO reductase arrangement, consisting of ferredoxin reductase (FRN) and C-terminal ferredoxin domains (6).

Using a nested PCR approach with a CBB5 genomic DNA library, we determined the full ndmE (orf8) sequence and an additional complete open reading frame, orf9, located 3′ to ndmE (Fig. 1A). ndmE encoded a 25.5-kDa protein comprised of 222 amino acids, which is slightly larger than the 22-kDa protein band observed in a highly enriched protein fraction containing NdmC, NdmD, and NdmE from CBB5 (6). As suggested by the partial gene sequence, the full-length NdmE protein was homologous to many GSTs of various classes (Fig. 1B).

The theoretical 284-amino-acid sequence of orf9 displayed significant homology to S-formylglutathione hydrolases (frmB). FrmB and FrmA, an orf2 homolog, catalyze the glutathione (GSH)-dependent conversion of formaldehyde to formate (12) and provide a genetic correlation for utilization of formaldehyde produced by N-demethylation, which has been confirmed experimentally (6). It should be noted that although FrmAB use GSH, no GST is known to be associated with these reactions.

Functional expression and characterization of ndmCDE. Although NdmD can be expressed alone as a partly (−5%) soluble protein in E. coli, all attempts to express either NdmC or NdmE individually in E. coli resulted only in production of inclusion bodies (see Table S1 in the supplemental material). In order to detect soluble NdmC and NdmE, both proteins must be coexpressed with NdmD. The ndmCDE genes were cloned for overexpression in E. coli with an N-terminal Hisα tag fused only to NdmC. When cell lysate containing Hisα-NdmCDE was passed through a Ni-NTA column, all three peptides bound to the column and eluted only with 125 mM imidazole. The Hisα-NdmCDE peptides also coeluted from a gel filtration column as a tight complex (Fig. 1C) with an estimated molecular mass of 360 kDa, suggesting that the complex is composed of trimers of each peptide (αβγ). Interestingly, NdmA and NdmB, which are also dependent on NdmD for activity, do not complex with NdmCDE (see Fig. S1).

Hisα-NdmCDE was highly specific for N7-demethylation of 7-methylxanthine; no activity was observed toward caffeine or theobromine (3,7-dimethylxanthine), both of which also contain the 7-methyl group (data not shown). The apparent Km and kcat values of Hisα-NdmCDE for 7-methylxanthine were 15.3 ± 2.3 μM and 9.4 ± 0.3 min−1, respectively, with a catalytic efficiency (kcat/Km) of 0.61 ± 0.09 min−1 μM−1, similar to the efficiency of the purified enzyme from wild-type bacterium (5, 6). One O2 molecule is consumed per N7-demethylation step (see Fig. S2 in the supplemental material), resulting in stoichiometric formation of xanthine and formaldehyde (Fig. 2). Addition of 50 μM exogenous reduced GSH did not alter N7-demethylation activity. Also,
His$_6$-NdmCDE displayed no 1-chloro-2,4-dinitrobenzene-dependent conjugation of GSH, an activity displayed by many, but not all, GSTs. These data suggest that the $N_7$-demethylation reaction occurs at the non-heme Fe(II) found in the binding pocket of NdmC, and that NdmE provides a structural, rather than catalytic, role.

**Individual gene knockout studies.** NdmCDE copurified from CBB5 under fairly stringent conditions, indicating that it exists as a complex. To provide additional proof that the NdmCDE complex is required for $N$-demethylation of 7-methylxanthine to xanthine, we individually knocked out each gene and assayed for $N$-demethylation activity. CBB5 resting cells completely $N$-demethylated 500 $\mu$M 7-methylxanthine (●) to xanthine within 75 min. (A) CBB5 strains with a single knockout of ndmC (○), ndmD (□), or ndmE (△) oxidized 60 to 125 $\mu$M 7-methylxanthine to 7-methyluric acid over 24 h. Each assay contained 0.5 mM caffeine or 7-methylxanthine, as well as freshly harvested and washed cells (OD$_{600}$ of 2.5) in 50 mM potassium phosphate buffer (pH 7.5). Concentrations are reported as means with standard deviations from triplicate results.

Phylogenetic analysis of ndmCDE. The ndmCDE genes were used as Blast queries to identify potential homologs, resulting in 18 organisms (see Table S2 in the supplemental material) containing either one or two ndmC homologs of unknown function (C-terminal oxygenase domain) adjacent to homologs of ndmD (Rieske/reductase fusion) and ndmE (GST) (Fig. 1B). Organisms with two ndmC homolog genes belong to the *Sinorhizobium* and *Mesorhizobium* genera. Interestingly, *Sinorhizobium medicae* WSM419 is the only member of these genera with this gene cluster containing one ndmC homolog. Neighbor-joining trees were prepared for the NdmCDE proteins to further investigate the evolutionary links between the various proteins (Fig. 4). The general morphology of the trees is very similar; the major difference is that some organisms have two NdmC homolog proteins. The proteins encoded by the ndmC gene closest to the ndmDE genes, labeled 2 and shown in blue in Fig. 4A, are positioned similarly to the NdmDE trees. The proteins encoded by the gene at the front of the cluster in the ndmCCDE arrangement (red in Fig. 4A) appear remote from their cluster counterparts. This suggests that these genes diverged after gene
duplication with a potentially new function, and that the gene duplication and divergence occurred prior to gene dissemination.

**DISCUSSION**

In this report, we have identified and characterized the genes of *P. putida* CBB5 responsible for \( N_7 \)-demethylation of 7-methylxanthine to xanthine, which is the final step in the caffeine \( N \)-demethylation pathway (Fig. 5). While *ndmC* was previously hypothesized to encode an \( N_7 \)-demethylase, this could not be demonstrated, because previous attempts to express NdmC resulted only in formation of inclusion bodies (6). Soluble NdmC is obtained only when coexpressed with an RO reductase, NdmD, and a GST, NdmE, to form a large multisubunit protein complex. The reaction stoichiometry for \( N_7 \)-demethylation by NdmCDE is consistent with reactions catalyzed by monoxygenases and is similar to reactions catalyzed by NdmA and NdmB. As with NdmA and NdmB, the \( N_7 \)-demethylation is likely catalyzed at the non-heme iron in NdmC. NdmC itself is composed only of the SRPBCC catalytic domain of a traditional RO. Although NdmC lacks a Rieske [2Fe-2S] cluster, NdmD carries the extra Rieske [2Fe-2S] cluster at its N terminus. In spite of the unusual arrangement of domains in the *ndmCDE* gene cluster, the flow of electrons (see Fig. S4 in the supplemental material) appears to be typical of an RO (13).

NdmE was identified as a member of the GST superfamily (Fig. 1B), which contains proteins with great mechanistic and functional diversity housed within a conserved three-dimensional thioredoxin structure (14,15). Catalytic functions within this superfamily include nucleophilic substitution, epoxide ring opening, ester thiolysis, isomerization, disulfide bond reduction, and hydroperoxidase reactions. GST superfamily members have also been found with noncatalytic functions, including intracellular transport, regulation of signaling cascades, transcriptional regulation, protein degradation and folding, and binding of a wide range of ligands (16–20). Newly identified functions of GSTs, including S-glutathionylation of proteins, indicate that other GST functionalities remain to be discovered (21).

Of the thousands of oxygenases known, this is the first report of the absolute requirement for a GST-like protein in any oxygenation reaction. Although GSH is not required for \( N_7 \)-demethylation of 7-methylxanthine by NdmCDE, it is possible that GSH is covalently bound to NdmE. GST proteins are almost exclusively found as dimers; however, the GST monomer has been hypothesized to be involved in protein-protein interactions (22). Recombinant protein biosynthesis, as a laboratory technique, often takes advantage of increased solubility when fused with GST (23). NdmE itself has no known redox groups. Although the exact function of the GST homolog NdmE has not been ascertained, tight complex formation found between NdmCDE and the inability to synthesize or purify these proteins individually in a soluble form suggest a key role for NdmE in complex formation and solubilization.

GST homologs of undetermined function have also been found in other operons, some containing RO enzymes. These operons are involved in the degradation of many xenobiotics and recalcitrant aromatic compounds, including gentisate (24), \( m \)-toluate and \( m \)-xylene (25), and the polycyclic aromatic hydrocarbons biphenyl, polychlorinated biphenyl, dibenzofuran, and dibenzo-\( p \)-dioxin (24–27). Unlike NdmE, these GST homologs do not appear to be essential for degradation of the target compounds.

**FIG 4** Neighbor-joining trees for homologous protein present in conserved gene clusters: NdmC (A), NdmD (B), and NdmE (C). Full strain names are listed in Table S2 in the supplemental material. Operons containing two NdmC proteins are colored: the first NdmC homolog is red, and the second is blue. NdmD and NdmE proteins in those strains are in purple.

September 2013 Volume 195 Number 17 jb.asm.org 3937 3937
of their operons. However, their presence is thought to confer advantages and allow function under specific conditions or with different substrates (20, 25, 27, 28).

Although other characterized bacterial GSTs are involved in degradation of xenobiotics and recalcitrant compounds, they either catalyze a specific reaction or are not necessary for the degradative pathway to function (20). To our knowledge, NdmE is the first characterized noncatalytic GST protein that is absolutely required for oxygenase activity and enables soluble NdmCDE expression in E. coli. NdmCDE also exists as a tight complex in the native organism, CBB5, and is required for bacteria to live on caffeine by enabling the final N$_7$-demethylation of 7-methylxanthine to xanthine. The exact structural role that NdmE plays is still under investigation. However, there might be a larger role for NdmE-like proteins in the degradation of other natural products, xenobiotics, and complex hydrocarbons by bacteria, such as those recently released in the Gulf of Mexico (29).

Resting cells of CBB5 strains containing a single deletion of ndmC, ndmD, or ndmE were unable to metabolize 7-methylxanthine via N$_7$-demethylation. Instead, these strains produced small amounts of 7-methyluric acid. Although oxidation of 7-methylxanthine at the C-8 position has not been previously observed in CBB5, CBB5 does oxidize theophylline and 1- and 3-methylxanthines to their respective methyluric acids as dead-end metabolites (4). Similarly, other bacterial strains have been shown to produce 7-methyluric acid from 7-methylxanthine (30, 31). This C-8 oxidation has previously been attributed to a broad-specificity xanthine dehydrogenase (4, 30, 31). That formation of 7-methyluric acid has not been previously observed in CBB5 is most likely due to the speed at which 7-methylxanthine is converted to xanthine by NdmCDE (Fig. 3E).

Further evidence that NdmCDE are required for N$_7$-demethylation comes from a metabolically engineered, caffeine-addicted E. coli strain. The addition of a plasmid containing ndmABCD from CBB5 to the guanine auxotroph E. coli $\Delta$guaB mutant failed to provide growth on caffeine. However, addition of an NdmE homolog gene from Jannichinobacterium sp. strain Marseille enabled a caffeine-plus phenotype. This metabolically engineered strain N-demethylated caffeine to xanthine, which was further converted to guanine (32). These experiments agree with the knockout and purified protein experiments described here and indicate that NdmE is essential for the final N$_7$-demethylation reaction and caffeine-plus growth phenotype.

We have also identified 18 additional organisms with a homologous ndmCDE gene cluster. Eleven of these organisms, which belong to either the Sinorhizobium or Mesorhizobium genus, actually contain an additional ndmC homolog lacking a Rieske center. This is intriguing, since the organisms with two ndmC genes have taken a different evolutionary route than CBB5. In both cases, multiple oxygenases utilize the same RO reductase (NdmD homologs), but in the case of caffeine degradation, only one of the oxygenase proteins, NdmC, utilizes the Rieske center attached to the reductase (NdmD); NdmAB have their own Rieske centers. In the ndmCCDE cases, both of the oxygenases likely utilize the same Rieske domain attached to the NdmD homolog. Currently, none of these genes have been characterized; however, their divergence from the caffeine operon genes and difference in gene organization suggest a different function.

Dissemination of these genes could have been due to linear or horizontal gene transfer or a combination of the two. A 16S rRNA phylogenetic tree placed Sinorhizobium and Mesorhizobium genera on adjacent branches (33), indicating that an ndmC gene duplication event occurred prior to divergence with gene loss in other Sinorhizobium and Mesorhizobium genomes. Alternatively, Rhizobium genomes are known for repeated regions, frequent genome rearrangements, and megaplasmids (34–37). The presence of 2 sets of ndmCCDE gene clusters in Sinorhizobium meliloti 1021 and their position in the phylogenetic tree suggests that these two sets of genes were acquired independently by genomic rearrangement and explain the limited number of organisms containing ndmCCDE genes. Octadecabacter arcticus 238 has transposase sequences flanking the gene region containing the ndmCCDE homologs, thereby providing additional means of genetic mobilization. The caffeine degradation phenotype in CBB5 has not been observed to be transferable, mobile, or lost after culturing in a variety of culture conditions. Therefore, these genes were acquired by a pseudomonal is not evident at this time.

In the case presented here, the Rieske domain has fused to the NdmD RO reductase protein. There are six additional cases lacking this fusion in which the oxygenase, composed of the SRPBBC domain and the Rieske domain, appear in separate proteins, with an adjacent GST and RO reductase. There are an additional 52 bacterial genomes in which the split oxygenase and GST cluster in close proximity without a close RO reductase. In fact, in the 76 present instances involving a split oxygenase, a GST protein is adjacent. Homologous regions in which the oxygenase appears as a single protein lack the GST. This implicates GST in functioning in some way to make the oxygenase, whose domains have been split into two separate proteins, into a functional oxygenase, which represents a completely new function for the GST superfamily.

In summary, P. putida CBB5 utilizes a new combination of RO and GST proteins for N$_7$-demethylation of 7-methylxanthine to
xanthine, which enables it to live on caffeine as the sole carbon and nitrogen source (Fig. 5). The traditional RO has been split in two: the catalytic (SRPBCC) domain is encoded by NdmC, while the Rieske domain has been fused to the N-terminal end of the RO reductase, NdmD. The NdmE GST protein is proposed to have a structural role, making it possible for an oxygenase that is split across two proteins to function. The tight NdmCDE complex acts as a highly specific 7-methylxanthine N7-demethylase, with one O2 molecule consumed per N7-methyl group removed as formaldehyde. This first report of a GST protein that is essential for an RO system adds more knowledge and a new function to the GST enzyme superfamily, and it suggests that biochemical studies of other RO systems need to include adjacent GST proteins for functionality.

REFERENCES