J|A|C|S COMMUNICATIONS

Published on Web 04/26/2010

In Vitro Reconstitution of Mycobacterial Ergothioneine Biosynthesis

Florian P. Seebeck*

Abteilung Physikalische Biochemie, Max Planck Institut für Molekulare Physiologie, Otto-Hahn Strasse 11, 44227 Dortmund, Germany

Received March 1, 2010; E-mail: florian.seebeck@mpi-dortmund.mpg.de

Aerobic cells depend on cysteine derivatives to control intracellular redox potential and metal homeostasis and to fend off electrophilic toxins.[1,2] In addition to their main intracellular thiols, such as glutathione and mycothiol, certain fungi and mycobacteria also produce ergothioneine (**1**, Figure 1), a histidine betaine derivative with a thiol group attached to the C_2 atom of the imidazole ring.[3–5] The function **1** plays in microbial cells is not well understood, but recent findings point to critical functions in human physiology. The human body absorbs **1** from dietary sources and concentrates it in specific tissues or cells such as liver, kidney, central nervous system, and red blood cells.[5] A cation transporter (OCTN1) with high specificity for **1** is responsible for this nonuniform distribution, and both OCTN1 hyperactivity and OCTN1 deficiency exert negative effects on human cells.[6–8] Despite these recent discoveries, the precise function of **1** in human tissue is still a matter of debate.[8,9]

Understanding sulfur metabolism and defense against oxidative stress is an important prerequisite for the development of therapies against infectious and inflammatory diseases and cancer. Therefore, elucidating the biochemistry of **1** is an important challenge.[10–12] To date, the genes for the production of **1** remain entirely unknown. Cell-free extracts from *Neurospora crassa* were shown to assemble **1** from histidine, cysteine, and methionine with hercynine (**2**) as a first intermediate and hercynylcysteine sulfoxide (**4**) as a second intermediate.[13,14] In these experiments, *S*-adenosyl methionine (SAM)-dependent methyltransferases, an iron(II)-dependent oxidase, and a pyridoxal 5-phosphate (PLP)-dependent β-lyase emerged as possible catalysts.[13,14] However, none of these enzymes has been identified or studied in detail.

To find the corresponding genes, I equated identifying the biosynthetic genes for **1** with identifying a SAM-dependent methyltransferase, an iron(II)-dependent oxidase, and a PLP-dependent β-lyase that are present in organisms that produce **1** but absent in nonproducers.[4] Because iron(II)-dependent enzymes are difficult to recognize from sequence data alone[15] and functional assignment of uncharacterized PLP-binding proteins is dicey,[16] I focused on methyltransferases. From a list of all 78 annotated methyltransferases encoded by *Mycobacterium avium* (Table S1 in the Supporting Information), all of the genes without a homologue in *N. crassa* were eliminated. The remaining set of 29 genes was cleared of genes with homologues in either *Escherichia coli* or *Bacillus subtilis*, neither which produces **1**.[4] One of the remaining 10 mycobacterial methyltransferases is encoded next to a PLP-binding protein in the context of a five-gene cluster (Figure 1). This cluster appeared to be a valid candidate for biosynthesis of **1** and was explored in detail. The following experimental data show that these five genes are in fact responsible for this activity and may therefore be designated *egtABCDE*.

To test whether EgtD is a histidine methyltransferase, I cloned the corresponding gene from *Mycobacterium smegmatis* for recombinant expression in *E. coli*. Purified EgtD was assayed for

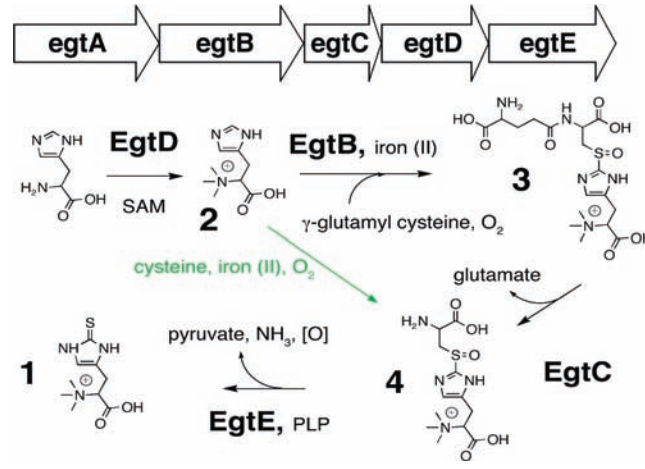

Figure 1. (top) The mycobacterial ergothioneine gene cluster codes for a γ-glutamyl cysteine synthetase (EgtA), an FGE-like protein (EgtB), a glutamine amidotransferase (EgtC), a methyltransferase (EgtD), and a PLP-binding protein (EgtE). (bottom) Reaction sequence of ergothioneine biosynthesis with hercynine (**2**), **3**, and hercynylcysteine sulfoxide (**4**) as intermediates. A previously suggested abbreviated pathway in *N. crassa* is shown in green.[13,14]

methylation activity with each of the proteinogenic amino acids as the substrate in the presence of 5 mM SAM and *S*-adenosylhomocysteine nucleosidase. In these reactions, **2** was the only methylation product detectable by ESI-MS (Figure S2 in the Supporting Information). Similarly, monitoring of SAM consumption using an enzyme-coupled UV assay showed that among the tested amino acids, histidine and α-*N*,*N*-dimethylhistidine are preferred substrates of ErgD (Figure S2).[17]

After EgtD was identified as a histidine methyltransferase, the question arose as to which of the remaining four gene products encoded by *egtABCDE* could possibly catalyze iron(II)-dependent oxidative sulfurization of **2**. EgtA, EgtC, and EgtE showed sequence similarity to a γ-glutamyl cysteine ligase,[18] a putative class-II glutamine amidotransamidase, and a PLP-binding protein, respectively. Inspection of their sequences suggested that their putative catalytic residues are indeed conserved. In contrast, no functional prediction was available for EgtB. This protein shares a common DUF323 domain with formylglycine-generating enzymes (FGEs),[19] which catalyze posttranslational oxidation of cysteine to formylglycine. However, two active-site cysteines that are crucial for O_2 activation in FGEs are absent in EgtB. Instead, EgtB carries an N-terminal domain with three strictly conserved nucleophilic residues (His-X_3-His-X-Glu) as a possible iron(II)-binding motif reminiscent of that in non-heme iron(II) enzymes (Figure S3).[20] Furthermore, fungal homologues of EgtB are encoded as fusion proteins with EgtC homologues, indicating that the two proteins form a functional unit.

10.1021/ja101721e

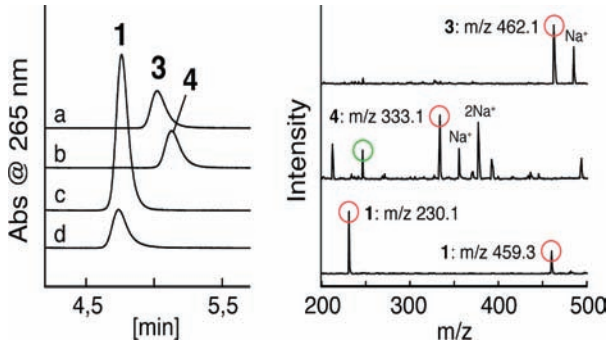

Figure 2. In vitro-reconstituted biosynthesis of **1**. (left) HPLC analysis of enzymatic reactions containing (a) EgtB; (b) EgtB and EgtC; and (c) EtgB, EgtC, and a putative PLP-dependent β-lyase from *E. tasmaniensis*. Authentic **1** coeluted with the final product (d). (right) ESI-MS analysis of isolated products (red circles). **4** partially fragments to the sulfenic acid of **1** during ionization (green circle).

To test the capability of EgtB to catalyze oxidative sulfurization, the recombinant protein was assayed in a buffer solution containing 5 mM **2**, 0.2 mM $FeSO_4$, and 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) (pH 7.6). γ-Glutamylcysteine was chosen as the sulfur donor because EgtA, the first enzyme encoded by the cluster, has been shown to supply this dipeptide.[18] HPLC analysis of this reaction revealed a single product with a strong absorption at 250 nm ($\varepsilon_{250} = 10\,600\ M^{-1}\ cm^{-1}$; Figure S4) and a molecular mass consistent with the formation of **3** (*m*/*z* 462.1; HRMS calcd. 462.16531, found 462.16501) (Figure 2).[13] The proposed structure of **3** (Figure S5) was confirmed by ^{1}H, COSY, and HSQC NMR spectra. The iron dependence of EgtB was confirmed by reconstituting the metal-depleted enzyme with 0.2 mM metal ion [either iron(II), copper(II), manganese(II), or zinc(II)] or 2 mM EDTA. Metal depletion led to drastically reduced protein activity, which completely vanished in the presence of EDTA. Among the tested metals, only iron(II) was able to reconstitute the EgtB activity (Figure S6).[13]

The same HPLC assay was applied to demonstrate that **2** rather than histidine is the preferred sulfur acceptor. Consistent with the earlier observation in cell-free lysates from *N. crassa*, histidine is a poor substrate for EgtB (Figure 1 and Figure S7).[13,14] The same previous study also stated that **2** is directly converted to **4** (Figure 1) using cysteine rather than γ-glutamylcysteine as the substrate. To address this possibility, EgtB was also assayed with cysteine, *N*-acetylcysteine, or glutathione substituting for γ-glutamylcysteine. None of these alternative thiols gave detectable products, supporting the present idea that **3** is a necessary intermediate in the mycobacterial biosynthesis of **1** (Figure S8).

The two final reactions, hydrolysis of the γ-glutamyl amide bond in **3** and conversion of **4** into **1**, are presumably catalyzed by EgtC and EgtE, respectively. Indeed, a reaction containing **2**, γ-glutamylcysteine, and recombinant EgtB and EgtC produced **4** (*m*/*z* 333.1; HRMS calcd. 333.12272, found 333.12287), confirming the role of EgtC as a glutamine amidotransferase (Figure 2). Because recombinant production of soluble EgtE failed, this enzyme was replaced with an unrelated β-lyase from *Erwinia tasmaniensis*. Addition of this β-lyase to the above reaction resulted in a compound (*m*/*z* 230.1; HRMS calcd. 230.09577, found 230.09567) that coeluted with authentic **1** (HRMS found 230.09571) (Figure 2 and Figure S9).

The data presented herein allow a role to be assigned for each of the five Egt enyzmes in the biosynthesis of **1**. The key reactions are catalyzed by a histidine-specific methyltransferase (EgtD) and an iron(II)-dependent enzyme (EgtB) that catalyzes oxidative sulfurization of **2**. A blastp search for this pair of enzymes suggests that production of **1** is a common trait among actinobacteria, cyanobacteria, pezizomycotina, and basidiomycota but also occurs in numerous bacteroidetes and α-, β-, γ-, and δ-proteobacteria (Table S3). The breadth of different lifestyles in which production of **1** is beneficial suggests that this molecule is of fundamental physiological importance. Since higher eukaryotes lack these genes, the biosynthetic pathway of **1** may represent a novel target for antimicrobial therapeutics.[10]

Acknowledgment. I thank R. S. Goody for financial support and valuable comments, B. Grievel for NMR measurements, and S. Gentz for technical assistance.

Supporting Information Available: Figures S1−S9, Tables S1−S3, detailed experimental procedures, and ^{1}H and 2D NMR spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Fahey, R. C. *Annu. Rev. Microbiol.* **2001**, *55*, 333.
(2) Newton, G. L.; Rawat, M.; La Clair, J. J.; Jothivasan, V. K.; Budiarto, T.; Hamilton, C. J.; Claiborne, A.; Helmann, J. D.; Fahey, R. C. *Nat. Chem. Biol.* **2009**, *5*, 625.
(3) Melville, D. B.; Eich, S.; Ludwig, M. L. *J. Biol. Chem.* **1957**, *224*, 871.
(4) Melville, D. B.; Genghof, D. S.; Inamine, E.; Kovalenko, V. *J. Biol. Chem.* **1956**, *223*, 9.
(5) Hartman, P. E. *Methods Enzymol.* **1990**, *186*, 310.
(6) Gründemann, D.; Harlfinger, S.; Golz, S.; Geerts, A.; Lazar, A.; Berkels, R.; Jung, N.; Rubbert, A.; Schömig, E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5256.
(7) Taubert, D.; Jung, N.; Goeser, T.; Schömig, E. *Gut* **2009**, *58*, 312.
(8) Paul, B. D.; Snyder, S. H. *Cell Death Differ.* **2009**, *47*, S149.
(9) Ey, J.; Schömig, E.; Taubert, D. *J. Agric. Food Chem.* **2007**, *55*, 6466.
(10) Bhave, D. P.; Muse, W. B., III.; Carroll, K. S. *Infect. Disord.: Drug Targets* **2007**, *7*, 140.
(11) Winyard, P. G.; Moody, C. J.; Jacob, C. *Trends Biochem. Sci.* **2005**, *30*, 453.
(12) Trachootham, D.; Alexandre, J.; Huang, P. *Nat. Rev. Drug Discovery* **2009**, *8*, 579.
(13) Ishikawa, Y.; Israel, S. E.; Melville, D. B. *J. Biol. Chem.* **1974**, *249*, 4420.
(14) Ishikawa, Y.; Melville, D. B. *J. Biol. Chem.* **1970**, *245*, 5967.
(15) Andreini, C.; Bertini, I.; Rosato, A. *Acc. Chem. Res.* **2009**, *42*, 1471.
(16) Eliot, A. C.; Kirsch, J. F. *Annu. Rev. Biochem.* **2004**, *73*, 383.
(17) Dorgan, K. M.; Wooderchak, W. L.; Wynn, D. P.; Karschner, E. L.; Alfaro, J. F.; Cui, Y. Q.; Zhou, Z. S.; Hevel, J. M. *Anal. Biochem.* **2006**, *350*, 249.
(18) Harth, G.; Maslesa-Galic, S.; Tullius, M. V.; Horwitz, M. A. *Mol. Microbiol.* **2005**, *58*, 1157.
(19) Dierks, T.; Dickmanns, A.; Preusser-Kunze, A.; Schmidt, B.; Mariappan, M.; von Figura, K.; Ficner, R.; Rudolph, M. G. *Cell* **2005**, *121*, 541.
(20) Hegg, E. L.; Que, L. *Eur. J. Biochem.* **1997**, *250*, 625.

JA101721E