ChemPubSoc Europe

DOI: 10.1002/cbic.201402522

Ergothioneine Biosynthetic Methyltransferase EgtD Reveals the Structural Basis of Aromatic Amino Acid Betaine Biosynthesis

Allegra Vit,[b, c] Laëtitia Misson,[a] Wulf Blankenfeldt,[b, c] and Florian P. Seebeck*[a]

Ergothioneine is an *N*-α-trimethyl-2-thiohistidine derivative that occurs in human, plant, fungal, and bacterial cells. Biosynthesis of this redox-active betaine starts with trimethylation of the α-amino group of histidine. The three consecutive methyl transfers are catalyzed by the *S*-adenosylmethionine-dependent methyltransferase EgtD. Three crystal structures of this enzyme in the absence and in the presence of *N*-α-dimethylhistidine and S-adenosylhomocysteine implicate a preorganized array of hydrophilic interactions as the determinants for substrate specificity and apparent processivity. We identified two active site mutations that change the substrate specificity of EgtD 10^7-fold and transform the histidine-methyltransferase into a proficient tryptophan-methyltransferase. Finally, a genomic search for EgtD homologues in fungal genomes revealed tyrosine and tryptophan trimethylation activity as a frequent trait in ascomycetous and basidomycetous fungi.

Introduction

Ergothioneine is a redox-active betaine of 2-thiohistidine occurring in bacteria, fungi, plants, and mammals (Scheme 1).[1] The precise physiological role of ergothioneine is not clear, but several observations indicate that this amino acid derivative might participate in cellular redox defense.[1c,2]

Mycobacteria such as the pathogenic *Mycobacterium tuberculosis* or the saprophytic *Mycobacterium smegmatis* produce ergothioneine in a five-step process starting with N-α-trimethylation of histidine catalyzed by the *S*-adenosylmethyltransferase (SAM) EgtD. The resulting histidine betaine is then subjected to oxidative sulfurization of the imidazole side chain at C2 by sequential action of the enzymes EgtB, EgtC, and EgtE (Scheme 1).

EgtD is the first known methyltransferase that converts a free amino acid directly into its corresponding betaine. In contrast, the common osmolyte and methyl donor glycine betaine is usually produced through a multi-enzyme pathway starting from the precursor choline (Scheme 1).[3] This pathway is specific for glycine betaine and is hardly adaptable to the production of betaines from other amino acids. Nevertheless, amino acid betaines constitute a diverse and ubiquitous class of natural products. The betaine of tryptophan (hypaphorine), for example, is a neuroactive component in peanuts and other Leguminosae species and consequently part of our daily nutrition.[4] In addition, hypaphorine is the most abundant soil alkaloid, because ectomycorrhizal fungi secrete hypaphorine as a suppressor for auxin (indol-3-acetic acid)-controlled root hair development by their host plants.[5] Isolation of hydroxylated or halogenated hypaphorine derivatives from marine invertebrates, or of trimethylated tyrosine or dihydroxyphenylalanine from fungal species, adds to the abundance and diversity of this class of natural products.[6] The biosynthetic origins of amino acid betaines other than trimethylglycine are unknown—until now.

In this report we show that the Methyltransf_33 protein family consists of synthases for histidine, tryptophan, and tyrosine betaine. We discuss the crystal structure and catalytic activity of EgtD in comparison with a double mutant of EgtD that displays specific and proficient tryptophan methyltransfer activity. On the basis of this analysis we identified tyrosine and tryptophan betaine synthases in ascomycetous and basidiomycetous fungi, thus paving the way for systematic and genome-based investigation of amino acid betaines in living systems.

Results and Discussion

EgtD catalyzes trimethylation of histidine with remarkable apparent processivity.[1b] An EgtD-catalyzed reaction involving one equivalent of histidine and one equivalent of *S*-adenosyl-L-methionine (SAM), for example, produces 88±4% trimethylhistidine, together with 13±9% dimethylhistidine, but less than 1% *N*-α-methylhistidine. This processivity allows EgtD to route a fraction of the intracellular histidine pool towards ergothioneine biosynthesis, without interfering with ribosomal protein synthesis. The following analysis suggests that the observed

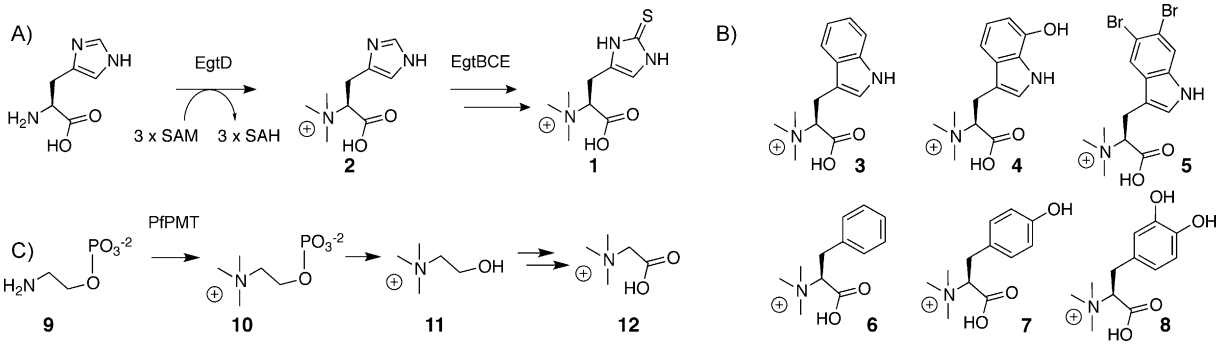
[a] L. Misson,[+] Prof. Dr. F. P. Seebeck
Department for Chemistry, University of Basel
St. Johanns-Ring 19, 4056, Basel (Switzerland)
E-mail: florian.seebeck@unibas.ch

[b] A. Vit,[+] Prof. Dr. W. Blankenfeldt
[3]Structure and Function of Proteins
Helmholtz Centre for Infection Research
Inhoffenstrasse 7, 38124 Braunschweig (Germany)

[c] A. Vit,[+] Prof. Dr. W. Blankenfeldt
Previous address: Department of Biochemistry, University of Bayreuth
Universitätsstrasse 30, 95447 Bayreuth (Germany)

[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201402522.



Scheme 1. A) The first and essential step in ergothioneine biosynthesis is catalyzed by EgtD, an S-adenosylmethionine-dependent methyltransferase.[1b] B) Examples of additional aromatic amino acid betaines isolated from fungi, plants, and marine invertebrates: hypaphorine (**3**),[5] hypaphorine derivatives **4** and **5**,[6c] phenylalanine betaine (**6**), tyrosine betaine (**7**),[6a] and sticticine (**8**).[6b] C) Biosynthesis of glycine betaine (**12**) by the malarial parasite *Plasmodium falciparum*: this synthesis starts with trimethylation of phosphatidylethanolamine (**9**) to phosphocholine (**10**),[3b] followed by dephosphorylation to choline (**11**) and two oxidation steps via the betaine aldehyde to glycine betaine.[3a]

Table 1. Kinetic parameters of the aromatic amino acid betaine synthases EgtD, Ybs, and $EgtD_{M252A,E282A}$.[a]

	EgtD		$EgtD_{E285A}$		Ybs		$EgtD_{M252A,E285A}$	
Substrates	k_{cat} [s^{-1}]	k_{cat}/K_M [$M^{-1}s^{-1}$]	k_{cat} [s^{-1}]	k_{cat}/K_M [$M^{-1}s^{-1}$]	k_{cat} [s^{-1}]	k_{cat}/K_M [$M^{-1}s^{-1}$]	k_{cat} [s^{-1}]	k_{cat}/K_M [$M^{-1}s^{-1}$]
His	5.8×10^{-1}	5.3×10^{3}	1.0×10^{-1}	4.0×10^{1}	2.0×10^{-2}	7.9×10^{2}	–	1.9×10^{0}
N-α-His	2.3×10^{-1}	1.3×10^{4}	–	–	–	–	–	–
N,*N*-α-His	4.3×10^{-1}	1.7×10^{4}	–	–	–	–	–	–
Phe	–	2.4×10^{0}	–	–	1.4×10^{-1}	2.5×10^{1}	3.4×10^{-1}	1.0×10^{2}
Tyr	–	1.1×10^{0}	–	–	1.1×10^{-1}	6.3×10^{3}	–	1.7×10^{1}
Dopa	–	2.2×10^{0}	–	–	1.0×10^{-1}	6.6×10^{2}	–	4.5×10^{0}
Trp	–	2.0×10^{0}	3.0×10^{-2}	4.0×10^{2}	–	7.2×10^{0}	1.1×10^{-1}	5.5×10^{3}

[a] Data represent averages from multiple measurements. The standard error is less than 20% of the average value. Reaction conditions: 50 mM Tris·HCl pH 8.0, 26 °C, 200 μM SAM, 1 u S-adenosylhomocysteine nucleosidase, 1 u adenine deaminase. Ybs: tyrosine betaine synthase from *A. nidulans*.

processivity is a result of increased affinity for the methylated intermediates, rather than increased catalytic efficiency of the second and third methylation steps (Table 1).

We employed a UV-based assay[7] to monitor the rate of EgtD-catalyzed consumption of SAM in the presence of histidine, *N*-α-methyl- or *N*-α-dimethylhistidine. All three EgtD-catalyzed reactions proceed with similar turnover numbers (k_{cat}), and with only two- and threefold increased catalytic efficiency in the presence of *N*-α-methyl- and *N*-α-dimethylhistidine, respectively (Table 1). These slight differences among the three substrates could hardly explain the observed processivity. On the other hand, isothermal titration calorimetry (ITC) showed that EgtD is a four- and 70-times stronger binder of α-*N*-methylhistidine and *N*-α-dimethylhistidine, respectively, than of histidine (Table 2). A similar but weaker trend was observed in the presence of saturating concentrations of *S*-adenosyl-L-homocysteine (SAH).

Table 2. Isothermal calorimetry titration of EgtD.[a]

Titrant	K_D [μM]	ΔH [kcal mol^{-1}]	ΔS [cal mol^{-1} deg^{-1}]
His	290 ± 14	−8 ± 1	−11
N-α-methyl	70 ± 30	−13 ± 3	−25
N-α-dimethyl	4 ± 2	−5 ± 3	8
His[b]	37 ± 1	−10 ± 1	−13
N-α-methyl[b]	14 ± 7	−11 ± 2	−15
N-α-dimethyl[b]	2 ± 1	−27 ± 2	−64
SAH	210 ± 20	−12 ± 4	−25
SAM	270 ± 20	−19 ± 1	−49

[a] Reaction conditions: 20 mM Tris·HCl pH 7.5, 150 mM NaCl, 25 °C, 100 μM EgtD in cell, 5 mM ligand in syringe. [b] EgtD solution contained 7 mM *S*-adenosyl-L-homocysteine (SAH).

To investigate the structural basis for this substrate selectivity we determined the structure of EgtD as an apoprotein (resolution: 1.75 Å), as a binary complex with *N*-α-dimethylhistidine (resolution: 1.9 Å), and as a ternary complex with *N*-α-dimethylhistidine and SAH (resolution 1.5 Å, Tables S1 and S2 in the Supporting Information). The asymmetric unit cells of these crystals each contained two protein chains arranged either in the $P2_12_12_1$ space group (apo form, binary complex) or in the $P2_1$ space group (ternary complex; Figure 2, below). The overall structure consists of a Rossmann-fold domain typical for SAM-dependent methyltransferases.[8] The first 60 residues and a 100-residue insertion after β-strand 5 of the central β-sheet combine to form a second domain. The cleft between the two domains harbors the binding site for histidine and SAM.

Phosphoethanolamine methyltransferase (PfPMT) from *Plasmodium falciparum* and EgtD both catalyze SAM-dependent trimethylation of primary amines in the context of amino acid betaine biosynthesis (Scheme 1). The two structures share a su-

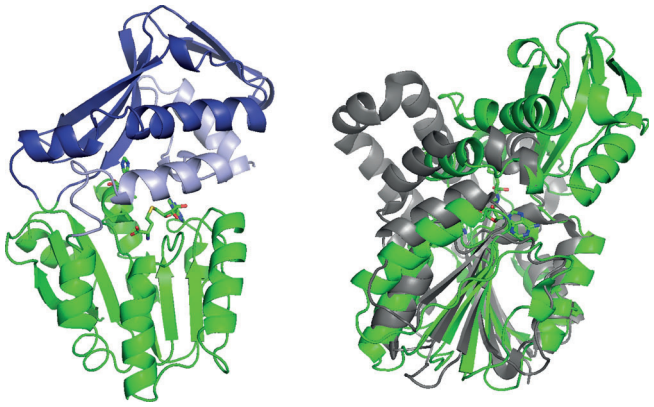


Figure 1. Left. The structure of EgtD in complex with SAH and *N*-α-dimethylhistidine. The substrate-binding site is located in a cleft between a Rossmann-fold domain (green) and a domain formed by residues 15–60 (light blue) and 196–286 (dark blue). Right: Superimposition of EgtD (green) and phosphoethanolamine methyltransferase (PfPMT) [EC:2.1.1.103] from *Plasmodium falciparum* (gray).[3b]

perimposable Rossman-fold domain and similar SAM-binding sites (Figure 1). The second domains, which define substrate specificity, are unrelated;[3b] this suggests that glycine and histidine betaine biosynthesis emerged through separate evolutionary channels.

Closer inspection of the active site in EgtD reveals a network of seven hydrogen bonds between the enzyme and the N-α-dimethylamino acid moiety of the substrate (Figure 2). The carboxylate group hydrogen-bonds to the side chains of Tyr206 (2.6 Å), Ser284 (2.8 Å), Tyr56 (2.6 Å), and Asn166 (2.9 Å). Asn166 also approaches the N-α-dimethylamino moiety of the substrate at hydrogen-bonding distance (2.8 Å), thus suggesting that the N-α-dimethylamino group is protonated.

The interactions between the active site and the methyl groups of the *N*-α-dimethylhistidine should give some indication as to why the methylated intermediate is a 70-times better EgtD ligand than histidine. In the ternary structure (Figure 2), one *N*-α-methyl group is in van der Waals contact with the sulfur atom of SAH (3.4 Å), the oxygen atoms in the side chains of Thr163 (3.5 Å) and Tyr39 (3.5 Å), and the phenyl ring of Phe47 (4.0 Å). The second *N*-α-methyl group makes close contact to the backbone carbonyl group of Gly161 (3.0 Å). An identical set of interactions is observed in the binary complex. This dense array of short contacts between nucleophilic side chains and the polarized *N*-α-methyl groups provides a viable explanation for preferential binding of *N*-α-methyl- and *N*-α-dimethylhistidine versus histidine.[9, 11a, 12]

The side chain of the substrate is clamped by two edge-to-face interactions to Phe216 and Tyr206 (Figure 2). Nτ of the imidazole ring connects to Thr213 through a water-mediated hydrogen bond, and Nπ shares a proton with Glu282 (2.5 Å). This last interaction is essential for efficient EgtD activity: mutation of Glu282 to alanine reduces EgtD activity 10^3-fold because of a 200-fold increase in K_M ($EgtD_{E282A}$, Table 1). Consistently this glutamate residue is highly conserved among all EgtD homologues that are co-encoded with EgtB and are therefore clearly involved in ergothioneine biosynthesis. However, BLAST searches for EgtD homologues in fungal genomes also revealed sequences with apolar residues at position 282. To us this was the first indication that distant EgtD homologues might have diversified to trimethylate amino acids other than histidine.

With the structure of EgtD—the first structure of any member of the Methyltransf_33 protein family—to hand we can now test this idea. EgtD homologues are present in in almost 700 sequenced species.[10] Bacterial genomes in this group usually each encode a single EgtD homologue; these are co-encoded with EgtB and share nearly identical sets of active site residues with EgtD from *M. smegmatis*. It is therefore safe to assume that these bacterial homologues are bona fide histidine methyltransferases involved in ergothioneine biosynthesis. Fungal genomes, on the other hand, encode on average two EgtD homologues. One homologue is usually encoded as a fusion protein with EgtB; this is consistent with the notion that ergothioneine biosynthesis is a highly conserved trait among fungi.[1b, 11] The second EgtD homologues are characterized by considerable variation at positions 213, 252, and 282

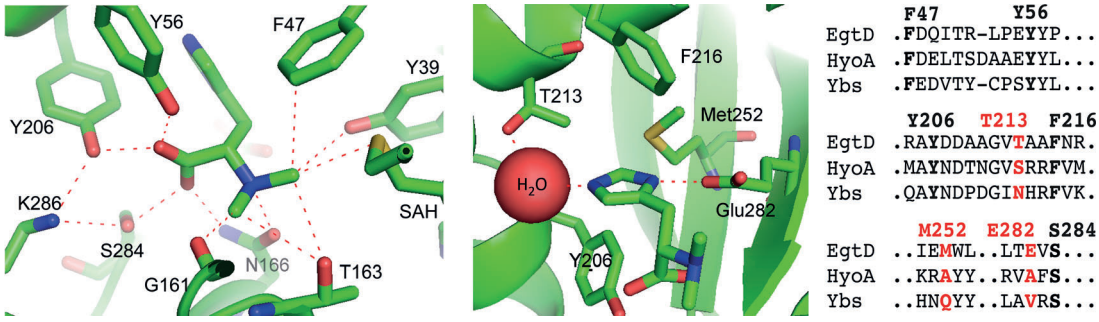


Figure 2. Left: The amino acid moiety of *N*-α-dimethylhistidine is recognized through an array of seven hydrogen bonds, three of which are C–H···O bonds to the protonated *N*-α-dimethyl moiety. Middle: Two additional hydrogen bonds to Glu282 and a Thr213-coordinated water molecule immobilize the imidazole side chain of the substrate. Right: Abbreviated sequence alignment of the histidine methyltransferase from *M. smegmatis* (EgtD), a tyrosine methyltransferase from *Aspergillus nidulans* with 28% sequence identity to $EgtD_{M.\ smegmatis}$ (Ybs, AN8594.2), and a putative tryptophan methyltransferase from *Dichomitus squalens* with 26% sequence identity to $EgtD_{M.\ smegmatis}$ (HyoA, DICSQ_157002). In red: the active sites of the three enzymes differ most significantly at positions 213, 252 and 282.

(EgtD numbering, Figure 2), thus indicating that these methyltransferases might be specific for alternative substrates.

One such subclass contains the uncharacterized protein AN8594.2 from the ascomycetous fungus *Aspergillus nidulans* (Figure 3, below). This protein consists of a methyltransferase domain with 28% sequence identity to $EgtD_{M.\ smegmatis}$ and a putative iron(II)-dependent hydroxylase domain (TauD-like domain). We produced the methyltransferase domain of AN8594.2 in *E. coli* and assayed its methyltransferase activity for each of the 20 standard amino acids.

This screen identified tyrosine as the preferred substrate and tyrosine betaine as the corresponding product. The fungal enzyme catalyzed tyrosine methylation with a catalytic efficiency comparable to that of EgtD-catalyzed histidine methylation (Table 1). Phenylalanine, dihydroxyphenylalanine (DOPA), and histidine are ten times less efficient substrates, and tryptophan or non-aromatic amino acids are at least 10^3 times poorer substrates. A structural model of AN8594.2 based on the EgtD structure suggests that the fungal enzyme recognizes tyrosine through hydrogen bonds between the phenol side chain and the asparagine at position 213 (EgtD numbering, Figure 2) and/or a glutamine side chain at position 252.

On the basis of the observed in vitro activity we suggest that this fungal EgtD homologue is a tyrosine betaine synthase (Ybs). Presently, we have no information about the function of the TauD-like domain. However, we note that the betaines of tyrosine and dihydroxyphenylalanine (sticticine, **8**, Scheme 1) are both known metabolites from ascomycetous fungi.[6a,12] It therefore seems possible that the TauD-like domain is responsible for the transformation of tyrosine betaine into sticticine.

A BLAST search for EgtD homologues in genomes from basidiomycetous fungi revealed a different subclass of potential amino acid betaine synthases (Figure 3). One member of this

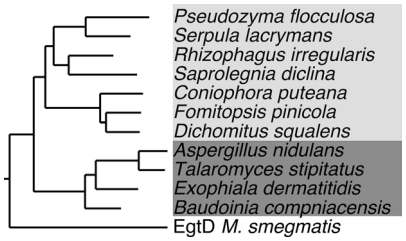


Figure 3. Basidomycetes (light gray) and Ascomycetes (gray) encode putative EgtD homologues with tryptophan (light gray) and tyrosine (gray) specificity. *Serpula lacrymans*, *Saprolegnia diclina*, *Coniophora puteana*, *Fomitopsis pinicola*, and *Dichomitus squalens* are saprophytic fungi, whereas *Rhizophagus irregularis* is a mycorrhizal fungus.

subclass is the uncharacterized protein DICSQ_157002 from the wood-rotting fungus *Dichomitus squalens*. Sequence alignment with $EgtD_{M.\ smegmatis}$ indicates that this protein contains smaller residues at positions 213, 252, and 282 (EgtD numbering, Figure 2). The corresponding substrate-binding pocket should be more spacious than that in EgtD and might accommodate tryptophan as a substrate. On the basis of this alignment we surmised that DICSQ_157002 might be a hypaphorine synthase (HyoA). In support of this prediction we were able to

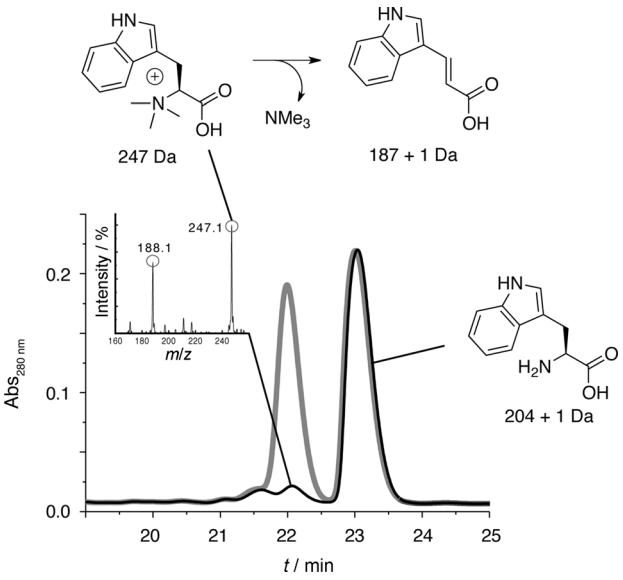


Figure 4. Detection of hypaphorine in malt-extract-grown *D. squalens* by ESI-MS and RP-HPLC (black line). Hypaphorine was identified by ESI-MS (*m/z* calcd 247.14; measured 247.13), by HRMS (*m/z* calcd 247.1441; measured 247.1441, consistent with an elemental composition $C_{14}H_{19}N_2O_2^+$), by co-elution with an authentic sample of hypaphorine (gray line), and by ESI-induced fragmentation to trimethylamine and indol-3-acrylic acid (*m/z* calcd 188.1; measured 188.1).

detect hypaphorine as a metabolite in *D. squalens* growing on maltose extract agarose (Figure 4). The isolated compound co-eluted with authentic hypaphorine on reversed-phase HPLC, the elemental composition as determined by HRMS (ESI) was according to expectation, and the compound fragmented to indol-3-acrylic acid during MS analysis. Discovery of hypaphorine in a saprophytic fungus is surprising and suggests that this redox-active alkaloid might have a physiological role in addition to supporting fungus-to-plant communication in the ectomycorrhiza.[13]

Because we were unable to produce HyoA from *D. squalens* in *E. coli* we engineered an EgtD variant with an active site configuration similar to that in the fungal protein ($EgtD_{M252A,E282A}$, Figure 2). The two mutations Met252 to Val and Glu282 to Ala increased tryptophan-specific activity more than 2×10^3-fold and reduced the histidine-specific activity 3×10^3-fold. Despite this dramatic change in substrate specificity $EgtD_{M252A,E282A}$ still discriminates against other possible substrates such as phenylalanine, tyrosine, or DOPA (Table 1) with strictness similar to those of the evolved enzymes EgtD and Ybs.

To explain the structural basis for this excellent substrate specificity we also determined the crystal structure of $EgtD_{M252A,E282A}$ in complex with tryptophan and SAH (resolution 1.8 Å, Figure 5). The complex structure contains weak but significant residual electron density due to tryptophan bound in the mutant active site (Figure S1). The binding mode is very similar to that described for the complex between EgtD and *N*-α-dimethylhistidine: the indole ring is pinched by the aromatic rings of Tyr206 and Phe216 (Figure 2) and engages in a water-mediated hydrogen bond to Thr213 (Figure 5). This combina-

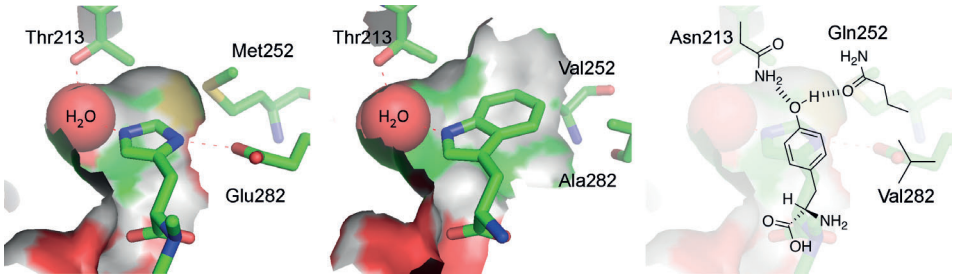


Figure 5. Comparison of substrate binding sites of EgtD wild type in complex with *N*-α-dimethylhistidine (left) and of $EgtD_{M252,E282A}$ in complex with tryptophan (middle). A model of the tyrosine methyltransferase from *A. nidulans* (right) suggests that tyrosine is recognized by hydrogen bonding to the residues Asn213 and Gln252 (residue numbering based on the EgtD sequence).

tion of hydrophobic and hydrophilic interactions provides the basis for the observed tryptophan specificity of $EgtD_{M252A,E282A}$. Tyrosine or phenylalanine could not establish a hydrogen bond to Thr213, and are therefore rejected as substrates.

The almost 10^7-fold change in enzyme specificity ($[k_{cat,His}/K_{M,His}]/[k_{cat,Trp}/K_{M,Trp}]$) as a result of two point mutations illustrates the ease with which new enzyme activity can emerge.[14] More specifically, this observation suggests that the Methyltransf_33 protein family might include many more amino acid betaine synthases.

Conclusion

We have described the crystal structure of EgtD, thus providing a first view into the ubiquitous Methyltransf_33 protein family. The structure in complex with the ligands *N*-α-dimethylhistidine and SAH revealed determinants for substrate selectivity and product specificity. On the basis of this analysis we identified tyrosine betaine synthases and tryptophan betaine synthases in fungal genomes; this suggests that aromatic amino acid betaines might be common metabolites in eukaryotic microorganisms. Finally, EgtD is an essential enzyme for ergothioneine biosynthesis in Mycobacteria.[2c] Inhibition of this pathway might present a target for new therapeutics against *Mycobacterium tuberculosis*. This structure provides the basis for rational design of ergothioneine biosynthesis inhibitors to test this concept.

Experimental Section

Protein production, crystallization, data collection, structure solution, and refinement: Native and seleno-L-methionine-labeled EgtD, as well as $EgtD_{M252A,E282A}$, were produced as described previously.[15] Crystallization of native and seleno-L-methionine-labeled EgtD was achieved as described previously.[15] For co-crystallization of native EgtD and $EgtD_{M252A,E282A}$ with substrates, new crystallization conditions had to be determined, and screens were set up with a Phoenix nano-dispensing robot (Art Robbins Instruments). Crystals of EgtD with *N*-α-dimethylhistidine (EgtD_AVI) were obtained by mixing protein and ligand at concentrations of 5 mM *N*-α-dimethylhistidine and 20 mg mL^{-1} EgtD in a drop (0.5 μL) mixed 1:1 with reservoir [sodium phosphate (1.6 M), dipotassium phosphate (0.4 M), sodium phosphate citrate (0.1 M), pH 4.2]. Crystals of EgtD with *N*-α-dimethylhistidine and SAH (EgtD_AVI_SAH) were obtained by mixing both compounds with native EgtD to a final concentration of 5 mM *N*-α-dimethylhistidine, 5 mM SAH, and 20 mg mL^{-1} EgtD. This solution was mixed 1:1 with reservoir [magnesium chloride (0.2 M), Tris·HCl (pH 8.5, 0.1 M), PEG 4000 (30%, *w*/*v*)] and equilibrated against reservoir (500 μL). Crystals of $EgtD_{M252A,E282A}$ with L-tryptophan and SAH ($EgtD_{M252A,E282A}$_TRP_SAH) were obtained by preincubation of $EgtD_{M252A,E282A}$ (15 mg mL^{-1}) with L-tryptophan (5 mM) and SAH (5 mM). This solution was mixed 1:1 with reservoir [PEG 3350 (20%, *w*/*v*) and magnesium acetate (0.2 M)]. All crystallization experiments were performed at room temperature. Prior to data collection, crystals of EgtD_AVI were cryoprotected in reservoir solution supplemented with glycerol (5%, *v*/*v*). No cryoprotection was required for crystals of EgtD_AVI_SAH. Crystals of $EgtD_{M252A,E282A}$_TRP_SAH were cryoprotected in reservoir solution supplemented with PEG 3350 (35%, *w*/*v*).

Diffraction data for seleno-L-methionine-labeled and native apo form crystals of EgtD were collected at beamline PXII of the Swiss Light Source (SLS at the Paul Scherrer Institute, Villingen, Switzerland) as described previously.[15] Diffraction data for EgtD_AVI, EgtD_AVI_SAH, and $EgtD_{M252A,E282A}$_TRP_SAH were collected at 100 K at a wavelength of $\lambda = 0.918$ Å on beamline BL 14.1 of the BESSY II synchrotron (Helmholtz-Zentrum Berlin, Germany).

The single anomalous dispersion (SAD) data for the seleno-L-methionine-labeled protein were processed as described[15] and led, with the aid of autoSHARP,[16] to a structure solution of the EgtD apo form. The model was then automatically built with ARP/wARP 7.0.[17]

Diffraction data for the EgtD complex structures (EgtD_AVI, EgtD_AVI_SAH, $EgtD_{M252A,E282A}$ _TRP_SAH) were indexed with XDS[18] and then processed with COMBAT and SCALA from the CCP4 program suite.[19] In the case of EgtD_AVI, structure solution was carried out by rigid body refinement in REFMAC5[19b] with use of the EgtD apo structure. The apo structure of EgtD was also used in MOLREP[20] to solve the structure of EgtD_AVI_SAH. This structure was used to phase the structure of $EgtD_{M252A,E282A}$ by rigid body refinement in REFMAC5.[19b]

For structure refinement, ligands were prepared with ProDRG[21] for REFMAC5[19b] and eLBOW[22] for phenix.refine.[23] Model building was done in COOT.[23] After refinement and model building, all structures were checked with MolProbity.[24] All figures were prepared with PyMOL.[25] Details of the data collection and refinement statistics can be found in Tables S1 and S2. Coordinates and structures factors have been deposited in the protein data bank with IDs 4PIM, 4PIN, 4PIO, and 4PIP.[26]

Isothermal titration calorimetry (ITC): The different compounds used for ITC measurements were dissolved in Tris/HCl (pH 7.5, 20 mM), NaCl (150 mM). Directly before the measurements, EgtD was eluted from a NAP5 Sephadex column (GE Healthcare) primed with the exact same buffer to enhance baseline stability. Protein solutions were then diluted to a concentration of 100 μM and transferred into the sample cell of an isothermal titration calorime-

ter (VP200-ITC system, MicroCal LLC). Compounds (L-histidine, *N*-methyl-L-histidine, *N*-α-dimethylhistidine, SAM, SAH) were used in different concentrations (5, 7, or 10 mM) and were added to the protein solution by syringe/stirrer in 2 μL steps every 2 min. All ITC measurements were conducted at 25 °C. Data were plotted as the power needed to maintain the reference and sample cell at the same temperature against time and as kcal mol^{-1} of injectant against the molar ratio of ligand and protein. The program Origin7 (OriginLab Corporation) was used to analyze the data.

Cloning and production of Ybs: The gene coding for the methyltransferase domain (Methyltransf_33) of the protein AN8594.2 from *Aspergillus nidulans* was purchased from GeneScript with codon adaptions for optimal production in *Escherichia coli*. This fragment was cloned into a pOPIN-expression vector by using restriction-free protocols. The resulting vector encodes Ybs as a fusion with an N-terminal hexahistidine tag. Ybs was produced in *E. coli* Bl21 cell by the same procedure as described for EgtD.[15]

Construction of EgtD variants—EgtD$_{E282A}$ and EgtD$_{M252A,E282A}$: The two variants were constructed by using mutagenesis primers: M252Vs: 5′-ATC GAG GTT TGG TTG CGT GCC CGC A-3′, M252Va: 5′-CAC GCA ACC AAA CCT CGA TGC GTT CCT-3′, E282As: 5′-GAT GCT CAC CGC AGT GTC CTG CAA GTT-3′, E282Aa: 5′-TGC AGG ACA CTG CGG TGA GCA TCT CCT-3′, EgtDs: 5′-ATA TCA TAT GGC GCT CTC ACT GGC CAA-3′, EgtDa: 5′-ATA TCT CGA GTC ACC GCA CCG CCA GCG ACA-3′. The gel-purified fragment was digested with NdeI and XhoI restriction enzymes and ligated to a modified pET19 expression vector. The proteins were produced and purified by the same protocols as used for the EgtD wild type.

Methyltransferase assay: Methyltransferase activity was determined by published protocols.[7] Reactions were monitored in a 2 mm quartz cuvette at 25 °C and 265 nm. The 200 μL reaction mixtures each contained Tris·HCl (pH 8, 50 mM), NaCl (50 mM), $MnBr_2$ (100 μM), SAM (200 μM), adenine deaminase (4.5 μM), AdoHcy nucleosidase (12 μM), methyltransferase (2–4 μM), and the appropriate amino acid at a concentration of 5–500 μM. The data were fitted to the function $v = k_{cat}[s]/(K_M + [s])$. Methyltransferase products were identified by ESI-MS (trimethylhistidine: *m/z* calcd: 198.12; found: 198.1; trimethyltyrosine: *m/z* calcd: 224.13; found: 224.1; trimethyltryptophan: *m/z* calcd: 247.14; found: 247.1; trimethylated dihydroxyphenylalanine: *m/z* calcd: 240.12; found: 240.1).

Adenosylhomocysteine nucleosidase: AdoHcy was produced in *E. coli* BL21 cells grown in LB medium [kanamycin (50 mg L^{-1}), chloramphenicol (34 mg L^{-1})] and induced with isopropyl-β-D-thiogalactopyranoside (IPTG; 1 mM) at 37 °C for 3 h. Cells were pelleted and resuspended in phosphate buffer (pH 8.0, 50 mM) and NaCl (300 mM). Cells were lysed at 4 °C by sonication, and the cleared lysate was mixed with Ni^{II}-NTA agarose (Qiagen) at 4 °C for 20 min. The agarose beads were washed with phosphate buffer containing imidazole (10 and 20 mM). The protein was eluted in an imidazole solution (250 mM). The purified eluted protein was dialyzed into Tris·HCl buffer (pH 8.0, 50 mM), NaCl (50 mM), β-mercaptoethanol (5 mM) and stored at −80 °C. To approximate the concentration of the prepared protein, we used a calculated molar absorption coefficient of ε_{280}(AdoNuc) = 5960 M^{-1} cm^{-1}.

Adenine deaminase: Adenine deaminase was produced in *E. coli* BL21 cells grown in LB medium [kanamycin (50 mg L^{-1}), chloramphenicol (34 mg L^{-1})] with bipyridine (50 mM) and induced with IPTG (1 mM) and $MnCl_2$ (1 mM) at 37 °C for 3 h. Cells were pelleted and resuspended in phosphate buffer (pH 8.0, 50 mM), Tween 20 (0.1%), and NaCl (300 mM). Cells were lysed at 4 °C by sonication, and the cleared lysate was mixed with Ni^{II}-NTA agarose (Qiagen) at 4 °C for 20 min. The agarose beads were washed with phosphate buffer containing imidazole (10 and 20 mM), glycerol (10%), and Tween 20 (0.1%). The protein was eluted in a solution of imidazole (250 mM), glycerol (10%), and Tween 20 (0.1%). The purified eluted protein was dialyzed into Tris·HCl buffer (pH 8.0, 50 mM), NaCl (50 mM), β-mercaptoethanol (5 mM) and stored at −80 °C (ε_{280}(deaminase) = 41 370 M^{-1} cm^{-1}).

Detection of hypaphorine production: *D. squalens* was purchased from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSM9615) as an actively growing culture. The fungus was grown on malt extract agar at 22 °C for three weeks. Mycelium (13 g) was extracted with a MeOH/H_2O (9:1) mixture. The concentrated extract was analyzed by HPLC (RP C_{18}).

Note Added in Proof

Jeong et al. published the crystal structure of EgtD in the apo form, in complex with histidine or histidine and SAH (J.-H. Jeong, H. J. Cha, S.-C. Ha, C. Rojviriya, Y.-G. Kim, *Biochem. Biophys. Res. Commun.* **2014**, *452*, 1098–1103, DOI: 10.1016/j.bbrc.2014.09.058). These structures provide a valuable complement to the EgtD complex with SAH and *N*-α-dimethylhistidine described here.

Acknowledgements

The authors are indebted to Roger Goody, who enabled the initiation of this project at the MPI for Molecular Physiology in Dortmund. The authors would like to thank the X-ray communities at the MPI in Heidelberg and Dortmund and the University of Bayreuth for help with data collection, and the Swiss Light Source (PSI, Villingen) and BESSY II (Helmholtz Centre Berlin, Berlin) for beamline access. This project was supported by the Swiss National Foundation. F.P.S. is supported by the "Professur für Molekulare Bionik". A.V. is supported by the HZI Graduate School for Infection Research.

Keywords: amino acids · betaines · biosynthesis · ergothioneine · hypaphorine · methyltransferases

[1] a) D. B. Melville, M. L. Ludwig, E. Inamine, J. R. Rachele, *J. Biol. Chem.* **1959**, *234*, 1195–1198; b) F. P. Seebeck, *J. Am. Chem. Soc.* **2010**, *132*, 6632–6633; c) M. H. Bello, V. Barrera-Perez, D. Morin, L. Epstein, *Fungal Genet. Biol.* **2012**, *49*, 160–172; d) C. Pfeiffer, B. Surek, E. Schöming, D. Gründemann, *Food Chem.* **2011**, *129*, 1766–1769; e) C. Tanret, *C. R. Acad. Sci.* **1909**, *149*, 222–224; f) D. Gründemann, *Prev. Med.* **2012**, *54*, S71–S74; g) I. K. Cheah, B. Halliwell, *Biochim. Biophys. Acta* **2012**, *1822*, 784–793; h) R. C. Fahey, *Biochim. Biophys. Acta* **2013**, *1830*, 3182–3198.

[2] a) I. K. Cheah, R. L. Ong, J. Gruber, T. S. Yew, L. F. Ng, C. B. Chen, B. Halliwell, *Free Radical Res.* **2013**, *47*, 1036–1045; b) B. D. Paul, S. H. Snyder, *Cell Death Differ.* **2009**, *16*, S149–S149; c) C. Sao Emani, M. J. Williams, I. J. Wiid, N. F. Hiten, A. J. Viljoen, R.-D. D. Pietersen, P. D. van Helden, B. Baker, *Antimicrob. Agents Chemother.* **2013**, *57*, 3202–3207.

[3] a) T. H. Chen, N. Murata, *Plant Cell Environ.* **2011**, *34*, 1–20; b) S. G. Lee, Y. Kim, T. D. Alpert, A. Nagata, J. M. Jez, *J. Biol. Chem.* **2012**, *287*, 1426–1434.

[4] a) A. Tsopmo, A. D. Muir, *J. Agric. Food Chem.* **2010**, *58*, 8715–8721; b) B. O. Keller, B. T. Wu, S. S. Li, V. Monga, S. M. Innis, *J. Agric. Food Chem.* **2013**, *61*, 7654–7660.

[5] T. Kawano, *Plant Cell Rep.* **2003**, *21*, 829–837.

[6] a) T. Bernard, G. Goas, J. Hamelin, M. Joucla, *Phytochemistry* **1981**, *20*, 2325–2326; b) T. Bernard, M. Joucla, G. Goas, J. Hamelin, *Phytochemistry*

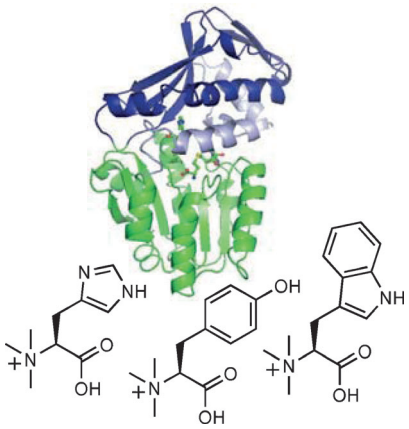
1980, *19*, 1967 – 1969; c) A. Mollica, M. Locatelli, A. Stefanucci, F. Pinnen, *Molecules* **2012**, *17*, 6083 – 6099.
[7] K. M. Dorgan, W. L. Wooderchak, D. P. Wynn, E. L. Karschner, J. F. Alfaro, Y. Q. Cui, Z. S. Zhou, J. M. Hevel, *Anal. Biochem*. **2006**, *350*, 249 – 255.
[8] R. Gana, S. Rao, H. Huang, C. Wu, S. Vasudevan, *BMC Struct. Biol.* **2013**, *13*, 6.
[9] a) S. Horowitz, L. M. A. Dirk, J. D. Yesselman, J. S. Nimtz, U. Adhikari, R. A. Mehl, S. Schreiner, R. L. Houtz, H. M. Al-Hashimi, R. C. Trievel, *J. Am. Chem. Soc.* **2013**, *135*, 15536 – 15548; b) S. Horowitz, R. C. Trievel, *J. Biol. Chem.* **2012**, *287*, 41576 – 41582.
[10] R. D. Finn, A. M. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, S. R. Eddy, A. Heder, K. Hetherington, J. Mistry, E. L. Sonnhammer, J. Tate, M. Punta, *Nucleic Acids Res.* **2014**, *42*, D222 – D230.
[11] G. W. Jones, S. Doyle, D. A. Fitzpatrick, *Gene* **2014**, *549*, 161 – 170.
[12] C. A. Carollo, A. L. Calil, L. A. Schiave, T. Guaratini, D. W. Roberts, N. P. Lopes, G. U. Braga, *Fungal Genet. Biol.* **2010**, *114*, 473 – 480.
[13] C. J. Gagliardi, R. A. Binstead, H. H. Thorp, T. J. Meyer, *J. Am. Chem. Soc.* **2011**, *133*, 19594 – 19597.
[14] a) M. D. Toscano, K. J. Woycechowsky, D. Hilvert, *Angew. Chem. Int. Ed.* **2007**, *46*, 3212 – 3236; *Angew. Chem.* **2007**, *119*, 3274 – 3300; b) H. W. Wilks, K. W. Hart, R. Feeney, C. R. Dunn, H. Muirhead, W. N. Chia, D. A. Barstow, T. Atkinson, A. R. Clarke, J. J. Holbrook, *Science* **1988**, *242*, 1541 – 1544.
[15] A. Vit, L. E. Misson, W. Blankenfeldt, F. P. Seebeck, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2014**, *70*, 676 – 680.
[16] C. Vonrhein, E. Blanc, P. Roversi, G. Bricogne, *Methods Mol. Biol.* **2007**, *364*, 215 – 230.
[17] G. Langer, S. X. Cohen, V. S. Lamzin, A. Perrakis, *Nat. Protoc.* **2008**, *3*, 1171 – 1179.
[18] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125 – 132.
[19] a) P. R. Evans, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 282 – 292; b) G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 355 – 367.
[20] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* **1997**, *30*, 1022 – 1025.
[21] A. W. Schüttelkopf, D. M. F. van Aalten, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 1355 – 1363.
[22] N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2009**, *65*, 1074 – 1080.
[23] P. D. Adams, P. V. Afonie, G. Bunkoczi, V. B. Chen, I. W. Davies, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213 – 221.
[24] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 12 – 21.
[25] W. L. DeLano, *The PyMol User's Manual*, DeLano Scientific, San Carlos, CA, USA; **2002**, 2002.
[26] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235 – 242.

Received: September 9, 2014
Published online on ■■ ■■, 0000

*A. Vit, L. Misson, W. Blankenfeldt, F. P. Seebeck**

■■ – ■■

Ergothioneine Biosynthetic Methyltransferase EgtD Reveals the Structural Basis of Aromatic Amino Acid Betaine Biosynthesis



Transferase transformations: We have determined crystal structures of the histidine methyltransferase EgtD. Two active-site mutations were shown to convert this enzyme into a proficient tryptophan methyltransferase. A bioinformatics search based on these substrate specificity determinants in EgtD identified tyrosine- and tryptophan-specific methyltransferases in many fungal genomes, suggesting that aromatic amino acid betaine biosynthesis is a frequent fungal trait.