A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Iterative L-tryptophan methylation in Psilocybe evolved by sub-domain duplication

Authors: Felix Blei, Janis Fricke, Jonas Wick, Jason Slot, and Dirk Hoffmeister

To be cited as: *ChemBioChem* 10.1002/cbic.201800336

Link to VoR: http://dx.doi.org/10.1002/cbic.201800336

A Journal of ChemPubSoc Europe

WILEY-VCH www.chembiochem.org

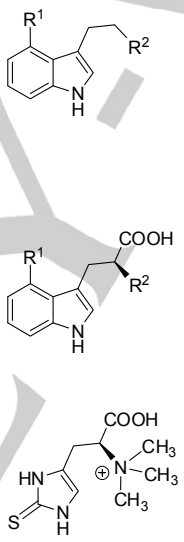
Iterative L-tryptophan methylation in *Psilocybe* evolved by sub-domain duplication

Felix Blei,[a] Janis Fricke,[a] Jonas Wick,[a] Jason C. Slot,*[b] and Dirk Hoffmeister*[a]

Abstract: *Psilocybe* mushrooms are best known for their L-tryptophan-derived psychotropic alkaloid psilocybin. The dimethylation of its precursor norbaeocystin by the enzyme PsiM is a critical step during its biosynthesis. However, the "magic" mushroom *Psilocybe serbica* also mono- and dimethylates L-tryptophan, which is incompatible with the specificity of PsiM. Here, a second methyltransferase, TrpM, was identified and functionally characterized. Mono- and dimethylation activity on L-tryptophan was reconstituted *in vitro*, while tryptamine was rejected as substrate. Therefore, we describe a second L-tryptophan-dependent pathway in *Psilocybe* which is not part of the psilocybin biosynthesis. TrpM is unrelated to PsiM, but originates from a retained ancient duplication event of a portion of the *egtDB* gene which encodes an ergothioneine biosynthesis enzyme. During mushroom evolution, this duplicated gene was widely lost, but re-evolved sporadically and independently in various genera. We propose a new secondary metabolism evolvability mechanism, in which weakly selected genes are retained through preservation in a widely distributed, conserved pathway.

The fungal genus *Psilocybe* comprises numerous mushroom species that biosynthesize tryptamine alkaloids, primarily psilocybin (4-phosphoryloxy-*N*,*N*-dimethyltryptamine, **1,** Figure 1).[1] It represents a prodrug to its dephosphorylated congener psilocin (**2**), which agonistically targets serotonin receptors in the central nervous system.[2] The colloquial term "magic mushrooms" for these fungi alludes to their psychotropic effects and to their use as recreational drug. However, clinical trials have recognized **1** as a valuable candidate to be developed into a medication against cancer-related anxiety and treatment-resistant depression.[3]

Relying on *Psilocybe cubensis*, *P. cyanescens*, and other species, evolutionary, ecological, and biochemical aspects of **1** biosynthesis have been investigated by our laboratories.[4] Evolutionarily, **1** production has been advantageous to alter the behaviour of mycophagous and wood-inhabiting insects that share their habitat with the fungus. For its biosynthesis, L-tryptophan (**3**) is fed into a sequential four-step pathway, catalyzed by the decarboxylase PsiD, the P_{450} monooxygenase PsiH, the kinase PsiK, and the *N*-methyltransferase PsiM (Scheme 1). Importantly, The preference of PsiM for norbaeocystin (Figure 1), i.e., a substrate with a 4-phosphoryloxy group, places this enzyme at the end of the **1** biosynthetic cascade.[4a,4c] However, during mass spectrometric analyses of *P. serbica* extracts, we surprisingly found signals that were consistent with the masses of L-abrine (=*N*-α-methyl-L-tryptophan, **4**) and *N*,*N*-α-dimethyl-L-tryptophan (**5**), respectively (Figure 2), which are known natural products, e.g., from the jequirity bean (*Abrus precatorius*).[5] Since PsiM does not accept **3** as substrate, occurrence of **4** and **5** in *P. serbica* indicates a PsiM-independent methylation by an as yet unknown enzyme.



psilocybin (**1**): R^1=O-PO_3H_2, R^2=$N(CH_3)_2$
psilocin (**2**): R^1=OH, R^2=$N(CH_3)_2$
norbaeocystin: R^1=O-PO_3H_2, R^2=NH_2
tryptamine: R^1=H, R^2=NH_2
N,*N*-dimethyltryptamine: R^1=H, R^2=$N(CH_3)_2$
4-hydroxytryptamine: R^1=OH, R^2=NH_2

L-tryptophan (**3**): R^1=H, R^2=NH_2
L-abrine (**4**): R^1=H, R^2=$NHCH_3$
N,*N*-dimethyl-L-tryptophan (**5**): R^1=H, R^2=$N(CH_3)_2$
4-hydroxy-L-tryptophan: R^1=OH, R^2=NH_2
L-hypaphorine (**6**): R^1=H, R^2=$N^+(CH_3)_3$

ergothioneine (**7**)

Figure 1. Structures of indole alkaloids isolated from *Psilocybe* and other fungal species, products of the methyltransferase TrpM, and of ergothioneine.

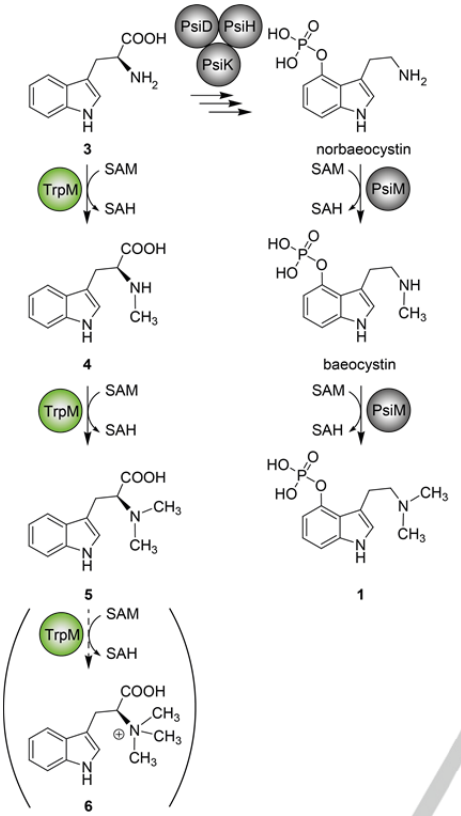
We here report on the functional and phylogenetic characterization of TrpM, an *S*-adenosyl-L-methionine (SAM)-dependent methyltransferase of *P. serbica*. This enzyme does not take part in **1** biosynthesis, but catalyzes a second pathway that originates from **3** and includes formation of a tertiary amine by iterative *N*-methylation.

PsiM-independent **3** methylation in *Psilocybe* mushrooms appeared plausible, given the report on EgtD, a mycobacterial enzyme that features a MT_33 methyltransferase superfamily domain. EgtD is required for ergothioneine (**7**) biosynthesis and processively trimethylates L-histidine.[6] **7** occurs in various fungi (and bacteria and plants alike) and is hypothesized to participate in cellular redox processes.[6a] Specifically, crystallographic work and subsequent targeted active site engineering of EgtD ($EgtD_{E282A,M252V}$) to emulate a fungal active site converted this enzyme into a **3** methyltransferase. Combined with surveys on fungal genomic data, this finding suggested that mushrooms use

[a] F. Blei, J. Fricke, J. Wick, Prof. Dr. D. Hoffmeister
Department Pharmaceutical Microbiology at the Hans Knöll Institute
Friedrich-Schiller-Universität
Beutenbergstrasse 11a, 07745 Jena (Germany)
E-mail: dirk.hoffmeister@leibniz-hki.de
[b] Prof. Jason C. Slot, Ph.D.
Department of Plant Pathology
Ohio State University
2021 Coffey Road
Columbus, OH 43210 (USA)
E-mail: slot.1@osu.edu

Supporting information for this article is given via a link at the end of the document.

EgtD-like methyltransferases to produce L-hypaphorine (*N*,*N*,*N*-α-trimethyl-L-tryptophan, **6,** Figure 1).[6a] This fungal product is ecologically relevant as an agonist of the phytohormone indole-3-acetic acid, which is e.g., secreted by the tree partner in ectomycorrhizal symbioses.[7]



Scheme 1. Methyl transfer steps during biosynthetic pathways to **5** and **1** in *Psilocybe serbica*. The parentheses and the dashed arrow indicate low turnover to **6** by TrpM.

The genomic sequences of a North American *P. cyanescens* isolate is published, as well as a Central European isolate of *P. serbica* (which falls into *P. cyanescens* species complex in the wide sense).[4] This sequence data was browsed for further genes encoding small molecule methyltransferases. In the European isolate (*P. serbica* FSU12416), two *egtD*-like genes were identified: one was fused to an *egtB* homolog, which is the expected gene for sulfoxide synthase required for putative **7** assembly. Fused *egtDB* genes for **7** biosynthesis are common in fungi. The second gene encoded a non-fusion methyltransferase, which we considered a candidate enzyme for **3** methylation. Standalone *egtD*-like genes rarely exist in basidiomycetes. Examples include HyoA,[6a] a putative **6** synthase of the mushroom *Dichomitus squalens* which shares 37% identical amino acids, and a hypothetical protein of *Galerina marginata* (protein ID: KDR66884.1, 70% identical aa).

In *P. serbica* FSU12416, the 1237 bp long candidate gene *trpM* is disrupted by three introns, as predicted by Augustus software. The 1077 bp reading frame thus encodes a 358 aa protein with a calculated pI=5.1 and a mass of 39.6 kDa. Notably, *trpM* is not encoded within the **1** biosynthesis gene cluster. Also, TrpM and PsiM are phylogenetically unrelated, as the former is a member of the methyltransferase family 33 whereas the latter belongs to family 10. The *trpM* cDNA was sequenced to confirm the predicted introns, while a synthetic codon-optimized gene was used to create expression plasmid pFB06. *N*-terminally tagged hexahistidine TrpM fusion protein was produced in *E. coli* KRX and purified by immobilized metal chelate affinity chromatography (Figure S1).

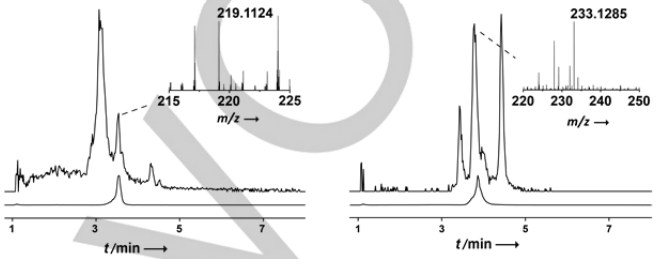


Figure 2. Mass spectrometric analysis of *P. serbica* mycelial extracts. Left panel: extracted ion chromatogram for *m*/*z* 219-220. At t_R=3.6 min, the mass of monomethylated **3**, e.g., L-abrine (**4,** calculated 219.1129 $[M+H]^+$, found 219.1124) was detected. Right panel: extracted ion chromatogram for *m*/*z* 233-234. Only the signal at t_R=3.9 min (*m*/*z* 233.1285 $[M+H]^+$) is consistent with the exact mass of dimethylated **3,** e.g., *N*,*N*-α-dimethyl-L-tryptophan (**5**, calculated *m*/*z* 233.1284 $[M+H]^+$). Ion chromatograms of **4** and **5** standards are shown below the sample traces.

Gel permeation chromatography confirmed that TrpM is a monomer under native conditions (Figure S2), which is consistent with previous findings by Seebeck and colleagues on EgtD.[8] Subsequently, TrpM was tested for activity *in vitro*. Hypothesizing that it catalyzes **3** methylation, we first used this substrate to test the activity of the recombinantly produced enzyme. Assays were run in TRIS-buffer (pH=8.0) for 15 min. For a time course of the reaction, samples were taken in 30 s intervals over 150 s, and subsequently analyzed by LC/MS (Figure 3, Figure S3). The analysis showed simultaneous formation of mono- and dimethylated **3** across this time course. One product was identical to an authentic standard of **4** regarding its retention time (t_R=16.6 min, Figure 3), UV/Vis spectrum, and molecular mass (*m/z* 219.1134 $[M+H]^+$; calculated 219.1129 $[M+H]^+$). The identity of the expected second product (**5**, t_R=15.6 min) was confirmed after purification by ^{1}H NMR spectroscopy (Figure S4). K_m values for **3** and **4** were 3.7 μM and 1.0 μM, respectively, as determined with a luminescence assay. The k_{cat} was 3×10^{-2} s^{-1} and 2.9×10^{-2} s^{-1}. Low K_m and k_{cat} values are well documented for other small molecule methyltransferases, e.g., for EgtD and $EgtD_{E282A,M252V}$,[6a-c] or the *C*-methyltransferase TylC3 for L-mycarose biosynthesis.[9] The k_{cat} values for the first and second TrpM-catalyzed methylation virtually do not diverge, which points to increased affinity of TrpM to the monomethylated intermediate, as indicated by the respective K_m values, which is also consistent with findings for EgtD.

Under standard conditions, **6** formation was not observed. In separate assays, *S*-adenosylhomocysteine (SAH) nucleosidase

and adenine deaminase were added to irreversibly remove SAH from the reaction. After prolonged incubation for 4 h, **6** was detectable as minor product as well (t_R=13.3 min), besides **4** and **5** (Figure 3). 4-Hydroxy-L-tryptophan and D-tryptophan did not serve as substrate (Figure 3).

Psilocybe mushrooms control the PsiM-catalyzed methyl transfer steps during **1** biosynthesis by requiring substrates that have a 4-phosphoryloxy group. Consequently, tryptamine is not accepted. This important property minimizes, or prevents altogether, that the reactive and instable **2** is made. Otherwise, **2** could be formed from *N*,*N*-dimethyltryptamine which may undergo subsequent 4-hydroxylation by the rather unspecific monooxygenase PsiH. Here, we tested whether tryptamine was a substrate for TrpM. Under the applied conditions, turnover was not observed by chromatographic analysis, and not even traces of *N*,*N*-dimethyltryptamine or of its *N*-monomethyl precursor were detected by mass spectrometry (Figure 3, traces g and h). 4-Hydroxytryptamine was not turned over either. An important feature of **1** biosynthesis is that it **2** formation is avoided. TrpM does not participate in **1** biosynthesis, but its substrate specificity prevents **2** formation as well, as the products of TrpM, **4**, **5**, and **6**, are not intermediates or shunt products of the **1** biosynthetic pathway. Both **1** and **5** biosynthesis originates from **3** (Scheme 1). However, cellular localization of the respective enzymes or mutually exclusive gene expression, e.g., dependent on the developmental stage, may prevent competition for **3** as principal building block.

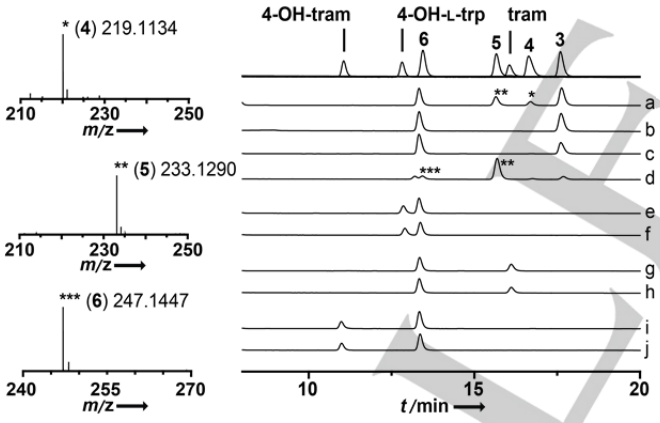


Figure 3. LC/MS analyses of TrpM *in vitro* activity tests. HR-ESIMS spectra were recorded in positive mode. Experimentally determined masses are indicated. Top trace: overlaid separate chromatograms of standard compounds. Trace a: TrpM reaction with **3**. Trace b: TrpM reaction with D-tryptophan. Trace c: control with heat-treated TrpM. Trace d: prolonged TrpM reaction with **3** in the presence of SAH nucleosidase and adenine deaminase. Traces e and f: TrpM reaction with 4-Hydroxy-L-tryptophan and negative control, respectively. Traces g and h: TrpM reaction with tryptamine and negative control. Traces i and j: TrpM reaction with 4-´hydroxytryptamine and negative control. The signal at t_R=12.9 min corresponds to *m*/*z* 298.0974 $[M+H]^+$, which is compatible with the mass of 5´-methylthioadenosine, a breakdown product of SAM. Tram: tryptamine, trp: tryptophan.

Given the activity of EgtD,[6] we further tested if L-histidine was converted by TrpM *in vitro* as well. ESI-MS showed mono- and dimethylation whereas trimethylation was not found. L-tyrosine and L-phenylalanine were also mono- and dimethylated *in vitro* (Figure S5). In mycelial extracts of *P. serbica*, mono- and dimethylated L-phenylalanine was identified by LC/MS (Figure S6) while not even traces of methyl derivatives of L-histidine and L-tyrosine were detected.

Seebeck et al. characterized the mycobacterial wild type EgtD (that is specific for L-histidine), the engineered variant $EgtD_{E282A,M252V}$ (specific for **3**), and Ybs, an *Aspergillus* methyltransferase, specific for L-tyrosine.[6] Our *in vitro* results indicate that TrpM is more flexible for acceptor substrates than the above methyltransferases. Yet, our data also support the notion that wild type EgtD is outstanding in that it efficiently trimethylates, while $EgtD_{E282A,M252V}$ and TrpM virtually stop at the dimethyl stage. Of note, the residues of EgtD known to interact with the (methylated) α-amino group of L-histidine, Gly161 and Asn166, are conserved in TrpM (Gly192 and Asn197). Asn 166 in EgtD is essential to make *N*,*N*-dimethyl-L-histidine a 100-fold stronger ligand than L-histidine.[6b] Crystallographic work would be warranted to elucidate the respective roles of active site residues for TrpM.

PsiM and TrpM share a functional relationship. Therefore, we hypothesized that the evolution of their respective pathways is correlated. To address this question, and to understand the evolutionary relationship between EgtD and **7** metabolism in fungi, we conducted a phylogenetic analysis of homologs of TrpM-coding genes in Agaricomycetes. A maximum likelihood tree places *trpM* in a clade with 12 orthologs from diverse agaricomycete species that is nested in a larger group of enzymes from a larger diversity of Agaricomycetes (Figure 4A, Figure S7). Homologs of *trpM* in this phylogeny are variable in structure; while all orthologs of *trpM* code for a single MT_33/EasF (EgtD) domain, the rest of the tree is composed mainly of EgtDB bifunctional proteins, with a C-terminally linked YfmG-superfamily sulfoxide synthase (EgtB) domain. Based on the topological congruence between the *trpM* clade and that of its mostly EgtDB sister clade (Figure 4A), which includes the *trpM* taxa, we infer that *trpM* originated by partial duplication of *egtDB* at an intermediate stage of agaricomycete diversification. The phylogeny also reveals at least ten recent (mostly species-specific) duplications that resulted in taxonomically diverse EgtD proteins. Since its origin by an ancient duplication, *trpM* has been periodically lost and is no longer present in most Agaricomycetes, and it is the only long-retained duplication in the phylogeny, suggesting EgtD-only paralogs have been instable during Agaricomycetes diversification. This instability is exemplified by the recent loss of EgtD in *P. cubensis* and pseudogenization in *P. cyanescens*, where it is truncated and contains an in-frame stop codon (Figure S8).

The functions of convergently evolved EgtD are only partly understood in fungi. The only other functionally characterized enzyme in the phylogeny (Figure 4A) is the *D. squalens* EgtD previously predicted to produce **6**,[6a] that is the product of a recent partial duplication of EgtDB. *Pisolithus tinctorius egtD* genes, which may also be involved in the production of **6**, are distantly related homologs, analyzed separately (Figure S9A), while genes for EgtDB from *P. tinctorius* and other Boletales are present, along with four recent *egtD* paralogs from *Scleroderma citrina*. The gene for EasF is another distant *egtD* homolog in the

Claviceps purpurea (ergot fungus, Ascomycota) ergotamine biosynthesis cluster, which is more similar to Ascomycota *egtDB* sequences than to *egtDB* sequences presented here (Figure S9B).

The phylogenetic distribution of gene families indicates ecological functions.[10] EgtD is important for mycorrhizal formation aided by **6** in *P. tinctorius*,[7a] which has four paralogs of a very distantly related *egtD*, all of which are co-orthologous with just one EgtD encoded in *Scleroderma citrinum* (Figure S9A). If the parallel recent expansion of *egtD* in *S. citrinum* (Figure 4A) represents convergent ectomycorrhiza-related metabolism, this is in contrast to the broader distribution of *egtD* homologs. The majority of Agaricomycetes with orthologs of *trpM* and/or recent *egtD* paralogs shown here are involved in lignocellulose decay, often of exposed substrates, which suggests their products are involved in competition with invertebrates or bacteria. It is also interesting to note that EgtD in the Saprolegniales (oomycetes) appears to have originated by partial horizontal transfer of an EgtDB from aquatic true fungi (Figure S9C).

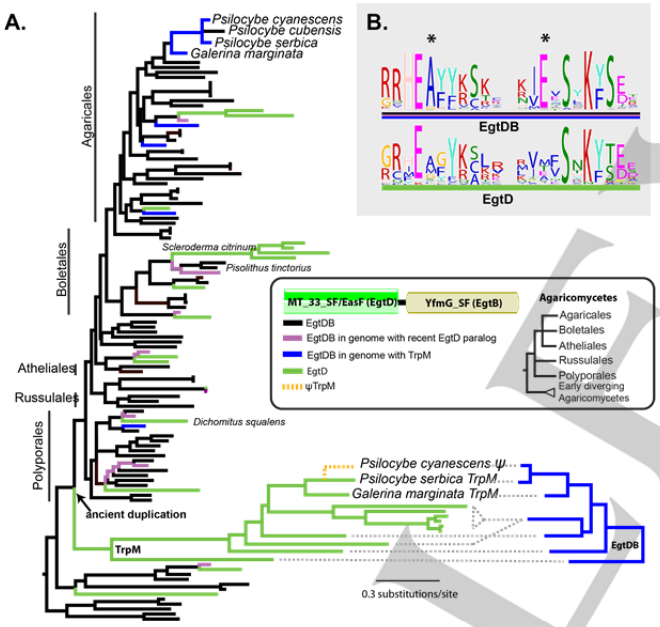


Figure 4. Evolutionary origin of TrpM and related enzymes. TrpM originated by partial duplication of a putative ergothioneine (**7**) bifunctional enzyme (EgtDB). A) Maximum Likelihood phylogeny of MT_33 Superfamily domains extracted from nearest PsiM homologs in Agaricomycetes (phylogeny with accession numbers in Figure S7A). EgtD enzymes (green branches) emerged once anciently (TrpM), and at least 10 times recently by partial duplication of bifunctional putative **7** synthesis enzyme EgtDB. There are few TrpM monofunctional proteins surviving since the ancient duplication in Agaricomycetes, and it has been lost in *P. cubensis* and pseudogenized in *P. cyanescens*. The *P. cyanescens* pseudogene (ψ) is shown for illustration, but was not included in phylogenetic analysis. A reversed phylogeny of EgtDB enzymes from genomes containing TrpM (see also Figure S7B) demonstrates topological congruence between the alternate paralogs. Species phylogeny in the key follows Hibbett et al.[11]. B) Motifs associated with specificity-determining residues (*) 252 (left) and 282 (right) of Agaricomycetes EgtD domains/proteins. Position number is relative to *M. smegmatis* EgtD (A0R5M8.1).

The evolutionary instability of EgtD is consistent with a transition from a broadly selected function in **7** biosynthesis[12] to one in secondary metabolism, which tends to be conditionally retained through diversifying ecological selection. In keeping with functional diversification of EgtD in Agaricomycetes, the substrate specificity-determining amino acid position 282 (relative to *Mycobacterium smegmatis*) is highly conserved as glutamic acid for histidine binding in 97 EgtDB sequences, but is variable as mainly Met, Thr, Glu, Ala, and Ser in 25 EgtD sequences (Figure 4B). Similarly, position 252 is highly conserved as alanine in EgtDB, but is variable as mainly Ala, Met, Gly, and Ser in EgtD.

It is also plausible that fitness costs to the organism account in part for the instability of recently evolved EgtD, by detrimentally interfering with **7** synthesis due to the irreversibility of histidine *N*-methylation; this hypothesis should ultimately be tested in a genetically tractable model fungus. Gene duplication has long been identified as a source of genetic novelty,[13] and has strongly influenced the metabolic diversity of fungi through the sub- and neofunctionalization of paralogs.[14] Further, gene fissions are important sources of novel functions in fungi and oomycetes.[15] We have characterized the function of TrpM, a methyltransferase generated by duplication of the first domain of a presumed **7** synthesis bifunctional enzyme, EgtDB. We further demonstrated several convergent origins of EgtD in Agaricomycetes by the same mechanism. Filamentous Ascomycota maintain their EgtD enzymes encoded in diverse secondary metabolism gene clusters, which can be specialized and retained in large pan-genomes, and acquired by horizontal gene transfer (HGT).[16] However, our data indicate that Basidiomycota benefit more from *de novo* neo-functionalization of vertically-inherited enzymes because of lower rates of clustering and HGT.

Our results highlight a secondary metabolism evolvability mechanism, in which weakly or sporadically selected functions remain accessible through their preservation in a widely distributed, structurally conserved bifunctional protein. Selection on within-protein interactions and broadly conserved functions possibly maintain these domains as a source of secondary metabolic diversity, where they otherwise are lost in transiently adaptive roles. In short, genes for these highly conserved bifunctional proteins may function as genetic templates from which less conserved functions can repeatedly evolve. Traditionally, the fungal genus *Psilocybe* impacted natural product chemistry due to its capacity to biosynthesize **1**. Our results demonstrate that the relevance of this genus reaches further and help address more fundamental aspects of the evolution of fungal small molecule-processing enzymes.

Experimental Section

General and microbiological procedures: *P. serbica* FSU12416 was grown at 25°C in the dark on malt extract peptone (MEP) agar plates or in MEP liquid medium shaken at 140 rpm for 20 d at room temperature. Plasmid isolation, DNA restriction and ligation followed the instructions of the manufacturers of kits and enzymes (NEB, Promega, Thermo, Zymo).

Chemicals, media ingredients and solvents were purchased from Roth, Sigma-Aldrich, VWR, and Deutero.

Cloning of *trpM* cDNA and construction of expression plasmid: mRNA was isolated from *P. serbica* mycelium using Promega's SV Total RNA isolation kit. First strand cDNA synthesis was primed with oligo-d(T)-primers and RevertAid reverse transcriptase. The first strand reaction served as template in a subsequent PCR with primers (0.2 µM each) oFB51 (5´-ATGCCGCGAATCCAGGTT-3´) and oFB52 (5´-TTAGCTTCGTCCGGTTACTTCG-3´). The reaction contained 0.2 mM (each) deoxynucleoside triphosphate, 2 mM $MgSO_4$, in HF buffer that was supplied with the enzyme (Phusion DNA polymerase, 1 U). Amplified DNA was purified by agarose gel electrophoresis and subsequent extraction from the gel. The PCR amplicon was ligated to pJET1.2 and several clones were sequenced to confirm predicted exon/intron junctions. For heterologous production of TrpM, *E. coli* KRX (Promega) and a codon-optimized gene (Invitrogen) were used that was inserted between the BamHI and XhoI sites of pET28a, to create expression plasmid pFB06. The codon-optimized sequence is provided in Figure S10, the genomic DNA sequence of *trpM* is deposited at GenBank under accession number MH423322.

Protein purification: Cultivation of *E. coli* and gene expression were performed as described.[17] TrpM was purified by metal chelate affinity chromatography on Protino Ni^{2+}-NTA agarose (Macherey-Nagel). Purification was verified by polyacrylamide gels (12% Laemmli gel), and protein concentrations were determined by Bradford's method.[18] TrpM was desalted on a PD-10 column (GE Healthcare), equilibrated with buffer (50 mM TRIS-HCl, pH 8.0) for enzyme assays. Production of *E. coli S*-adenosylhomocysteine nucleosidase and adenine deaminase[6] using pET28-based expression plasmids followed the same protocol, except that the *E. coli* culture to express the adenine deaminase gene was supplemented with 50 µM 2,2'-bipyridyl and 1 mM $MnCl_2$. For size-exclusion chromatography, TrpM was eluted in phosphate buffer (10 mM sodium dihydrogen phosphate, 140 mM NaCl, pH 7.4). Size-exclusion chromatography was performed by FPLC (Äkta Pure 25, GE Healthcare) and a Superdex 200 increase 10/300 GL column with 24 mL bed volume. Signals were referenced to the GE Healthcare high molecular weight standard.

Product formation by TrpM: *In vitro* TrpM assays were carried out in triplicates in a volume of 200 µL, buffered in 50 mM TRIS-HCl, pH=8.0, the enzyme concentration was 200 nM. The methyl acceptor (**3** or others) was added at 500 µM final concentration, SAM at 1.5 mM. The reactions were incubated at 25°C for 15 min. The time course with **3** as acceptor substrate included additional sampling in 30 s intervals from 0-150 s. The reaction with SAH-nucleosidase and adenine deaminase was incubated for 4 h. Reactions were stopped in liquid nitrogen and then lyophilized, dissolved in MeOH and centrifuged. The supernatant was collected for chemical analysis. TrpM kinetics were recorded using the MTase-Glo luminescence assay (Promega).

Chemical analysis of TrpM assays: LC/ESIMS-analyses of TrpM *in vitro* assays were carried out on an Agilent Infinity 1260 instrument with a Zorbax Eclipse XDB-C18 column (150 × 4.6 mm, 5 µm particle size, flow: 1 mL min^{-1}), coupled to an Agilent 6130 Single Quadrupole mass detector. Diode array detection was between λ=200–400 nm. Chromatograms were extracted at λ=280 nm. To analyze indolic compounds, the following linear gradient was applied with 0.1% trifluoroacetic acid (TFA) in H_2O (solvent A) and methanol (MeOH, solvent B): initially 5% B, increased to 15% within 5 min, to 37% B within further 13 min, and to 100% B within 1 min. LC/HRESIMS was performed on a Thermo Accela liquid chromatograph coupled to an Exactive Orbitrap spectrometer operated in positive and negative mode and equipped with a C18 column. Solvent A was 0.1% (v/v) formic acid in H_2O, solvent B was acetonitrile (ACN).To analyze mycelial extracts for methylated L-phenylalanine, a Grom-Sil 100 ODS-0 AB, 250 × 4.6 mm, 3 µm particle size and the following conditions were used: initial hold at 5% B for 1 min, linear gradient to 100% B within 15 min at a flow rate of 1 mL min^{-1}. For other analyses, a Betasil C18 column (150 × 2.1 mm, 3 µm particle size) and these conditions were used: initial hold at 5% B for 1 min, a linear gradient of 5-98% B within 15 min at a flow of 0.2 mL min^{-1}.

Chemical analysis of *P. serbica* extracts: To prepare mycelium for LC/HRESIMS-analyses, it was collected from liquid cultures by filtration, washed with distilled water, blotted dry, snap-frozen in liquid nitrogen, homogenized with mortar and pestle, and shaken in 200 mL MeOH for 18 h at 240 rpm, before the biomass was removed by filtration and dried over sodium sulfate. Residual solvent was evaporated under reduced pressure. The dry residue was dissolved in MeOH, filtered, and subjected to LC/HRESIMS measurements (above).

Purification and NMR spectroscopy: From *in vitro* assays, **5** was purified in two steps on the above instrument. The first step was accomplished with an Eclipse XDB-C18 column (250 × 9.4 mm, 5 µm particle size). H_2O + 0.1% TFA was solvent A and MeOH solvent B. A linear gradient (2 mL min^{-1}) was applied, beginning with 5% B, increased to 15% in 5 min, to 37% B within 18 min, and to 100% B within 1 min. The second purification step included a Hypercarb column (150 × 4.6 mm, 5 µm), 0.1% TFA in H_2O as solvent A and ACN as solvent B. A linear gradient (1 mL min^{-1}) was applied, beginning with 21% B and increased to 100% B in 9 min. The ^{1}H NMR spectrum of **5** was recorded at 300 K on a Bruker Avance III spectrometer at 500 MHz in D_2O. Chemical shifts were referenced to residual non-deuterated solvent (δ_H 4.79 ppm).

Bioinformatic analysis: Initially**,** Blast[19] and Augustus (v.3.3.1)[20] software was used to browse the *P. serbica* genome.[4a] Homologs of *trpM* were obtained from a local database of 540 fungal and oomycete proteomes using usearch (v. 8.0.1517)[21] with TrpM (KDR66884.1) and EgtDB (KDR82860.1) from *G. marginata* as queries, an e-value of $1e^{-3}$, and protein identity of 30%. Domain structure of proteins was determined using rpsblast (NCBI BLAST v. 2.6.0+)[19] to search the NCBI Conserved Domain Database (v. 3.16)[22] with an e-value of 0.001. Absence of YfmG superfamily domains in proteins was confirmed by tblastn against the respective genome assemblies using the nearest EgtDB sequence from phylogenies (see below) as a query. MT_33 superfamily domains were extracted from protein sequences using hmmer (v. 3.1b2).[23] MT_33 phylogenies were constructed by first aligning with mafft (v. 7.221),[24] removing ambiguously aligned characters with Trimal (v. 1.4)[25] using the -automated1 method. Exploratory phylogenetic analyses were performed using fasttree (v. 2.1.7).[26] A well-supported clade consisting exclusively of Agaricomycetes was selected for subsequent analyses. For analyses of distant relatives of TrpM, hmmer was used to search for MT_33 domains (952 total) across the proteome database. Figure S9A (*P. tinctorius* EgtD) combined the same four additional Agaricomycetes rooting sequences along with near and distant clades of Boletales EgtDB/EgtD. Figure S9B (EasF) combined two clades of Pezizomycotina MT_33 homologs with two orthologs each encoding Agaricomycetes TrpM and EgtDB for rooting. Figure S9C (oomycetes EgtD) combines a clade of early diverging fungi that includes oomycete EgtD sequences and the Agaricomycetes root. Datasets were re-aligned and curated as above, followed by maximum likelihood analysis using RAxML (v. 8.2.9)[27] with the model (LG) selected automatically according to Bayesian information criterion. Protein domains identified by rpsblast were mapped to the Agaricomycetes PsiM2/EgtDB phylogeny using the ETE Toolkit (v. 3.1.1).[28]

Acknowledgments

We thank A. Perner H. Heinecke, F. Baldeweg, and H. Schoeler (Hans Knöll Institute Jena) for recording high-resolution mass

spectra, for recording and interpreting NMR spectra, and for luminescence measurements, respectively. M. Gressler (Friedrich Schiller University Jena) is acknowledged for helpful comments on the manuscript. We also thank F.J. Ruiz-Dueñas, Ph.D. (Consejo Superior de Investigaciones Científicas, Madrid, Spain), J.M. Barrasa, Ph.D. (Universidad de Alcalá, Alcalá de Henares, Spain), D.S. Hibbett, Ph.D. (Clark University, Worcester, MA, USA), D. Lindner, Ph.D. (US Dept. of Agriculture, Forest Service, Madison, WI, USA), and L. Nagy, Ph.D. (Hungarian Academy of Sciences, Szeged, Hungary), for access to unpublished genome sequences. These data were produced by the US Department of Energy Joint Genome Institute in collaboration with the user community. Evolutionary analyses were performed using the resources of the Ohio Supercomputer Center. This work was supported by the National Science Foundation grant (grant DEB-1638999 to JCS) and by the Deutsche Forschungsgemeinschaft (DFG, grant HO2515/7-1 to DH). Work in DH's laboratory is also funded by the Collaborative Research Center ChemBioSys (SFB1127).

Keywords: biosynthesis • enzymes • methyltransferase • psilocybin • tryptophan

[1] a) A. Hofmann, R. Heim, A. Brack, H. Kobel, A. Frey, H. Ott, T. Petrzilka, F. Troxler, *Helv. Chim. Acta* **1959**, *42*, 1557-1572; b) A.Y. Leung, A.H. Smith, A.G. Paul, *J. Pharm. Sci.* **1965**, *54*, 1576-1579.

[2] a) D.E. Nichols, *Pharmacol. Ther.* **2004**, *101*, 131-181; b) F. Hasler, U. Grimberg, M.A. Benz, T. Huber, F.X. Vollenweider, *Psychopharmacol.* **2004**, *172*, 145-156.

[3] a) C.S. Grob, A.L. Danforth, G.S. Chopra, M. Hagerty, C.R. McKay, A.L. Halberstadt, G.R. Greer, *Arch. Gen. Psychiat.* **2011**, *68*, 71-78; b) M.W. Johnson, A. Garcia-Romeu, M.P. Cosimano, R.R. Griffiths, *J. Psychopharmacol.* **2014**, *28*, 983-992; c) A.M. Sherwood, T.E. Prisinzano, *Exp. Rev. Clin. Pharmacol.* **2018**, *11,* 1-3.

[4] a) J. Fricke, F. Blei, D. Hoffmeister, *Angew. Chem. Intl. Ed.* **2017**, *56*, 12352-12355; b) H.T. Reynolds, V. Vijayakumar, E. Gluck-Thaler, H.B. Korotkin, P.B. Matheny, J.C. Slot, *Evol. Lett.* (doi: 10.1002/evl3.42); c) F. Blei, F. Baldeweg, J. Fricke, D. Hoffmeister, *Chem. Eur. J.* **2018**, *24*, 10028-10031; d) J. Fricke, F. Blei, D. Hoffmeister, *Angew. Chem.* **2017**, *129*, 12524-12527.

[5] S. Ghosal, S.K. Dutta, *Phytochemistry* **1971**, *10*, 195-198.

[6] a) A. Vit, L. Misson, W. Blankenfeldt, F. P. Seebeck, *ChemBioChem* **2015**, *16*, 119-125; b) L. Misson, R. Burn, A. Vit, J. Hildesheim, M.A. Beliaeva W. Blankenfeldt, F.P. Seebeck, *ACS Chem. Biol.* **2018**, *13*, 1333-1342; c) F.P. Seebeck, *J. Am. Chem. Soc.* **2010**, *132*, 6632–6633.

[7] a) T. Kawano, *Plant Cell Rep*. **2003**, *21*, 829-837; b) A. Jambois, A. Dauphin, T. Kawano, F.A. Ditengou, F. Bouteau, V. Legue, F. Lapeyrie, *Physiol. Plantarum* **2005**, *123*, 120-129.

[8] A. Vit, L. Misson, W. Blankenfeldt, F.P. Seebeck, *Acta Crystallogr. F* **2014**, *70*, 676-680.

[9] H. Chen, Z. Zhao, T.M. Hallis, Z. Guo, H.-w. Liu, *J. Am. Chem. Soc.* **2001**, *40*, 607-610.

[10] D. Floudas, M. Binder, R. Riley, K. Barry, R.A. Blanchette, B. Henrissat, et al. *Science*, **2012**, *336*, 1715-1719.

[11] D.S. Hibbett, M. Binder, J.F. Bischoff, M. Blackwell, P.F. Cannon, O.E. Eriksson, et al., *Mycol. Res*. **2007**, *111*, 509-547.

[12] G.W. Jones, S. Doyle, D.A. Fitzpatrick, *Gene*, **2014**, *549*, 161-170.

[13] S. Ohno, *Evolution by gene duplication,* 1st ed., Springer-Verlag, Heidelberg, **1970**.

[14] a) I. Wapinski, A. Pfeffer, N. Friedman, A. Regev, *Nature*, **2007**, *449*, 54-61; b) J.H. Wisecaver, J.C. Slot, A. Rokas, *PLoS Genetics*, **2014**, *10*, e1004816.

[15] G. Leonard, T.A. Richards, *Proc. Natl. Acad. Sci*., **2012**, *109*, 21402-21407.

[16] M. Marcet-Houben, T. Gabaldón, *Fungal Genet. Biol*., **2016**, *86*, 71-80.

[17] D. Kalb, T. Heinekamp, S. Schieferdecker, M. Nett, A. A. Brakhage, D. Hoffmeister, *ChemBioChem* **2016**, *17*, 1813-1817.

[18] M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248-254.

[19] S. F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, *Nucleic Acids Res*. **1997**, *25*, 3389-3402.

[20] M. Stanke, R. Steinkamp, S. Waack, B. Morgenstern, *Nucleic Acids Res.* **2004**, *32*, W309-312.

[21] R.C. Edgar, *Bioinformatics*, **2010**, *26*, 2460-2461.

[22] A. Marchler-Bauer, M.K. Derbyshire, M.R. Gonzales, S. Lu, F. Chitsaz, L.Y. Geer, et al., *Nucleic Acids Res*., **2015**, *43*, D222-226.

[23] S.R. Eddy, *Bioinformatics*, **1998**, *14*, 755-763.

[24] K. Katoh, D.M. Standley, *Mol. Biol. Evol*. **2013**, *30*, 772-780.

[25] S. Capella-Gutierrez, J.M. Silla-Martinez, T. Gabaldón, *Bioinformatics*, *25*, **2009**, 1972-1973.

[26] M.N. Price, P.S. Dehal, A.P. Arkin, *PLoS One*, **2010**, *5*, e9490.

[27] A. Stamatakis, *Bioinformatics*, **2014**, *30*, 1312-1313.

[28] J.F. Huerta-Cepas, Serra, P. Bork, *Mol. Biol. Evol*., **2016**, *33*, 1635-1638.

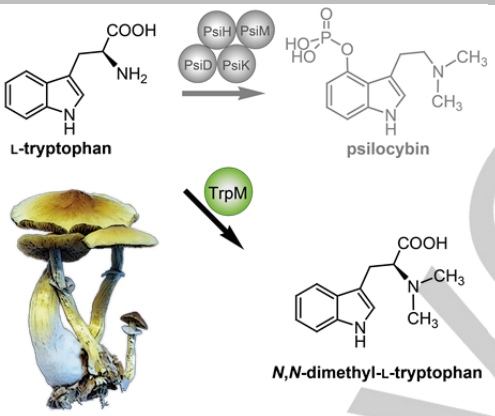
10.1002/cbic.201800336

WILEY-VCH

Entry for the Table of Contents

COMMUNICATION

Magic evolution - The methyltransferase TrpM of the "magic" mushroom and psilocybin producer *Psilocybe serbica* was characterized. TrpM catalyzes iterative L-tryptophan *N*-methylation, its specificity profile prevents interference with the psilocybin pathway. The *trpM* gene evolved from an ergothioneine biosynthetic gene. Its phylogenetic history provides insight how mushrooms evolve natural product biosyntheses.



Felix Blei, Janis Fricke, Jonas Wick, Jason C. Slot, and Dirk Hoffmeister**

Page No. – Page No.

Iterative L-tryptophan methylation in *Psilocybe* evolved by sub-domain duplication