Structural diversity of the coenzyme methylofuran and identification of enzymes for the biosynthesis of its polyglutamate side chain

Methylofuran (MYFR) is a formyl-carrying coenzyme essential for the oxidation of formaldehyde in most methylotrophic bacteria. In Methylotherubium extorquens, MYFR contains a large and branched polyglutamate side chain of up to 24 glutamates. These glutamates play an essential role in interfacing the coenzyme with the formyltransferase/hydrolase complex, an enzyme that generates formate. To date, MYFR has not been identified in other methylotrophs, and it is unknown whether its structural features are conserved. Here, we examined nine bacterial strains for the presence and structure of MYFR using high-resolution liquid chromatography–mass spectrometry (LC-MS). Two of the strains produced MYFR as present in M. extorquens, while a modified MYFR containing tyramine instead of tyrosine in its core structure was detected in six strains. When M. extorquens was grown in the presence of tyramine, the compound was readily incorporated into MYFR, indicating that the biosynthetic enzymes are unable to discriminate tyrosine from tyramine. Using gene deletions in combination with LC-MS analyses, we identified three genes, orf5, orfY, and orf17 that are essential for MYFR biosynthesis. Notably, the orfY and orf5 mutants accumulated short MYFR intermediates with only one and two glutamates, respectively, suggesting that these enzymes catalyze glutamate addition. Upon homologous overexpression of orf5, a drastic increase in the number of glutamates in MYFR was observed (up to 40 glutamates), further corroborating the function of Orf5 as a glutamate ligase. We thus renamed OrfY and Orf5 to MyfA and MyfB to highlight that these enzymes are specifically involved in MYFR biosynthesis.

The majority of methylotrophic bacteria use a tetrahydrofolate (THF)-dependent pathway for the oxidation and conversion of the one-carbon unit originating from methanol or methane (1–4). In addition to the one-carbon carrier THF, this pathway requires the presence of a second coenzyme (5), which has recently been structurally elucidated and termed methylofuran (MYFR; Fig. 1A) (6). MYFR acts as a formyl carrier and is tightly bound by formyltransferase/hydrolase complex (Fhc) that generates formate from formyl-THF (7–9). THF/MYFR-dependent formaldehyde oxidation is closely related to methanogenesis—a type of energy and carbon metabolism found in archaea (10)—and many of the involved enzymes are conserved (1, 5, 11). MYFR is a structural and functional analog of the archaeal coenzyme methanofuran (MFR) (12), with which it shares a similar core structure. In the model methylotroph Methylotherubium extorquens, the core structure of MYFR differs from MFR only in the presence of a tyrosine moiety in place of the tyramine residue (6). However, the most important distinction between MYFR and MFR is the large polyglutamate side chain. In the case of MYFR from M. extorquens, this polyglutamate chain contains up to 24 glutamate residues, which are both α- and γ-linked and form a branched chain (6, 9). These features are unique to MYFR, as the five known types of archaeal MFR (MFR-a–MFR-e) contain exclusively γ-linked glutamates arranged in a linear chain of a maximum of 12 units (12–15). The different types of MFR are distinguished by the composition of the side chain, i.e., the number of glutamate residues and the presence of linkers (as in MFR-e) or terminal residues (as in MFR-a/c). At present, MYFR has only been identified in M. extorquens, and it is currently unknown if a similar structural diversity exists within bacteria.

The unique structure of the polyglutamate side chain of MYFR also raises the question of its biosynthesis. In archaea, six enzymes (MfnA–F) have been identified that catalyze the biosynthetic steps required to form the MFR core structure linked to the first γ-glutamate (16–19). However, the enzymes responsible for the ligation of additional glutamates (all MFRs contain at least two γ-glutamates) as well as the enzymes that introduce the linker or terminal moieties into the side chain of MFR remain to be identified. In M. extorquens and other proteobacteria, the proteins Orf22, Orf21, and Orf9 are homologs of the archaeal enzymes MfnB, MfnE, and MfnF respectively. The corresponding bacterial genes are located in close proximity in the genome and are part of the so-called “archaeal-like” gene cluster (5, 20). This cluster contains genes involved in the THF/MYFR-linked pathway including fhcABCD and genes involved in H4MPT biosynthesis. For the transaminase MfnC—responsible for the formation of the amine functionality in the furan moiety (17)—five different homologs are present.
Diversity and biosynthesis of MYFR

in M. extorquens, and it is unclear which of them fulfills the role of MfnC. As expected and noted before (6), no homolog of MfnA—the L-tyrosine decarboxylase that forms tyramine (16)—is present in M. extorquens. Potentially related to that finding is the fact that no homolog of MfnD—the tyramine::glutamate ligase involved in the formation of γ-glutamyltyramine (18)—was found either. Analogous to MFR, the enzymes involved in the biosynthesis of the polyglutamate side chain of MYFR have not yet been identified. However, several of the genes in the “archaeal-like” cluster in M. extorquens have an unknown function and could therefore play a role in MYFR biosynthesis. Notably, all genes in the cluster are essential for growth on methanol (5, 21). However, prior genetic analysis was complicated by the emergence of secondary mutations and not all mutants could be complemented (21). Further, the identity of MYFR has only been described recently (6), which makes it now possible to directly measure MYFR in mutants of biosynthesis pathways.

Here, we first assessed the structural diversity of MYFR by analyzing different proteobacterial strains by LC-MS to determine whether the structure of MYFR—and consequently its biosynthesis—is likely conserved. We then analyzed mutants of genes from the “archaeal-like” cluster for their ability to biosynthesize MYFR and identified three genes essential for MYFR biosynthesis. Our data suggest that two of these genes, orfY and orf5, are involved in the elongation of the polyglutamate side chain of MYFR. We thus renamed these genes to myfA and myfB, respectively, and will use these names throughout the text.

Results

Two types of MYFR are present in proteobacteria

To determine the structural diversity of MYFR, we selected several methylotrophic strains (Table S1, Fig. S1). The nine strains covered a diversity of classes within the Alpha-, Beta-, and Gammaproteobacteria (Fig. 1B) and included a methylotrophic strain from the Arabidopsis thaliana leaf microbiome (At-LSPHERE) (22). After cultivation of the strains, MYFR was enriched from cell extracts using anion-exchange solid-phase extraction, assuming that MYFR in these strains would be negatively charged as in M. extorquens (6). Using LC-MS, we were able to detect MYFR, as it is present in M. extorquens, only in the two strains Methylobacterium radiotolerans and Methylobacterium album. For the remaining strains, we then searched the LC-MS data for various derivatives of MYFR. Surprisingly, a series of features corresponded to a decarboxylated form of MYFR (i.e., with masses corresponding to MYFR–CO₂) in the six strains Hyphomicrobiurn zavarzinii, Starkeya novella, Variorovax paradoxus, Burkholderia sp. Leaf177, Methylphilus methylotrophus, and Methylobacillus flagellatus. Most likely, such a modification would be localized in the core structure of MYFR. The only available functionality for decarboxylation of the core is the carboxylic acid group that is part of the tyrosine residue (Fig. 1A). Indeed, MS/MS fragmentation showed the presence of fragments consistent with a modified core structure containing tyramine instead of tyrosine (Table S2). The core structure of this novel MYFR type is thus identical to all known core structures of MFRs from archaea, which exclusively contain tyramine. To distinguish between the MYFR derivatives, the notation MYFRTyrosine and MYFRTyramine will be used (Fig. 1A). Notably, all strains tested that belong to the Betaproteobacteria produced MYFRTyramine, while strains with either type of MYFR were found in the Alphaproteobacteria (Fig. 1B).

In Methylosinus trichosporium, none of the two MYFR derivatives could be detected. To try to identify MYFR with an unknown core structure, we used an untargeted approach to identify features in the LC-MS data that are separated by the mass of one or multiple glutamates from each other. This approach assumed that any unknown MYFR derivatives would be present as a distribution of species with different number of glutamates. Using this method, we identified the polyglutamate-containing compound p-aminobenzoyl-Gluₙ with 3 to 5 glutamates in extracts of the bacterium, demonstrating that the approach was working in principle. However, we did not identify a novel MYFR derivative and the coenzyme might therefore be present in concentrations below our detection limit.

For the strains in which MYFR was identified, the number of glutamates attached to the core structure was evaluated (Fig. 2). While most strains showed similar glutamate distributions to the one from M. extorquens (most abundant MYFR contained 16–20 glutamates), MYFR from V. paradoxus and Burkholderia sp. Leaf177 had less than 14 glutamates (most abundant MYFR contained 12 and 10 glutamates respectively). Strikingly, the distribution of MYFR from Burkholderia sp. Leaf177 (and to some extent also the one from V. paradoxus)
showed a distinct gap between MYFR-Glu\textsubscript{10} and MYFR-Glu\textsubscript{12}, i.e., MYFR-Glu\textsubscript{11} had a much lower abundance.

**Generation of tyramine and incorporation into MYFR**

The discovery of MYFR\textsubscript{Tyramine} prompted us to investigate the genetic and biosynthetic basis for the incorporation of tyrosine or tyramine into MYFR. In the biosynthesis of MFR, tyramine is generated from tyrosine by MfnA, a pyridoxal phosphate-dependent L-tyrosine decarboxylase first identified in the methanogenic archaeon *Methanocaldococcus jannaschii* (16). To determine whether a similar enzyme is present in bacteria, we used BLAST (23) to search for homologs of MfnA in all strains analyzed above (Table S3). As expected, no homolog of MfnA could be identified for the three strains producing MYFR\textsubscript{Tyrosine}. However, for three of the six strains producing MYFR\textsubscript{Tyramine}, potential *mfnA* genes were identified: Mfla\textsubscript{2033} in *M. flagellatus* (adjacent to the methanol dehydrogenase cluster *mxa*), Snov\textsubscript{0063} in *S. novella*, and VAPA\_RS28950 in *V. paradoxus*. All three genes belong to the Pfam protein family PF00282, which contains PLP-dependent decarboxylases. It thus seems likely that these enzymes generate tyramine, which can then be integrated into MYFR. For the remaining three MYFR\textsubscript{Tyramine}-producing strains (*M. methylotrophus*, *H. zavarzinii*, and *Burkholderia* sp. Leaf177), no MfnA homolog was found, suggesting that other tyrosine decarboxylases with weak or no homology to MfnA exist. Alternatively, tyramine might also be produced from a different precursor than tyrosine. An MfnA homolog was also found for *M. trichosporium*, a strain for which we have not been able to identify MYFR so far, thus predicting the presence of MYFR\textsubscript{Tyramine} in this strain.

In each strain producing MYFR, exclusively one type of MYFR was detected and mixtures of MYFR\textsubscript{Tyrosine} and MYFR\textsubscript{Tyramine} were never observed. While tyramine is a specific metabolite that does not seem to be produced in every strain, tyrosine is an indispensable amino acid that must always be present. The incorporation of tyramine into MYFR\textsubscript{Tyramine} (and also MFR) must therefore be specific and no promiscuity of the responsible enzyme be tolerated. The specificity is consistent with MFR biosynthesis, for which the tyramine-L-glutamate ligase MfnD accepts only tyramine as substrate *in vitro* (18). For strains that do not synthesize tyramine—and consequently are only able to produce MYFR\textsubscript{Tyrosine}—the biosynthetic enzyme might be more relaxed. To test whether the MYFR\textsubscript{Tyrosine}-producing strains would principally be able to incorporate tyramine, we cultivated *M. extorquens* in minimal medium supplemented with 5 mM tyramine and analyzed the cell extract for the presence of the two types of MYFR by LC-MS. Notably, *M. extorquens* was indeed able to take up tyramine and incorporate it into MYFR, as 64% of the total MYFR pool was made up by MYFR\textsubscript{Tyramine} (Fig. 3). This result demonstrates that the biosynthetic enzyme responsible for tyrosine incorporation is unable to discriminate between tyrosine and tyramine in *M. extorquens*. Likely, this is also the case for the other strains producing MYFR\textsubscript{Tyramine}.
**Diversity and biosynthesis of MYFR**

*The genes mYfA, mYfB, and orf17 are involved in MYFR biosynthesis*

Our analysis of MYFR from different methylotrophic strains revealed that its structure is largely conserved, since MYFR in all cases consisted of a core structure (containing either tyrosine or tyramine) and a polyglutamate side chain of at least ten glutamates. The biosynthesis of MYFR is thus likely also conserved, although some of the enzymes must have different specificities to distinguish between tyrosine and tyramine. To identify additional proteins involved in the biosynthesis of MYFR, we focused on the genes with unknown function in the “archaeal-like” gene cluster (5, 21). We analyzed strains with deletions in the three genes mYfB (orf5; in *M. extorquens* AM1), orf17 (in *M. extorquens* PA1; harboring additionally an mxaF deletion, see Experimental procedures), and mYfA (orfY; in *M. flagellatus* KT) to determine whether MYFR biosynthesis is affected by the mutation. As the ΔmYfA mutant could not be obtained in *M. extorquens* and appears to be lethal, a previously described mutant in *M. flagellatus* was used (20), assuming that the biosynthesis will be similar in both strains. After cultivation of the strains and fractionation of the cell extracts by anion-exchange chromatography, we used LC-MS to detect MYFR. While a WT control contained MYFR as expected (with 6 up to 24 glutamates), MYFR could not be detected in all three mutants. The absence of MYFR in the deletion strains indicates the involvement of mYfA, mYfB, and orf17 in MYFR biosynthesis.

Both mYfA and mYfB are—based on their sequence—predicted to be members of the ATP-grasp superfamily (24, 25), which contains among others ribosomal protein S6 modification enzyme (RimK) (26, 27), d-alanine:d-alanine ligase (28), and glutathione synthetase (28). All members of this family share the unusual ATP-grasp fold for nucleotide binding and one or multiple glutamates with tyrosine or tyramine. Interconversion of the glutamates of MYFR is catalyzed by MfnB/Orf22 (19), the function of Orf17 might be unrelated to these activities. *M. jannaschii* possesses with MJ_0703 a homolog of Orf17 (26% sequence identity), which has been suggested to be involved in H₄SPT biosynthesis (32). However, we were able to detect methenyl-dH₄MPT in the Δorf17 *M. extorquens* mutant (Fig. S2), thus invalidating that hypothesis at least in bacteria.

To verify that Orf17 is indeed involved in MYFR biosynthesis, we complemented the mutant by expressing orf17 from a plasmid. This strain produced MYFR with a WT-like glutamate distribution (Fig. S3), thus confirming that the deletion of orf17 is responsible for the lack of MYFR in the mutant.

**Accumulation of MYFR biosynthesis intermediates in ΔmYfA and ΔmYfB**

Even though the deletion strains described above no longer produced MYFR, we expected that intermediates of the interrupted MYFR biosynthesis pathway would continue to be present or even accumulate. We thus analyzed the cell extracts (or purified fractions thereof) by LC-MS for the presence of MYFR intermediates. Surprisingly, in the extract of the ΔmYfA (=ΔorfY) *M. flagellatus* strain, the two intermediates MYFR(Tyrosine-Glu2) (MYFR(Tyrosine) without glutamates, i.e., just the core) and MYFR(Tyramine-Glu1) (MYFR(Tyramine) with one glutamate unit) were accumulating (Fig. 4A). The identity of these two compounds was confirmed by MS/MS fragmentation. In a WT cell extract, MYFR(Tyramine-Glu0) (MYFR(Tyramine) without glutamates, i.e., just the core) and MYFR(Tyramine-Glu1) (MYFR(Tyramine) with one glutamate unit) were accumulating (Fig. 4A). The identity of these two compounds was confirmed by MS/MS fragmentation. In a WT cell extract, MYFR(Tyramine-Glu0) was absent while the peak area of MYFR(Tyramine-Glu1) was 39-fold lower than that in the deletion strain. The accumulation of these two intermediates suggests that mYfA plays a role in the elongation of the polyglutamate side chain, most likely by adding glutamates to the observed intermediates.

To determine biosynthetic intermediates in the ΔmYfB (=Δorf5) *M. extorquens* mutant, anion-exchange purified fractions of the cell extract were analyzed by LC-MS. In a fraction eluting earlier than the full-length MYFR(Tyrosine) peak with a mass corresponding to MYFR(Tyrosine)-Glu₂ was observed (Fig. 4B). MS/MS fragmentation further verified the identity of this compound (Fig. S4). In addition, the same fraction contained Glu2-Tyr and Glu1-Tyr (Fig. 4B). All three compounds were not observed in a strain that was producing full-length MYFR(Tyrosine), thus suggesting that these are biosynthesis intermediates that are only present when mYfB is deleted. The accumulation of the intermediate MYFR(Tyrosine)-Glu₂ indicates that mYfB acts as a glutamate ligase. mYfB is thus likely involved in the addition of glutamates to MYFR.

A putative function of Orf17 in the biosynthesis of MYFR could not be determined based on its sequence. Orf17 belongs to the “histidine biosynthesis protein” family (Pfam PF00977), which is based on ProFAR isomerase (HisA) and imidazole glycerol phosphate synthase (HisF) (31). However, *M. extorquens* PA1 already contains two adjacent genes—Mext_2551 and Mext_2552—annotated with the functions of HisA and HisF. As there is no obvious need for a potential isomerase (HisA) or cyclase (HisF) in the biosynthesis of MYFR (the formation of the furan ring is catalyzed by MfnB/Orf22 (19)), the function of Orf17 might be unrelated to these activities. *M. jannaschii* possesses with MJ_0703 a homolog of Orf17 (26% sequence identity), which has been suggested to be involved in H₄SPT biosynthesis (32). However, we were able to detect methenyl-dH₄MPT in the Δorf17 *M. extorquens* mutant (Fig. S2), thus invalidating that hypothesis at least in bacteria.

To verify that Orf17 is indeed involved in MYFR biosynthesis, we complemented the mutant by expressing orf17 from a plasmid. This strain produced MYFR with a WT-like glutamate distribution (Fig. S3), thus confirming that the deletion of orf17 is responsible for the lack of MYFR in the mutant.
Strikingly, MYFR from this strain had a larger polyglutamate overexpression of the strain was then analyzed by LC-MS to determine whether the overexpression strain was broader compared with the WT. The increased expression level of MyfB thus results in a higher number of glutamates in the chain, confirming that MyfB is a crucial enzyme for the elongation of the polyglutamate side chain.

Next, we partially purified the tagged MyfB from M. extorquens by affinity purification and size-exclusion chromatography (Fig. SS5, A and B). To determine whether MyfB was binding MYFR as previously observed for Fhc (9), we heat-denatured a sample of MyfB and analyzed the released MYFR by LC-MS. Surprisingly, purified MyfB had a broad distribution of MYFR bound, containing between 10 and 30 glutamates (Fig. 5B). MYFR was still bound to MyfB after two chromatographic steps, indicating a rather strong, yet non-covalent (heat-labile) association.

The glutamate distribution of MYFR in the cell lysate principally corresponds to the sum of the distribution of unbound MYFR and MYFR bound to MyfB and other proteins (such as Fhc). The observation that the MyfB-bound glutamate distribution (Fig. 5B) was distinct from the distribution of MYFR in the cell lysate (Fig. 5A) suggests that MyfB-bound MYFR represents only a minor subset of the total MYFR pool.

**MyfB catalyzes the synthesis of polyglutamates from L- and D-glutamate in vitro**

To further investigate the function of MyfB (Orf5), we performed in vitro assays using purified enzyme produced heterologously in Escherichia coli (Fig. SS5C). Based on the above findings, the putative function of MyfB is to extend the polyglutamate side chain in MYFR, starting from the N-terminus of MYFR-Glu1, the resulting MYFR-Glu2 intermediate can then be elongated C-terminally by MyfB.

In the extract of the Δorf17 M. extorquens mutant, no MYFR-Gluu intermediates were observed by LC-MS. Orf17 might thus be involved in an initial step of MYFR biosynthesis, e.g., in the formation of the core structure, which results in intermediates not detectable by our LC-MS method.

**MyfB binds and elongates MYFR in vivo**

The similarity of MyfB (Orf5) to known glutamate ligases and the accumulation of the intermediate MYFRtyrosine-Gluu2 in the M. extorquens mutant prompted us to further characterize this enzyme. To study its function, we first homologously overexpressed Strep-tagged myfB under control of the strong mxaF promoter in M. extorquens. The cell lysate of this strain was then analyzed by LC-MS to determine whether the overexpression of myfB affected the structure of MYFR. Strikingly, MYFR from this strain had a larger polyglutamate side chain and MYFR with up to 40 glutamates was observed (Fig. 5A). Further, the distribution of the number of glutamates was broader compared with the WT. The increased expression level of myfB thus results in a higher number of glutamates in the chain, confirming that MyfB is a crucial enzyme for the elongation of the polyglutamate side chain.

![Figure 4. LC-MS extracted ion chromatograms showing the presence of MYFR biosynthesis intermediates in the deletion strains.](image)

![Figure 5. Distribution of the number of glutamates in MYFR extracted from a myfB (orf5) overexpressing M. extorquens strain.](image)
intermediate MYFR-Glu2. However, this intermediate was present only in very low amounts in the ΔmyfB M. extorquens strain, and we were therefore limited to more available substrates for in vitro assays. Assuming that MyfB can form polyglutamates even in the absence of a precursor (similarly as the MyfB-homolog RimK (33)), we performed assays using only glutamate as substrate. When MyfB was incubated with L/D-glutamate and ATP/GTP, indeed the synthesis of polyglutamates with 2 up to 11 units (Glu2–Glu11) was observed by LC-MS (Fig. 6). Assays performed with either L- or D-glutamate alone also resulted in the synthesis of polyglutamates, demonstrating that both enantiomers can act as a substrate (Fig. S6).

The observed polyglutamate synthesis activity confirms the role of MyfB as a glutamate ligase. However, the amounts of polyglutamates produced were small and required around 1 day of incubation to appear. It is thus likely that the de novo formation of polyglutamates is merely a side activity that occurs in the absence of an MYFR precursor. This assumption is consistent with the finding that cell extracts of M. extorquens Δ- and D-glutamate, ATP, and GTP in the presence of 5 mM of MgCl2, MnCl2, and KCl. A mass tolerance of 5 ppm was used to generate the chromatograms.

**Discussion**

Our analysis of MYFR from different methylotrophic proteobacteria revealed that there are two main types of MYFR: the previously described MYFR-Tyrosine (6) and the herein identified MYFR-Tyramine that contains tyramine instead of tyrosine in its core structure. Interestingly, the predominant type of MYFR in the surveyed strains is MYFR-Tyramine (six out of nine strains), which has the same core structure as archaeal MFR. Due to the predominance of MYFR-Tyramine, the missing carboxylic acid group is not expected to significantly affect the function of MYFR. Besides these variances in the core structure, the analysis also highlights differences in the number of glutamates attached to the core (Fig. 2). Notably, MYFR in all strains contained at least ten glutamates, suggesting that this might be the minimal number of units required for MYFR to function as a prosthetic group of Fhc. Based on the crystal structure of Fhc from M. extorquens (9), we previously estimated that about five glutamates are necessary to span the mere distance between the MYFR binding site and the two active sites, while a similar number of glutamates might be required to create the branched structure needed for tight binding of MYFR to Fhc. The structure further showed that the branched polyglutamate side chain of MYFR associates with Fhc via a large number of electrostatic interactions between the negatively charged glutamates and numerous positive residues of Fhc (9). These amino acids are conserved in methylotrophic bacteria (9), including the ones tested here, indicating a similar binding mode of MYFR in all strains. It thus remains puzzling what purpose the observed structural diversity serves and whether the differences in MYFR also require structural adaptions of Fhc.

Variances in the number of glutamates (both within and between strains) are also observed for the coenzymes tetrahydrofolate (34) and F420 (35). The glutamate chains in these coenzymes are catalytically not required and their physiological role is still not entirely clear. However, several functions have been proposed, including increased intracellular retention, enhanced recognition by enzymes, alteration of kinetic parameters, and facilitation of channeling (34, 36). In the case of MYFR, the polyglutamate chain serves a distinct function: not only does it allow tight binding to Fhc, but it also acts as a flexible linker that enables MYFR to reach both active sites of the bifunctional enzyme for the shuttling of formyl units (9).

The structural diversity of MYFR also affects its biosynthesis. In strains producing MYFR-Tyramine, MfnA is required to decarboxylate tyrosine to tyramine, thus adding a step to the biosynthesis. The selective incorporation of tyramine further requires a biosynthetic enzyme of high specificity to prevent cross-reactivity with tyrosine. In contrast, we showed that the incorporation of tyrosine in MYFR-Tyrosine-producing strains is nonspecific, as M. extorquens was able to synthesize MYFR-Tyramine when supplemented with tyramine. At least in M. extorquens, the biosynthetic enzymes thus seem to be agnostic to the type of the core structure.

Focusing further on the biosynthesis of the polyglutamate side chain of MYFR, we identified OrfY and Orf5 as crucial enzymes for the elongation of the chain. We therefore renamed these proteins to MyfA and MyfB, respectively, highlighting that these are the first enzymes discovered to be specifically involved in the biosynthesis of MYFR. Based on the biosynthetic pathway of archaeal MFR and incorporating the findings of this work, a putative pathway for MYFR biosynthesis is proposed (Fig. 7). The biosynthesis starts with

![Figure 6. LC-MS extracted ion chromatograms showing the in vitro formation of polyglutamates by MyfB (Orf5). Polyglutamates with 2 up to 11 units (Glu2–Glu11) were observed after incubation of MyfB for 23 h with 100 μM of l- and d-glutamate, ATP, and GTP in the presence of 5 mM of MgCl2, MnCl2, and KCl. A mass tolerance of 5 ppm was used to generate the chromatograms.](image-url)
Diversity and biosynthesis of MYFR

Figure 7. Proposed biosynthetic pathway for the production of MYFR in bacteria. The conversion of tyrosine to tyramine only takes place in strains that produce MYFR_Tyramine and thus carry the *myfA* gene. Note that the types of the linkages between the glutamates are unknown; here, they are shown arbitrarily as γ-linkages. F1-PP, 5-(aminomethyl)-3-furanmethanol-pyrophosphate.

tyrosine, which can either be decarboxylated to tyramine by MfnA (in strains producing MYFR_Tyramine) or directly be used for the addition of a glutamate unit. The ligation of tyramine with glutamate might be catalyzed by homologs of archaeal MfnD (18), although no homolog was found for the MYFR_Tyramine-producing strain *H. savarzinii*. The enzyme catalyzing the ligation of tyrosine with glutamate is still unknown, as most MYFR_Tyrosine-producing strains do not have a homolog of MfnD (Table S4). The intermediate F1-PP (5-(aminomethyl)-3-furanmethanol-pyrophosphate) is assumed to be formed analogously as in MFR biosynthesis (17, 19) using MfnB (Orf22), a yet to be identified transaminase (MfnC), and MfnE (Orf21). In the next step, MfnF (Orf9) presumably catalyzes the ligation of F1-PP with the glutamate-tyrosine/tyramine dipeptide (Glu-Tyr), resulting in MYFR-Glu1. The addition of the second glutamate is catalyzed by MyfA, since MYFR-Glu1 is accumulating in the M. agellatus strain. The enzyme catalyzing the ligation of tyrosine with glutamate is still unknown, as most MYFR_Tyrosine-producing strains do not have a homolog of MfnD (Table S4). The intermediate F1-PP (5-(aminomethyl)-3-furanmethanol-pyrophosphate) is assumed to be formed analogously as in MFR biosynthesis (17, 19) using MfnB (Orf22), a yet to be identified transaminase (MfnC), and MfnE (Orf21). In the next step, MfnF (Orf9) presumably catalyzes the ligation of F1-PP with the glutamate-tyrosine/tyramine dipeptide (Glu-Tyr), resulting in MYFR-Glu1. The addition of the second glutamate is catalyzed by MyfA, since MYFR-Glu1 is accumulating in the *M. agellatus* strain. It is still unclear, why additionally MYFR-Glu1 (i.e., just the MYFR core) was observed in this strain. This observation also makes it conceivable that tyrosine/tyramine is first ligated with F1-PP and only then the first glutamate is added (potentially still by MfnD). The further elongation of the glutamate chain of MYFR-Glu1 is at least partly catalyzed by MyfB, whose deletion in *M. extorquens* leads to the accumulation of MYFR-Glu2, while its overexpression results in a drastic increase in the number of glutamates in MYFR (Fig. 5A). *In vitro* assays further confirmed its role as a glutamate ligase (Fig. 6). These assays also revealed that MyfB can use both enantiomers of glutamate as a substrate, raising the possibility for the presence of D-glutamate in MYFR. Other, yet to be identified, enzymes might additionally be required for the maturation of the polyglutamate chain, which especially includes the addition of branching sites. The role of Orf17 in MYFR biosynthesis is still unclear and hence not part of the proposed pathway.

In all analyzed strains, MYFR was present as a mixture of species with a varying number of glutamates. Thus, an interesting question concerns the regulation of the length of the polyglutamate chain. The ability of MyfB to noncovalently bind MYFR (Fig. 5B) points toward a processive mechanism (37), where the polyglutamate moiety stays bound for multiple rounds of elongation and is only released once its destined length is reached. However, the finding that the overexpression of *myfB* was enough to drastically increase the number of glutamates in MYFR suggests that MyfB does not intrinsically limit the length of the chain, but that the length is rather determined by the expression level of *myfB*. It will thus be interesting to investigate the interaction between MYFR and MyfB and elucidate the mechanism by which the enzyme elongates the chain.

In general, amide bond forming ligases, including ATP-grasp enzymes, are a versatile class of proteins with unique properties for synthetic applications (38–40). The glutamate ligase activities of MyfA and MyfB thus provide interesting prospects for biocatalysis, e.g., for the production of polyglutamate, a nontoxic biopolymer of industrial relevance (41, 42). Our finding that MyfB is able to synthesize polyglutamates in *vitro* using both L- and D-glutamate demonstrates the potential of this enzyme. Using enzyme engineering, the activity and substrate scope might further be fine-tuned for a given application. Lastly, the identification of the enzymes that introduce the branching into the polyglutamate side chain of MYFR would open up even more potential for biocatalysis, as such enzymes have not yet been described.

**Experimental procedures**

**Cultivation of bacterial strains**

When possible, the strains used for the MYFR screen (Table S1) were grown in baffled flasks in a minimal medium with 0.5% (v/v) methanol (MMM). MMM is composed of mineral salts (30.3 mM NH₄Cl, 0.8 mM MgSO₄), buffer (13.8 mM K₂HPO₄, 6.9 mM NaH₂PO₄), iron solution (40.3 μM Na₂EDTA, 10.8 μM FeSO₄), and trace elements (15.7 μM ZnSO₄, 12.6 μM CoCl₂, 5.1 μM MnCl₂, 16.2 μM H₃BO₃, 1.9 μM CuSO₄, 0.3 μM NiCl₂).
Diversity and biosynthesis of MYFR

1.65 μM Na₂MoO₄, 1.2 μM CuSO₄, 20.4 μM CaCl₂). The strains V. paradoxus 351, and Burkholderia sp. Leaf177 harbor a lantanide-dependent methanol dehydrogenase and additionally required LaCl₃ (60 μM) for growth on methanol. The strain Burkholderia sp. Leaf177 did not grow well in liquid minimal medium and was thus cultured on solid MMM agar plates (1.5% agar). S. novella DSM 506 was grown in R2A medium (0.5 g/l yeast extract, 0.5 g/l proteose peptone, 0.5 g/l casamino acids, 0.5 g/l glucose, 0.5 g/l soluble starch, 2.7 mM sodium pyruvate, 1.2 mM K₂HPO₄, 0.2 mM MgSO₄) supplemented with 0.5% (v/v) methanol.

M. extorquens mutants were grown in minimal medium with succinate (MMS), which is identical to MMM except for the replacement of the buffer component with 9.1 mM K₂HPO₄ and 11.5 mM NaH₂PO₄ and for the addition of 30.8 mM disodium succinate instead of methanol. The M. flagellatus mutant was grown in MMM.

All strains were incubated at 28 °C, except for M. flagellatus KT, which was grown at 37 °C.

**Extraction and enrichment of MYFR**

For the identification of MYFR in the different proteobacterial strains, cells were harvested from liquid cultures or, in case of Burkholderia sp. Leaf177, from an agar plate. Cell pellets were extracted three times using 5 ml of 60% methanol in a boiling water bath as described before (6). The cell extracts were purified using weak anion-exchange solid-phase extraction (SPE) columns (Strata X-AW, 33 μm, 30 mg, Phenomenex). After equilibration of the column with two-time 1 ml methanol and two-time 1 ml water, the methanolic extract was applied. The column was washed two times with 1 ml water and 1 ml methanol, before elution with first 1 ml of 0.05% ammonia in 50% methanol and then with 1 ml of 0.25% ammonia in 50% methanol was performed. The first eluate contained mainly compounds with only a few negative charges (less than eight), while the second eluate contained highly negatively charged compounds including MYFR. Eluates were dried in vacuo (using a SpeedVac), and the second eluate was usually used for LC-MS analysis (in a few cases, both eluates were combined).

For M. flagellatus, the methanolic cell extract was purified slightly differently. After drying in vacuo and redissolving with water, the extract was loaded on a strong anion-exchange column (HiTrap Q HP, 1 ml, GE Healthcare) attached to an FPLC system (AKTA Purifier, GE Healthcare). A gradient from 200 mM to 2 M ammonium bicarbonate was applied within 15 min (at 1 ml/min) for elution. MYFR eluted at ~75 mS/cm (~1.4 M ammonium bicarbonate). The corresponding fractions were dried in vacuo, redissolved with water, and heated at ~80 °C for the removal of the remaining ammonium bicarbonate before they were analyzed by LC-MS. Due to the ability for parallelization and the savings in time, the SPE method was usually preferred for MYFR enrichment.

To determine the presence of MYFR in the deletion strains, the cell pellets were lysed by either passing them three times through a French press cell (ΔmyfB and Δorf17) or by using boiling methanol/water as described above (ΔmyfA). In the former case, proteins were removed from the lysate by heat denaturation at 95 °C for 10 min followed by centrifugation. The cell extracts were then purified using strong anion-exchange chromatography as described above. The MYFRTyrosine-Glu₂ intermediate observed in the ΔmyfB M. extorquens strain was eluting at ~38 mS/cm (~0.8 M ammonium bicarbonate). The MYFRTyrosine-Glu₀/₁ intermediates observed in the ΔmyfA M. flagellatus strain were detected by LC-MS directly in the cell extract.

For the extraction of MYFR from the myfB overexpressing M. extorquens strain, a small amount of a cell lysate obtained by French press was heated at 99 °C for 15 min to denature proteins. After centrifugation, the sample was diluted with 1 ml of water and purified using SPE as described above.

To extract MYFR bound to purified MyfB, a small protein sample was heat-denatured at 100 °C for 10 min and diluted to 10 ng/μl for LC-MS analysis.

**LC-MS measurements**

For the measurement of MYFR, nanoscale ion-pair reversed-phase LC-MS (using tributylamine as ion-pair reagent) was used (43). Measurements were performed on a Q Exactive Plus orbitrap mass spectrometer (Thermo Fisher Scientific) operated in negative mode and coupled to a nano-2D Ultra LC system (Eksigent/AB SCIEX) as described before (9). The cell extracts were diluted to 1 to 10 μg/μl cell dry weight (cdw; assuming ~300 μg/ml cdw at OD₆₀₀ = 1) with solvent A (230 μM tributylamine, 230 μM acetic acid, and 3% methanol in water, adjusted to pH 9.0 with ammonia) and 1 μl was injected. The low-abundant MYFRTyrosine-Glu₂ biosynthesis intermediate was also measured using this method.

For the detection of the MYFRTyrosine-Glu₀/₁ intermediates and the polyglutamates produced in vitro, a less sensitive method was sufficient. Here, the mass spectrometer was coupled with a Dionex UltiMate 3000 UHPLC (Thermo Fisher Scientific). LC separation took place on a Kinetex Polar C18 column (2.6 μm, 100 Å, 2.1 × 100 mm, Phenomenex). Solvent A was 1.1 M formic acid with 0.2 M ammonium, solvent B was acetonitrile, and solvent C was water. The following gradient was applied at 500 μl/min (% A/B/C): 0 min, 5/3/92; 0.5 min, 5/3/92; 3.5 min, 5/50/45; 4 min, 5/95/0; 7 min, 5/95/0; 7 min, 5/3/92; 10 min, 5/3/92. Heated electrospray ionization (HESI) was performed at 380 °C with a spray voltage of 3 kV or ~2.5 kV in positive or negative mode respectively. The sheath and aux gas flow rates were set to 60 and 20, respectively, the capillary was heated to 275 °C, and the S-lens RF level was set to 50. The maximum injection time was 50 ms (AGC target 1e6) and spectra were recorded as centroids using a resolution of 35,000 in the range of 120 to 1200 m/z (270–1200 m/z for improved sensitivity). For cell extracts, a volume corresponding to 10 μg cdw was injected. The samples from the in vitro assays were analyzed at a concentration of 2 to 4 μM of (initial) substrate and 25 μl were injected.
MS/MS fragmentation was generally performed using the same method as for MS1 detection; however, an additional scan event was added for fragmentation of the precursor. For the fragmentation of MYFR_{Tyramine}-Glu_{16} from *M. flagellatus*, the UHPLC LC-method was used instead of the nanoscale ion-pair method.

**LC-MS data analysis**

For basic data analysis, Xcalibur Qual Browser (Thermo Fisher Scientific) was used. The untargeted search for unknown MYFR derivatives was implemented as a Python script using eMZed2 (44). After feature finding, mass differences corresponding to one or multiple glutamate units determined using eMZed2 as described before (6, 9). Extracted ion chromatograms of the MYFR species with different numbers of glutamates (core + ion) were generated by homologous recombination using electrocompetent ΔmxaF *M. extorquens* PA1 cells. MYFR from the complemented strain was extracted using boiling methanol/water, enriched by solid-phase extraction, and analyzed by LC-MS as described above.

**Cloning of myfB for overexpression in M. extorquens and E. coli**

For the tagging and homologous overexpression of *myfB* (orf5, Mext_1832) in *M. extorquens* PA1, *myfB* was amplified together with the region 22 bp upstream of the gene (containing the native ribosomal binding site) from genomic DNA by PCR (using primers Mext_orf5_fwd/Mext_orf5_rev; Table S5). The PCR product was restricted with PstI/Ncol, while the vector pCM80 (46) was cut with PstI/BamHI. A C-terminal Strept-tag II (47) was introduced by ordering both strands of the tag sequence as a primer (Strep_1/Strep_2; Table S5) and ligating them together with restricted *myfB* and vector. The final plasmid was amplified in *E. coli* DH5α and verified by sequencing, before it was transformed into electrocompetent *M. extorquens* PA1 cells.

**Complementation of the Δorf17 M. extorquens mutant**

To complement the Δorf17 mutant, the *orf17* gene (Mext_1835) was amplified from genomic DNA by PCR (using primers Orf17_fwd/Orf17_rev; Table S5). The forward primer further encoded a strong ribosomal binding site. The restriction sites HindIII and BamHI were used to insert the gene into pCM80 (46). The final plasmid was amplified in *E. coli* DH5α and verified by sequencing, before it was transformed into electrocompetent Δorf17ΔmxaF *M. extorquens* PA1 cells.

**Production and purification of MyfB**

For the production of Strep-tagged MyfB (Orf5) in *M. extorquens*, transformed cells were grown in MMS medium containing 0.5% to 1% methanol and supplemented with 10 μg/ml tetracycline. The cell pellets were resuspended with 3 ml Tris buffer (25 mM Tris, 300 mM NaCl, pH 7.6) and the cells were lysed by passing them four times through a French press cell. After ultracentrifugation, MyfB was affinity-purified using a StrepTrap HP column (1 ml, GE Healthcare) equilibrated with the buffer used for resuspension. Elution took place using 2.5 mM desthiobiotin in the same buffer. Fractions containing protein were pooled and concentrated in a 10 kDa centrifugal filter (Amicon Ultra, Merck Millipore). SDS-PAGE analysis revealed the presence of two contaminating proteins (Fig. S5A), which were identified by in-gel digestion and LC-MS as propionyl-CoA carboxylase (Pcc) A and B. To further purify MyfB, size-exclusion chromatography was performed using a Superose 6 Increase column (24 ml, GE Healthcare).
Diversity and biosynthesis of MYFR

that was run with the same buffer used for affinity purification. PccAB separated from MyfB during that step; however, the protein was still not entirely pure (Fig. S5B).

For the production of MyfB in E. coli, transformed cells were grown in LB medium supplemented with 100 μg/μl ampicillin. After an OD₆₀₀ of about 0.8 was reached, gene expression was induced by the addition of 0.1 mM IPTG and the culture was switched to 28 °C for overnight (23 h) expression. Cell lysis and StrepTrap purification took place as described above, but using a buffer containing 25 mM Tris, 150 mM NaCl, and 2 mM TCEP at pH 7.8 for resuspension of the pellet and washing of the column and a buffer containing 25 mM Tris, 150 mM NaCl, and 2.5 mM desthiobiotin at pH 7.5 for elution. Concentrated protein was washed once with the resuspension buffer in a 10 kDa centrifugal filter to remove desthiobiotin. SDS-PAGE analysis revealed that MyfA was mostly pure (Fig. S5C).

In vitro assays with MyfB

The in vitro assays were performed in a total volume of 10 μl and contained 1 μl of 100 μM L- and/or D-glutamate, ATP, and GTP as well as 0.5 μl of 100 mM MgCl₂, MnCl₂, and KCl. A volume of 2 μl of 2.8 mg/ml MyfB purified from E. coli was added and the remaining volume (3.5 μl or 2.5 μl, depending on the presence of L- and/or D-glutamate) was filled up with Tris buffer (25 mM Tris, 150 mM NaCl, 2 mM TCEP, pH 7.8).

As a negative control, the enzyme was replaced with buffer. The assays were incubated at 28 °C and 2 μl-samples were taken after 0 h and ~1 day and quenched using 50 μl of 60:20:20 acetonitrile:methanol:0.5 M formic acid. The samples were dried in vacuo, dissolved with 50 μl of water, and analyzed by LC-MS for the presence of polyglutamates.

Data availability

All data are available upon request.

Supporting information—This article contains supporting information. (49)

Acknowledgments—We thank Dr Ludmila Chistoserdova (University of Washington, Seattle, WA, USA) for providing the Δorf5 M. extorquens AM1 and ΔorfY M. flagellatus KT mutants. We also thank the Blockcourse students Viktoria Brunner, Rafael Ruch, Jacob Haninnann, and Timothy Bradley for their help with cloning and characterization of MyfB. We further appreciate the support of Philipp Christen with LC-MS measurements and acknowledge the Functional Genomics Center, Zurich, for the identification of proteins.

This work was supported by Swiss National Science Foundation Grants 31003A-173094 and by ETH Zurich.

Author contributions—J. L. H., M. R. B., and M. B. M. performed the experiments and constructed the deletion strains. J. L. H. analyzed all the data. J. L. H. and J. A. V designed the research and wrote the paper with inputs from all the authors.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: Fhc, formyltransferase/hydrolase complex; H₂MPT, tetrahydromethanopterin; LC-MS, liquid chromatography–mass spectrometry; MFR, methanofuran; MYFR, methylofuran.

References

1. Vorholt, J. A., Chistoserdova, L., Stolyar, S. M., Thauer, R. K., and Lidstrom, M. E. (1999) Distribution of tetrahydromethanopterin-dependent enzymes in methylotrophic bacteria and phylogeny of methenyl tetrahydromethanopterin cyclohydrolases. J. Bacteriol. 181, 5750–5757
2. Vorholt, J. A. (2002) Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. Arch. Microbiol. 178, 239–249
3. Chistoserdova, L. (2011) Modularity of methylotrophy, revisited. Environ. Microbiol. 13, 2603–2622
4. Chistoserdova, L. (2016) Wide distribution of genes for tetrahydromethanopterin/methanofuran-linked C1 transfer reactions argues for their presence in the common ancestor of bacteria and archaea. Front. Microbiol. 7, 1425
5. Chistoserdova, L., Vorholt, J. A., Thauer, R. K., and Lidstrom, M. E. (1998) C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic archaea. Science 281, 99–102
6. Hemmann, J. L., Saurel, O., Ochsner, A. M., Stodden, B. K., Kiefer, P., Milon, A., and Vorholt, J. A. (2016) The one-carbon carrier methylofuran from Methylbacterium extorquens AM1 contains a large number of α- and γ-linked glutamic acid residues. J. Biol. Chem. 291, 9042–9051
7. Pomer, B. K., and Vorholt, J. A. (2001) Characterization of the formyltransferase from Methylbacterium extorquens AM1. Eur. J. Biochem. 268, 4769–4775
8. Pomer, B. K., Saurel, O., Milon, A., and Vorholt, J. A. (2002) Generation of formate by the formyltransferase/hydrolase complex (Fhc) from Methylbacterium extorquens AM1. FEBS Lett. 523, 133–137
9. Hemmann, J. L., Wagner, T., Shima, S., and Vorholt, J. A. (2019) Methylofuran is a prosthetic group of the formyltransferase/hydrolase complex and shuttles one-carbon units between two active sites. Proc. Natl. Acad. Sci. U. S. A. 116, 25583–25590
10. Thauer, R. K. (1998) Biochemistry of methanogenesis: A tribute to Marjory Stephenson. Microbiology 144, 2377–2406
11. Chistoserdova, L., Jenkins, C., Kaluzhnaya, M. G., Marx, C. J., Lapidus, A., Vorholt, J. A., Staley, J. T., and Lidstrom, M. E. (2004) The enigmatic planctomycetes may hold a key to the origins of methanogenesis and methylation. Mol. Biol. Evol. 21, 1234–1241
12. Leigh, J. A., Rinehart, K. L., and Wolfe, R. S. (1984) Structure of methanofuran, the carbon dioxide reduction factor of Methanobacterium thermoautotrophicum. J. Am. Chem. Soc. 106, 3636–3640
13. Bobik, T. A., Donnelly, M. L., Rinehart, K. L., Jr., and Wolfe, R. S. (1987) Structure of a methanofuran derivative found in cell extracts of Methanosarcina barkeri. Arch. Biochem. Biophys. 254, 430–436
14. White, R. H. (1988) Structural diversity among methanofuranas from different methanogenic bacteria. J. Bacteriol. 170, 4594–4597
15. Allen, K. D., and White, R. H. (2014) Identification of structurally diverse methanofuran coenzymes in methanococcales that are both N-formylated and N-acetylated. Biochemistry 53, 6199–6210
16. Kezmarovsky, N. D., Xu, H., Graham, D. E., and White, R. H. (2005) Identification and characterization of a L-tyrosine decarboxylase in Methanocaldococcus jannaschii. Biochim. Biophys. Acta 1722, 175–182
17. Miller, D., Wang, Y., Xu, H., Harich, K., and White, R. H. (2014) Biosynthesis of the 5-(aminomethyl)-3-furanmethanol moiety of methanofuran. Biochemistry 53, 4635–4647
18. Wang, Y., Xu, H., Harich, K. C., and White, R. H. (2014) Identification and characterization of a tyrosine–glutamate ligase (MfnD) involved in methanofuran biosynthesis. Biochemistry 53, 6220–6230
19. Wang, Y., Xu, H., Jones, M. K., and White, R. H. (2015) Identification of the final two genes functioning in methanofuran biosynthesis in Methanocaldococcus jannaschii. J. Bacteriol. 197, 2850–2858
20. Kaluzhnaya, M. G., Korotkova, N., Crowther, G., Marx, C. J., Lidstrom, M. E., and Chistoserdova, L. (2005) Analysis of gene islands involved in...
methylotetrahydrofolate reductase, which catalyzes the reduction of methanopterin-linked Co transfer reactions reveals new functions and provides evolutionary insights. J. Bacteriol. 187, 4607–4614.

21. Chistoserdova, L., Rasche, M. E., and Lidstrom, M. E. (2005) Novel dephospho-redundant methylotetrahydrofolate reductase genes discovered via mutagenesis in Methylococcus extorquens AM1. J. Bacteriol. 187, 2508–2512.

22. Bai, Y., Müller, D. B., Srinivas, G., Garrido-Oter, R., Potthoff, E., Rott, M., Dombrowski, N., Münch, P. C., Spaepen, S., Remus-Emsemann, M., Hüttel, B., McHardy, A. C., Vorholt, J. A., and Schulze-Lefert, P. (2015) Functional overlap of the Arabidopsis leaf and root microbiota. Nature 528, 364–369.

23. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

24. Galperin, M. Y., and Koonin, E. V. (1997) A diverse superfamily of enzymes with ATP-dependent carboxylate—amine/thiol ligase activity. Protein Sci. 6, 2639–2643.

25. Fawaz, M. V., Topper, M. E., and Firestone, S. M. (2011) The ATP-grasp enzymes. Bioorg. Chem. 39, 185–191.

26. Kang, W.-K., Icho, T., Isono, S., Kitakawa, M., and Isono, K. (1989) Characterization of the gene rimK responsible for the addition of glutamic acid residues to the C-terminus of ribosomal protein S6 in Escherichia coli K12. Mol. Gen. Genet. 217, 281–288.

27. Zhao, G., Jin, Z., Wang, Y., Allewell, N. M., Tuchman, M., and Shi, D. (2013) Structure and function of Escherichia coli RimK, an ATP-grasp fold, L-glutamyl ligase enzyme. Proteins 81, 1847–1854.

28. Fan, C., Moews, P. C., Shi, Y., Walsh, C. T., and Knox, J. R. (1995) A common fold for peptide synthetases cleaving ATP to ADP: Glutathione synthetase homologs encode α-L-glutamate ligases for methanogenic ATP-grasp enzymes. J. Biol. Chem. 270, 359–365.

29. Afjeh, P., Fan, R., Li, P., Lazzaro, A., Bazzicalupo, M., Carlosmagno, M. S., and Bruni, C. B. (1996) Histidine biosynthetic pathway and genes: Structure, regulation, and evolution. Microbiol. Rev. 60, 44–69.

30. White, R. H. (2011) The conversion of a phenol to an aniline occurs in the biochemical formation of the 1-(4-aminophenyl)-1-deoxy-D-ribitol moiety in methanotroph. Biochemistry 50, 6041–6052.

31. Kino, K., Arai, T., and Arimura, Y. (2011) Poly-a-glutamic acid synthesis using a novel catalytic activity of RimK from Escherichia coli K-12. Appl. Environ. Microbiol. 77, 2019–2025.

32. Schirch, V., and Strong, W. B. (1989) Interaction of polyglycylglutamates with enzymes in one-carbon metabolism. Arch. Biochem. Biophys. 269, 371–380.

33. Schäfer, A., and Schmidt, T. G. M. (2000) Use of the Strep-tag and streptavidin for detection and purification of recombinant proteins. In Methods Enzymol., Academic Press, San Diego, CA: 271–304.

34. Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., and Minh, B. Q. (2016) W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 44, W232–W235.

35. Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kaufl, F., Wilczynski, B., and de Hoon, M. J. (2009) Biopython: Free and open source Python tools for computational molecular biology and bioinformatics. Bioinformatics 25, 1422–1423.