Methylofuran is a prosthetic group of the formyltransferase/hydrolase complex and shuttles one-carbon units between two active sites

Jethro L. Hemmann[a], Tristan Wagner[b,c], Seigo Shima[b], and Julia A. Vorholt[a,1]

[a]Institute of Microbiology, Eidgenössische Technische Hochschule Zurich, 8093 Zurich, Switzerland; [b]Max Planck Institute for Terrestrial Microbiology, 35043 Marburg, Germany; and [c]Max Planck Institute for Marine Microbiology, 28359 Bremen, Germany

Edited by Mary E. Lidstrom, University of Washington, Seattle, WA, and approved November 3, 2019 (received for review July 6, 2019)

Methylotrophy, the ability of microorganisms to grow on reduced one-carbon substrates such as methane or methanol, is a feature of various bacterial species. The prevailing oxidation pathway depends on tetrahydromethanopterin (H_4MPT) and methylofuran (MYFR), an analog of methanofuran from methanogenic archaea. Formyltransferase/hydrolase complex (Fhc) generates formate from formyl-H_4MPT in two consecutive reactions where MYFR acts as a carrier of one-carbon units. Recently, we chemically characterized MYFR from the model methylotroph *Methylorubrum extorquens* and identified an unusually long polyglutamate side chain of up to 24 glutamates. Here, we report on the crystal structure of Fhc to investigate the function of the polyglutamate side chain in MYFR and the relatedness of the enzyme complex with the orthologous enzymes in archaea. We identified MYFR as a prosthetic group that is tightly, but noncovalently, bound to Fhc. Surprisingly, the structure of Fhc together with MYFR revealed that the polyglutamate side chain of MYFR is branched and contains glutamates with amide bonds at both their α- and γ-carboxyl groups. This negatively charged and branched polyglutamate side chain interacts with a cluster of conserved positively charged residues of Fhc, allowing for strong interactions. The MYFR binding site is located equidistantly from the active site of the formyltransferase (FhcD) and metallo-hydrolase (FhcA). The polyglutamate serves therefore an additional function as a swinging linker to shuttle the one-carbon carrying amine between the two active sites, thereby likely increasing overall catalysis while decreasing the need for high intracellular MYFR concentrations.

methylotrophy | one-carbon metabolism | coenzyme | prosthetic group | polyglutamate

The metabolism of methanogenic, methanotrophic, and sulfate-reducing archaea and methylotrophic bacteria hinges on coenzymes that bind, convert, and release one-carbon units (1–4). Two of these coenzymes, methanofuran (MFR) and tetrahydromethanopterin (H_4MPT), are present in both the domains of Archaea and Bacteria (3, 5). Oxidation levels of coenzyme-bound one-carbon intermediates are shared in anaerobic methanogenesis and aerobic methylotrophy (Fig. 1*A*), and the respective interconverting enzymes are proposed to have existed in the last common ancestor of Bacteria and Archaea (6–8).

The analog of MFR present in the methylotrophic bacterium *Methylorubrum extorquens* (formerly *Methylobacterium extorquens*; ref. 9) has recently been identified and termed methylofuran (MYFR) (5). Both, MFR and MYFR contain a furan moiety with a formylatable aminomethyl group. However, compared to the five known archaeal MFR derivatives (10–13), the core structure of MYFR contains a tyrosine residue instead of the tyramine and has a polyglutamic acid side chain consisting of between 12 and 24 glutamates attached (5), significantly more than what has been described for MFRs (up to 12 γ-linked glutamates in the case of MFR-d) (13). In addition, the glutamic acid residues of MYFR were identified as a mixture of α-linked and γ-linked amino acids.

In aerobic methylotrophy, the conversion of formyl-H_4MPT to formate is catalyzed by the formyltransferase/hydrolase complex (FhcABCD) and proceeds via the intermediate formyl-MYFR (14, 15) (Fig. 1*B*). Formate can then be further oxidized to CO_2 by one of multiple formate dehydrogenases (16, 17). This is in contrast to methanogens growing hydrogenotrophically (on H_2 and CO_2). Here, CO_2 is metabolized in the reverse direction (Fig. 1*A*). The enzyme complex formyl-MFR dehydrogenase (FwdABCDFG) catalyzes the reduction of CO_2 to formate as well as the subsequent binding of formate to MFR as a formyl group (18). A separate enzyme, formyltransferase (Ftr), then transfers the formyl group from formyl-MFR to H_4MPT (19). In methanogens growing on methanol or acetate, however, the archaeal enzymes described above run in the oxidative direction, too (2).

The Fhc subunits A, B, and C of *M. extorquens* show sequence identities of 35%, 24%, and 32%, respectively, with the corresponding subunits of FwdABC from the hydrogenotrophic methanogen *Methanothermobacter wolfeii*. FhcD exhibits a sequence identity of 45% with Ftr from the same methanogen and its formyltransferase activity has been demonstrated (14). FhcA shares the sequence signature for amidohydrolases and is attributed as formyl-MYFR hydrolase (15). FhcB lacks the sequence motifs for the binding of the tungstopterin cofactor and the iron-sulfur cluster present in FwdB. FhcB was therefore

Significance

Methylotrophs play a crucial role in the global carbon cycle as they oxidize reduced one-carbon compounds such as methanol or methane to CO_2. The step-wise conversion of these substrates generally takes place while the one-carbon units are bound to diffusible carrier molecules. However, our crystal structure of the formyltransferase/hydrolase complex from *Methylorubrum extorquens* demonstrates that the one-carbon carrier methylofuran tightly binds to the enzyme via its extended and branched polyglutamate chain. A swinging motion of the coenzyme then shuttles one-carbon units between the two active sites of the enzyme complex over a distance of 50 Å. The structure further highlights how the bacterial formate generation system relates to the archaeal CO_2 fixation system.

Author contributions: J.L.H., T.W., S.S., and J.A.V. designed research; J.L.H. and T.W. performed research; J.L.H. and T.W. analyzed data; and J.L.H., T.W., S.S., and J.A.V. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB), http://www.wwpdb.org (ID code 6S6Y).

[1]To whom correspondence may be addressed. Email: jvorholt@ethz.ch.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1911595116/-/DCSupplemental.

First published November 27, 2019.

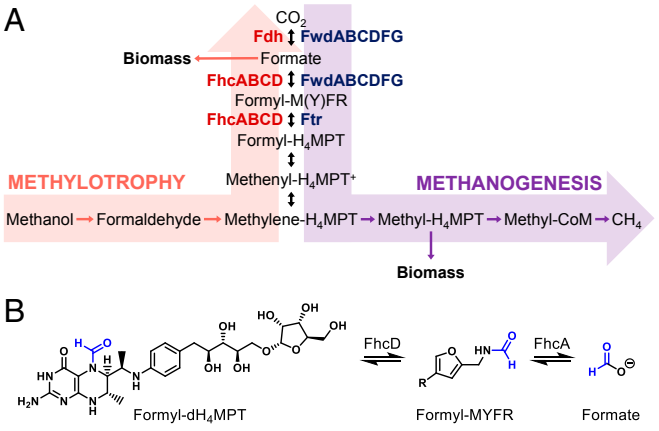

Fig. 1. Intermediates involved in the conversion of methanol to CO_2 in methylotrophy. (*A*) Comparison of tetrahydromethanopterin-dependent methylotrophy with methanogenesis. Both pathways use the one-carbon carriers tetrahydromethanopterin (H_4MPT) and methanofuran/methylofuran (M[Y]FR). In methylotrophs, the conversion of formyl-H_4MPT to formate via the intermediate formyl-MYFR is catalyzed by the formyltransferase/hydrolase complex (FhcABCD). Separate formate dehydrogenases (Fdh) oxidize the formate to CO_2. In methanogens, the reduction of CO_2 to formate and the subsequent binding of the formyl unit to MFR is catalyzed by FwdABCDFG. A separate formyltransferase (Ftr) is responsible for the transfer of the formyl-group from MFR to H_4MPT. (*B*) Chemical reaction catalyzed by FhcABCD. Formyl-H_4MPT is shown here as dephospho-H_4MPT (dH_4MPT), which is the form that is present in *M. extorquens* (3). For a complete structure of MYFR, see Fig. 4*B*.

assumed to have lost its dehydrogenase activity, which is in line with the release of formate by the Fhc complex (15). The function of FhcC is unknown.

The unusually large polyglutamate side chain of MYFR prompted us to determine its biological function and particularly its interaction with Fhc, the only known bacterial enzyme that uses the coenzyme MYFR. Here, we present the crystal structure of FhcABCD with bound MYFR. The structural analysis combined with biochemical experiments elucidates the molecular basis by which Fhc generates formate using the coenzyme MYFR and further highlights important differences to the methanogenic CO_2 fixation system.

Results

MYFR as a Prosthetic Group of FhcABCD. For biochemical characterization and crystallization, the *fhcBADC* gene cluster was cloned with a Strep-tag at the C terminus of *fhcC* into a plasmid and homologously expressed in *M. extorquens* PA1. A three-step purification yielded pure and homogenous FhcABCD (*SI Appendix*, Fig. S1).

Both of the expected activities of purified Fhc were verified. We assayed formyltransferase activity in the direction of formyl-MYFR formation from formyl-H_4MPT using the cytochrome *c*-dependent assay described before (14). No addition of MYFR was required for activity (0.9 U/mg), indicating that MYFR was already bound to purified Fhc (Fig. 2*A*). Next, the formylhydrolase activity was assayed using archaeal formyl-MFR-a as substrate. Both substrate depletion and MFR-a formation were monitored using liquid chromatography-mass spectrometry (LC-MS) and the activity (0.3 U/mg) was confirmed (Fig. 2*B*).

To validate that MYFR was bound to Fhc, purified enzyme was heat-denatured and the supernatant analyzed by LC-MS for the presence of MYFR. We detected MYFR after each step of the purification, indicating a strong, yet noncovalent binding of MYFR to Fhc. MYFR released from purified Fhc was a mixture of species containing between 12 and 20 glutamates (*SI Appendix*, Fig. S2). The observed distribution of glutamate residues had a maximum at 16 glutamates and was slightly skewed compared to the more symmetrical distribution centered at 19 glutamates that was observed in cell extracts from WT *M. extorquens* (5). The shifted glutamate distribution is likely due to the overexpression of the *fhc* genes in this recombinant strain, presumably because the higher intracellular concentration of Fhc might result in early binding of MYFR that has not been fully extended to a higher number of glutamates.

To determine whether all of the bound MYFRs with different numbers of glutamates were active, the formyl hydrolysis reaction was run in the direction of formyl-MYFR generation by incubating purified Fhc with excess of formate. About 40% of MYFR was converted to formyl-MYFR when the reaction was run to equilibrium, as determined by LC-MS (Fig. 2*C*). Interestingly, the relative abundance of generated formyl-MYFR was independent of the number of glutamates in MYFR, indicating that all bound MYFRs are able to participate in the formylation reaction (a slight drop in the amount of formyl-MYFR was observed for MYFRs with >18 glutamates).

The observed copurification of MYFR with Fhc indicated that at least a part of the total MYFR pool is protein-bound. To further quantify this proportion, we fractionated an *M. extorquens* cell lysate using size exclusion chromatography into a protein fraction (>10 kDa) and a metabolite fraction (<10 kDa) and detected MYFR in the fractions by LC-MS (*SI Appendix*, Fig. S3). Ninety-seven percent of MYFR was present in the protein fraction and therefore must be protein-bound. Consequently, almost no free intracellular MYFR pool exists, confirming that MYFR is a prosthetic group.

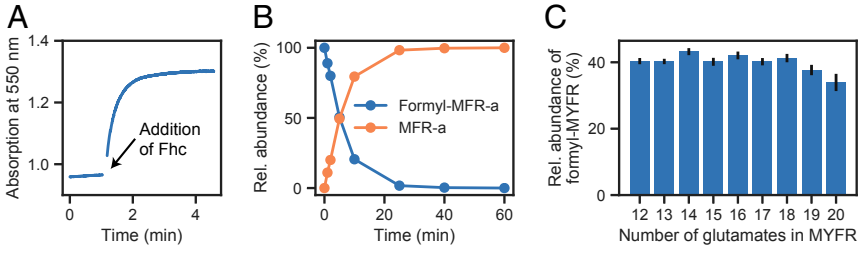

Fig. 2. Activity measurements performed with purified Fhc. (*A*) Reaction progress of the formyltransferase reaction (formyl-H_4MPT + MYFR → H_4MPT + formyl-MYFR) followed using an assay where released H_4MPT spontaneously reduces cytochrome *c*, which is monitored at 550 nm. No MYFR was added in this assay, indicating that MYFR must be already bound to Fhc. The activity was 0.9 U/mg. (*B*) Reaction progress of the formylhydrolase reaction (formyl-MFR-a + H_2O → MFR-a + formate) using archaeal formyl-MFR-a as a substrate followed using LC-MS. The activity was 0.3 U/mg. (*C*) Relative abundance of formyl-MYFRs with different number of glutamates that were produced upon incubation of Fhc (with bound MYFR) with 0.45 M formate, as determined by LC-MS. The error bars denote the SD of individual LC-MS measurements (n = 6).

Overall Structure of FhcABCD. To determine the binding site of MYFR and its interaction with Fhc, we crystallized the enzyme complex. The X-ray structure was solved by molecular replacement at 3.1 Å (20) (*SI Appendix*, Table S1) using FwdABC from *M. wolfeii* (PDB ID code 5T5I) (21) and Ftr from *Methanopyrus kandleri* (PDB ID code 2FHJ) (22) as templates. The overall structure of FhcABCD is a dimer of heterotetramers in which FhcD is the core dimer and FhcABC are the peripheral units (Fig. 3). This oligomeric state also corresponds to the native structure of Fhc, as shown by clear-native PAGE (*SI Appendix*, Fig. S1). The subunit structures of FhcABC are similar to the corresponding structures of FwdABC from *M. wolfeii* (21) with rmsd values of 0.85 Å (for 374 Cα aligned), 1.51 Å (for 224 Cα aligned), and 0.91 Å (for 172 Cα aligned), respectively. The structure of FhcD is close to Ftr from *M. kandleri* with an rmsd of 0.89 Å (for 510 Cα aligned) (*SI Appendix*, Fig. S4).

FhcA (58 kDa), the formylhydrolase, belongs to the amidohydrolase superfamily (containing e.g., urease, dihydroorotase, phosphotriesterase, and dihydropyrimidinase) (23) and shows a characteristic binuclear metal center containing two zinc ions and a conserved arrangement of four histidines (His57/59/229/258), an aspartic acid (Asp377), and a carboxylated lysine (carboxy-Lys176). Due to this fully conserved zinc-ion coordination motif and the close similarity to other metallo-hydrolases, the two metal ions in the active site were modeled as zinc ions. However, the second zinc ion of FhcA is not coordinated with the conserved Asp377, which resulted in a slight shift of this second zinc ion (*SI Appendix*, Fig. S5).

FhcB (37 kDa) and FhcC (27 kDa without the Strep-tag) serve as structural scaffolds and hold the two catalytic subunits (FhcA and D) together. FhcB is structurally related not only to FwdB, but also to other molybdo(tungsto)pterin-containing enzymes such as formate dehydrogenase (24), DMSO reductase (25), and periplasmic nitrate reductase (26). However, FhcB lacks the features of a functioning dehydrogenase (*SI Appendix*, Fig. S6, see also below). FhcC has a characteristic right-handed β-helix fold and shows high structural similarity to FwdC and also to the C-terminal domain of glutamate synthases (27); however, the C-terminal part of Fhc adopts a helix–loop–helix motif instead of a β-hairpin (residues 216–246). FhcC further serves as a binding platform for FhcD, the formyltransferase that is a stand-alone enzyme in methanogens (*SI Appendix*, Fig. S7).

FhcD (32 kDa) shows only minor structural differences to Ftr, as mentioned above. Its conserved active site is constituted at the homodimer interface (22), explaining the obligate dimeric structure of FhcD at the core of the Fhc complex. Most of the active site residues important for MFR and H_4MPT binding in Ftr are conserved in FhcD (*SI Appendix*, Fig. S8), and the reaction likely proceeds with the same mechanism as described for Ftr (22). This mechanism includes the formation of a tetrahedral oxyanion intermediate upon nucleophilic attack of the formyl-carbon by the amine of MYFR. The oxyanion is presumably stabilized by the conserved and protonated Glu252.

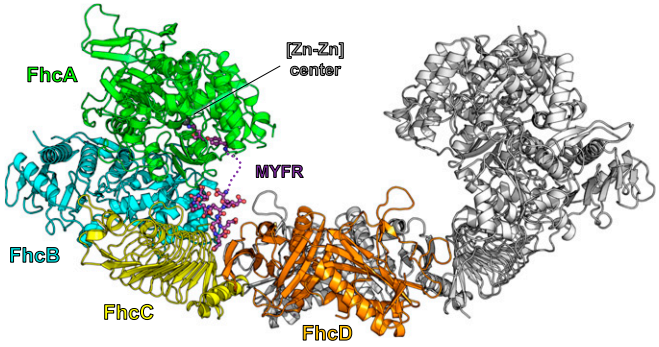

Fig. 3. The structure of the FhcABCD dimer. Fhc forms a dimer of heterotetramers containing a dinuclear metal center (shown as gray spheres) in the formylhydrolase FhcA. The active site for the formyltransferase is located at the interface of the FhcD dimer. The bound MYFR, consisting of the branched polyglutamate side chain and the core structure (the aminomethylfuran and the tyrosine moiety, bound in the active site of FhcA), is shown in purple. The dashed line indicates the part of the polyglutamate chain for which no electron density was observed.

MYFR and Its Branched Polyglutamate Side Chain. Notably, we observed extra electron density at the center of each of the four FhcABCD tetramers in the asymmetric unit, which could be attributed to the bound MYFR. One tetramer presented a particularly pronounced electron density, comparable to that of the amino acid backbone of the protein and, thus, indicating high occupancy (*SI Appendix*, Fig. S9). We then used this electron density to model the polyglutamate side chain of MYFR. To our surprise, the electron density best fitted a branched chain, where some glutamates are connected with both their α-carboxyl and γ-carboxyl groups to additional glutamates (Fig. 4 *A* and *B*). This is especially striking as MYFR and MFRs have been assumed to contain a linear glutamate chain (5, 11, 13). The density allowed the modeling of nine glutamates in the polyglutamate chain. Due to the heterogeneity of bound MYFR in terms of number of glutamates (*SI Appendix*, Fig. S2) and possibly even in terms of multiple branching patterns, an unambiguous modeling of MYFR was, however, not possible. The presented arrangement of the glutamates in the polyglutamate chain (Fig. 4) should therefore exemplify a possible structure. For the remaining 3–11 glutamates (Fhc-bound MYFR has 12–20 glutamates) no electron density was observed, suggesting high spatial flexibility and/or large heterogeneity. Closer inspection of the active site of FhcA revealed electron density that matched the core structure of MYFR (the aminomethylfuran and the tyrosine moiety) (*SI Appendix*, Fig. S9). To connect the terminal part of the polyglutamate chain with the separately modeled core structure of MYFR, two additional γ-linked glutamates were assumed (Fig. 4*B*).

For structure elucidation of organic molecules, nuclear magnetic resonance (NMR) and mass spectrometry are commonly used. In the case of MYFR, however, the large number of glutamates and the above-mentioned heterogeneity limited the structural information that could be obtained by NMR and no sequential assignment of the glutamates in the chain was possible (5). Nevertheless, the previous NMR measurements showed that both α- and γ-linkages are present in the polyglutamate chain, in agreement with a branched structure. To further investigate the structure of the polyglutamate chain in MYFR, we used tandem mass spectrometry (MS/MS). However, since the mass of MYFR is independent of any branching present in the chain and MS/MS fragmentation does not provide information about where in the chain the dissociation occurs, MYFR first had to be derivatized. As MYFR is a modified peptide, it contains a unique N-terminal glutamate residue carrying a primary α-amino group. To identify the fragment ions, which still contain this glutamate with the primary amine, we chemically methylated the amine. The N-methylated MYFR was then subjected to MS/MS fragmentation, and the smallest fragment that still contained the methylated amine was identified (*SI Appendix*, Fig. S10). Unexpectedly, this analysis showed that the glutamate with the primary amine is the second glutamate residue in the chain (counting from the core of MYFR), rather than being located toward the end of the chain. All further glutamates thus have to be connected via the carboxyl group of the second glutamate, which therefore has both its carboxyl groups involved in amide bonds with other glutamates. Additionally, the data excluded any branching at the first glutamate. Together with the electron density, this analysis indicates a revised chemical structure of MYFR (Fig. 4*B*).

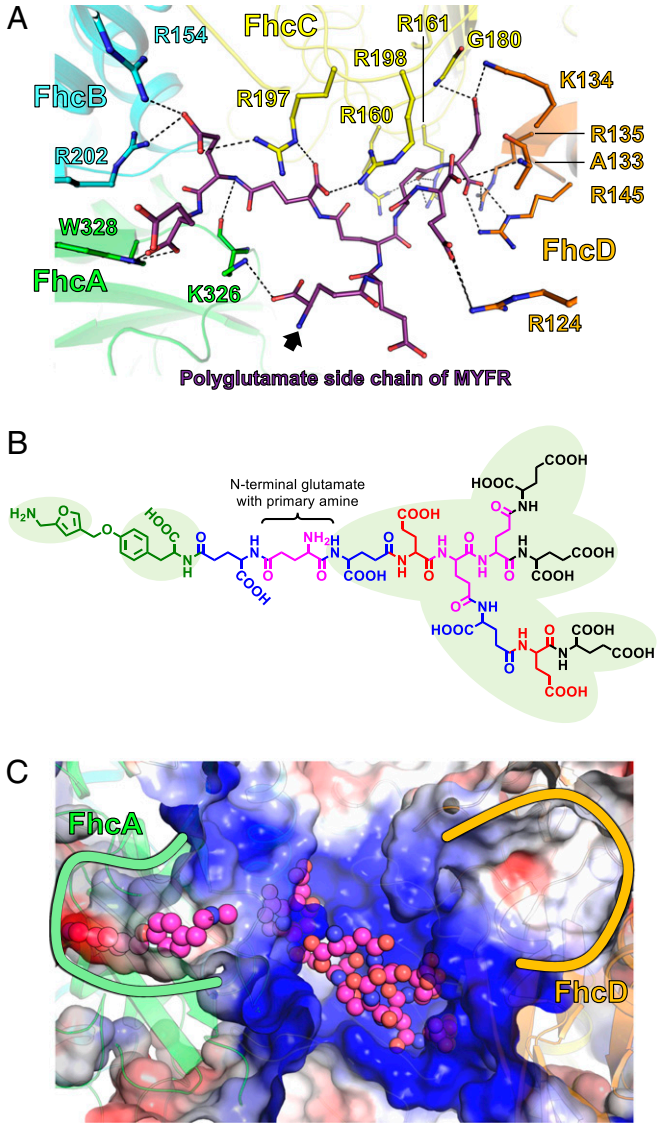

Fig. 4. Description of the MYFR binding site and the modeled MYFR structure. (*A*) Interactions between the branched polyglutamate side chain of MYFR and FhcABCD. The arrow indicates the amine where the carboxyl group of the additional two glutamates connecting the polyglutamate chain to the core structure is assumed to be attached. As no electron density was observed for these glutamates, they could not be modeled. (*B*) Potential structure of MYFR, which matches the observed electron density and considers the fact that the N-terminal primary amine of the polyglutamate is located at the second glutamate in the chain, as demonstrated by MS/MS (*SI Appendix*, Fig. S10). The core structure of MYFR is shown in green, while γ-linked, α-linked, branched, and C-terminal glutamates are shown in blue, red, pink, and black, respectively. For all green shaded parts of MYFR, electron density was observed (*SI Appendix*, Fig. S9). (*C*) Electrostatic surface potential of the MYFR binding site, color-coded from red for negative charge to blue for positive charge. The modeled parts of MYFR are represented as balls. Carbon, nitrogen, and oxygen atoms are colored pink, blue, and red, respectively. The active site cavities of FhcA and FhcD are highlighted by the green and yellow lines.

Additional evidence for an unusual arrangement of the glutamates in the chain was obtained when we tried to cleave MYFR using peptidases with specific activity against α- or γ-linked glutamates. The peptidases carboxypeptidase G, γ-glutamyltranspeptidase, carboxypeptidase Y, and endoproteinase Glu-C were tested for activity with MYFR. However, none of them were able to cleave MYFR (*SI Appendix*, Table S2). The inability of these enzymes to cleave MYFR, despite being active with synthetic linear peptides consisting of 20 α- or γ-linked L-glutamates connected with an L-tyrosine ($E_{20}Y$, $\gamma E_{20}Y$), further indicates a nonlinear polyglutamate chain, which might render MYFR inaccessible for peptidases.

Anchoring of MYFR to Fhc by an Extended Polyglutamate Side Chain. The binding site for the branched polyglutamate side chain of MYFR consists of more than 10 positively charged arginine and lysine residues (Fig. 4 *A* and *C*). These are mainly contributed by FhcC, but notably also by the other three subunits. The branched polyglutamate chain interacts with the protein core through electrostatic interactions between the carboxylates of the glutamates in MYFR and the guanidinium and ammonium groups of the arginines and lysines of Fhc, respectively. The many interacting amino acids allow for a large interaction area and the branched polyglutamate chain interlocks with the enzyme surface, thus explaining the strong binding of MYFR.

To determine whether an extended polyglutamate chain is required for binding to Fhc, we tested the binding ability of apo-Fhc toward archaeal MFR-a, which has two glutamates and a terminal tetracarboxylic acid residue attached to the core structure. Since Fhc purified from *M. extorquens* always had MYFR bound, apo-Fhc was obtained by expression of the *fhc* genes in *Escherichia coli* (*SI Appendix*, Fig. S11). To determine binding, a qualitative assay was used where Fhc was incubated with MFR-a (in 5-fold excess) and with a small amount of MYFR allowing for 10% saturation of the MYFR binding site. The added MYFR was used as an internal positive control, as it should tightly bind to apo-Fhc. After incubation, any unbound molecules were removed by ultrafiltration and excessive washing. LC-MS analysis revealed that the entire distribution of MYFR was binding to apo-Fhc as expected, while almost no binding was observed for MFR-a, even when provided in large excess (*SI Appendix*, Fig. S12). This result demonstrates that neither the short glutamate chain present in MFR-a nor its core structure are sufficient for significant binding to Fhc. Fhc therefore seems to specifically recognize the extended and branched polyglutamate side chain of MYFR and prevents the unspecific binding of smaller negatively charged compounds, such as MFR-a. With the terminal part of the polyglutamate side chain of MYFR being anchored to Fhc, the remaining glutamates might function as a flexible linker, allowing the formyl-carrying 2-aminomethylfuran moiety of MYFR to swing between the two active sites on FhcD and FhcA (Fig. 4*C*). The active sites are separated by ∼50 Å from each other and by ∼34 Å from the main anchoring point of the side chain of MYFR on FhcC and are therefore easily reachable by the furan core of MYFR.

Comparison of the Bacterial Formate Generation Complex with the Archaeal CO_2 Fixation System. The two multienzyme complexes Fhc and Fwd show high primary and tertiary structural similarity, in line with a proposed common evolutionary origin (3, 6, 7). However, the two systems show fundamental differences in function: While Fwd is a CO_2 reducing and fixation system, Fhc is a formate generation system (Fig. 1). By comparing the two structures, the difference in function can be explained at a molecular level.

A major difference between the two systems originates from the fact that FhcB lacks the molybdo(tungsto)pterin binding motif typical for formate dehydrogenases. The catalytic Cys118 from FwdB is replaced by a Ser62 in FhcB, and other catalytically important residues in FwdB (His119, Arg288) are also not conserved. Importantly also, the classic so-called fourth domain of formate dehydrogenase (24) and the DMSO reductase family (25), which is represented by FwdD in the Fwd complex, is absent in the structure of Fhc (*SI Appendix*, Fig. S6). This fourth domain is critical to cap the tungstopterin site and properly coordinate the cofactor. Moreover, the superposition of FhcB with

FwdB shows three loops (residues 93–97, 214–220, and 270–277) that are clashing with the tungstopterin from FwdB (*SI Appendix*, Fig. S6). In addition, the N-terminal domain of FwdB that harbors the [4Fe-4S] cluster to transfer electrons to/from the tungstopterin center is missing in FhcB. Together, these features clearly show the loss of the dehydrogenase function, as has been previously assumed from the primary structure (14). FhcB therefore mainly serves as a structural scaffold.

The structure of archaeal Fwd revealed the presence of three channels: a hydrophobic channel that connects the bulk solvent with the dehydrogenase active site for passing on CO_2, a hydrophilic channel for supplying protons, and another hydrophilic channel that connects the dehydrogenase with the active site of FwdA for tunneling of formate (21). We thus examined the Fhc structure for the presence of similar tunnels and identified only one extended channel in FhcAB (Fig. 5). This mostly hydrophilic channel has a length of 29 Å and connects the hydrolase active site to the bulk solvent, allowing the generated formate to leave the enzyme.

Another functional difference between the two enzymatic systems is the presence of the formyltransferase FhcD as a homodimer in the Fhc complex. FhcD is attached to the complex via mainly the C-terminal part of FhcC (*SI Appendix*, Fig. S7). Notably, FhcC has in its C-terminal part two successive α-helices separated by a loop (residues 216–246) in place of the β-hairpin present in FwdC (residues 229–241). This extension allows for additional intersubunit interactions between FhcC and FhcD and might be crucial for the association.

Apart from providing a binding site to FhcD, FhcC has yet another function that is absent in FwdC: It also serves as the main anchoring point for MYFR. The positively charged residues involved in MYFR binding seem to be conserved in bacteria, while only a few of the residues are positively charged in archaea (*SI Appendix*, Fig. S13). This is in agreement with the fact that MFR does not associate with FwdC and has to diffuse between the two enzymes Fwd and Ftr.

Discussion

The interconversion of carrier-bound one-carbon units catalyzed by Fhc is a critical step in formaldehyde metabolism in methylotrophic bacteria. While the hydrolysis of the formyl group to formate is exergonic ($\Delta G^{\circ\prime} = -22.9$ kJ/mol, ref. 28), the formyltransfer from formyl-H_4MPT to MYFR is slightly endergonic ($\Delta G^{\circ\prime} = +4.4$ kJ/mol, ref. 29). One way to turn endergonic reactions favorable is to maintain low product concentrations, which is applied in enzymatic pathways through substrate channeling between active sites (30, 31). Similarly, the fusion of formyltransferase (FhcD) and formylhydrolase (FhcA) as part of a single-enzyme complex might be advantageous for thermodynamic reasons. Here, we show that the shuttling of formyl units between the two active sites is further optimized through the permanent anchoring of MYFR to Fhc via its branched polyglutamate side chain. We propose that MYFR uses the flexibility of the polyglutamate chain to swing between the two active sites. The exergonic hydrolysis of the formyl group to formate by FhcA ensures low local concentrations of formyl-MYFR and drives the overall reaction. Additionally, the observed negligible fraction of unbound MYFR in combination with the tight binding to Fhc might minimize the need for MYFR synthesis.

Based on these results, we propose a sequence of coordinated reactions for the activity of Fhc (Fig. 6). In a first step, formyl-H_4MPT binds to the active site on the formyltransferase FhcD, followed by the swinging of MYFR to the same active site. After transfer of the formyl group, the formylated aminomethylfuran moiety of formyl-MYFR changes to the second active site at the formylhydrolase FhcA and H_4MPT is released. Subsequently, the formyl group is hydrolyzed and formate leaves the enzyme through the formate channel (Fig. 5).

The use of permanently bound coenzyme linkers that swing between active sites in multifunctional enzymes has precedent (32). For example, in the case of pyruvate dehydrogenase, biotin-dependent carboxylases, and acyl carrier proteins (ACPs), the swinging arm is a covalently attached lipoic acid, biotin, or phosphopantetheine, respectively (32). Fhc is similar to these systems with respect to intersubunit transfer of coenzyme-bound intermediates; it is however an example of an enzyme where the swinging arm is noncovalently attached, yet still stays permanently bound during the reaction sequence.

Modeling of the MYFR structure showed that in principle 4 to 5 linearly connected glutamates would suffice to cover the distance from the MYFR binding site to the two active sites. A few additional glutamates at branching points are required to provide the anchoring interface to Fhc. There is, however, still a considerable number of glutamates left (Fhc-bound MYFR has up to 20 glutamates) for which we do not know where they might be located in the polyglutamate chain. Considering that there is an optimal length to cover the distance to the active sites, it seems reasonable that most of the remaining glutamates would be found at branching sites, while maintaining the length of the main chain constant.

The presence of a branched polyglutamate chain in a biological compound has to our knowledge not been reported so far. The short polyglutamates found in MFRs (11, 13), coenzyme

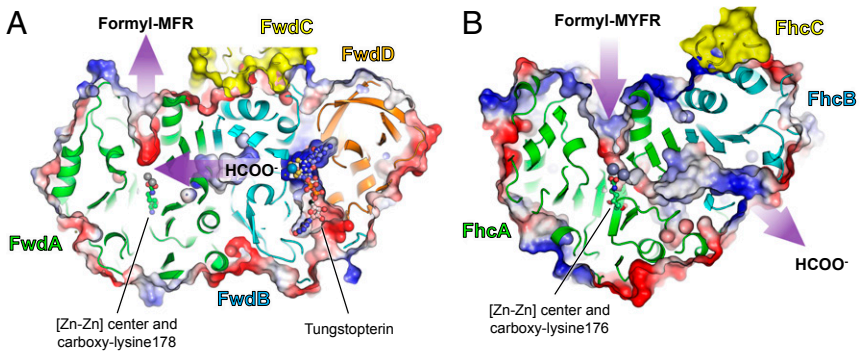

Fig. 5. Internal formate channel in the Fwd and Fhc complex. Both representations show a cutoff where the subunits are drawn as cartoon and surface. The surface corresponds to the electrostatic surface potential, from red for negative charge to blue for positive charge. The formyl-M(Y)FR cavity and the formate channel are highlighted by purple arrows. (*A*) In FwdABCD from *M. wolfeii* (PDB ID code 5T5I), formate is generated at the tungstopterin center contained in FwdBD and then transferred to FwdA by an internal cavity. Formyl-MFR is produced at the [Zn-Zn] center and released. (*B*) In FhcABC, the produced formate can leave through a channel similar to the one in FwdAB; however, here it is exposed to the solvent.

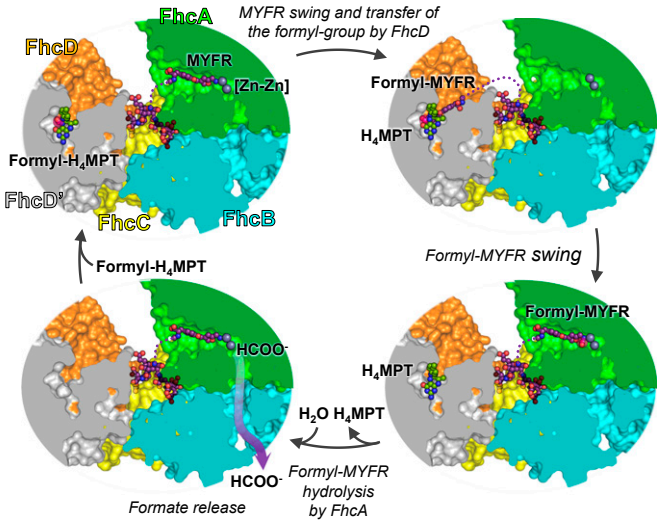

Fig. 6. Proposed reaction cycle for the generation of formate from formyl-H_4MPT, involving a swinging motion of MYFR to transfer formyl units between the active sites. MYFR and the pterin of H_4MPT shown in the active site of FhcD were modeled based on PDB ID code 2FHJ.

F_{420} (33) or tetrahydrofolate (34), and the high molecular weight poly-(γ-glutamic acid) produced by some bacteria (35) have all been described as linear chains. Consequently, no enzyme catalyzing the production of branched polyglutamates is known yet. Thus, the identification and characterization of the glutamate ligase(s) involved in MYFR biosynthesis will be of interest and might explain not only the regulation of the number of glutamates, but also the branching pattern found in the polyglutamate side chain of MYFR.

Our structural data elucidated structure/function relationships of Fhc at the molecular level. They further revealed potential evolutionary trajectories how nature transformed a probably most ancient CO_2 fixation machinery (7, 36) into a formate production system via three steps: 1) loss of the dehydrogenase activity through the loss of iron-sulfur cluster- and tungstopterin-binding residues and means to release formate from the enzyme, 2) association of the formyltransferase with FhcC, and 3) evolution of a MFR binding site together with the elongation and branching of the polyglutamate side chain of MFR, resulting in MYFR.

Materials and Methods

Cloning of *fhcABCD* for Overexpression in *M. extorquens*. The four genes *fhcBADC* (Mext_1824–Mext_1827), including the region 23 bp upstream of *fhcB* (containing the native ribosomal binding site), were amplified from genomic DNA of *M. extorquens* PA1 by PCR and subcloned into the plasmid pUCBM21 (Boehringer Mannheim) by restriction enzyme cloning. Finally, the genes were cloned together with a Strep-tag II (37) C-terminally on *fhcC* (added amino acid sequence of Strep-tag: HGSAWSHPQFEK) between the HindIII and BamHI restriction sites into the plasmid pCM80 (38). This plasmid harbors the constitutive *mxaF* promoter and a tetracycline resistance gene. After verification by sequencing, the final plasmid was transformed into electrocompetent *M. extorquens* PA1 cells.

Small-Scale Fhc Production in Culture Flasks. For activity assays, small-scale production of Fhc was performed in baffled flasks filled with minimal medium (39) with 1% methanol as the carbon source and 10 μg/mL tetracycline. The medium was inoculated with a preculture to OD_{600} of about 0.1 and was incubated at 28 °C. The cells were harvested by centrifugation at OD_{600} of 5–10.

Large-Scale Fhc Production in a Bioreactor. For protein crystallization, large-scale production of Fhc was performed in a 3.6-L Labfors bioreactor (Infors AG). The reactor was filled with 1.4 L of a minimal medium as described before (5) (containing increased amounts of trace elements, iron, $CaCl_2$, $(NH_4)_2SO_4$ and $MgSO_4$ compared to the medium used for small-scale expression, see ref. 5). Methanol was used as the carbon source and the medium was supplemented with 10 μg/mL tetracycline. The reactor was stirred with 1,500 rpm, aerated with 5 L/min of air, and the pH was kept constant at 7.0 through the addition of 2 M ammonia. The temperature was set to 28 °C, and methanol was constantly fed to the culture. Antifoam C (Sigma-Aldrich) was added to prevent foam formation. Once the culture reached an OD_{600} of 80–100, the cells were harvested by centrifugation, while leaving a small amount of the culture in the reactor as inoculum for the next batch. The reactor was then immediately refilled with 1 L of fresh medium and the cultivation continued. This process was repeated two more times, resulting in a combined biomass of about 800 g of cell wet weight.

Purification of Fhc. Cell pellets were resuspended in phosphate-buffered saline (PBS; 20 mM phosphate, 280 mM NaCl, 6 mM KCl, pH 7.3) and lyzed by passing four times through a French press cell. For large-scale purifications, the lysate was additionally sonicated to decrease its viscosity before it was ultracentrifuged. For affinity purification, the clarified lysate was loaded on a StrepTrap HP column (1 or 5 mL, GE Healthcare) equilibrated with PBS buffer. Elution was performed with a gradient from 0 to 10% of PBS buffer containing 2.5 mM desthiobiotin. Fractions containing Fhc were pooled and diluted with 1 volume of a Tris buffer (25 mM Tris, 6 mM KCl, 2 mM dithiothreitol, pH 7.7) to lower the salt content. The sample was then loaded on a HiTrap Q HP anion exchange column (1 mL, GE Healthcare) equilibrated with Tris buffer containing 100 mM NaCl. Elution was performed using a gradient from 0 to 40% of Tris buffer containing 1 M NaCl. Fractions containing Fhc were pooled and further purified using size exclusion chromatography with a Superose 6 Increase column (24 mL, GE Healthcare) and a buffer consisting of 25 mM Tris, 200 mM NaCl, 6 mM KCl, 2 mM dithiothreitol, and 10% glycerol at pH 7.7. Fractions containing pure Fhc were pooled and concentrated in a centrifugal filter with a molecular mass cutoff of 10 kDa (Amicon Ultra, Merck Millipore).

Detection of MYFR/MFR-a by LC-MS. MYFR and MFR-a species were detected using nanoscale ion-pair reversed-phase LC-MS (40). (Formyl-)MFR-a was measured on an LTQ Orbitrap XL mass spectrometer (Thermo Fischer Scientific) with an LC-MS method described before (5). (Formyl-)MYFR was measured on a Q Exactive Plus orbitrap mass spectrometer (Thermo Fischer Scientific) coupled to a nano-2D Ultra LC system (Eksigent Technologies). Chromatographic separation was performed using a C18 column (Reprosil-Gold 120 C18 3 μm, 0.1 × 100 mm, Dr. Maisch GmbH). Solvent A was 230 μM tributylamine, 230 μM acetic acid, and 3% methanol in water at pH 9.0 (adjusted with ammonia), and solvent B was 50:50 isopropanol:methanol. The following gradient was applied at a flow rate of 400 nL/min: 0 min, 0% B; 3 min, 0% B; 35 min, 16% B; 36 min, 90% B; 48 min, 90% B, 49 min, 0% B; 60 min, 0% B. Samples were dissolved in solvent A, and 1 μL was injected. For nanospray ionization, the column was connected to a silica tip emitter (PicoTip Emitter, tip: 10 μm, New Objective). The mass spectrometer was operated in negative mode with a spray voltage of 2.1 kV. Capillary temperature was set to 250 °C, the S-lens RF level to 50. Mass spectra were recorded as centroids with a resolution of 70,000 in the *m/z* range of 150–1,500 using four microscans.

Data analysis was performed using custom made scripts developed with eMZed (41). Briefly, mass traces of MYFRs with different number of glutamates were generated with 5 ppm tolerance for charge states with $1 \leq z \leq 8$ and for all natural isotopologs with at least 9% abundance. The mass traces were integrated over the elution time window and peak areas for the different charge states and isotopologs were summarized for each MYFR species. MFR-a was analyzed similarly, but only charge states up to $z = 3$ and no isotopologs were considered.

Activity Assays. The formyltransferase was assayed similarly to a method described before (14) by indirectly measuring H_4MPT formation via its spontaneous reduction of cytochrome *c*. The assay mixture contained 50 mM tricine, 600 mM NaCl, 20 μM methenyl-H_4MPT^+, 50 μM cytochrome *c*, 13 μg/mL methenyl-H_4MPT^+ cyclohydrolase (Mch) and was at pH 8.0. A total assay volume of 70 μL was used, and the temperature was set to 28 °C. The reaction was started through the addition of 2.5 μL of 1.2 mg/mL Fhc (affinity purified) and followed photometrically at 550 nm.

Mch (Mext_1831) was cloned from *M. extorquens* PA1 with a C-terminal *Strep*-tag II into pCM80, similarly as described for Fhc. Expression and affinity purification were performed in small-scale as described for Fhc.

Methenyl-H_4MPT$^+$ was purified from *Methanothermobacter marburgensis* (42) and methenyl-dephosphoH_4MPT was generated (43). LC-MS analysis showed that the sample was a ∼1:1 mixture of methenyl-dephosphoH_4MPT and methenyl-H_4MPT, both coenzymes work however as a substrate for Fhc.

The formylhydrolase was assayed using archaeal formyl-MFR-a as a substrate. Formyl-MFR-a was a gift from R. K. Thauer, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany. The assay mixture contained 100 μM formyl-MFR-a in PBS buffer (20 mM phosphate, 280 mM NaCl, 6 mM KCl, pH 7.3). The total assay volume was 20 μL, and the reaction was started through the addition of 1 μL of 0.7 mg/mL Fhc (affinity purified). Before the addition of Fhc and after 1, 2, 5, 10, 25, 40, and 60 min of incubation at 28 °C, 1-μL samples were taken and quenched with 9 μL of ice-cold quenching solution (60:20:20 acetonitrile:methanol:0.5 M formic acid). The samples were dried and measured by LC-MS for the detection of remaining substrate and the generated MFR-a.

The formylhydrolase assay in direction of formyl-MYFR formation from formate was performed by adding Fhc (size exclusion purified) with a final concentration of 0.13 mg/mL to a solution containing 0.45 M formic acid at pH 7.5 (adjusted with ammonia). The mixture was incubated at 28 °C and after 20 s, 40 s, 60 s, 120 s, 390 s, and 930 s, a sample was taken and quenched by adding 20 volumes of ice-cold quenching solution. The samples were dried and formyl-MYFR was measured by LC-MS. After 20 s, the reaction was already in equilibrium and, therefore, all of the time points were used to average the data. A negative control, performed analogously by mixing heat-denatured Fhc with the formic acid and incubating for 15 min, did not show any nonenzymatic formation of formyl-MYFR.

Determination of Protein-Bound MYFR in a Cell Lysate. A protein and a metabolite fraction was generated by separating an *M. extorquens* PA1 cell lysate (cultivated as described above, lysed using French press) using size exclusion chromatography with a Superose 12 column (GE Healthcare) and PBS (described above) as buffer. Standard solutions (blue dextran, ferritin, cytochrome *c*, aprotinin, and vitamin B_{12}) were used to determine the fractions that correspond to molecules larger (protein fraction) and smaller (metabolite fraction) than ∼10 kDa. Fractions were pooled and heat-denatured (99 °C, 20 min) to release bound MYFR. The samples were spiked with a custom-synthesized peptide (by Thermo Fisher Scientific) consisting of 20 γ-linked L-glutamic acids connected with an L-tyrosine ($\gamma E_{20}Y$) as an internal standard. This peptide behaves chemically very similar to MYFR (5) and allowed to control for matrix effects. For desalting and enrichment of MYFR, the samples were diluted with water and concentrated ∼10-fold with a centrifugal filter with a molecular mass cutoff of 1 kDa (Microsep Advance with Omega membrane, Pall Life Sciences). This step was repeated four more times. The retentate was dried and measured by LC-MS. MYFR peak areas were normalized to the peak area of the internal standard $\gamma E_{20}Y$.

Crystallization of Fhc. Crystallization screens were performed using the sitting-drop vapor diffusion method. A volume of 0.6 μL of protein solution (10 mg/mL, size exclusion purified in 25 mM Tris pH 7.7, 10% glycerol, 2 mM DTT, 6 mM KCl, and 200 mM NaCl) was mixed with an equal amount of well solution and incubated at 18 °C on 96-well 2-drop MRC Crystallization Plates (Molecular Dimensions). First crystals were obtained after a few days to a few weeks. The most promising condition contained 0.2 M $CaCl_2$ and 20% (wt/vol) polyethylene glycol 3350. The best diffracting crystals were obtained by reproducing this condition with 2 μL of protein and 1 μL of precipitant in a CombiClover Junior Plate (Jena Bioscience).

X-ray Data Collection and Analysis. The crystals of Fhc were flash-frozen after soaking (3–5 s) in a solution containing 0.2 M $CaCl_2$ and 20% (wt/vol) polyethylene glycol 3350 and 20% (vol/vol) glycerol at 18 °C. First screenings were performed at the BM30A beamline (European Synchrotron Radiation Facility, ESRF), and the best dataset came from beamline PXII-X10SA at the Swiss Light Synchrotron (SLS) equipped with a PILATUS 6M detector. All X-ray diffraction measurements were performed at 100 K. The data were processed with XDS (44) and scaled with SCALA from the CCP4 suite (45). The structure was determined by molecular replacement with PHASER (46) by using the subunits A, B, and C of the tungsten formylmethanofuran dehydrogenase from *M. wolfeii* (PDB ID code 5T5I) in combination with the formylmethanofuran:tetrahydromethanopterin formyltransferase from *M. kandleri* (PDB ID code 2FHJ) as template. The initial model was first refined by the LORESTR (47) pipeline from the CCP4 suite and further built with the BUCCANEER program (48). The model was corrected and manually built with COOT (49) and refined with BUSTER (50). The final refinement cycles were performed with PHENIX (51) by using noncrystallographic symmetry and translation/libration/screw. The final models were validated by using the MolProbity server (52). Data collection, refinement statistics, and PDB ID code for the deposited model are listed in *SI Appendix*, Table S1. The figures were generated and rendered with PyMOL (version 1.7, Schrödinger, LLC).

LC-MS/MS Analysis of *N*-Methylated MYFR. A pure sample of MYFR was obtained by extraction from purified Fhc. For removal of buffer and salts, Fhc was first washed several times with water in a centrifugal filter with a molecular mass cutoff of 10 kDa (Amicon Ultra, Merck Millipore). To release MYFR, Fhc was then heat-denatured (95 °C, 10 min).

N-methylation of the two primary amines in MYFR was performed using reductive methylation as described (53). For the reaction, 5 μL of MYFR, corresponding to the amount extracted from 11 μg of Fhc, were mixed with 35 μL of water, 5 μL of 300 mM borane pyridine complex (diluted with methanol from an 8 M solution, Sigma-Aldrich) and 5 μL of 200 mM formaldehyde (Sigma-Aldrich). Following brief sonication, the reaction was carried out for 2 h at room temperature. After drying, the methylated sample was measured by LC-MS/MS.

LC-MS measurements were performed similarly as described above. For MS/MS fragmentation, a parallel reaction monitoring experiment was added to the method. For the MS/MS scans, the resolution was set to 35′000, the isolation window to 2 *m/z*, the fixed first mass to 100 *m/z* and the normalized collision energy to 30. To increase signal intensity, multiple charge states of the MYFR with 12 glutamates were collected and fragmented simultaneously in the higher-energy collisional dissociation cell (multiplexing). The MSX count was set to 4, and an inclusion list containing the four precursor ions 947.10 ($z = 2$), 631.23 ($z = 3$), 473.17 ($z = 4$), and 378.33 ($z = 5$) was created. MS/MS was performed in the retention time window between 25 min and 36 min. The resulting spectra were analyzed using Xcalibur Qual Browser (Thermo Fischer Scientific).

Cloning and Expression of *fhcABCD* in *E. coli*. The four *fhc* genes (including the C-terminal Strep-tag on *fhcC*) were amplified from the plasmid used for the expression in *M. extorquens* (see above) by PCR. The following restriction sites were added during PCR and used for sequential restriction enzyme cloning into the vector pCDFDuet-1 (Merck Millipore): NdeI+KpnI for *fhcD*, NcoI+BamHI for *fhcB*, XhoI+PacI for *fhcC*, and SacI+HindIII for *fhcA*. Ribosomal binding sites (RBSs) of the sequence AAGAAGGAGATATACC were additionally added upstream of *fhcA* and *fhcC* during the PCR, as the vector only provided two multiple cloning sites that included an RBS. As the resulting plasmid did not lead to equal expression of all four subunits, the C-terminal Strep-tag was relocated from *fhcC* to *fhcA* by PCR and restriction enzyme cloning, as this seemed to be the least abundant subunit. This plasmid was then transformed into *E. coli* BL21 cells. Cells were cultured in lysogeny broth medium at 37 °C until an OD_{600} of 0.8 was reached, thereafter protein expression was induced by the addition of 0.1 mM IPTG. After induction, the cells were transferred to 28 °C and harvested after 20 h. Cell lysis and affinity purification of Fhc were performed as described above.

Determination of Ligand Binding to Apo-Fhc. To determine the binding ability of MFR-a to apo-Fhc produced in *E. coli*, 1 μL of affinity-purified apo-Fhc (3 mg/mL, 19 μM) was mixed with 3 μL of Tris buffer (20 mM Tris, 150 mM NaCl, pH 7.6) and 5 μL of 20 μM formyl-MFR-a (5x excess over Fhc). As an internal positive control, 1 μL of MYFR, corresponding to the amount extracted (by heat-denaturation) from 0.3 μg of size-exclusion purified Fhc from *M. extorquens*, was added. The added MYFR should allow for 10% saturation of the MYFR binding site on Fhc. The mixture was then incubated at 28 °C for 1 h. Afterward, the samples were diluted to 500 μL with Tris buffer, and unbound molecules were removed by concentrating the samples ∼20-fold with a centrifugal filter with a molecular mass cutoff of 30 kDa (Amicon Ultra 0.5 mL, Merck Millipore) at 28 °C. This process of dilution and concentration was repeated four more times, in order to make sure that all unbound ligands were completely removed. The Fhc sample was then heat-denatured (99 °C, 20 min) to release bound ligands and measured by LC-MS on the LTQ Orbitrap XL. The LC method was the same as described above for the detection of MYFR. Extracted ion chromatograms were generated in Xcalibur Qual Browser (Thermo Fischer Scientific). MYFR distributions were analyzed with eMZed as described above using 10 ppm mass tolerance.

ACKNOWLEDGMENTS. We thank Dr. Ulrich Ermler (Max Planck Institute for Biophysics, Frankfurt) for providing us synchrotron time at the SLS for the X-ray diffraction experiments and the staff of the BM30A beamline at the ESRF and of the PXII at the SLS for their advice during data collection. This work was supported by Swiss National Science Foundation Grant 31003A-173094 (to J.A.V.) and by ETH Zurich.

1. R. K. Thauer, J. Kunow, "Sulfate-reducing archaea" in *Sulfate-Reducing Bacteria*, L. L. Barton, Ed. (Springer, Boston, MA, 1995), pp. 33–48.
2. R. K. Thauer, Biochemistry of methanogenesis: A tribute to Marjory Stephenson. 1998 Marjory Stephenson prize lecture. *Microbiology* **144**, 2377–2406 (1998).
3. L. Chistoserdova, J. A. Vorholt, R. K. Thauer, M. E. Lidstrom, C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. *Science* **281**, 99–102 (1998).
4. L. Chistoserdova, J. A. Vorholt, M. E. Lidstrom, A genomic view of methane oxidation by aerobic bacteria and anaerobic archaea. *Genome Biol.* **6**, 208 (2005).
5. J. L. Hemmann *et al*., The one-carbon carrier methylofuran from *Methylobacterium extorquens* AM1 contains a large number of α- and γ-linked glutamic acid residues. *J. Biol. Chem.* **291**, 9042–9051 (2016).
6. L. Chistoserdova, 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 (2016).
7. M. C. Weiss *et al*., The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1**, 16116 (2016).
8. W. Nitschke, M. J. Russell, Beating the acetyl coenzyme A-pathway to the origin of life. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120258 (2013).
9. P. N. Green, J. K. Ardley, Review of the genus *Methylobacterium* and closely related organisms: A proposal that some *Methylobacterium* species be reclassified into a new genus, *Methylorubrum* gen. nov. *Int. J. Syst. Evol. Microbiol.* **68**, 2727–2748 (2018).
10. J. A. Leigh, K. L. Rinehart, R. S. Wolfe, Structure of methanofuran, the carbon dioxide reduction factor of *Methanobacterium thermoautotrophicum*. *J. Am. Chem. Soc.* **106**, 3636–3640 (1984).
11. T. A. Bobik, M. I. Donnelly, K. L. Rinehart, Jr, R. S. Wolfe, Structure of a methanofuran derivative found in cell extracts of Methanosarcina barkeri. *Arch. Biochem. Biophys.* **254**, 430–436 (1987).
12. R. H. White, Structural diversity among methanofurans from different methanogenic bacteria. *J. Bacteriol.* **170**, 4594–4597 (1988).
13. K. D. Allen, R. H. White, Identification of structurally diverse methanofuran coenzymes in *methanococcales* that are both N-formylated and N-acetylated. *Biochemistry* **53**, 6199–6210 (2014).
14. B. K. Pomper, J. A. Vorholt, Characterization of the formyltransferase from *Methylobacterium extorquens* AM1. *Eur. J. Biochem.* **268**, 4769–4775 (2001).
15. B. K. Pomper, O. Saurel, A. Milon, J. A. Vorholt, Generation of formate by the formyltransferase/hydrolase complex (Fhc) from *Methylobacterium extorquens* AM1. *FEBS Lett.* **523**, 133–137 (2002).
16. L. Chistoserdova, M. Laukel, J.-C. Portais, J. A. Vorholt, M. E. Lidstrom, Multiple formate dehydrogenase enzymes in the facultative methylotroph *Methylobacterium extorquens* AM1 are dispensable for growth on methanol. *J. Bacteriol.* **186**, 22–28 (2004).
17. L. Chistoserdova *et al*., Identification of a fourth formate dehydrogenase in *Methylobacterium extorquens* AM1 and confirmation of the essential role of formate oxidation in methylotrophy. *J. Bacteriol.* **189**, 9076–9081 (2007).
18. T. Wagner, U. Ermler, S. Shima, "Formyl-methanofuran dehydrogenase" in *Encyclopedia of Inorganic and Bioinorganic Chemistry*, R. A. Scott, Ed. (John Wiley & Sons, Ltd., 2018), pp. 1–18.
19. M. I. Donnelly, R. S. Wolfe, The role of formylmethanofuran: Tetrahydromethanopterin formyltransferase in methanogenesis from carbon dioxide. *J. Biol. Chem.* **261**, 16653–16659 (1986).
20. J. L. Hemmann, T. Wagner, S. Shima, J. A. Vorholt, X-ray crystal structure of the formyltransferase/hydrolase complex from *Methylorubrum extorquens* in complex with methylofuran. Protein Data Bank. https://www.rcsb.org/structure/6S6Y. Deposited 4 July 2019.
21. T. Wagner, U. Ermler, S. Shima, The methanogenic CO_2 reducing-and-fixing enzyme is bifunctional and contains 46 [4Fe-4S] clusters. *Science* **354**, 114–117 (2016).
22. P. Acharya, E. Warkentin, U. Ermler, R. K. Thauer, S. Shima, The structure of formylmethanofuran: Tetrahydromethanopterin formyltransferase in complex with its coenzymes. *J. Mol. Biol.* **357**, 870–879 (2006).
23. C. M. Seibert, F. M. Raushel, Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry* **44**, 6383–6391 (2005).
24. J. C. Boyington, V. N. Gladyshev, S. V. Khangulov, T. C. Stadtman, P. D. Sun, Crystal structure of formate dehydrogenase H: Catalysis involving Mo, molybdopterin, selenocysteine, and an Fe_4S_4 cluster. *Science* **275**, 1305–1308 (1997).
25. H. Schindelin, C. Kisker, J. Hilton, K. V. Rajagopalan, D. C. Rees, Crystal structure of DMSO reductase: Redox-linked changes in molybdopterin coordination. *Science* **272**, 1615–1621 (1996).
26. S. Najmudin *et al*., Periplasmic nitrate reductase revisited: A sulfur atom completes the sixth coordination of the catalytic molybdenum. *J. Biol. Inorg. Chem.* **13**, 737–753 (2008).
27. C. Binda *et al*., Cross-talk and ammonia channeling between active centers in the unexpected domain arrangement of glutamate synthase. *Structure* **8**, 1299–1308 (2000).
28. A. Flamholz, E. Noor, A. Bar-Even, R. Milo, eQuilibrator–The biochemical thermodynamics calculator. *Nucleic Acids Res.* **40**, D770–D775 (2012).
29. R. K. Thauer, R. Hedderich, R. Fischer, "Reactions and enzymes involved in methanogenesis from CO_2 and H_2" in *Methanogenesis*, J. G. Ferry, Ed. (Springer, Boston, MA, 1993), pp. 209–252.
30. X. Huang, H. M. Holden, F. M. Raushel, Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annu. Rev. Biochem.* **70**, 149–180 (2001).
31. I. Wheeldon *et al*., Substrate channelling as an approach to cascade reactions. *Nat. Chem.* **8**, 299–309 (2016).
32. R. N. Perham, Swinging arms and swinging domains in multifunctional enzymes: Catalytic machines for multistep reactions. *Annu. Rev. Biochem.* **69**, 961–1004 (2000).
33. H. Li, M. Graupner, H. Xu, R. H. White, CofE catalyzes the addition of two glutamates to F_{420}-0 in F_{420} coenzyme biosynthesis in *Methanococcus jannaschii*. *Biochemistry* **42**, 9771–9778 (2003).
34. B. Shane, "Folylpolyglutamate synthesis and role in the regulation of one-carbon metabolism" in *Vitamins & Hormones*, G. D. Aurbach, D. B. McCormick, Eds. (Academic Press, 1989), vol. 45, pp. 263–335.
35. I.-L. Shih, Y.-T. Van, The production of poly-(γ-glutamic acid) from microorganisms and its various applications. *Bioresour. Technol.* **79**, 207–225 (2001).
36. P. S. Adam, G. Borrel, S. Gribaldo, An archaeal origin of the Wood-Ljungdahl H_4MPT branch and the emergence of bacterial methylotrophy. *Nat. Microbiol.* (2019)., in press.
37. A. Skerra, T. G. M. Schmidt, "Use of the Strep-tag and streptavidin for detection and purification of recombinant proteins" in *Methods in Enzymology*, J. Thorner, S. D. Emr, J. N. Abelson, Eds. (Academic Press, 2000), vol. 326, pp. 271–304.
38. C. J. Marx, M. E. Lidstrom, Development of improved versatile broad-host-range vectors for use in methylotrophs and other Gram-negative bacteria. *Microbiology* **147**, 2065–2075 (2001).
39. A. M. Ochsner *et al*., Transposon sequencing uncovers an essential regulatory function of phosphoribulokinase for methylotrophy. *Curr. Biol.* **27**, 2579–2588.e6 (2017).
40. P. Kiefer, N. Delmotte, J. A. Vorholt, Nanoscale ion-pair reversed-phase HPLC-MS for sensitive metabolome analysis. *Anal. Chem.* **83**, 850–855 (2011).
41. P. Kiefer, U. Schmitt, J. A. Vorholt, eMZed: An open source framework in Python for rapid and interactive development of LC/MS data analysis workflows. *Bioinformatics* **29**, 963–964 (2013).
42. S. Shima, R. K. Thauer, "Tetrahydromethanopterin-specific enzymes from Methanopyrus kandleri" in *Methods in Enzymology*, R. M. Kelly, M. W. W. Adams, Eds. (Academic Press, 2001), vol. 331, pp. 317–353.
43. M. I. Donnelly, J. C. Escalante-Semerena, K. L. Rinehart, Jr, R. S. Wolfe, Methenyltetrahydromethanopterin cyclohydrolase in cell extracts of *Methanobacterium*. *Arch. Biochem. Biophys.* **242**, 430–439 (1985).
44. W. Kabsch. *XDS*. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).
45. M. D. Winn *et al*., Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
46. A. J. McCoy *et al*., Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
47. O. Kovalevskiy, R. A. Nicholls, G. N. Murshudov, Automated refinement of macromolecular structures at low resolution using prior information. *Acta Crystallogr. D Struct. Biol.* **72**, 1149–1161 (2016).
48. K. Cowtan, The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002–1011 (2006).
49. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
50. G. Bricogne *et al*., BUSTER (Version 2.10.1, Global Phasing Ltd., Cambridge, UK, 2016).
51. P. D. Adams *et al*., PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
52. V. B. Chen *et al*., MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
53. C. J. Krusemark, J. T. Ferguson, C. D. Wenger, N. L. Kelleher, P. J. Belshaw, Global amine and acid functional group modification of proteins. *Anal. Chem.* **80**, 713–720 (2008).