Expanding the possibilities of crystallomics towards environmental samples

Dissertation zur Erlangung des

Doktorgrades der Naturwissenschaften

Dr. rer. nat.

Dem Fachbereich 2 Biologie/Chemie vorgelegt von:

Marie-Caroline Müller

Bremen, Dezember 2024

Universität Bremen

To understand proteins is to understand life.

Die vorliegende Doktorarbeit wurde in der Zeit von Juni 2020 bis November 2024 im Rahmen der International Max Planck Research School of Marine Microbiology (IMPRS MarMic) in der Arbeitsgruppe Mikrobielle Metabolismen am Max-Planck-Institut für Marine Mikrobiologie unter der Aufsicht von Dr. Tristan Wagner angefertigt.

This doctoral thesis was prepared from June 2020 to November 2024 as part of the International Max Planck Research School of Marine Microbiology (IMPRS MarMic) in the Microbial Metabolism research group at the Max Planck Institute for Marine Microbiology under the supervision of Dr. Tristan Wagner.

Gutachter: Dr. Tristan Wagner

Gutachter: Prof. Dr. Michael Friedrich

Gutachter: Dr. Guillaume Borrel

Prüfungsvorsitzender: Prof. Dr. Tilmann Harder

Prüfer: Dr. Tristan Wagner

Prüfer: Prof. Dr. Michael Friedrich

Prüfer: Dr. Guillaume Borrel

Datum des Promotionskolloquiums: 22.01.25

Summary

The majority of microbes cannot be isolated, representing the "microbial dark matter" that harbors a wealth of undiscovered enzymes inaccessible to traditional microbiological techniques. Studying these enzymes could reveal novel pathways, reactions, and mechanisms, while insights at the molecular level could be harnessed for biotechnological applications and pave the way for understanding the unexplored microbial world. **This thesis aimed** to optimize and apply "crystallomics" – a native shotgun protein purification and crystallization approach – to samples containing a heterogeneous microbial community. Methanogenic and methanotrophic archaea were chosen as biological models as they have a global impact on the climate, harbor enzymes with major biotechnological potential, and are less well-studied than bacterial systems. The approach was first tested on pure cultures. **Chapter II** describes the blind purification of multiple enzymes from the hyperthermophile methanogen *Methanocaldococcus infernus* and their *de novo* structural elucidation. In **Chapter III**, native purification was performed to study two archaeal glutamine synthetases that exhibit different regulation mechanisms: an allosteric activation by 2-oxoglutarate and a feedback inhibition by glutamine. **Chapter IV** describes the native isolation of the methyl-coenzyme M reductase (MCR) from the methoxydotrophic methanogen *Methermicoccus shengliensis*. The work describes an MCR containing the lowest number of post-translational modifications (PTMs) described to date. The native purification approach was then applied to more complex microbial samples. In **Chapter V**, MCR was purified from marine and freshwater methanotrophic enrichment cultures. The structures obtained at (sub)atomic resolution revealed an unexpectedly high content of PTMs. Furthermore, a novel modification on a different subunit could be detected. The technique was further challenged by a sample with a very high microbial diversity: activated sludge from a wastewater treatment plant. **Chapter VI** describes the purification of two MCRs from *Methanothrix* species using a physical enrichment protocol that could be successfully reproduced with a purification performed directly

on the sludge sample. **Chapter VII** gives an outlook about other proteins that could potentially be captured utilizing the shotgun approach on the heterogeneous activated sludge, including promising results regarding the acetate processing machinery of *Methanothrix*. **In summary**, the presented results highlight the strengths and weaknesses of the approach. Ultimately, this work supports the feasibility of direct protein purification from heterogeneous samples as a step toward environmental structural biology.

Zusammenfassung

Bis heute bleibt der Großteil aller Mikroben unkultiviert. Diese so genannte "mikrobielle dunkle Materie" birgt eine Fülle unentdeckter Enzyme die für herkömmliche mikrobiologische Techniken unzugänglich sind. Molekulare Erkenntnisse über Stoffwechselwege, enzymatische Reaktionen und Mechanismen könnten biotechnologische Anwendungen inspirieren und zum Verständnis der unerforschten mikrobiellen Welt beitragen. Das **Ziel** dieser Arbeit war es, die Anwendung von „Crystallomics“ - ein nativer ungezielter Ansatz zur Proteinaufreinigung und Kristallisierung - an Proben mit heterogener mikrobieller Zusammensetzung zu testen. Methanogene und methanotrophe Archaeen wurden aufgrund ihres globalen Klimaeinflusses, ihres bio-technologischen Potenzials und ihrer geringeren Erforschung im Vergleich zu Bakterien als biologisches System ausgewählt. Der Ansatz wurde zunächst an Reinkulturen getestet. **In Kapitel II** wird die blinde Aufreinigung mehrerer Enzyme aus dem hyperthermophilen Methanogen *Methanocaldococcus infernus* beschrieben, gefolgt von drei verschiedenen Phasenbestimmungsansätzen zur *de novo* Strukturaufklärung. In **Kapitel III** wurde ein natives Protokoll zur Untersuchung von zwei archaealen Glutaminsynthetasen verwendet die auf verschiedene Weisen reguliert werden: allosterische Aktivierung mit 2-Oxoglutarat und Glutaminrückkopplungs-hemmung. **Kapitel IV** beschreibt die native Isolierung der Methyl-Coenzym-M-Reduktase (MCR) aus dem methoxydotrophen Methanogen *Methermicoccus shengliensis.* Das Enzym wies die geringste Anzahl an posttranslationalen Modifikationen (PTM) unter beschriebenen Homologen auf. Der Ansatz wurde dann auf komplexere mikrobielle Proben angewendet. In **Kapitel V** wurde MCR aus marinen und Süßwasser-Anreicherungskulturen aufgereinigt. Strukturen mit (sub)atomarer Auflösung zeigten einen unerwartet hohen Anteil an PTMs und eine neuartige Modifikation an einer unüblichen Untereinheit. Schließlich wurde die Technik auf eine Probe mit sehr hoher mikrobieller Vielfalt ausgeweitet: Klärschlamm. **Kapitel VI** beschreibt die Isolierung

von zwei *Methanothrix* MCRs anhand eines physikalischen Trennungsprotokolls und direkt aus der Probe. **Kapitel VII** gibt einen Ausblick auf andere Proteine, die mit Hilfe des ungezielten Ansatzes aus Klärschlamm gewonnen werden könnten, einschließlich vielversprechender Ergebnisse bezüglich der Acetat-Aktivierungsmaschinerie in *Methanothrix.* **Zusammenfassend** zeigen die vorgestellten Ergebnisse die Stärken und Schwächen des Ansatzes. Letztendlich bekräftigt diese Arbeit die Durchführbarkeit der direkten Proteinaufreinigung aus komplexen mikrobiellen Proben als Schritt in Richtung Umweltstrukturbiologie.

List of original publications

Kurth, J.M.*, Müller, M.-C.*, Welte, C.U., Wagner, T. Structural insights into the methane-generating enzyme from a methoxydotrophic methanogen reveal a restrained gallery of post-translational modifications. *Microorganisms* **2021**, *9*, 837. Doi:10.3390/microorganisms9040837 (Open Access, CC BY 4.0)

Müller, M.-C., Lemaire, O.N., Kurth, J.M., Welte, C., Wagner, T. Differences in regulation mechanisms of glutamine synthetases from methanogenic archaea unveiled by structural investigations. *Commun. Biol.* **2024**, 7, 111. Doi:10.1038/s42003-023-05726-w (Open Access, CC BY 4.0)

Müller, M.-C., Wissink, M., Mukherjee, P., von Possel, N., Laso-Pérez, R., Engilberge, S., Carpentier, P., Kahnt, J., Wegener, G., Welte, C., Wagner, T. Atomic resolution structures of methane-fixing enzymes from marine and freshwater ANME-2 depict a conserved organisation with numerous post-translational modifications. *In preparation.*

Müller*, M.-C., Jespersen*, M., Maslać, N., D'Angelo, G., Lonsing, A., Dunemann, S., Huettel, B., Engilberge, S., Liebeke, M., Harder, J., Wagner, T. Native isolation of the CH_4-generating enzyme from a mesocosm – steps towards microbial environmental structural biology. *In preparation.*

*Authors have contributed equally

TABLE OF CONTENTS

Summary ..... **I**

Zusammenfassung ..... **III**

List of original publications ..... **V**

Abbreviations ..... **XI**

CHAPTER I. Introduction **1**

Enzymes: structure, function and biotechnological applications 1

A brief history of protein purification and crystallization 2

The microbial dark matter 7

Recombinant protein expression 10

Protein structure prediction 11

The concept of crystallomics 13

Aim of the thesis: Expanding crystallomics to environmental samples 15

Archaea and their global impact 17

Methanogens and methanogenesis 20

Methanothrix 23

Alkane oxidizing archaea 27

The methyl-coenzyme M reductase 29

Energy saving strategies of methanogens 31

Specific aims and presented chapters 34

References 36

CHAPTER II. Crystallomics on *Methanocaldococcus infernus* **51**

Main 51

The 2-oxoglutarate:ferredoxin oxidoreductase of *Methanocaldococcus infernus* 58

Supplementary material 64

References 73

CHAPTER III. Differences in regulation mechanisms of glutamine synthetases from methanogenic archaea unveiled by structural investigations **77**

Abstract 78
Introduction 79
Results 82
*Mt*GS activity is strictly dependent on 2OG in contrast to *Ms*GS. 82
Structural snapshots of archaeal GS 84
A 2OG allosteric site localized at the inter-subunit junction 87
2OG-dependent motions and active site remodeling in *Mt*GS. 89
*Ms*GS is feedback inhibited by glutamine 94
Direct regulation by 2OG and glutamine within the domain *Archaea* 94
Discussion 99
Material and Methods 102
Acknowledgments, author contributions and data availability 112
Supplementary information 114
Supplementary data 133
References 134
CHAPTER IV. Structural insights into the methane-generating enzyme from a methoxydotrophic methanogen reveal a restrained gallery of post-translational modifications 139
Abstract 140
Introduction 141
Results 143
Purification and crystallization of *Ms*MCR obtained under methoxydotrophic methanogenesis 143
A conserved overall structure and active site 145
The smallest post-translational modification gallery observed in methanogens 147
Discussion 149
Material and Methods 152
Acknowledgments, author contributions, data availability and funding 160
Supplementary materials 161

References..... 167

CHAPTER V. Atomic resolution structures of methane-fixing enzymes from marine and freshwater ANME-2 depict a conserved organization with numerous post-translational modifications. 173

Abstract..... 174

Introduction..... 175

Results 178

Isolation of the core catabolic enzyme from marine and freshwater ANME-2..... 178

MCR identification and a well-conserved overall architecture 180

A sealed, strictly conserved catalytic chamber covered by seven post-translational modifications..... 183

Discussion..... 188

Material and Methods..... 192

Acknowledgements and author contributions 200

Supplementary materials..... 201

References..... 219

CHAPTER VI. Native isolation of the CH_4-generating enzyme from a mesocosm – steps towards microbial environmental structural biology..... 223

Abstract..... 224

Introduction..... 225

Results 227

Enrichment of a *Methanothrix* population from active sludge..... 227

Isolation and identification of the main catabolic enzyme of *Methanothrix* from the physical enrichment..... 229

An overlay of two MCRs from *Methanothrix* arranged as a shamrock in the crystal. 233

A conserved organization with an active site equipped by five post-translational modifications..... 235

MCR isolation directly from active sludge collected months later..... 236

Discussion..... 237

Material and Methods 241
Acknowledgements, author contributions and data availability 252
Extended data 254
Supplementary materials 264
References 278
CHAPTER VII. Crystallomics on activated sludge from wastewater treatment 283
Main 283
Supplementary material 293
References 307
CHAPTER VIII. Discussion 309
Scientific insights gained from the native approach 309
Under control – studying different GS regulation systems 309
Snapshots of archaeal C-assimilation enzymes 311
MCRs from species utilizing three different methane metabolisms 312
Considerations for the application of crystallomics 317
The effects of increasing sample heterogeneity 322
Complementary techniques 325
Reproducibility and strategies for specific targets 327
Transferability to other microbial samples 330
Studying new biocatalysts from extreme environments 332
Conclusion and outlook 334
References 336
Acknowledgements 343
Appendix 344
Versicherung an Eides Statt 344
Erklärungen zur elektronischen Version und zur Überprüfung einer Dissertation 345
Declaration on the contribution of the candidate for the presented chapters 346

For this work, AI-assisted technology has been utilized for grammar checking and translations. No intellectual content or continuous passages have been AI-generated.

Abbreviations

(reoccurring in multiple chapters)

2OG	2-oxoglutarate, α-ketoglutarate
ACN	Acetonitrile
ANME	Anaerobic methanotrophic archaea
AOM	Anaerobic oxidation of methane
BLAST	Basic Local Alignment Search Tool
CoB-SH	Coenzyme B, 7-mercaptoheptanoylthreonine phosphate
CODH/ACS	CO-dehydrogenase/acetyl-CoA decarbonylase
CoM-S-S-CoB	Heterodisulfide (of CoM and CoB)
cryoEM	Cryo-electron microscopy
DIET	Direct interspecies electron transfer
DTT	Dithiothreitol
ECR	Ethyl-coenzyme M reductase
FISH	Fluorescence in situ hybridization
GOGAT	Glutamate synthase
GS	Glutamine synthetase
GTDB	Genome Taxonomy Database
hrCN PAGE	High-resolution clear native polyacrylamide gel electrophoresis
HS-CoM	Coenzyme M, 2-mercaptoethylsulfonate
MAG	Metagenome assembled genome
MCR	Methyl-coenzyme M reductase
MDM	Microbial dark matter
microED	Microcrystal electron diffraction
MS	Mass spectrometry
OGOR	2-oxoglutarate:ferredoxin oxidoreductase
PDB	Protein Data Bank
PSM	Peptide spectral match
PTM	Post-translational modification
rmsd	Root mean square deviation
RT	Room temperature
SAD	Single-wavelength anomalous diffraction
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SRB	Sulfate reducing bacteria
TPP	Thiamine pyrophosphate
UV/Vis	Ultraviolet–visible

CHAPTER I. Introduction

This introduction will establish the characteristics of enzymes and their wide range of applications, briefly outline the history of crystallography, address the challenge of studying enzymes from uncultured microorganisms of the microbial dark matter and touch on how scientist have attempted to overcome this issue. The concept of crystallomics will be introduced with the proposal to apply this technique to heterogeneous samples. Finally, it will be explained why archaea were chosen as the focus of this thesis, which includes experiments performed on methanogenic and methanotrophic archaea.

Enzymes: structure, function and biotechnological applications

Proteins are essential building blocks of Life, involved in all biological processes. They have various different functions like contributing to the cellular structure[1-3], regulating cellular processes[4-6] and facilitating the transport of molecules across the cell membrane[7]. Enzymes are proteins specialized to catalyze metabolic reactions[8-10]. The function of an enzyme is directly related to its structure. The folding of the amino acid backbone results in a three-dimensional structure, often containing discrete functional regions (domains). In some cases, multiple polypeptides (subunits) will assemble into a functional complex. Catalysis is performed at the active site(s), formed by folds/pockets lined with chemically reactive groups. Often non-protein chemical groups are incorporated to facilitate catalysis by acting as redox centers such as transition metals (*e.g.*, Fe, Cu, Ni, Co) or cofactors (*e.g.*, flavins, hemes). In addition, individual amino acids can harbor post-translational modifications that affect their reactivity. Catalysis involves the binding of a substrate molecule to the active site through specific interactions. Subsequently, structural adjustments occur and different conformational states can be adopted before, after and throughout catalysis. By lowering the activation

energy, enzymes significantly accelerate the reaction rate compared to the spontaneous reaction. The diversity of reactions carried out by enzymes is astounding[11, 12].

The study of enzymes is of great importance for the scientific community and society in general. Mechanistic insights allow researchers to decipher reactions and metabolic networks ultimately contributing to understanding Life at the cellular level. Furthermore, natural features like enzyme structures, functions and mechanisms serve as inspiration for the development of new technologies. Today, enzymes are regularly employed as biocatalysts in various industries, as they are highly efficient, selective and operate under relatively moderate conditions opposed to purely chemical approaches. Their use reduces costs and putative toxic waste products and increases production rates as well as sustainability[13-16]. Biocatalysts are employed in a variety of applications like the degradation of stains (detergent industry)[17], improving paper brightness (paper industry)[18], generating lactose-free dairy products (food industry)[19], softening fabrics (textile industry)[20], degrading plant material for bioethanol production[21], and synthesis of antimicrobials and chiral intermediates (pharma industry)[22]. Furthermore, enzymes can be applied as therapeutic agents for the direct treatment of medical conditions, studies on enzyme therapy for infectious diseases and cancer are ongoing[23]. In contrast, enzymes associated with diseases can be targeted by drug discovery, an approach to design inhibitory molecules utilizing the enzyme structure [24].

As all these approaches rely on functional understanding of enzymes, structural studies are of high importance. The foundation for structural biology was laid when scientists began to crystallize proteins, and X-ray crystallography remains the most widely utilized approach for structure determination to date.

A brief history of protein purification and crystallization

Proteins have the ability to self-assemble into crystals. Under suitable conditions, supersaturated protein molecules will form stable nuclei which serve as the starting

point of crystal growth. Proteins will form non-covalent interactions (crystalline contacts), depending on their inherent structural properties. This process selectively incorporates a single molecular species (*e.g.*, single proteins, protein-protein complexes or nucleic acid-protein complexes), while excluding other molecules with features that favor different interaction patterns. Crystals are therefore a highly pure and concentrated form of a protein. The crystalline contacts are rather weak and the average solvent content of 50 % causes protein crystals to be softer compared to their mineral counterparts. Crystals themselves consist of a repetition of building blocks (unit cell) that generate a crystal lattice through translational symmetry. The asymmetric unit is the smallest part of the crystal structure and contains all unique structural information, consisting of *e.g.*, a single polypeptide, multiple protein copies or a larger protein complex. The application of distinct symmetry operators (defined by 65 chiral space groups) results in the formation of the unit cell. Different assemblies of the same protein can generate distinct crystalline forms[25].

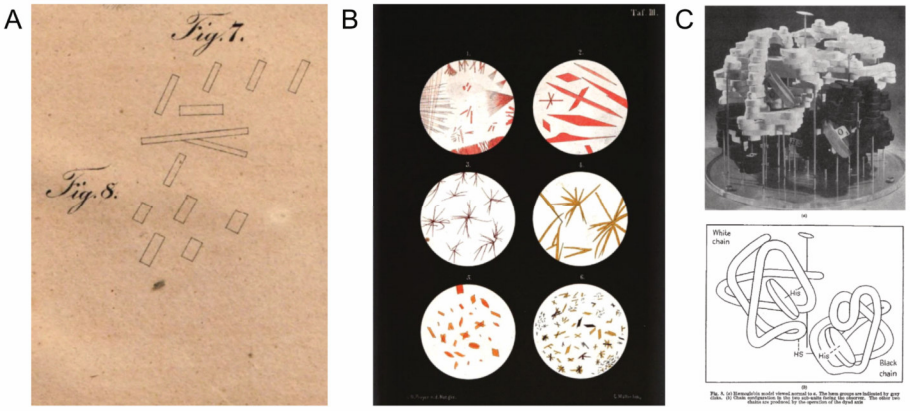


Figure 1. **First steps of crystallography.** (**A**) First report of protein crystals by Hünefeld (1840)[26]. (**B**) First published photographs of protein crystals (hemoglobin) by Preyer (1871)[27]. (**C**) First protein structure (myoglobin) by Perutz *et al*. (1960)[28]. Physical model on top and schematic drawing on the bottom. © *Reproduced with permission from Springer Nature*. *License Nr. 5912381495557*.

The first report of protein crystals in the literature is more than 180 years old. While working on the exploration of blood from different organisms, Friedrich Hünefeld noticed the formation of well-defined red crystals in a sample of earthworm blood that dried out due to fractures in the sealing varnish (Fig. 1A)[26]. Through mere coincidence, he demonstrated the formation of crystals through evaporation and the possibility of crystallizing proteins from an impure preparation. His findings were followed by a period of extensive studies on blood samples from various species by numerous scientists[29] (Fig. 1B), including the first deliberate crystallization of hemoglobin by Funke in 1851[30]. Crystallization was soon understood to be a useful technique for obtaining a pure preparation, and the process was applied to other samples, such as globulins from plant seeds[31, 32]. Studies on the specifics of crystallization began by investigating the effects of variables like salts, differing solvents, pH, temperature and even metal ions[29, 33, 34]. In 1917, James Sumner attempted, for the first time, to crystallize an enzyme in order to prove that enzymes are proteins, a controversial question at the time. As a target, he chose urease from the jack bean, and it took him nine years to obtain crystals. However, even though urease activity was detected for the crystals[35] his proposal received criticism from the community. After further studies it could finally be confirmed that the catalytic activity of enzymes lies within the protein[36] and Sumner, together with colleagues J. Northrop and W. Stanley, received the Nobel prize in 1946 for the crystallization of pure proteins and viruses. During this period, the field of X-ray crystallography emerged[37], starting with the discovery that crystals diffract X-rays[38] (1912), to the first diffraction pattern from a protein crystal[39] (1934) culminating in the first published protein structures awarded with a Nobel prize (1958)[28, 40] (Fig. 1C). The underlying principle for structure determination is that photons are scattered by the atoms electrons. The resulting diffraction pattern, representing a two-dimensional view of the crystal lattice, can be captured by a detector. Data from different angles is combined to generate a three-dimensional electron density map in which the protein

structure can be modelled. The novel technique soon led to the solving of more protein structures and the need for better documentation, as well as organized scientific exchange arose. In 1971, the Protein Data Bank (PDB) was founded as the first open access digital data resource in the biological sciences, storing protein crystallographic data (https://www.rcsb.org)[41].

In the following decades, the technique has undergone drastic improvements. Screening crystallization conditions (mostly defined by precipitant, salt , pH, temperature) got easier with the development of commercially available crystallization screens[42, 43]. Weaker laboratory-based X-ray vacuum tubes have been replaced by synchrotron radiation sources that have the advantage of a wide energy range, high photon flux and good beam coherence and are available at national facilities staffed with experts[44]. The tedious construction of physical models has been made obsolete by modern computing softwares. So, while the first structures took decades to solve, today, the process has been shortened to minutes from diffraction to 3D structure[45]. Ongoing upgrades of synchrotrons and X-ray detectors are improving the quality and quantity of datasets[46-48]. Nowadays, X-ray crystallography is more accessible than ever for multidisciplinary researchers[45], still being the dominating technique in structural biology. In fact, 83 % of all experimental structures deposited on the PDB have been determined via X-ray crystallography[49] and resolutions as high as 0.48 Å have been reached[50, 51] allowing individual atoms, even hydrogens, to be distinguished, revealing the most intricate features of the studied enzymes. However, the technique also has some drawbacks: Radiation damage can decrease the data quality, often high-quality crystals are required, and the data represents only a static snapshot. Additionally, structural artefacts can result from interactions in the crystalline packing.

Alternatively, new methods have emerged, expanding the field and complementing existing techniques. Microcrystal electron diffraction (microED) strongly lowers the

barrier for crystal size as sub-micron samples can be exploited[52]. Furthermore, dynamic processes can be studied by serial femtosecond crystallography (SFX), especially in combination with X-ray free electron lasers (XFEL). High-intensity femtosecond pulses are used to collect diffraction before radiation damage occurs[53]. This requires the use of multiple microcrystals and, therefore, allows the start of enzymatic reactions shortly before the precisely timed data collection. Results can resemble a series of stop-motion images of a protein performing its function[54]. Furthermore, crystal-independent techniques have been developed. Nuclear magnetic resonance (NMR) spectroscopy is an established approach utilizing the response of nuclei with non-zero spins (*e.g.*, ^{1}H, ^{15}N, ^{13}C), placed in a magnetic field and excited by radiofrequency irradiation[55, 56]. This allows the study of protein dynamics like conformational changes, ligand binding and protein interactions[56]. Six percent of structures on the PDB have been determined with NMR. However, the technique is limited by the molecular weight of the studied protein (< 50 kDa)[57] and requires high stability in solution[58], which has led to a stagnation in released structures per year[59]. Single-particle cryo-electron microscopy (cryoEM) can be carried out directly on purified proteins[60]. Freezing the sample at an ultra-low temperature increases the resistance against higher electron doses. 2D images of individual particles can be collected at different orientations and averaged to generate a 3D map. The technique is especially useful for investigating larger complexes or membrane proteins that usually evade crystallization[61-63]. Available X-ray or NMR substructures can be fitted into the cryoEM map to enhance the level of detail in larger molecular structures. While the majority of structures are obtained at 3-4 Å resolution, the development of direct electron detectors has allowed to overcome previous technical limitations, enabling single-particle cryoEM at near-atomic resolution[64]. Recent studies have been pushing the resolution limit, presenting datasets with resolutions as high as 1.22 Å[65, 66].

The microbial dark matter

Initially, the majority of scientific studies about microbes and their proteins relied on cultivated organisms. However, when the first genomic studies of environmental samples were performed[67] it was revealed that there is a large unknown diversity of microbes. A majority of current Life's diversity is made up of archaeal and bacterial phyla without cultivated representatives[68] (Fig. 2), a total of 85-99 % of microbes are estimated to remain uncultured[69]. Therefore, our knowledge about the microbial world is strongly limited and skewed by the few cultivated organisms[70]. The conglomerate of uncultured microbes is also referred to as microbial dark matter (MDM)[71-73]. While they comprise a large fraction of Earth's biodiversity, their basic metabolic and ecological properties are not known yet[74].

Various studies have revealed the MDM to be a treasure trove for novel biological insights and new findings have significantly impacted our understanding of the microbial world[68, 75]. For example, the discovery of *Candidatus* Lokiarchaeota and their genomic investigation have allowed scientists to place core eukaryotic genes within the archaeal domain, which has challenged the three-domain theory[76, 77]. This sparked high interest in the field and inspired scientists to dedicate a decade-long effort to isolating these microbes, resulting in the development of the entangle-engulf-endogenize model for eukaryogenesis[78]. Another instance with great impact was when Hug *et al.*[68] reconstructed a tree of life utilizing genomes from public databases and newly constructed genomes and noticed a major cluster of uncultivated bacterial lineages. The members of this Candidate Phyla Radiation (CPR) share relatively small genomes and reduced metabolisms and show an enormous extent of evolution making them a very interesting target for studying evolutionary history in bacteria[68, 79].

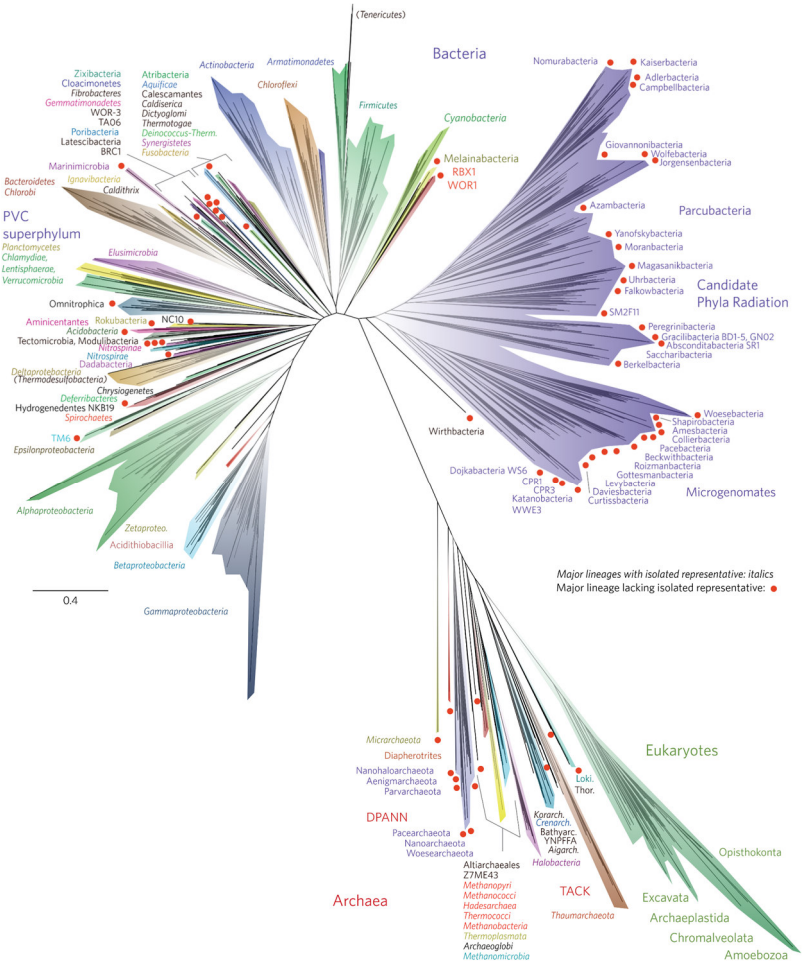


Figure 2. Microbial taxa without cultivated members. Figure from Hug *et al.*2016[68]. The tree includes 92 named bacterial phyla, 26 archaeal phyla, and all five eukaryotic supergroups. Major lineages are named and assigned arbitrary colors; well-characterized names are in italics. Lineages without isolated representatives are marked with non-italicized names and red dots. For methods on taxon sampling and tree generation, see the original publication. *© 2016 Macmillan Publishers Limited. Licensed under a Creative Commons Attribution 4.0 International License* (http://creativecommons.org/licenses/by/4.0/).

Besides challenging previous taxonomy and revealing new general concepts, studies on the MDM also reveal enzyme functions or metabolic pathways in new lineages, reshaping our understanding of their phylogenetic distribution. Genomic tools can provide insights to inspire hypotheses, which may then be experimentally validated. For example, CO_2 fixation through the nucleoside pathway was previously regarded as exclusive to archaea, until genes encoding RuBisCo forms II and III were detected in bacteria of the CPR[80]. In another study, a new aminopeptidase with no close homologs among cultured organisms was identified in *Thaumarchaeota archaeon* SCGC AB-539-E09, fueling theories about the role of archaea in protein remineralization in anoxic marine environments[81, 82]. Moreover, metagenome assembled genomes (MAGs) from *Ca.* Bathyarchaeota were found to contain genes coding for divergent homologs of enzymes involved in the methane metabolism, challenging the previously held belief that methanogenesis is exclusive to *Euryarchaeota* and raising the question of to what extent non-euryarchaeal lineages contribute to the global methane cycle[83]. Similarly, genes involved in both anaerobic methane and dissimilatory sulfate metabolisms were found in a MAG belonging to the deep-rooted archaeal phylum *Ca.* Korarchaeota suggesting for the first time a potential coupling within a single organism[84].

The isolation of organisms from the MDM is highly difficult and it is suspected that cultivation often does not succeed because these microbes engage in metabolite exchanges with other species in the environment and resulting co-dependencies are difficult to emulate under laboratory conditions[85]. Some success has been made through the establishment of enrichment cultures[86-91], allowing the cultivation of organisms in the presence of microbial partners. However, these approaches often require painstaking optimization processes that can take decades[78, 92]. The study of MDM members is therefore mostly carried out through metagenomics and single-cell genomics - powerful techniques that shed light on new microbial groups, their environmental distribution, and their genomic potential. Moreover, modern techniques like metatranscriptomics[93],

metaproteomics[94], and metabolomics[95] provide information on gene expression, protein abundance, and metabolite concentrations and identities, respectively. However, omics-based hypotheses about enzyme functions are limited by functional annotations. These are derived from similarity to studied enzymes from cultivated organisms, which biases the annotation process and potentially overlooks homologs that are too divergent. Strikingly, more than 50 % of genes from MDM genomes can still not be assigned to any function[96]. So, while these methods help answering the questions: "Who are they? Where are they? What can they do?" [96], the study of enzymes is still lacking and the question, "How are they doing it?" remains unanswered.

Recombinant protein expression

While sequence analysis can fuel hypotheses, characteristics that define enzymatic features like catalysis, specificity, turnover, kinetics, protein stability, and more are out of reach. Therefore, recombinant expression has been instrumental for the study of enzymes from MDM members. The gene of interest is introduced into genetically tractable host organisms that can be cultivated under laboratory conditions for protein expression on demand. This efficient approach has greatly contributed to the study of many proteins from uncultivated microorganisms. However, some drawbacks exist[97]. The number of genetically accessible host systems is limited. Expression can be lowered due to low mRNA stability[98] or rare codon usage[99]. If proteins are produced, many factors can influence their solubility and functionality, and often, misfolding[100] or the generation of insoluble inclusion bodies occurs[101]. High expression levels during overexpression might overload the cellular machinery, exceeding chaperone availability and increased amounts of incorrectly folded proteins will activate protein quality control systems initiating protein degradation[97]. The lack of protein partners or specific cofactors can prevent to reach thermodynamic stability resulting in the lack of physiological enzyme activity[97]. Furthermore, correct folding is a prerequisite for installing post-translational modifications (PTMs), which play an important role in

structural changes and functionality of the enzyme[102]. Multiple techniques exist to optimize recombinant expression[97]: Codon optimization facilitates synthesis, aggregated proteins can be unfolded and refolded, correct folding can be supported by changing expression conditions like temperature and inducer concentration as well as adding fusion proteins or chaperones. However, method optimization cannot yield a functional protein product in every case.

Protein structure prediction

An alternative way of studying protein structures has emerged in recent years. Already decades ago, biophysicists became interested in protein folding and ways to computationally derive a protein structure from the amino acid sequence[103, 104]. Since 1994 advancements in *ab initio* protein structure modeling have been presented every second year at the Critical Assessment of Structure Prediction (CASP) conference[105]. Participating teams are testing their algorithms by modeling structures that have been experimentally solved but not yet deposited in the PDB, making it a blind test. In 2018, Deep Mind entered the contest (CASP13) with their deep learning algorithm AlphaFold and achieved a jump in modeling accuracy[106]. However, a major turning point for the field was reached in the subsequent CASP14 as AlphaFold 2 showed major improvement in modeling accuracy scoring close to experimentally determined structures[107] (Fig. 3).

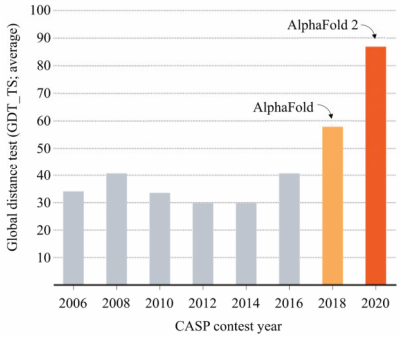


Figure 3. Highest modeling accuracy at CASP. Modified from Callaway *et al.* 2020**[107].** A score above 90 is considered equivalent to the experimentally determined structure. *© Reproduced with permission from Springer Nature. License Nr. 5912390359810*

AlphaFold 2 is trained on experimental models from the PDB but also on high-confidence models generated by itself[106, 108]. While previous modeling strategies have often focused on subsectioning the sequence and harnessing the physical concept of the native structure being the most thermodynamically stable with the lowest free energy[109], AlphaFold 2 approaches the issue differently. From the input sequence, a multiple sequence alignment (MSA) is generated to identify patterns in conserved regions. This approach is complemented by deep learning techniques to predict the pairwise distance between amino acids. The predicted model is folded as a whole, followed by iterative refinement steps, and finally multiple output models are ranked by confidence[106]. The model interpretation and evaluation are facilitated by the per-residue estimate of reliability in the form of predicted local-distance difference test (plDDT) values. Another noteworthy contributor to protein structure prediction is the RoseTTAFold deep learning-based framework (by the Baker lab), which is particularly good at providing accurate protein-protein complex predictions[110]. The development of accurate computational protein modeling has greatly impacted the field and was recently recognized with a Nobel prize[111]. Today, computed structure models have surpassed experimental structures on the PDB (227,344 experimental vs 1,068,577 computated[49]). Modeling algorithms are continuously advancing. AlphaFold 3 has improved its modeling accuracy further by including the prediction of complexes between proteins, nucleic acids, small molecules, ions and modified residues[112]. Great expectations and hopes lie in these methods, especially in the medical field of drug discovery and *de novo* protein design[109]. However, algorithms still struggle with certain features like allosteric mechanisms[108] or the depiction of flexibility through modeling active and inactive states[108, 112]. Experimentally determined structures remain essential for the training and evaluation of the algorithms. While it is possible to model previously unknown folds[106, 113] it would be near impossible to predict novel features like post-translational modifications or unknown cofactors solely through modeling.

The concept of crystallomics

In contrast to recombinant expression, native protein purification allows to study proteins in a nearly unbiased way, correctly folded and with all required partners, cofactors and modifications, able to adopt biologically relevant conformational or functional states. On the basis of native purification, the Microbial Metabolism lab developed "Crystallomics", a systematic "shotgun" approach to purifying and crystallizing the most abundant (often catabolic) proteins from a native organism (Fig. 4).

The first step of this procedure is cell lysis, either via sonication or, in the case of a tougher cell envelope, via French press. The total cell extract is centrifuged to separate the cell debris from the soluble protein extract. At this point, two optional steps can be included: Ultracentrifugation to isolate smaller membrane fragments and membrane proteins from the soluble extract for further experiments and ultrafiltration to harvest the metabolite fraction. Afterward, the soluble protein extract is subjected to multiple liquid chromatography steps in order to separate as many distinguishable protein fractions as feasible. Steps can include rough and fine fractionation

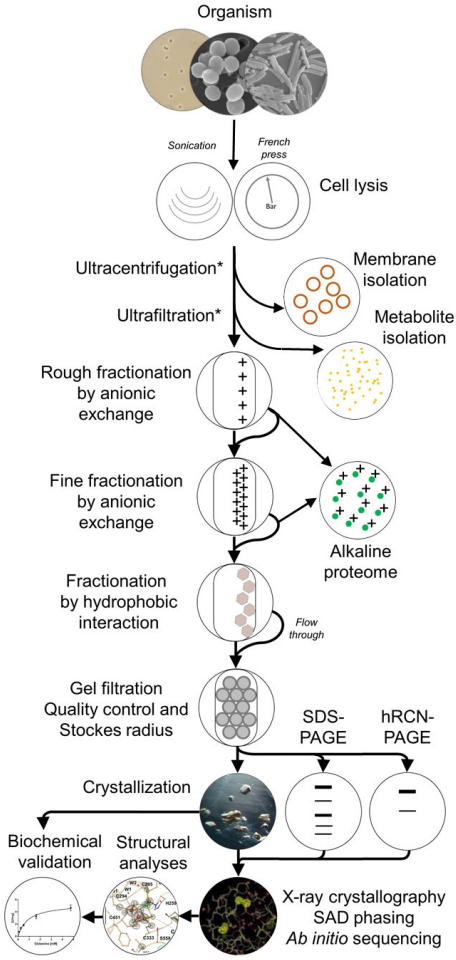


Figure 4. Crystallomics pipeline. Overview of the different steps for simultaneous native protein purification and subsequent analysis. Courtesy of Dr. Tristan Wagner.

via anionic exchange chromatography, followed by fractionation through hydrophobic interaction chromatography, and finally size exclusion chromatography. During all steps, the flow-through containing the unbound proteins (*e.g.*, Alkaline proteome) can be kept and used for further purification steps allowing the utilization of the entirety of the soluble extract. Throughout the purification, proteins are followed via spectrophotometry at multiple wavelengths: 280 nm for absorbance of aromatic amino acids, and other specific wavelengths corresponding to particular cofactors (*e.g.*, 420 and 550 nm). Additionally, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) provides information about the sample composition, including purity and size of single protein chains. After each chromatography step, the major protein populations are combined into separate pools, and each is passed on to the next column. Ideally, through

each chromatography step, the proteins can be separated more, ultimately yielding multiple fractions enriched in a single protein(complex). The final protein fractions are concentrated into amounts suitable for protein crystallization (4-40 $mg.mL^{-1}$) and are then crystallized using the sitting drop vapor diffusion method. Preferably, crystallization plates are prepared with a crystallization robot to increase volume accuracy, reproducibility and to allow spotting of smaller sample volumes therefore increasing the number of crystallization conditions to be tested from a limited amount of sample. Crystallization acts as the final purification step as the formation of crystal contacts select for a single molecular species while excluding others. Though it is possible, that different enzymes from a mix crystallize simultaneously into separate crystals.

The protein structures are solved by X-ray crystallography. For this, the crystals have to diffract X-rays to a sufficient resolution, and the phase must be solved. A related structure can be used for molecular replacement if the sequence of the crystallized

sample is known. However, as this approach aims to explore previously unstudied enzymes, it might well occur that there are no exploitable related structures or that the protein sequence is unknown. In these cases, experimental phasing can be performed through single-wavelength anomalous diffraction (SAD) utilizing anomalous scattering properties of some elements (see Ch. II) to determine the phase and to build a protein backbone[114]. This can then be combined with *ab initio* sequencing where the identity of the protein is determined by examination of the electron density of the side chains protruding from the protein backbone model if the resolution is below 3 Å.

The goal of crystallomics is not only to solve structures but to decipher enzyme functions and mechanisms. Obtained crystals can be used for a variety of analyses, providing information about the enzyme characteristics and reaction. Crystals can be dissolved to yield pure protein preparations. SDS-PAGE or high-resolution clear native (hrCN) PAGE can identify the composition and oligomerization state. Furthermore, biochemical validation through activity assays on dissolved crystals can reveal or confirm the enzyme function[115]. Ultraviolet-visible (UV/Vis)-spectra can be measured in solution and *in cristallo*[116] to detect the absorption of ligands or cofactors and to identify oxidation states of redox-active centers. This approach of native purification enables the study of enzymes in their most physiological state and uncover unique features that might not be captured by other techniques like -omics, computational modeling or recombinant expression.

Aim of the thesis: Expanding crystallomics to environmental samples

At this point, an obvious question arises: Doesn't native purification require lab grown cultures and is therefore unsuitable for studying proteins of MDM members? However, throughout the history of protein research, protein crystals were often obtained from relatively crude samples and although the field has shifted its focus towards pure cultures and high-grade purified samples, the success of these early, less refined

approaches should not be overlooked. The crystallomics pipeline is based on an untargeted approach of iteratively purifying abundant proteins from the sample, not necessarily limited to pure cultures. It may as well be applied to more heterogeneous samples or potentially even environmental samples. Crystallomics on uncultured microorganisms from natural environments could reveal novel metabolic pathways and enzymes.

Furthermore, the approach might contribute to the better understanding of microbial life in its predominant physiological state[117]. Cultivated microbes are growing under optimal conditions and nutrient supply and are mostly studied in the narrow window of exponential growth with maximum reproduction rates. In contrast microbes growing in environmental conditions are often adapted to nutrient and energy limitations and interact with the surrounding microbial community[118]. Phenotypic responses to energy limitation include the use of transport systems with altered substrate affinities, shifts of the energy metabolism and differential expression of catabolic and anabolic enzymes[117]. The resulting metabolic landscape of a "wild" microbe might be a more accurate reflection of the microbe's ecological niche and life cycle than one observed in its "domesticated" counterpart[118].

A proof of concept, for the study of an enzyme directly from an environmental community, has been shown by Shima *et al.* who have purified and crystallized the enzyme methyl-coenzyme M reductase (MCR) from Black Sea microbial mats containing up to 55 % ANME-1 (based on catalyzed reporter deposition fluorescence *in situ* hybridization (CARD-FISH))[119, 120]. Furthermore, the ethyl-coenzyme M reductase (ECR) was recently purified from an enrichment culture containing a consortium of *Ca.* Ethanoperedens thermophilum (50 % of 16S rRNA reads) and a sulfate-reducing partner bacterium[121, 122]. These impressive achievements have inspired this project, raising the question of whether this approach is reproducible on other

targets and to what extent sample heterogeneity can be increased. Therefore, the aim of this work is "Expanding the possibilities of crystallomics towards environmental samples" and the core concept that will be tested is the possibility of purifying and studying microbial proteins directly from mixed populations.

To apply and evaluate this method, a microbial target with global environmental and resulting high societal impact should be selected. Therefore, the presented studies focus on methanogenic and methanotrophic archaea for their major influence on the global carbon and nitrogen cycle.

Archaea and their global impact

Archaea are one of the most ancient lineages of Life on Earth with fossils that have been proposed to be ~3.5 billion years old[123]. Initially, archaea were grouped together with bacteria into the kingdom of prokaryotes[124]. At the time, taxonomy was based on morphological and physiological features. This changed when genomic tools became available. Universal small-subunit ribosomal RNA (rRNA) analysis led to the proposal of *Archaebacteria* as the third domain[125], comprising two main groups: the *Crenarchaeaota* - containing extremophilic organisms, and *Euryarchaeota* - composed of methanogens and some relatives[126]. This view was further expanded when environmental 16S rRNA amplicon sequencing and metagenomics revealed a previously unknown diversity and ubiquity of archaea around the globe[127]. In the past 40 years the archaeal tree of Life has been continuously expanded and revised. There are ongoing efforts to reach a clear and consistent taxonomic framework by merging traditional cultivation-based taxonomy with phylogenetics in the Genome Taxonomy Database (GTDB)[128].

There are currently 19 phyla and 5869 species of *Archaea* listed in the GTDB[129]. The phylum *Euryarchaeota* still harbors a majority of methanogens but has also been expanded by phylogenetically diverse new groups while other groups have been

recognized as their own phyla. Most cultured archaeal representatives are *Euryarchaeota*[127, 130]. Furthermore, three superphyla have been proposed[127, 130]: The TACK superphylum including *Ca.* Thaumarchaeota, *Ca.* Aigarchaeota, *Crenarchaeota*, *Ca.* Korarchaeota and more; the DPANN superphylum consisting of *Ca.* Diapherotrites, *Ca.* Parvarchaeota, *Ca.* Aenigmarchaeota, *Ca.* Nanohaloarchaeota and *Nanoarchaeota* that share small cell sizes, reduced genomes and likely a symbiotic lifestyle; the Asgard superphylum including *Ca.* Lokiarchaeota, *Ca.* Thorarchaeota and other phyla with some eukaryotic genomic features that have stimulated ongoing research on eukaryogenesis[78, 131].

Archaea are a diverse and abundant group of microorganisms constituting up to 20 % of prokaryotes in ocean waters[132] and dominating geothermal marine subsurface sediments[133]. They have a significant impact on the global nutrient cycle[134] (Fig. 5). The impact of archaea on the C-cycle is immense. It has been proposed that autotrophic *Ca.* Thaumarchaeota might contribute 1 % (0.6–0.7 Gt C) of the ocean's annual primary production[135]. Other groups live organotrophically, *e.g.*, *Ca.* Bathyarchaeota are highly abundant in anoxic marine and freshwater as well as hot springs and likely play a role in the degradation of detrital proteins[81]. Furthermore, marine groups II and III (*Ca.* Poseidoniales and *Ca.* Pontarchaea, *Euryarchaeota*) represent up to 40 % of the marine prokaryotic community and are able to recycle organic compounds[136]. The largest archaeal impact is on biogenic methane (CH_4) production by methanogens[137]. Methane is the second most important greenhouse gas[138], surpassing CO_2 in its potency with a global warming potential of 86 (for a period of 20 years)[139]. In recent years atmospheric methane growth rates have reached the highest level since the start of systematic measurements[140]. It is estimated that archaeal methanogenesis is responsible for the annual release of 1 Gt methane[137] and 74 % of net annual emissions to the atmosphere[141]. Two-thirds of methane is generated from acetate by acetoclastic methanogens[142], while the other third is generated from CO_2 reduction and only minor amounts through the

dismutation of methyl-compounds[134]. On the other hand, methanotrophic archaea control methane emissions by consuming >70 % of the methane produced in marine benthic systems before it enters the watercolumn[143]. Furthermore, they facilitate carbon precipitation, contributing to long-term carbon storage[134].

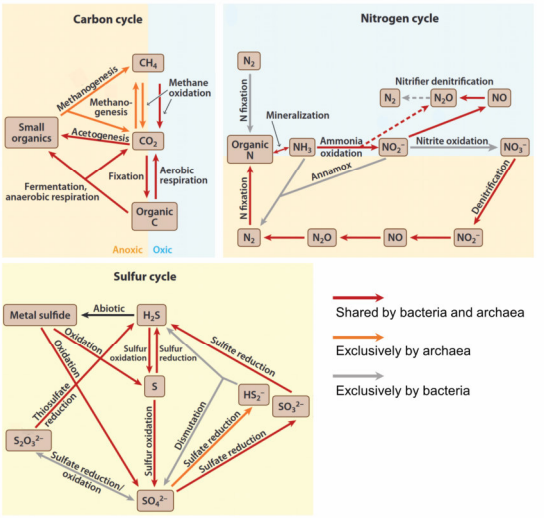


Figure 5. Archaea in biochemical cycles. Adapted from Offre *et al.* [134]. © *Reproduced with permission from Annual reviews. License Nr. 1547484-1*

In addition to the C-cycle, there are also large contributions to the N-cycle[134]. Most archaea obtain nitrogen by assimilating NH_3 or organic compounds while others are able to fix N_2. Especially NH_3 oxidation to NO_2^- (the first step of nitrification) is of greater importance as ammonia-oxidizing archaea outcompete ammonia-oxidizing bacteria in low substrate environments like ocean waters[144] and constitute large fractions of acidic soil communities[145]. Furthermore, *Ca.* Thaumarchaeota are among the most abundant archaea on the planet and account for most of the oceanic N_2O production[146].

The impact of archaeal processes on the S-cycle has also not yet been resolved, however it is known that representatives are able to reduce sulfur and sulfate[147, 148]. Furthermore,

some groups are able to oxidize sulfite minerals and H_2S [134]. The marine group II might be able to reduce dimethyl sulfoxide (DMSO) to dimethyl sulfide (DMS)[149], a compound with global climate relevance and an impact on cloud formation[150].

In summary, archaea are significant at the planetary level, with a particularly strong impact on methane emissions. Their extremophilic properties push the boundaries of Life and novel lineages and new metabolic pathways are continuously discovered unveiling unique enzymes with high biotechnological potential. Therefore, archaea are of significant scientific interest

Methanogens and methanogenesis

Methanogens are exclusive to *Archaea* and have previously been classified as six orders within the *Euryarchaeota*: *Methanobacteriales, Methanococcales, Methanopyrales, Methanosarcinales, Methanomicrobiales and Methanocellales*[151]. However, further studies revealed additional lineages like *Methanonatronarchaeales*[152] and *Methanomassiliicoccales*[153], and more recent metagenomics studies have described various MAGs containing methanogenesis related genes in non-euryarchaeal groups like *Ca.* Nezhaarchaeota[154], *Ca.* Korarchaeota[84] , *Ca.* Verstraetarchaeota[155] ,*Ca.* Thaumarchaeota[156], and *Ca.* Bathyarchaeota[157]. Recently a methyl reducing archaeon of the class *Methanosuratincolia* (*Thermoproteota*) has been isolated[158]. These findings expand our view about methane related metabolisms in archaea.

To date there are six known methane-involving pathways and two related processes that utilize multi-carbon alkanes found in archaea[151] (Table 1). While these pathways share a set of core enzymes, they differ in the substrate activation and energy conservation steps. Ultimately, all rely on generating a chemical H^+/Na^+ gradient across the membrane that is utilized for ATP synthesis via the ATP synthase.

Table 1. Methanogenesis pathways. In the case of various possible substrates an example reaction was chosen. The metabolic pathway for methanogenesis with methoxylated compounds is not fully resolved and therefore greyed out.

Metabolism		**Reaction**	**Standard Gibbs free energy change $\Delta G^{\circ\prime}$**
CO_2-reducing hydrogenotrophic methanogenesis		$4\ H_2 + CO_2 \rightarrow CH_4 + 2\ H_2O$	-131 kJ.mol^{-1} of CH_4 [160]
Acetoclastic methanogenesis		$CH_3COO^- + H^+ \rightarrow CH_4 + CO_2$	-36 kJ.mol^{-1} of CH_4 [160]
Methylotrophic methanogenesis	Methyl-reducing hydrogenotrophic methanogenesis	$CH_3OH + H_2 \rightarrow CH_4 + H_2O$	-112.5 kJ.mol^{-1} of CH_4 [160]
	Methyl-disproportionating methanogenesis	$4\ CH_3OH \rightarrow 3\ CH_4 + CO_2 + 2\ H_2O$	-106.5 kJ.mol^{-1} of CH_4 [160]
Methoxyl-disproportionating methanogenesis		4 2-methoxyphenol + 2 $H_2O \rightarrow 3\ CH_4$ + 4 2-hydroxyphenol + CO_2	-90 kJ.mol^{-1} of CH_4 [161]
Methanotrophy		$CH_4 + SO_4^{2-} + 2\ H^+ \rightarrow CO_2 + H_2S + 2\ H_2O$	-21 kJ.mol^{-1} of CH_4 [162]
Alkanotrophy		$4\ C_2H_6 + 7\ SO_4^{2-} + 14\ H^+ \rightarrow 8\ CO_2 + 7\ H_2S + 12\ H_2O$	-73.2 kJ.mol^{-1} of C_2H_6 [92]
Alkanotrophic methanogenesis		$4\ C_{16}H_{34} + 30\ H_2O \rightarrow 49\ CH_4 + 15\ CO_2$	-339.2 kJ.mol^{-1} of $C_{16}H_{34}$ [89]

Hydrogenotrophic methanogenesis is carried out through the following steps[159]: CO_2 is reduced and fixed as a formyl group to the C_1-carrier methanofuran (MFR) by the formylmethanofuran dehydrogenase (FWD/FMD). The formyl group is transferred to tetrahydromethanopterin (H_4MPT) by the formyltransferase (FTR) and subsequently dehydrated to a methenyl- H_4MPT by the methenyl-H_4MPT cyclohydrolase (MCH). Methenyl-H_4MPT is then successively reduced to methylene-H_4MPT and methyl-H_4MPT by the F_{420}-dependent or H_2-forming methylene-H_4MPT dehydrogenase

(MTD/HMD) and F_{420}-dependent methylene-H_4MPT reductase (MER), respectively. The methyl group is then transferred to coenzyme M (HS-CoM, 2-mercaptoethylsulfonate) by the methyl-H_4MPT: coenzyme M methyltransferase (MTR). This exergonic step is coupled to a Na^+-translocation across the membrane. Finally, methyl-CoM is consumed by the MCR using the coenzyme B (CoB-SH, 7-mercaptoheptanoylthreonine phosphate) releasing methane and forming a heterodisulfide of both coenzymes (CoM-S-S-CoB). CoB-S-S-CoM is reduced back to HS-CoM and CoB-SH by the soluble heterodisulfide reductase (HDR, composed of HdrA, HdrB, HdrC). This enzyme receives electrons from a donor (*e.g.*, H_2, $F_{420}H_2$, or formate) and employs an electron bifurcating mechanism that reduces the CoM-S-S-CoB and provides electrons of low redox potential to FMD/FWD either directly through physical contact or indirectly through *e.g.*, ferredoxin reduction.

Methylotrophic methanogenesis differs as the methyl group of the substrate (*e.g.*, methanol, methylamines) is first transferred to a corrinoid protein by a methyltransferase with high substrate specificity and then further to CoM by a second methyltransferase[151]. The methyl-CoM is then further processed by MCR to generate methane and CoM-S-S-CoB. In the methyl-reducing hydrogenotrophic methanogenesis, H_2 is utilized for CoM-S-S-CoB reduction with the concomitant ferredoxin reduction, allowing energy conservation. In the methyl disproportionating pathway, the reducing equivalents originate from the oxidation of the methyl-CoM to CO_2 through the reverse methanogenesis described above. The generated reducing equivalents allow for energy conservation by reducing the CoM-S-S-CoB. In total, this pathway generates three methanes for one CO_2 [151].

During acetoclastic methanogenesis, acetate is activated by the acetate kinase/phosphoacetyl-transferase system (AK/PTA). The CO-dehydrogenase/acetyl-CoA decarbonylase (CODH/ACS) reduces ferredoxin, releases CO_2 and transfers the

methyl group onto tetrahydrosarcinapterin H_4SPT (an H_4MPT analog). As presented in the hydrogenotrophic pathway, the methyl group is transferred to HS-CoM through MTR and released as methane by MCR. CoM-S-S-CoB is reduced by the proton translocating membrane-bound HDR (composed of HdrD, HdrE), utilizing the membrane-associated electron carrier methanophenazine (MPH). Reduction of MPH coupled to energy conservation is carried out by different systems like the energy-converting hydrogenase (ECH) and MPH-reducing hydrogenase (VHO) utilizing H_2, the Rnf-complex utilizing reduced ferredoxin or the F_{420}:phenazine oxidoreductase (FPO) utilizing $F_{420}H_2$ [163].

Methanothrix

There are only two genera of acetoclastic methanogens: *Methanothrix* (previously *Methanosaeta*[164]) and *Methanosarcina*[165]. While *Methanosarcina* is metabolically more flexible and able to utilize various substrates, *Methanothrix* specializes in utilizing acetate solely as carbon and energy source[165-167]. Furthermore, *Methanothrix* has a distinct morphology of elongated cells forming long filaments[168] of several hundred cells divided by septa and encased by a paracrystalline sheath[1].

The acetoclastic pathway in *Methanothrix* is similar to the well-studied *Methanosarcina*, but still, some key differences exist (Fig. 6). Instead of a two-step acetate activation by the acetate-kinase (converting ATP to ADP) and phosphotransacetylase, *Methanothrix* utilizes an acetyl-CoA synthetase that hydrolyzes ATP to AMP and pyrophosphate[169]. Comparison of the kinetic parameters revealed that the acetyl-CoA synthetase of *Methanothrix* spp. has a higher affinity but lower maximum rate (K_m 0.8 mM, V_{max} 55 U.mg^{-1}) compared to the acetate kinase of *Methanosarcina thermophila* (K_m 22 mM, V_{max} 660 U.mg^{-1}), explaining the high acetate affinity but slow growth of *Methanothrix* (K_s 0.5 mM, doubling time 7 days)[170].

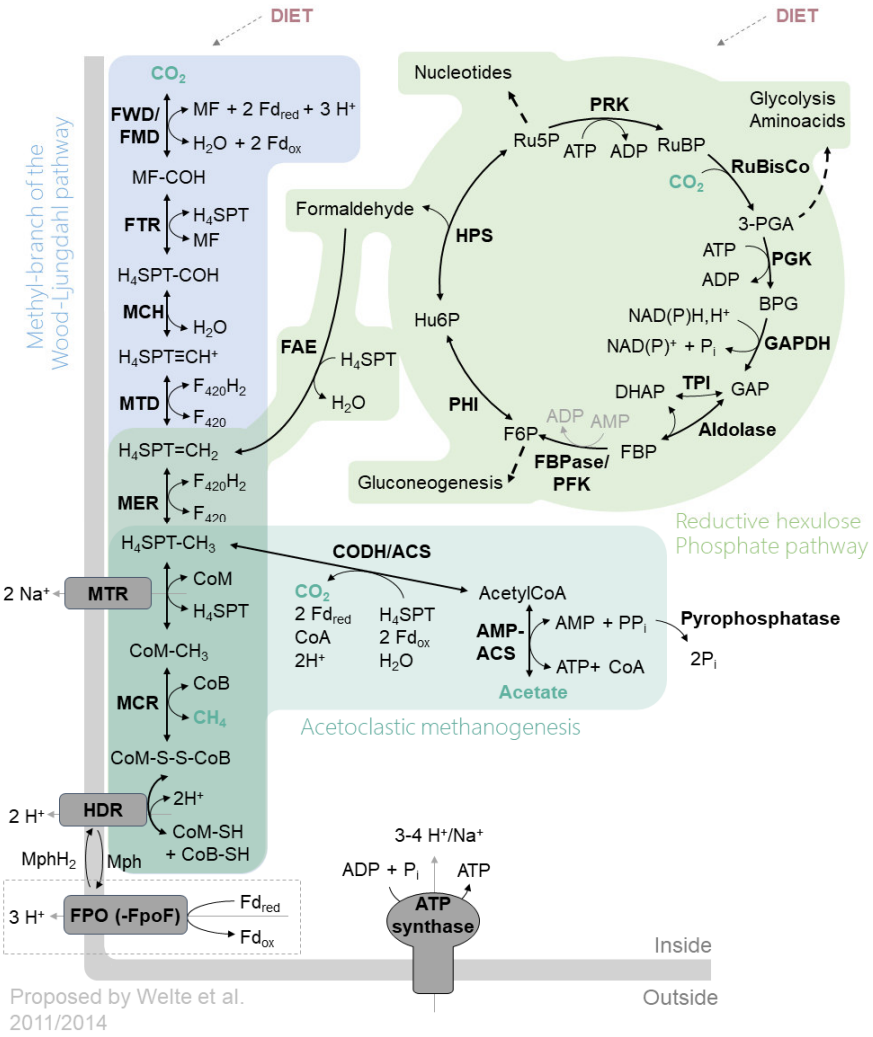


Figure 6. C1-metabolism in *Methanothrix*. Abbreviations: FWD/FMD formyl-methanofuran dehydrogenase, FTR formyl-methanofuran-tetrahydromethanopterin formyl-transferase, MCH methenyl-tetrahydromethanopterin cyclohydrolase, MTD methylene-tetrahydromethanopterin dehydrogenase MER methylene-tetrahydromethanopterin reductase, MTR tetrahydromethanopterin S-methyl-transferase, MCR methyl-coenzyme M reductase,

Figure 6. cont. HDR dihydromethanophenazine:CoB-CoM heterodisulfide reductase, FPO F_{420}:phenazine dehydrogenase, CODH/ACS CO-dehydrogenase/ acetyl-CoA decarbonylase, AMP-ACS acetyl-CoA synthetase/acetate-CoA ligase, MF Methanofuran, H_4SPT tetrahydrosarcinapterin, MPH methanophenazine, CoM coenzyme-M, CoB coenzyme B, CoM-S-S-CoB heterodisulfide. RuBisCo ribulose-bisphosphate carboxylase, 3-PGA 3-phosphoglycerate, PGK phosphoglycerate kinase, BPG 1,3-diphosphoglycerate, GAPDH glyceraldehyde-3-phosphate dehydrogenase, GAP glyceraldehyde-3-phosphate, TPI triose-phosphate isomerase, DHAP dihydroxyacetone phosphate, Aldolase fructose-bisphosphate aldolase, FBP fructose-1,6-bisphosphate, FBPase fructose-bisphosphatase, PFK ADP-specific phosphofructokinase, F6P fructose-6-phosphate, PHI 6-phospho-3-hexuloisomerase, Hu6P D-arabino-3-hexulose-6-phosphate, HPS 3-hexulose-6-phosphate synthase, FAE formaldehyde-activating enzyme, Ru5P ribulose-5-phosphate, PRK phosphoribulokinase, RuBP ribulose-1,5-bisphosphate. Based on Welte *et al.* [171, 172] and Kono *et al.*[175].

Another difference between *Methanothrix* and *Methanosarcina* is the mechanism of energy conservation[171]. The HDR complex of *Methanothrix thermoacetophila* has been shown to reduce CoM-S-S-CoB solely with ferredoxin and no other electron donors (H_2, $F_{420}H_2$, NADH)[171]. However, typical ferredoxin utilizing complexes (ECH and Rnf-complex) are not present in *Methanothrix*. Furthermore, the FPO complex was found to lack the subunit responsible for $F_{420}H_2$ binding and oxidation (FpoF)[171]. It has been proposed that the incomplete FPO complex utilizes reduced ferredoxin instead of $F_{420}H_2$ for the reduction of MPH coupled to the translocation of three protons[171, 172]. Energy conservation by MTR, HDR, and the incomplete FPO would then theoretically provide two ATP and one extra translocated ion. Accounting for the acetate activation cost, the yield per generated methane would be one translocated ion. Therefore, *Methanothrix* can be considered thriving at the energetical limit[171]. Further studies will be required in order to confirm this.

It has been shown that, in addition to acetoclastic methanogenesis, *Methanothrix* is able to reduce CO_2 to CH_4 via the methyl branch of the Wood-Ljungdahl pathway[173] by accepting electrons through direct interspecies electron transfer (DIET) from partners

like *Geobacter* species who utilize conductive pili[173, 174]. The components of electrical connections in *Methanothrix* remain unclear. Furthermore, it has been proposed, based on genetic potential and expression levels, that another CO_2 fixation path exists via the reductive hexulose pathway (RHP)[175] coupled to DIET. Here, CO_2 could be fixed through multiple enzymes of the Calvin Benson cycle and then funneled into biosynthesis or be converted into formaldehyde, which is then converted to methylene-H_4SPT entering the methanogenesis pathway[175]. While acetoclastic methanogenesis yields little energy ($\Delta G^{\circ\prime} = -24.8$ kJ.mol^{-1}), complete acetate reduction might be achieved via the combination with DIET fueled classical CO_2 reduction ($\Delta G^{\circ\prime} = -87.0$ kJ.mol^{-1}) or RHP ($\Delta G^{\circ\prime} = -54.0$ kJ.mol^{-1})[176]. High transcriptional activity of all three pathways in *Methanothrix soehngenii* has been shown[176].

Methanothrix is highly abundant in major biogenic methane sources like rice paddies[177] and anaerobic waste digesters[178]. Especially, in low acetate environments, the high acetate affinity allows them to outcompete *Methanosarcina* species[165]. As the final step of anaerobic decomposition, after fermentative and acetogenic bacteria, methanogens consume the resulting H_2, formate and acetate [179]. This places *Methanothrix* at a critical juncture, preventing acetate accumulation and maintaining the flow of biological matter degradation. Recent studies have discussed the presence and influence of *Methanothrix* in environments like arctic permafrost soils[180], oxygenated wetland soils[181], and tidal flat sediments[182], which have a high climate relevance. It has been suggested that *Methanothrix* might have an even larger role in global methane production than previously assumed[182, 183] especially as interacting with other microbes through DIET would increase their methane output, allowing *Methanothrix* members to grow faster than predicted based on lab-grown cultures. Furthermore, *Methanothrix* has been proposed as “the predominant methane producer on Earth”[184].

Besides its environmental relevance, *Methanothrix* also holds major biotechnological potential. Their degradation capabilities are harnessed in wastewater treatment, and the genus is ubiquitously found in anaerobic reactors[185]. Furthermore, they are of high interest in the interlinked topic of biomethane production, especially since DIET allows for the decrease of unwanted CO_2 outputs and the increase of valuable methane production[186]. Studying the electron transfer mechanism (*e.g.*, conducting sheath amyloids[186]) or specialized enzymes (*e.g.*, Acetyl-CoA synthetase[170, 187]) of *Methanothrix* might inspire the development of new chemical processes or biotechnological applications. While biochemical characterizations of main enzymes have been conducted before[169] only a limited amount of protein structures, obtained through recombinant expression, have been reported[188-193] and native studies have been limited by low growth rates and low yields during cultivation[165, 184]. *Methanothrix*'s environmental relevance and biotechnological potential in combination with a not fully resolved metabolism, make it a very interesting target for crystallomics.

Alkane oxidizing archaea

In deep sediments, buried organic matter is transformed by pressure and heat to a complex hydrocarbon mixture[194, 195]. Natural gas and crude oil are formed which include a major fraction of methane and *n*-alkanes[196]. These hydrocarbons can migrate to surface sediments and enter the water column through cold seeps and hydrothermal vents. And while humans have only "recently" begun to use hydrocarbons as an energy source, microbes have been using them much longer. Aerobic oxidation of methane is performed by bacteria transforming methane or alkanes to alcohols via monooxygenases[197] and further oxidize them to CO_2. In the absence of oxygen archaea have been identified to oxidize alkanes and constitute key players for methane oxidation[198, 199]. Hydrocarbon-degrading archaea are found all across the globe in habitats like hydrothermal vents[88, 200], cold seeps[198, 201], mud volcanoes[202], shallow coastal sandy sediments[202], and freshwater systems[203, 204].

Anaerobic oxidation of methane (AOM) is carried out via "reversed methanogenesis"[205] utilizing the majority of methanogenesis enzymes[206]. Various compounds can be used as electron sinks (nitrate[207, 208], iron[209], manganese oxide[210]) but sulfate is the most commonly reported for marine environments[198]. Anaerobic methanotrophic archaea (ANME) and anaerobic alkanotrophic archaea (ANKA) often form distinct consortia with sulfate reducing bacteria (SRB) as observed by FISH[86, 89, 199, 211, 212] (Fig. 7). Different mechanisms of electron transfer have been proposed including the diffusion of small molecules[213-215] and DIET[216-218]. Genomic analyses have revealed a high abundance of membrane associated small multiheme cytochromes (MHC), large MHCs containing up to 80 heme binding sites and homologs of nanowire-forming proteins in various ANME[202].

Currently, there are no pure cultures of any anaerobic archaeal alkane degraders. However, some labs have managed to obtain enrichment cultures through dedicated, decade-long work. These include archaea able to oxidize methane[88, 219], ethane[92, 121], propane[220], butane[220], pentane to tetradecane[86] and even long-chain alkanes[87]. Since accessible biomass of these organisms is scarce, only few structural studies have been carried out on enzymes of alkanotrophs. These include MCR from an uncultured ANME-1 that was purified from a consortium of microorganisms collected from Black Sea mats[119] and the first structure of an ECR obtained from *Ca.* Ethanoperedens thermophilum[122].

Furthermore, a cluster of divergent *mcrA* genes was identified in members of *Ca.* Bathyarchaeota[83], *Ca.* Hadesarchaeota[154, 156], *Ca.* Helarchaeota[221], *Ca.* Methanoliparia[89, 222] and *Archaeoglobales*[154, 223] which have been proposed to mediate alkanotrophy. The low sequence similarity to canonical MCRs might reflect evolutionary changes to accommodate larger alkanes[220]. These findings suggest that anaerobic methanotrophs

and short chain alkane oxidizers may have an even wider phylogenetic and environmental distribution than previously thought[222].

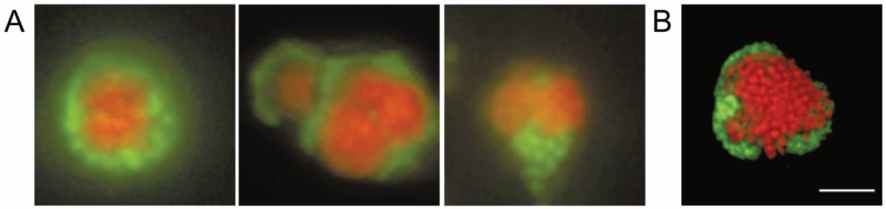


Figure 7. ANME-SRB-consortia. Figure from Boetius *et al*. 2000[211]. Epifluorescence micrographs (**A**) and Confocal laser scanning micrograph (**B**) of cells hybridized with fluorescent probes with archaea in red and SRB in green. For details on the probes see original publication. *© Reproduced with permission from Springer Nature. License Nr. 5912391092956*

The methyl-coenzyme M reductase

MCR is the key enzyme for methanogens and methanotrophs as it catalyzes the strictly anaerobic generation or fixation of methane and is shared by all presented pathways (Table 1). The forward reaction (generation of CH_4 and CoM-S-S-CoB) is exergonic with a standard free energy change (ΔG^o) of ~ −30 kJ.mol^{-1} [137]. The currently proposed reaction mechanism is illustrated in Fig. 8A. It starts at the active Ni(I) state (MCR_{red1}) with methyl-CoM and then CoB-SH entering the active site. Methyl-CoM binds to Ni(I), and the S-C bond is homolytically cleaved, generating a methyl-radical and a Ni(II)-thiolate ($MCR_{ox1\text{-}silent}$). The methyl radical abstracts a hydrogen from HS-CoB. One methane is generated and the formed CoB radical reacts with the Ni(II)-thiolate complex to form a heterodisulfide anion radical. Finally, this intermediate reduces Ni(II) back to Ni(I) and forms CoM-S-S-CoB[224].

The first structural studies were carried out on MCR from *Methanothermobacter marburgensis*[225] and revealed that the enzyme forms a dimer of heterotrimers $(\alpha\beta\gamma)_2$ of ~250 kDa (Fig. 8B). Two active sites, 50 Å apart, are formed by the α-, α′-, β- and γ-subunit (mirrored on the other side). Each harbors a nickel-containing porphinoid F_{430} cofactor that is tightly but not covalently bound. A 30 Å wide channel reaches from the surface to the F_{430} cofactor allowing substrate entry. The uniqueness of MCR is not only in its cofactor, but also through the multiple PTMs in the surroundings of its active site with: 1-*N*-methylhistidine, *S*-methylcysteine, 5-C-(*S*)-methylarginine, 2-C-(*S*)-methylglutamine, thioglycine and didehydroaspartate reported in *M. marburgensis*[225-227]. Currently, only the inactive Ni(II) state can be observed in crystallographic structures, and intense efforts are being made to trap the active state[159].

Further studies have expanded the available pool of structures to other species: *Methanosarcina barkeri* and *Methanopyrus kandleri*[226], an uncultured ANME-1[119], *Methanothermobacter wolfeii*[227], *Methanotorris formicicus,* and *Methanothermococcus thermolithotrophicus*[228] and *Methanosarcina acetivorans*[229]. These studies have shown an overall high sequence and structural conservation with only minor changes on the surface of the protein. Most differences were found in the PTM patterns: while some MCRs lacked certain modifications, others were found to include additional PTMs like 7-hydroxy-L-tryptophane and S-oxymethionine in ANME-1 and 6-hydroxytryptophan in *M. formicicus*[230]. Additionally, MCR from ANME-1 was also found to contain a modified methylthio-F_{430} cofactor[119]. Notably, the structure for the related ECR obtained from *Ca*. Ethanoperedens thermophilum[122] contained multiple unique features like a widened catalytic chamber and hydrophobic tunnel to accommodate the larger substrate ethane as well as a dimethylated-F_{430}. Furthermore, an extensive PTM pattern containing a unique 3-methylisoleucine and a N^2-methylhistidine was observed.

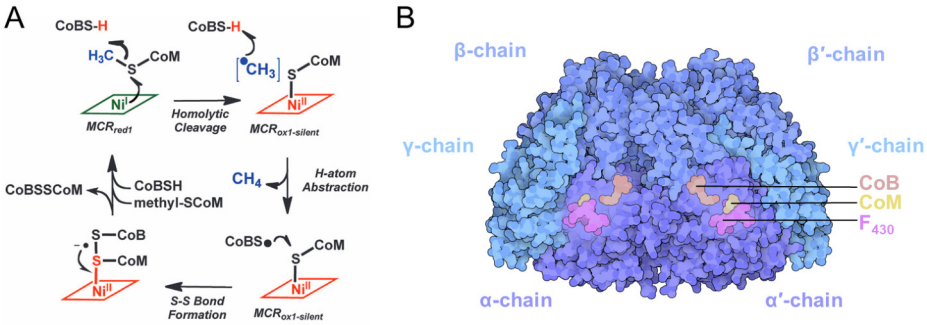


Figure 8. MCR reaction mechanism and structure. (**A**) Reaction mechanism reprinted from Wongnate *et al*. 2016[224] *Reprinted with permission from AAAS.* (**B**) Structure of MCR with the α-, β- and γ-chains containing coenzyme M (CoM), coenzyme B (CoB) and cofactor F_{430} in different shades, adapted from PDB molecule of the month[231].

Energy saving strategies of methanogens

Methanogenesis only provides limited amounts of energy. The energetic cost of the synthesis of one molecule of ATP from ADP and inorganic phosphate in a living cell is estimated to be around −50 kJ.mol−1 232. Therefore, based on thermodynamics, it is estimated that under standard conditions, a hydrogenotrophic methanogen should be able to generate up to three ATP per formed methane (−131 kJ.mol−1)232. However, in most habitats, H2 partial pressures are very low, resulting in a reduced energy yield between −17 to −40 kJ.mol−1 for each generated methane, which is insufficient to synthesize a single mole of ATP per mole of methane232. Therefore, methanogens have found ways to deal with this energy limitation through molecular tricks, allowing them to derive sufficient chemical energy through metabolic reactions and to save cellular energy for anabolic purposes. The most dramatic energy saving system in hydrogenotrophic methanogens is the CO2-fixation apparatus. Here, the enzyme FMD/FWD relies on the electrons of low redox potential and a sophisticated channeling system to fix CO2 in an ATP-independent fashion. A narrow hydrophobic channel

selectively allows CO2 to enter the first active site. After the reaction, a hydrophilic internal cavity facilitates the accumulation of formate, the resulting high local concentration being proposed to allow the thermodynamic equilibrium to shift towards formyl-MFR formation233. Another example is the HdrABC complex which couples the favorable reduction of CoM-S-S-CoB (E°′=−140 mV) through the oxidation of H2 (E°′=−414 mV) with the unfavorable formation of the low redox potential electron for CO2 fixation (E°′~−500 mV) through an elegant electron bifurcation230.

In addition to catabolic energy-saving strategies, methanogens have also developed other systems to optimize the cellular energy budget. One example is the fine-tuning of nitrogen (N) assimilation. Here, a large role is played by the small regulatory proteins of the P_{II} family[5, 234] and the metabolite 2-oxoglutarate (2OG, α-ketoglutarate)[235] (Fig. 9). Methanogens can utilize different N-sources with the main one being NH_3. The ammonium transporter AmtB is responsible for the uptake and has been shown to be inactivated by the binding of the P_{II} family regulatory protein GlnK which sterically blocks the channel[236-238]. GlnK is itself regulated by 2OG. The binding of the compound in combination with Mg^{2+} and ATP leads to a stiffening of three elongated surface loops, preventing the binding to AmtB[239]. The binding site and deactivation mechanism of P_{II} members is conserved in various methanogens[240]. NH_3-assimilation is carried out by the glutamine synthetase (GS), which catalyzes the ATP-dependent condensation of glutamate and ammonia to form glutamine, followed by further processing by the glutamate synthase (GOGAT), turning glutamine and 2OG into two molecules of glutamate of which one can enter the N-metabolism[5, 241]. Archaeal GS has been shown to be regulated by direct interaction with GlnK but is also stimulated by 2OG[242, 243].

Some methanogens are diazotrophs, utilizing atmospheric nitrogen (N_2) as sole N-source[244, 245]. The enzyme nitrogenase fixes N_2 with a high energy cost of 16 ATP[246]. The P_{II} member $NifI_{12}$ was recently found to inactivate the complex through binding to

the MoFe subunit, preventing the binding of the Fe protein required for complex assembly[247, 248]. In all cases, 2OG functions as a sensor for the cellular nitrogen availability[235]. At a low nitrogen availability, the intracellular 2OG pool will increase, resulting in deactivated GlnK and $NifI_{12}$ allowing ammonia import or nitrogen fixation and also boosting GS activity. If the nitrogen availability is high, the 2OG pool will decrease, resulting in the blocking of ammonia transport and nitrogen fixation and slowed GS activity, contributing to lowered ATP costs. This clever system allows to precisely manage energy expenses versus nitrogen acquisition, which is crucial for an energy-limited organism.

Future studies of more methanogens will likely unveil many more of these intricate regulation and energy-saving strategies, as well as enzymes harboring molecular tricks to promote and optimize reactions.

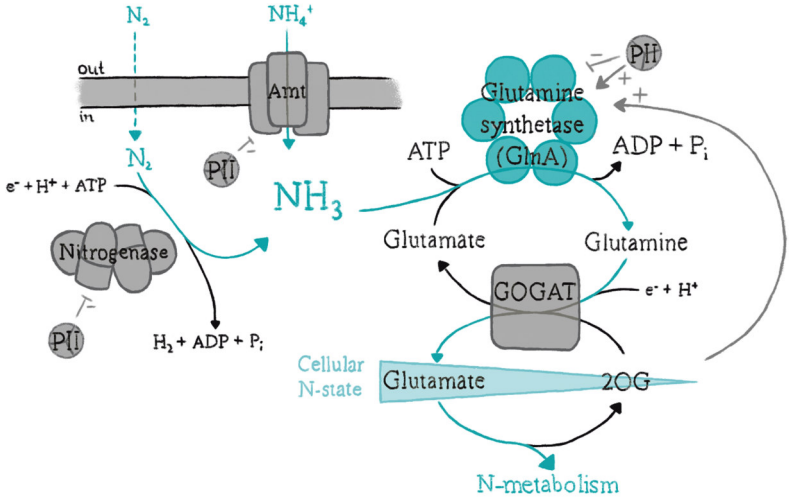


Figure 9. Nitrogen assimilation regulation. Adapted from Müller *et al.* 2024[249].The path of nitrogen is colored in blue, as well as the glutamine synthetase (focus of Ch. III). *Licensed under a Creative Commons Attribution 4.0 International License* (http://creativecommons.org/licenses/by/4.0/)

Specific aims and presented chapters

The specific aim of this work, titled "Expanding the possibilities of crystallomics towards environmental samples" is to investigate the feasibility of expanding the approach to environmental samples in order to study previously inaccessible enzymes. To evaluate the challenges and strengths of the technique, the following chapters will describe incremental steps in moving from pure cultures towards heterogeneous samples.

Chapter II explores the untargeted simultaneous purification of a pure culture of the hyperthermophilic hydrogenotrophic methanogen *Methanocaldococcus infernus*. Preliminary results, including the number of enriched protein fractions, crystals obtained, and resulting structures solved via three different SAD approaches, illustrate that the method is not restricted to catabolic enzymes but can also capture regulatory and anabolic systems.

In **Chapter III,** a targeted and improved purification was performed based on previous untargeted crystallomics from pure cultures. The enzyme glutamine synthetase was studied in the two archaea, *Methanothermococcus thermolithotrophicus* and *Methermicoccus shengliensis*, revealing a novel regulatory mechanism and its molecular basis.

Chapter IV describes an adaptive targeted purification where a target of interest was followed throughout the shotgun purification based on its specific features. As a result, the first MCR from a methoxydotrophic methanogen has been structurally characterized, presenting the lowest amount of post-translational modifications currently described.

In **Chapter V**, the approach was then applied to a complex microbial sample. Here, three different microbial enrichments containing marine and freshwater ANME-2 have

been studied. The adapted targeted purification setup of chapter IV was performed to isolate the MCR from these methanotrophs. Three MCR structures at (sub)atomic resolution were obtained, revealing a high content of post-translational modifications, with one of them being novel.

Chapter VI extended the technique to an even more heterogeneous sample: wastewater treatment sludge. MCR was identified as a major protein fraction in the soluble extract and selected for adaptive targeted purification. A physical enrichment protocol increased the fraction of *Methanothrix* filaments to facilitate the purification. In parallel, direct purification was also performed on the sample. In both cases, MCR from two *Methanothrix* subspecies could be purified and crystallized. This proof of concept highlights the feasibility of the approach on heterogeneous environment-like samples.

Finally, in **Chapter VII** the untargeted simultaneous purification was performed on the sludge sample. Preliminary results include an overview of the obtained fractions and their mass spectrometry identification. Additionally, some proteins could be successfully extracted from the membrane fraction. These promising results indicate that native protein purification on heterogeneous samples is worth further pursuit.

References

1. Dueholm, M. S., Larsen, P., Finster, K., Stenvang, M. R., Christiansen, G., Vad, B. S., Boggild, A., Otzen, D. E., Nielsen, P. H. The Tubular Sheaths Encasing Filaments Are Functional Amyloids. *J. Biol. Chem.* **290**, 20590-20600, doi:10.1074/jbc.M115.654780 (2015).
2. Albers, S. V., Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414-426, doi:10.1038/nrmicro2576 (2011).
3. Jarrell, K. F., Ding, Y., Nair, D. B., Siu, S. Surface appendages of archaea: Structure, function, genetics and assembly. *Life* **3**, 86-117, doi:10.3390/life3010086 (2013).
4. Prasse, D., Thomsen, J., De Santis, R., Muntel, J., Becher, D., Schmitz, R. A. First description of small proteins encoded by spRNAs in *Methanosarcina mazei* strain Gö1. *Biochimie* **117**, 138-48, doi:10.1016/j.biochi.2015.04.007 (2015).
5. Leigh, J. A., Dodsworth, J. A. Nitrogen regulation in bacteria and archaea. *Annu. Rev. Microbiol.* **61**, 349-77, doi:10.1146/annurev.micro.61.080706.093409 (2007).
6. Kennelly, P. J. Archaeal protein kinases and protein phosphatases: Insights from genomics and biochemistry. *Biochem J.* **370**, 373-89, doi:10.1042/BJ20021547 (2003).
7. Albers, R. W., Siegel, G. J., Xie, Z.-J. Membrane Transport. In *Basic Neurochemistry* (eds Brady, S. T.; Siegel, G. J.; Albers, R. W.; Price, D. L.) 40-62 (Academic Press, 2012).
8. Ellermann, J., Rospert, S., Thauer, R. K., Bokranz, M., Klein, A., Voges, M., Berkessel, A. Methyl-coenzyme-M reductase from *Methanobacterium thermoautotrophicum* (strain Marburg). Purity, activity and novel inhibitors. *Eur. J. Biochem.* **184**, 63-8, doi:10.1111/j.1432-1033.1989.tb14990.x (1989).
9. Andersson, I., Backlund, A. Structure and function of Rubisco. In *Plant physiology and biochemistry* Vol. 46 (ed Dumas, R.) 275-291 (Elsevier, 2008).
10. Hoffman, B. M., Lukoyanov, D., Yang, Z. Y., Dean, D. R., Seefeldt, L. C. Mechanism of nitrogen fixation by nitrogenase: The next stage. *Chem. Rev.* **114**, 4041-62, doi:10.1021/cr400641x (2014).
11. Copeland, R. A. *Enzymes: a practical introduction to structure, mechanism, and data analysis*. 3 edn, Vol. 1 1-550 (John Wiley & Sons, 2023).
12. Fuchs, G. *Allgemeine Mikrobiologie*. 8 edn, (Georg Thieme Verlag, 2007).
13. Chapman, J., Ismail, A. E., Dinu, C. Z. Industrial applications of enzymes: recent advances, techniques, and outlooks. *Catalysts* **8**, doi:10.3390/catal8060238 (2018).
14. Mesbah, N. M. Industrial biotechnology based on enzymes from extreme environments. *Front. Bioeng. Biotechnol.* **10**, 870083, doi:10.3389/fbioe.2022.870083 (2022).
15. Singh, R., Kumar, M., Mittal, A., Mehta, P. K. Microbial enzymes: Industrial progress in 21st century. *3 Biotech* **6**, doi:10.1007/s13205-016-0485-8 (2016).
16. van Beilen, J. B., Li, Z. Enzyme technology: An overview. *Curr. Opin. Biotechnol.* **13**, 338-44, doi:10.1016/s0958-1669(02)00334-8 (2002).
17. Kumar, D., Savitri, T. N., Verma, R., Bhalla, T. C. Microbial proteases and application as laundry detergent additive. *Res. J. Microbiol.* **3**, 661-672, doi:jm.2008.661.672 (2008).
18. Bajpai, P. Application of enzymes in the pulp and paper industry. *Biotechnol. Prog.* **15**, 147-57, doi:10.1021/bp990013k (1999).
19. Schulz, P., Rizvi, S. S. H. Hydrolysis of lactose in milk: Current status and future products. *Food Rev. Internat.* **39**, 2875-2894, doi:10.1080/87559129.2021.1983590 (2023).
20. Araújo, R., Casal, M., Cavaco-Paulo, A. Application of enzymes for textile fibres processing. *Biocatal. Biotransfor.* **26**, 332-349, doi:10.1080/10242420802390457 (2008).
21. Rezania, S., Oryani, B., Cho, J., Talaiekhozani, A., Sabbagh, F., Hashemi, B., Rupani, P. F., Mohammadi, A. A. Different pretreatment technologies of lignocellulosic biomass for bioethanol production: an overview. *Energy* **199**, 117457, doi:10.1016/j.energy.2020.117457 (2020).
22. Zheng, G. W., Xu, J. H. New opportunities for biocatalysis: driving the synthesis of chiral chemicals. *Curr. Opin. Biotechnol.* **22**, 784-92, doi:10.1016/j.copbio.2011.07.002 (2011).
23. Meghwanshi, G. K., Kaur, N., Verma, S., Dabi, N. K., Vashishtha, A., Charan, P. D., Purohit, P., Bhandari, H. S., Bhojak, N., Kumar, R. Enzymes for pharmaceutical and therapeutic applications. *Biotechnol. Appl. Biochem.* **67**, 586-601, doi:10.1002/bab.1919 (2020).

24. Rufer, A. C. Drug discovery for enzymes. *Drug Discov. Today* **26**, 875-886, doi:10.1016/j.drudis.2021.01.006 (2021).
25. Rupp, B. *Biomolecular crystallography: Principles, practice, and application to structural biology*. Vol. 1 (Garland Science, 2009).
26. Hünefeld, F. L. *Der Chemismus in der thierischen Organisation: Physiologisch-chemische Untersuchungen der materiellen Veränderungen oder des Bildungslebens im thierischen Organismus, insbesondere des Blutbildungsprocesses, der Natur der Blutkörperchen und ihrer Kernchen.*, (Brockhaus, 1840).
27. Preyer, W. T. *Die Blutkrystalle: Untersuchungen*. (Mauke, 1871).
28. Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., North, A. C. Structure of haemoglobin: A three-dimensional fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis. *Nature* **185**, 416-22, doi:10.1038/185416a0 (1960).
29. McPherson, A. A brief history of protein crystal growth. *J. Cryst. Growth* **110**, 1-10, doi:10.1016/0022-0248(91)90859-4 (1991).
30. Funke, O. *Über das Milzvenenblut*. (1848).
31. Hartig, T. Ueber das Klebermehl. *Botanische Zeitung* **13**, 881-882 (1855).
32. Ritthausen, H. *Die Eiweisskörper der Gereidearten, Hülsenfrüchte und Ölsamen: Beiträge zur Physiologie der Samen der Culturgewächse der Nahrungs-und Futtermittel*. (1872).
33. Hofmeister, F. *Ueber die Darstellung von krystallisirtem Eieralbumin und die Krystallisirbarkeit colloider Stoffe*. (1890).
34. Giegé, R. A historical perspective on protein crystallization from 1840 to the present day. *FEBS J.* **280**, 6456-6497, doi:10.1111/febs.12580 (2013).
35. Sumner, J. B. The isolation and crystallization of the enzyme urease. Preliminary paper. *J. Biol. Chem.* **69**, 435-441, doi:10.4159/harvard.9780674366701.c115 (1926).
36. Sumner, J. B., Dounce, A. L. Crystalline catalase. *Science* **85**, 366-367, doi:10.1126/science.85.2206.366 (1937).
37. Rathore, I., Mishra, V., Bhaumik, P. Advancements in macromolecular crystallography: From past to present. *Emerg. Top. Life Sci.* **5**, 127-149, doi:10.1042/Etls20200316 (2021).
38. Laue, M. v. Eine Quantitative Prüfung der Theorie für die Interferenzerscheinungen bei Röntgenstrahlen. *Annalen der Physik* **346**, 989-1002, doi:10.1002/andp.19133461005 (1913).
39. Bernal, J. D., Crowfoot, D. X-ray photographs of crystalline pepsin. *Nature* **133**, 794-795, doi:10.1038/133794b0 (1934).
40. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., Phillips, D. C. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662-6, doi:10.1038/181662a0 (1958).
41. Crystallography: Protein Data Bank. *Nat. New Biol.* **233**, 223-223, doi:10.1038/newbio233223b0 (1971).
42. Jancarik, J., Kim, S.-H. Sparse matrix sampling: A screening method for crystallization of proteins. *J. Appl. Cryst.* **24**, 409-411, doi:10.1107/S0021889891004430 (1991).
43. McPherson, A., Gavira, J. A. Introduction to protein crystallization. *Acta Cryst. F* **70**, 2-20, doi:10.1107/S2053230X13033141 (2014).
44. Westneat, M. W., Socha, J. J., Lee, W.-K. Advances in biological structure, function, and physiology using synchrotron X-ray imaging. *Annu. Rev. Physiol.* **70**, 119-142, doi:10.1146/annurev.physiol.70.113006.100434 (2008).
45. Schaffer, J. E., Kukshal, V., Miller, J. J., Kitainda, V., Jez, J. M. Beyond X-rays: an overview of emerging structural biology methods. *Emerg. Top. Life Sci.* **5**, 221-230, doi:10.1042/ETLS20200272 (2021).
46. Susini, J., Cassagne, J., Gagey, B., Nadji, A., Taleb, A., Thompson, A., Daillant, J. A brief introduction to the Synchrotron SOLEIL and its upgrade programme. *Eur. Phys. J. Plus* **139**, 80, doi:10.1140/epjp/s13360-024-04872-2 (2024).
47. Chenevier, D., Joly, A. ESRF: Inside the extremely brilliant source upgrade. *Synchrotron Radiat. News* **31**, 32-35, doi:10.1080/08940886.2018.1409562 (2018).

48. Xiao, Q., Wu, T., Bao, K., Tang, J., Zhang, Y., Zhang, W., Zhu, Z., Gu, Y., Zhou, S., Li, C., et al. Upgrade of crystallography beamline BL19U1 at the Shanghai Synchrotron Radiation Facility. *J. Appl. Crystallogr.* **57**, 630-637, doi:10.1107/S1600576724002188 (2024).
49. *RCSB Protein Data Bank. PDB data distribution by experimental method and molecular type*, <https://www.rcsb.org/stats/summary> (accessed 15.11.2014).
50. Hirano, Y., Takeda, K., Miki, K. Charge-density analysis of an iron-sulfur protein at an ultra-high resolution of 0.48 Å. *Nature* **534**, 281-4, doi:10.1038/nature18001 (2016).
51. Schmidt, A., Teeter, M., Weckert, E., Lamzin, V. S. Crystal structure of small protein crambin at 0.48 Å resolution. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **67**, 424-8, doi:10.1107/S1744309110052607 (2011).
52. Shi, D., Nannenga, B. L., Iadanza, M. G., Gonen, T. Three-dimensional electron crystallography of protein microcrystals. *eLife* **2**, e01345, doi:10.7554/eLife.01345 (2013).
53. McNeil, B. W. J., Thompson, N. R. X-ray free-electron lasers. *Nat. Photonics* **4**, 814-821, doi:10.1038/nphoton.2010.239 (2010).
54. Orville, A. M. Recent results in time resolved serial femtosecond crystallography at XFELs. *Curr. Opini. Str. Biol.* **65**, 193-208, doi:10.1016/j.sbi.2020.08.011 (2020).
55. Hu, Y., Cheng, K., He, L., Zhang, X., Jiang, B., Jiang, L., Li, C., Wang, G., Yang, Y., Liu, M. NMR-based methods for protein analysis. *Anal. Chem.* **93**, 1866-1879, doi:10.1021/acs.analchem.0c03830 (2021).
56. Reif, B., Ashbrook, S. E., Emsley, L., Hong, M. Solid-state NMR spectroscopy. *Nat. Rev. Methods Primers* **1**, doi:10.1038/s43586-020-00002-1 (2021).
57. Bonvin, A. M., Boelens, R., Kaptein, R. NMR analysis of protein interactions. *Curr. Opin. Chem. Biol.* **9**, 501-8, doi:10.1016/j.cbpa.2005.08.011 (2005).
58. Zhou, P., Wagner, G. Overcoming the solubility limit with solubility-enhancement tags: Successful applications in biomolecular NMR studies. *J. Biomol. NMR* **46**, 23-31, doi:10.1007/s10858-009-9371-6 (2010).
59. *RCSB protein data bank. Number of Released PDB Structures per Year*, <https://www.rcsb.org/stats/all-released-structures> (accessed 15.11.2024).
60. Lyumkis, D. Challenges and opportunities in cryo-EM single-particle analysis. *J. Biol. Chem.* **294**, 5181-5197, doi:10.1074/jbc.REV118.005602 (2019).
61. Liao, M., Cao, E., Julius, D., Cheng, Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **504**, 107-12, doi:10.1038/nature12822 (2013).
62. Wei, X., Su, X., Cao, P., Liu, X., Chang, W., Li, M., Zhang, X., Liu, Z. Structure of spinach photosystem II-LHCII supercomplex at 3.2 Å resolution. *Nature* **534**, 69-74, doi:10.1038/nature18020 (2016).
63. Amunts, A., Brown, A., Bai, X. C., Llacer, J. L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S. H. W., Ramakrishnan, V. Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343**, 1485-1489, doi:10.1126/science.1249410 (2014).
64. Wu, S., Armache, J. P., Cheng, Y. Single-particle cryo-EM data acquisition by using direct electron detection camera. *Microscopy* **65**, 35-41, doi:10.1093/jmicro/dfv355 (2016).
65. Yip, K. M., Fischer, N., Paknia, E., Chari, A., Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157-161, doi:10.1038/s41586-020-2833-4 (2020).
66. Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P. M. G. E., Grigoras, I. T., Malinauskaite, L., Malinauskas, T., Miehling, J., et al. Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152-156, doi:10.1038/s41586-020-2829-0 (2020).
67. Baker, B. J., Dick, G. J. Omic approaches in microbial ecology: Charting the unknown. *Microbe Mag.* **8**, 353-360 (2013).
68. Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., Butterfield, C. N., Hernsdorf, A. W., Amano, Y., Ise, K., et al. A new view of the tree of life. *Nat. Microbiol.* **1**, 16048, doi:10.1038/nmicrobiol.2016.48 (2016).
69. Lok, C. Mining the microbial dark matter. *Nature* **522**, 270-3, doi:10.1038/522270a (2015).

70. Hedlund, B. P., Dodsworth, J. A., Murugapiran, S. K., Rinke, C., Woyke, T. Impact of single-cell genomics and metagenomics on the emerging view of extremophile "microbial dark matter". In *Extremophiles* Vol. 18 865-875 (Springer, 2014).
71. Zha, Y., Chong, H., Yang, P., Ning, K. Microbial dark matter: from discovery to applications. *Genom. Proteom. Bioinf.* **20**, 867-881, doi:10.1016/j.gpb.2022.02.007 (2022).
72. Schultz, J., Modolon, F., Peixoto, R. S., Rosado, A. S. Shedding light on the composition of extreme microbial dark matter: alternative approaches for culturing extremophiles. *Front. Microbiol.* **14**, 1167718, doi:10.3389/fmicb.2023.1167718 (2023).
73. Zhang, Y., Wang, Y., Tang, M., Zhou, J., Zhang, T. The microbial dark matter and "wanted list" in worldwide wastewater treatment plants. *Microbiome* **11**, 59, doi:10.1186/s40168-023-01503-3 (2023).
74. Solden, L., Lloyd, K., Wrighton, K. The bright side of microbial dark matter: Lessons learned from the uncultivated majority. *Curr. Opin. Microbiol.* **31**, 217-226, doi:10.1016/j.mib.2016.04.020 (2016).
75. Iqbal, S., Begum, F., Ullah, I., Jalal, N., Shaw, P. Peeling off the layers from microbial dark matter (MDM): Recent advances, future challenges, and opportunities. *Crit. Rev. Microbiol.*, 1-21, doi:10.1080/1040841X.2024.2319669 (2024).
76. Williams, T. A., Embley, T. M. Archaeal "dark matter" and the origin of eukaryotes. *Genome Biol. Evol.* **6**, 474-481, doi:10.1093/gbe/evu031 (2014).
77. Spang, A., Saw, J. H., Jørgensen, S. L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A. E., van Eijk, R., Schleper, C., Guy, L., Ettema, T. J. G. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173-179, doi:10.1038/nature14447 (2015).
78. Imachi, H., Nobu, M. K., Nakahara, N., Morono, Y., Ogawara, M., Takaki, Y., Takano, Y., Uematsu, K., Ikuta, T., Ito, M., et al. Isolation of an archaeon at the prokaryote-eukaryote interface. *Nature* **577**, 519-525, doi:10.1038/s41586-019-1916-6 (2020).
79. Brown, C. T., Hug, L. A., Thomas, B. C., Sharon, I., Castelle, C. J., Singh, A., Wilkins, M. J., Wrighton, K. C., Williams, K. H., Banfield, J. F. Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523**, 208-11, doi:10.1038/nature14486 (2015).
80. Wrighton, K. C., Castelle, C. J., Varaljay, V. A., Satagopan, S., Brown, C. T., Wilkins, M. J., Thomas, B. C., Sharon, I., Williams, K. H., Tabita, F. R., et al. RubisCO of a nucleoside pathway known from Archaea is found in diverse uncultivated phyla in bacteria. *ISME J.* **10**, 2702-2714, doi:10.1038/ismej.2016.53 (2016).
81. Lloyd, K. G., Schreiber, L., Petersen, D. G., Kjeldsen, K. U., Lever, M. A., Steen, A. D., Stepanauskas, R., Richter, M., Kleindienst, S., Lenk, S., et al. Predominant archaea in marine sediments degrade detrital proteins. *Nature* **496**, 215-218, doi:10.1038/nature12033 (2013).
82. Michalska, K., Steen, A. D., Chhor, G., Endres, M., Webber, A. T., Bird, J., Lloyd, K. G., Joachimiak, A. New aminopeptidase from "microbial dark matter" archaeon. *FASEB J.* **29**, 4071-4079, doi:10.1096/fj.15-272906 (2015).
83. Evans, P. N., Parks, D. H., Chadwick, G. L., Robbins, S. J., Orphan, V. J., Golding, S. D., Tyson, G. W. Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science* **350**, 434-438, doi:10.1126/science.aac7745 (2015).
84. McKay, L. J., Dlakić, M., Fields, M. W., Delmont, T. O., Eren, A. M., Jay, Z. J., Klingelsmith, K. B., Rusch, D. B., Inskeep, W. P. Co-occurring genomic capacity for anaerobic methane and dissimilatory sulfur metabolisms discovered in the Korarchaeota. *Nat. Microbiol.* **4**, 614-622, doi:10.1038/s41564-019-0362-4 (2019).
85. Kaeberlein, T., Lewis, K., Epstein, S. S. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **296**, 1127-9, doi:10.1126/science.1070633 (2002).
86. Zehnle, H., Laso-Pérez, R., Lipp, J., Riedel, D., Merino, D. B., Teske, A., Wegener, G. *Candidatus* Alkanophaga archaea from Guaymas Basin hydrothermal vent sediment oxidize petroleum alkanes. *Nat. Microbiol.* **8**, 1199-1212, doi:10.1038/s41564-023-01400-3 (2023).
87. Zhou, Z., Zhang, C. J., Liu, P. F., Fu, L., Laso-Pérez, R., Yang, L., Bai, L. P., Li, J., Yang, M., Lin, J. Z., et al. Non-syntrophic methanogenic hydrocarbon degradation by an archaeal species. *Nature* **601**, 257-262, doi:10.1038/s41586-021-04235-2 (2022).

88. Holler, T., Widdel, F., Knittel, K., Amann, R., Kellermann, M. Y., Hinrichs, K. U., Teske, A., Boetius, A., Wegener, G. Thermophilic anaerobic oxidation of methane by marine microbial consortia. *ISME J.* **5**, 1946-56, doi:10.1038/ismej.2011.77 (2011).
89. Laso-Pérez, R., Hahn, C., van Vliet, D. M., Tegetmeyer, H. E., Schubotz, F., Smit, N. T., Pape, T., Sahling, H., Bohrmann, G., Boetius, A., et al. Anaerobic degradation of non-methane alkanes by "*Candidatus* Methanoliparia" in hydrocarbon seeps of the Gulf of Mexico. *mBio* **10**, doi:10.1128/mBio.01814-19 (2019).
90. Hu, H., Natarajan, V. P., Wang, F. Towards enriching and isolation of uncultivated archaea from marine sediments using a refined combination of conventional microbial cultivation methods. *Mar. Life Sci. Technol.* **3**, 231-242, doi:10.1007/s42995-021-00092-0 (2021).
91. Luo, F., Devine, C. E., Edwards, E. A. Cultivating microbial dark matter in benzene-degrading methanogenic consortia. *Environ. Microbiol.* **18**, 2923-36, doi:10.1111/1462-2920.13121 (2016).
92. Chen, S. C., Musat, N., Lechtenfeld, O. J., Paschke, H., Schmidt, M., Said, N., Popp, D., Calabrese, F., Stryhanyuk, H., Jaekel, U., et al. Anaerobic oxidation of ethane by archaea from a marine hydrocarbon seep. *Nature* **568**, 108-111, doi:10.1038/s41586-019-1063-0 (2019).
93. Sorek, R., Cossart, P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat. Rev. Genet.* **11**, 9-16, doi:10.1038/nrg2695 (2010).
94. Schneider, T., Riedel, K. Environmental proteomics: Analysis of structure and function of microbial communities. *Proteomics* **10**, 785-798, doi:10.1002/pmic.200900450 (2010).
95. Aguiar-Pulido, V., Huang, W. R., Suarez-Ulloa, V., Cickovski, T., Mathee, K., Narasimhan, G. Metagenomics, metatranscriptomics, and metabolomics approaches for microbiome analysis. *Evol. Bioinf.* **12**, 5-16, doi:10.4137/Ebo.S36436 (2016).
96. Jiao, J. Y., Liu, L., Hua, Z. S., Fang, B. Z., Zhou, E. M., Salam, N., Hedlund, B. P., Li, W. J. Microbial dark matter coming to light: Challenges and opportunities. *Natl. Sci. Rev.* **8**, doi:10.1093/nsr/nwaa280 (2021).
97. Beygmoradi, A., Homaei, A., Hemmati, R., Fernandes, P. Recombinant protein expression: Challenges in production and folding related matters. *Int. J. Biol. Macromol.* **233**, doi:10.1016/j.ijbiomac.2023.123407 (2023).
98. Sørensen, H. P., Mortensen, K. K. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J. Biotechnol.* **115**, 113-28, doi:10.1016/j.jbiotec.2004.08.004 (2005).
99. Parret, A. H., Besir, H., Meijers, R. Critical reflections on synthetic gene design for recombinant protein expression. *Curr. Opin. Struct. Biol.* **38**, 155-62, doi:10.1016/j.sbi.2016.07.004 (2016).
100. Baneyx, F., Mujacic, M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.* **22**, 1399-408, doi:10.1038/nbt1029 (2004).
101. de Groot, N. S., Espargaró, A., Morell, M., Ventura, S. Studies on bacterial inclusion bodies. *Future Microbiol.* **3**, 423-35, doi:10.2217/17460913.3.4.423 (2008).
102. Jenkins, N., Murphy, L., Tyther, R. Post-translational modifications of recombinant proteins: Significance for biopharmaceuticals. In *Mol. Biotechnol.* Vol. 39 113-8 (Springer, 2008).
103. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223-230, doi:10.1126/science.181.4096.223 (1973).
104. Levitt, M., Warshel, A. Computer simulation of protein folding. *Nature* **253**, 694-698, doi:10.1038/253694a0 (1975).
105. Moult, J. A decade of CASP: Progress, bottlenecks and prognosis in protein structure prediction. In *Curr Opin Struc Biol* Vol. 15 285-289 (Elsevier, 2005).
106. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zídek, A., Potapenko, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
107. Callaway, E. 'It will change everything': Deepmind's AI makes gigantic leap in solving protein structures. *Nature* **588**, 203-204, doi:10.1038/d41586-020-03348-4 (2020).
108. Nussinov, R., Zhang, M. Z., Liu, Y. L., Jang, H. AlphaFold, artificial intelligence (AI), and allostery. *J. Phys. Chem. B.*, doi:10.1021/acs.jpcb.2c04346 (2022).

109. Gao, W. H., Mahajan, S. P., Sulam, J., Gray, J. J. Deep learning in protein structural modeling and design. *Patterns* **1**, doi:10.1016/j.patter.2020.100142 (2020).
110. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Schaeffer, R. D., et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871-876, doi:10.1126/science.abj8754 (2021).
111. Nobel Prize Outreach, A. B., Press release. 2024.
112. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, doi:10.1038/s41586-024-07487-w (2024).
113. Durairaj, J., Waterhouse, A. M., Mets, T., Brodiazhenko, T., Abdullah, M., Studer, G., Tauriello, G., Akdel, M., Andreeva, A., Bateman, A., et al. Uncovering new families and folds in the natural protein universe. *Nature* **622**, 646-653, doi:10.1038/s41586-023-06622-3 (2023).
114. Rose, J. P., Wang, B.-C., Weiss, M. S. Native SAD is maturing. *IUCrJ* **2**, 431-440, doi:10.1107/S2052252515008337 (2015).
115. Vögeli, B., Engilberge, S., Girard, E., Riobé, F., Maury, O., Erb, T. J., Shima, S., Wagner, T. Archaeal acetoacetyl-CoA thiolase/HMG-CoA synthase complex channels the intermediate via a fused CoA-binding site. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3380-3385, doi:10.1073/pnas.1718649115 (2018).
116. Royant, A., Carpentier, P., Ohana, J., McGeehan, J., Paetzold, B., Noirclerc-Savoye, M., Vernède, X., Adam, V., Bourgeois, D. Advances in spectroscopic methods for biological crystals. Fluorescence lifetime measurements. *J. Appl. Crystallogr.* **40**, 1105-1112, doi:10.1107/S0021889807044196 (2007).
117. Lever, M. A., Rogers, K. L., Lloyd, K. G., Overmann, J., Schink, B., Thauer, R. K., Hoehler, T. M., Jørgensen, B. B. Life under extreme energy limitation: A synthesis of laboratory- and field-based investigations. *FEMS Microbiol. Rev.* **39**, 688-728, doi:10.1093/femsre/fuv020 (2015).
118. Palkova, Z. Multicellular microorganisms: Laboratory versus nature. *EMBO Rep.* **5**, 470-6, doi:10.1038/sj.embor.7400145 (2004).
119. Shima, S., Krueger, M., Weinert, T., Demmer, U., Kahnt, J., Thauer, R. K., Ermler, U. Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature* **481**, 98-101, doi:10.1038/nature10663 (2012).
120. Krüger, M., Blumenberg, M., Kasten, S., Wieland, A., Känel, L., Klock, J. H., Michaelis, W., Seifert, R. A novel, multi-layered methanotrophic microbial mat system growing on the sediment of the Black Sea. *Environ. Microbiol.* **10**, 1934-47, doi:10.1111/j.1462-2920.2008.01607.x (2008).
121. Hahn, C. J., Laso-Pérez, R., Vulcano, F., Vaziourakis, K. M., Stokke, R., Steen, I. H., Teske, A., Boetius, A., Liebeke, M., Amann, R., et al. "*Candidatus* Ethanoperedens," a thermophilic genus of *Archaea* mediating the anaerobic oxidation of ethane. *mBio* **11**, doi:10.1128/mBio.00600-20 (2020).
122. Hahn, C. J., Lemaire, O. N., Kahnt, J., Engilberge, S., Wegener, G., Wagner, T. Crystal structure of a key enzyme for anaerobic ethane activation. *Science* **373**, 118-121, doi:10.1126/science.abg1765 (2021).
123. Schopf, J. W. Fossil evidence of Archaean life. *Phil. Trans. R. Soc.* **361**, 869-85, doi:10.1098/rstb.2006.1834 (2006).
124. Stanier, R. Y., Van Niel, C. B. The concept of a bacterium. *Arch. Mikrobiol.* **42**, 17-35, doi:10.1007/BF00425185 (1962).
125. Woese, C. R., Fox, G. E. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *P. Natl. Acad. Sci. U. S. A.* **74**, 5088-5090, doi:10.1073/pnas.74.11.5088 (1977).
126. Woese, C. R., Kandler, O., Wheelis, M. L. Towards a natural system of organisms - proposal for the domains Archaea, Bacteria, and Eucarya. *P. Natl. Acad. Sci. U. S. A.* **87**, 4576-4579, doi:10.1073/pnas.87.12.4576 (1990).
127. Tahon, G., Geesink, P., Ettema, T. J. G. Expanding archaeal diversity and phylogeny: Past, present, and future. *Annu. Rev. Microbiol.* **75**, 359-381, doi:10.1146/annurev-micro-040921-050212 (2021).
128. Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., Hugenholtz, P. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996-1004, doi:10.1038/nbt.4229 (2018).
129. *Genome taxonomic database. Release 09-RS220*, <https://gtdb.ecogenomic.org/> (accessed 15.11.2024).

130. Baker, B. J., De Anda, V., Seitz, K. W., Dombrowski, N., Santoro, A. E., Lloyd, K. G. Diversity, ecology and evolution of *Archaea*. *Nat. Microbiol.* **5**, 887-900, doi:10.1038/s41564-020-0715-z (2020).
131. Eme, L., Spang, A., Lombard, J., Stairs, C. W., Ettema, T. J. G. Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* **16**, 120, doi:10.1038/nrmicro.2017.133 (2017).
132. Karner, M. B., DeLong, E. F., Karl, D. M. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**, 507-10, doi:10.1038/35054051 (2001).
133. Lipp, J. S., Morono, Y., Inagaki, F., Hinrichs, K. U. Significant contribution of Archaea to extant biomass in marine subsurface sediments. *Nature* **454**, 991-994, doi:10.1038/nature07174 (2008).
134. Offre, P., Spang, A., Schleper, C. Archaea in biogeochemical cycles. *Annu. Rev. Microbiol.* **67**, 437-457, doi:10.1146/annurev-micro-092412-155614 (2013).
135. Ingalls, A. E., Shah, S. R., Hansman, R. L., Aluwihare, L. I., Santos, G. M., Druffel, E. R. M., Pearson, A. Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. *P. Natl. Acad. Sci. U. S. A.* **103**, 6442-6447, doi:10.1073/pnas.0510157103 (2006).
136. Needham, D. M., Fuhrman, J. A. Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nat. Microbiol.* **1**, doi:10.1038/Nmicrobiol.2016.5 (2016).
137. Thauer, R. K. Methyl (alkyl)-coenzyme M reductases: Nickel F-430-containing enzymes involved in anaerobic methane formation and in anaerobic oxidation of methane or of short chain alkanes. *Biochemistry* **58**, 5198-5220, doi:10.1021/acs.biochem.9b00164 (2019).
138. Zhu, J. X., Zheng, H. J., Ai, G. M., Zhang, G. S., Liu, D., Liu, X. L., Dong, X. Z. The genome characteristics and predicted function of methyl-group oxidation pathway in the obligate aceticlastic methanogens, *Methanosaeta* spp. *PLoS One* **7**, doi:10.1371/journal.pone.0036756 (2012).
139. Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., et al. Anthropogenic and natural radiative forcing. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, 2013).
140. Feng, L., Palmer, P. I., Parker, R. J., Lunt, M. F., Bösch, H. Methane emissions are predominantly responsible for record-breaking atmospheric methane growth rates in 2020 and 2021. *Atmos. Chem. Phys.* **23**, 4863-4880, doi:10.5194/acp-23-4863-2023 (2023).
141. Liu, Y., Whitman, W. B. Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Ann. N. Y. Acad. Sci.* **1125**, 171-89, doi:10.1196/annals.1419.019 (2008).
142. Ferry, J. G. How to make a living by exhaling methane. *Annu. Rev. Microbiol.* **64**, 453-473, doi:10.1146/annurev.micro.112408.134051 (2010).
143. Gao, Y. H., Wang, Y., Lee, H. S., Jin, P. K. Significance of anaerobic oxidation of methane (AOM) in mitigating methane emission from major natural and anthropogenic sources: a review of AOM rates in recent publications. *Environ. Sci. Adv.* **1**, 401-425, doi:10.1039/d2va00091a (2022).
144. Martens-Habbena, W., Berube, P. M., Urakawa, H., de la Torre, J. R., Stahl, D. A. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* **461**, 976-U234, doi:10.1038/nature08465 (2009).
145. Zhang, L. M., Hu, H. W., Shen, J. P., He, J. Z. Ammonia-oxidizing archaea have more important role than ammonia-oxidizing bacteria in ammonia oxidation of strongly acidic soils. *ISME J.* **6**, 1032-1045, doi:10.1038/ismej.2011.168 (2012).
146. Santoro, A. E., Buchwald, C., McIlvin, M. R., Casciotti, K. L. Isotopic signature of N_2O produced by marine ammonia-oxidizing archaea. *Science* **333**, 1282-1285, doi:10.1126/science.1208239 (2011).
147. Liu, Y., Beer, L. L., Whitman, W. B. Sulfur metabolism in archaea reveals novel processes. *Environ. Microbiol.* **14**, 2632-44, doi:10.1111/j.1462-2920.2012.02783.x (2012).
148. Kletzin, A. Metabolism of inorganic sulfur compounds in archaea. In *Archaea: evolution, physiology, and molecular biology* 261-74 (Blackwell Publ., 2007).
149. Martin-Cuadrado, A. B., Rodriguez-Valera, F., Moreira, D., Alba, J. C., Ivars-Martínez, E., Henn, M. R., Talla, E., López-García, P. Hindsight in the relative abundance, metabolic potential and genome dynamics of uncultivated marine archaea from comparative metagenomic analyses of bathypelagic plankton of different oceanic regions. *ISME J.* **2**, 865-886, doi:10.1038/ismej.2008.40 (2008).

150. Lee, P. A., de Mora, S. J. A review of dimethylsulfoxide in aquatic environments. *Atmos Ocean* **37**, 439-456, doi:10.1080/07055900.1999.9649635 (1999).
151. Garcia, P. S., Gribaldo, S., Borrel, G. Diversity and evolution of methane-related pathways in Archaea. *Annu. Rev. Microbiol.* **76**, 727-755, doi:10.1146/annurev-micro-041020-024935 (2022).
152. Sorokin, D. Y., Merkel, A. Y., Abbas, B., Makarova, K. S., Rijpstra, W. I. C., Koenen, M., Sinninghe Damsté, J. S., Galinski, E. A., Koonin, E. V., van Loosdrecht, M. C. M. *Methanonatronarchaeum thermophilum* gen. nov., sp. nov. and '*Candidatus* Methanohalarchaeum thermophilum', extremely halo(natrono)philic methyl-reducing methanogens from hypersaline lakes comprising a new euryarchaeal class *Methanonatronarchaeia* classis nov. *Int. J. Syst. Evol. Microbiol.* **68**, 2199-2208, doi:10.1099/ijsem.0.002810 (2018).
153. Borrel, G., O'Toole, P. W., Harris, H. M., Peyret, P., Brugère, J. F., Gribaldo, S. Phylogenomic data support a seventh order of methylotrophic methanogens and provide insights into the evolution of methanogenesis. *Genome Biol. Evol.* **5**, 1769-80, doi:10.1093/gbe/evt128 (2013).
154. Wang, Y., Wegener, G., Hou, J., Wang, F., Xiao, X. Expanding anaerobic alkane metabolism in the domain of Archaea. *Nat. Microbiol.* **4**, 595-602, doi:10.1038/s41564-019-0364-2 (2019).
155. Vanwonterghem, I., Evans, P. N., Parks, D. H., Jensen, P. D., Woodcroft, B. J., Hugenholtz, P., Tyson, G. W. Methylotrophic methanogenesis discovered in the archaeal phylum *Verstraetearchaeota*. *Nat. Microbiol.* **1**, 16170, doi:10.1038/nmicrobiol.2016.170 (2016).
156. Hua, Z. S., Wang, Y. L., Evans, P. N., Qu, Y. N., Goh, K. M., Rao, Y. Z., Qi, Y. L., Li, Y. X., Huang, M. J., Jiao, J. Y., et al. Insights into the ecological roles and evolution of methyl-coenzyme M reductase-containing hot spring Archaea. *Nat. Commun.* **10**, 4574, doi:10.1038/s41467-019-12574-y (2019).
157. Evans, P. N., Boyd, J. A., Leu, A. O., Woodcroft, B. J., Parks, D. H., Hugenholtz, P., Tyson, G. W. An evolving view of methane metabolism in the archaea. *Nat. Rev. Microbiol.* **17**, 219-232, doi:10.1038/s41579-018-0136-7 (2019).
158. Wu, K. J., Zhou, L., Tahon, G., Liu, L. Y., Li, J., Zhang, J. C., Zheng, F. F., Deng, C. P., Han, W. H., Bai, L. P., et al. Isolation of a methyl-reducing methanogen outside the Euryarchaeota. *Nature*, doi:10.1038/s41586-024-07728-y (2024).
159. Shima, S., Huang, G., Wagner, T., Ermler, U. Structural basis of hydrogenotrophic methanogenesis. *Annu. Rev. Microbiol.* **74**, 713-733, doi:10.1146/annurev-micro-011720-122807 (2020).
160. Thauer, R. K. Biochemistry of methanogenesis: A tribute to Marjory Stephenson. 1998 Marjory Stephenson prize lecture. *Microbiology* **144**, 2377-2406, doi:10.1099/00221287-144-9-2377 (1998).
161. Kurth, J. M., Op den Camp, H. J. M., Welte, C. U. Several ways one goal—methanogenesis from unconventional substrates. *Appl. Microbiol. Biotechnol.* **104**, 6839-6854, doi:10.1007/s00253-020-10724-7 (2020).
162. Scheller, S., Ermler, U., Shima, S. Catabolic pathways and enzymes involved in anaerobic methane oxidation. In *Anaerobic Utilization of Hydrocarbons, Oils, and Lipids* (ed Boll, M.) 31-59 (Springer International Publishing, 2020).
163. Stams, A. J. M., Teusink, B., Sousa, D. Z. Ecophysiology of acetoclastic methanogens. In *Biogenesis of Hydrocarbons. Handbook of Hydrocarbon and Lipid Microbiology.* Vol. 2 (eds Stams, A.; Sousa, D.) 1-14 (Springer, 2019).
164. Tindall, B. J. The genus name *Methanothrix Huser et al.* 1983 and the species combination *Methanothrix soehngenii* Huser et al. 1983 do not contravene Rule 31a and are not to be considered as rejected names, the genus name *Methanosaeta* Patel and Sprott 1990 refers to the same taxon as *Methanothrix soehngenii* Huser et al. 1983 and the species combination *Methanothrix thermophila* Kamagata et al. 1992 is rejected: Supplementary information to Opinion 75. Judicial Commission of the International Committee on Systematics of Prokaryotes. *Int. J. Sys. Evol. Microbiol.* **64**, 3597-3598, doi:10.1099/ijs.0.069252-0 (2014).
165. Jetten, M. S. M., Stams, A. J. M., Zehnder, A. J. B. Methanogenesis from acetate: A comparison of the acetate metabolism in *Methanothrix soehngenii* and *Methanosarcina* spp. *FEMS Microbiol. Rev.* **8**, 181-197, doi:10.1111/j.1574-6968.1992.tb04987.x (1992).

166. Huser, B. A., Wuhrmann, K., Zehnder, A. J. B. *Methanothrix soehngenii* gen. nov. sp. mov., a new acetotrophic non-hydrogen-oxidizing methane bacterium. *Arch. Microbiol.* **132**, 1-9, doi:10.1007/Bf00690808 (1982).
167. Touzel, J. P., Prensier, G., Roustan, J. L., Thomas, I., Dubourguier, H. C., Albagnac, G. Description of a new strain of *Methanothrix soehngenii* and rejection of *Methanothrix concilii* as a synonym of *Methanothrix soehngenii*. *Int. J. Sys. Bacteriol.* **38**, 30-36, doi:10.1099/00207713-38-1-30 (1988).
168. Zehnder, A. J. B., Huser, B. A., Brock, T. D., Wuhrmann, K. Characterization of an acetate-decarboxylating, non-hydrogen-oxidizing methane bacterium. *Arch. Microbiol.* **124**, 1-11, doi:10.1007/Bf00407022 (1980).
169. Jetten, M. S. M. *Acetate metabolism in Methanothrix soehngenii* Dissertation. Wageningen University and Research, (1991).
170. Jetten, M. S., Stams, A. J., Zehnder, A. J. Isolation and characterization of acetyl-coenzyme A synthetase from *Methanothrix soehngenii*. *J. Bacteriol.* **171**, 5430-5, doi:10.1128/jb.171.10.5430-5435.1989 (1989).
171. Welte, C., Deppenmeier, U. Membrane-bound electron transport in *Methanosaeta thermophila*. *J. Bacteriol.* **193**, 2868-2870, doi:10.1128/Jb.00162-11 (2011).
172. Welte, C., Deppenmeier, U. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochim. Biophys. Acta* **1837**, 1130-47, doi:10.1016/j.bbabio.2013.12.002 (2014).
173. Rotaru, A. E., Shrestha, P. M., Liu, F. H., Shrestha, M., Shrestha, D., Embree, M., Zengler, K., Wardman, C., Nevin, K. P., Lovley, D. R. A new model for electron flow during anaerobic digestion: direct interspecies electron transfer to *Methanosaeta* for the reduction of carbon dioxide to methane. *Energ Environ Sci* **7**, 408-415, doi:10.1039/c3ee42189a (2014).
174. Lovley, D. R. Syntrophy goes electric: Direct interspecies electron transfer. *Annu. Rev. Microbiol.* **71**, 643-664, doi:10.1146/annurev-micro-030117-020420 (2017).
175. Kono, T., Mehrotra, S., Endo, C., Kizu, N., Matusda, M., Kimura, H., Mizohata, E., Inoue, T., Hasunuma, T., Yokota, A., et al. A RuBisCO-mediated carbon metabolic pathway in methanogenic archaea. *Nat. Commun.* **8**, 14007, doi:10.1038/ncomms14007 (2017).
176. Yang, P. X., Tan, G. Y. A., Aslam, M., Kim, J., Lee, P. H. Metatranscriptomic evidence for classical and RuBisCO-mediated CO_2 reduction to methane facilitated by direct interspecies electron transfer in a methanogenic system. *Sci. Rep.* **9**, doi:10.1038/s41598-019-40830-0 (2019).
177. Fey, A., Conrad, R. Effect of temperature on carbon and electron flow and on the archaeal community in methanogenic rice field soil. *Appl. Environ. Microbiol.* **66**, doi:10.1128/Aem.66.11.4790-4797.2000 (2000).
178. Raskin, L., Poulsen, L. K., Noguera, D. R., Rittmann, B. E., Stahl, D. A. Quantification of methanogenic groups in anaerobic biological reactors by oligonucleotide probe hybridization. *Appl. Environ. Microbiol.* **60**, 1241-1248, doi:10.1128/Aem.60.4.1241-1248.1994 (1994).
179. Ferry, J. G. Methane from acetate. *J. Bacteriol.* **174**, 5489-5495, doi:10.1128/Jb.174.17.5489-5495.1992 (1992).
180. Tveit, A. T., Urich, T., Frenzel, P., Svenning, M. M. Metabolic and trophic interactions modulate methane production by Arctic peat microbiota in response to warming. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E2507-16, doi:10.1073/pnas.1420797112 (2015).
181. Angle, J. C., Morin, T. H., Solden, L. M., Narrowe, A. B., Smith, G. J., Borton, M. A., Rey-Sanchez, C., Daly, R. A., Mirfenderesgi, G., Hoyt, D. W., et al. Methanogenesis in oxygenated soils is a substantial fraction of wetland methane emissions. *Nat. Commun.* **8**, doi:10.1038/s41467-017-01753-4 (2017).
182. Mori, K., Iino, T., Suzuki, K., Yamaguchi, K., Kamagata, Y. Aceticlastic and NaCl-requiring methanogen "*Methanosaeta pelagica*" sp. nov., isolated from marine tidal flat sediment. *Appl. Environ. Microbiol.* **78**, 3416-23, doi:10.1128/AEM.07484-11 (2012).
183. Holmes, D. E., Shrestha, P. M., Walker, D. J. F., Dang, Y., Nevin, K. P., Woodard, T. L., Lovley, D. R. Metatranscriptomic evidence for direct interspecies electron transfer between *Geobacter* and *Methanothrix* species in methanogenic rice paddy soils. *Appl. Environ. Microbiol.* **83**, doi:10.1128/AEM.00223-17 (2017).

184. Smith, K. S., Ingram-Smith, C. *Methanosaeta*, the forgotten methanogen? *Trends Microbiol.* **15**, 150-5, doi:10.1016/j.tim.2007.02.002 (2007).
185. Vítezová, M., Kohoutová, A., Vítez, T., Hanisáková, N., Kushkevych, I. Methanogenic microorganisms in industrial wastewater anaerobic treatment. *Processes* **8**, doi:10.3390/pr8121546 (2020).
186. Liu, C. Q., Sun, D. Z., Zhao, Z. Q., Dang, Y., Holmes, D. E. *Methanothrix* enhances biogas upgrading in microbial electrolysis cell via direct electron transfer. *Bioresour. Technol.* **291**, doi:10.1016/j.biortech.2019.121877 (2019).
187. Berger, S., Welte, C., Deppenmeier, U. Acetate activation in *Methanosaeta thermophila*: Characterization of the key enzymes pyrophosphatase and acetyl-CoA synthetase. *Archaea* **2012**, 315153, doi:10.1155/2012/315153 (2012).
188. Duggin, I. G., Aylett, C. H. S., Walsh, J. C., Michie, K. A., Wang, Q., Turnbull, L., Dawson, E. M., Harry, E. J., Whitchurch, C. B., Amos, L. A., et al. CetZ tubulin-like proteins control archaeal cell shape. *Nature* **519**, 362-365, doi:10.1038/nature13983 (2015).
189. Fajardo, A. S., Legrand, P., Payá-Tormo, L., Martin, L., Pellicer Martínez, M. T., Echavarri-Erasun, C., Vernède, X., Rubio, L. M., Nicolet, Y. Structural insights into the mechanism of the radical SAM carbide synthase NifB, a key nitrogenase cofactor maturating enzyme. *J. Am. Chem. Soc.* **142**, 11006-11012, doi:10.1021/jacs.0c02243 (2020).
190. Kunichika, K., Nakamura, R., Fujishiro, T., Takahashi, Y. The structure of the dimeric state of IscU harboring two adjacent [2Fe–2S] clusters provides mechanistic insights into cluster conversion to [4Fe–4S]. *Biochemistry* **60**, 1569-1572, doi:10.1021/acs.biochem.1c00112 (2021).
191. Srivastava, S. S., Jamkhindikar, A. A., Raman, R., Jobby, M. K., Chadalawada, S., Sankaranarayanan, R., Sharma, Y. A transition metal-binding, trimeric βγ-crystallin from methane-producing thermophilic archaea, *Methanosaeta thermophila*. *Biochemistry* **56**, 1299-1310, doi:10.1021/acs.biochem.6b00985 (2017).
192. Verschueren, K. H. G., Blanchet, C., Felix, J., Dansercoer, A., De Vos, D., Bloch, Y., Van Beeumen, J., Svergun, D., Gutsche, I., Savvides, S. N., et al. Structure of ATP citrate lyase and the origin of citrate synthase in the Krebs cycle. *Nature* **568**, 571-575, doi:10.1038/s41586-019-1095-5 (2019).
193. Verstraete, K., Verschueren, K. H. G., Dansercoer, A., Savvides, S. N. Acetyl-CoA is produced by the citrate synthase homology module of ATP-citrate lyase. *Nat. Struct. Mol. Biol.* **28**, 636-638, doi:10.1038/s41594-021-00624-3 (2021).
194. Kissin, Y. V. Catagenesis and composition of petroleum: Origin of *n*-alkanes and isoalkanes in petroleum crudes. *Geochim. Cosmochim.* **51**, 2445-2457, doi:10.1016/0016-7037(87)90296-1 (1987).
195. Simoneit, B. R. T. Petroleum generation, an easy and widespread process in hydrothermal systems: An overview. *Appl. Geochem.* **5**, 3-15, doi:10.1016/0883-2927(90)90031-Y (1990).
196. Claypool, G. E., Kvenvolden, K. A. Methane and other hydrocarbon gases in marine sediment. *Annu. Rev. Earth Planet. Sci.* **11**, 299, doi:10.1146/annurev.ea.11.050183.001503 (1983).
197. Wang, V. C., Maji, S., Chen, P. P., Lee, H. K., Yu, S. S., Chan, S. I. Alkane oxidation: methane monooxygenases, related enzymes, and their biomimetics. *Chem. Rev.* **117**, 8574-8621, doi:10.1021/acs.chemrev.6b00624 (2017).
198. Hinrichs, K. U., Hayes, J. M., Sylva, S. P., Brewer, P. G., DeLong, E. F. Methane-consuming archaebacteria in marine sediments. *Nature* **398**, 802-5, doi:10.1038/19751 (1999).
199. Orphan, V. J., House, C. H., Hinrichs, K. U., McKeegan, K. D., DeLong, E. F. Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science* **293**, 484-7, doi:10.1126/science.1061338 (2001).
200. Teske, A., Hinrichs, K. U., Edgcomb, V., de Vera Gomez, A., Kysela, D., Sylva, S. P., Sogin, M. L., Jannasch, H. W. Microbial diversity of hydrothermal sediments in the Guaymas Basin: Evidence for anaerobic methanotrophic communities. *Appl. Environ. Microbiol.* **68**, 1994-2007, doi:10.1128/AEM.68.4.1994-2007.2002 (2002).
201. Michaelis, W., Seifert, R., Nauhaus, K., Treude, T., Thiel, V., Blumenberg, M., Knittel, K., Gieseke, A., Peterknecht, K., Pape, T., et al. Microbial reefs in the Black Sea fueled by anaerobic oxidation of methane. *Science* **297**, 1013-5, doi:10.1126/science.1072502 (2002).

202. Chadwick, G. L., Skennerton, C. T., Laso-Pérez, R., Leu, A. O., Speth, D. R., Yu, H., Morgan-Lang, C., Hatzenpichler, R., Goudeau, D., Malmstrom, R., et al. Comparative genomics reveals electron transfer and syntrophic mechanisms differentiating methanotrophic and methanogenic archaea. *PLoS Biol.* **20**, e3001508, doi:10.1371/journal.pbio.3001508 (2022).
203. Cui, M., Ma, A., Qi, H., Zhuang, X., Zhuang, G. Anaerobic oxidation of methane: An "active" microbial process. *Microbiologyopen* **4**, 1-11, doi:10.1002/mbo3.232 (2015).
204. Welte, C. U., Rasigraf, O., Vaksmaa, A., Versantvoort, W., Arshad, A., Op den Camp, H. J., Jetten, M. S., Lüke, C., Reimann, J. Nitrate- and nitrite-dependent anaerobic oxidation of methane. *Environ. Microbiol. Rep.* **8**, 941-955, doi:10.1111/1758-2229.12487 (2016).
205. Timmers, P. H., Welte, C. U., Koehorst, J. J., Plugge, C. M., Jetten, M. S., Stams, A. J. Reverse methanogenesis and respiration in methanotrophic archaea. *Archaea* **2017**, 1654237, doi:10.1155/2017/1654237 (2017).
206. Scheller, S., Goenrich, M., Boecher, R., Thauer, R. K., Jaun, B. The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature* **465**, 606-8, doi:10.1038/nature09015 (2010).
207. Haroon, M. F., Hu, S., Shi, Y., Imelfort, M., Keller, J., Hugenholtz, P., Yuan, Z., Tyson, G. W. Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature* **500**, 567-70, doi:10.1038/nature12375 (2013).
208. Arshad, A., Speth, D. R., de Graaf, R. M., Op den Camp, H. J., Jetten, M. S., Welte, C. U. A metagenomics-based metabolic model of nitrate-dependent anaerobic oxidation of methane by *Methanoperedens*-like archaea. *Front. Microbiol.* **6**, 1423, doi:10.3389/fmicb.2015.01423 (2015).
209. Cai, C., Leu, A. O., Xie, G. J., Guo, J., Feng, Y., Zhao, J. X., Tyson, G. W., Yuan, Z., Hu, S. A methanotrophic archaeon couples anaerobic oxidation of methane to Fe(III) reduction. *ISME J.* **12**, 1929-1939, doi:10.1038/s41396-018-0109-x (2018).
210. Leu, A. O., Cai, C., McIlroy, S. J., Southam, G., Orphan, V. J., Yuan, Z., Hu, S., Tyson, G. W. Anaerobic methane oxidation coupled to manganese reduction by members of the *Methanoperedenaceae*. *ISME J.* **14**, 1030-1041, doi:10.1038/s41396-020-0590-x (2020).
211. Boetius, A., Ravenschlag, K., Schubert, C. J., Rickert, D., Widdel, F., Gieseke, A., Amann, R., Jørgensen, B. B., Witte, U., Pfannkuche, O. A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* **407**, 623-6, doi:10.1038/35036572 (2000).
212. Wegener, G., Krukenberg, V., Ruff, S. E., Kellermann, M. Y., Knittel, K. Metabolic capabilities of microorganisms involved in and associated with the anaerobic oxidation of methane. *Front. Microbiol.* **7**, 46, doi:10.3389/fmicb.2016.00046 (2016).
213. Nauhaus, K., Treude, T., Boetius, A., Kruger, M. Environmental regulation of the anaerobic oxidation of methane: A comparison of ANME-I and ANME-II communities. *Environ. Microbiol.* **7**, 98-106, doi:10.1111/j.1462-2920.2004.00669.x (2005).
214. Moran, J. J., Beal, E. J., Vrentas, J. M., Orphan, V. J., Freeman, K. H., House, C. H. Methyl sulfides as intermediates in the anaerobic oxidation of methane. *Environ. Microbiol.* **10**, 162-173, doi:10.1111/j.1462-2920.2007.01441.x (2008).
215. Milucka, J., Ferdelman, T. G., Polerecky, L., Franzke, D., Wegener, G., Schmid, M., Lieberwirth, I., Wagner, M., Widdel, F., Kuypers, M. M. M. Zero-valent sulphur is a key intermediate in marine methane oxidation. *Nature* **491**, 541-546, doi:10.1038/nature11656 (2012).
216. McGlynn, S. E., Chadwick, G. L., Kempes, C. P., Orphan, V. J. Single cell activity reveals direct electron transfer in methanotrophic consortia. *Nature* **526**, 531-5, doi:10.1038/nature15512 (2015).
217. Wegener, G., Krukenberg, V., Riedel, D., Tegetmeyer, H. E., Boetius, A. Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. *Nature* **526**, 587-590, doi:10.1038/nature15733 (2015).
218. Meyerdierks, A., Kube, M., Kostadinov, I., Teeling, H., Glöckner, F. O., Reinhardt, R., Amann, R. Metagenome and mRNA expression analyses of anaerobic methanotrophic archaea of the ANME-1 group. *Environ. Microbiol.* **12**, 422-39, doi:10.1111/j.1462-2920.2009.02083.x (2010).

219. Nauhaus, K., Albrecht, M., Elvert, M., Boetius, A., Widdel, F. *In vitro* cell growth of marine archaeal-bacterial consortia during anaerobic oxidation of methane with sulfate. *Environ. Microbiol.* **9**, 187-196, doi:10.1111/j.1462-2920.2006.01127.x (2007).
220. Laso-Pérez, R., Wegener, G., Knittel, K., Widdel, F., Harding, K. J., Krukenberg, V., Meier, D. V., Richter, M., Tegetmeyer, H. E., Riedel, D., et al. Thermophilic archaea activate butane via alkyl-coenzyme M formation. *Nature* **539**, 396-401, doi:10.1038/nature20152 (2016).
221. Seitz, K. W., Dombrowski, N., Eme, L., Spang, A., Lombard, J., Sieber, J. R., Teske, A. P., Ettema, T. J. G., Baker, B. J. Asgard archaea capable of anaerobic hydrocarbon cycling. *Nat. Commun.* **10**, 1822, doi:10.1038/s41467-019-09364-x (2019).
222. Borrel, G., Adam, P. S., McKay, L. J., Chen, L. X., Sierra-García, I. N., Sieber, C. M. K., Letourneur, Q., Ghozlane, A., Andersen, G. L., Li, W. J., et al. Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nat. Microbiol.* **4**, 603-613, doi:10.1038/s41564-019-0363-3 (2019).
223. Boyd, J. A., Jungbluth, S. P., Leu, A. O., Evans, P. N., Woodcroft, B. J., Chadwick, G. L., Orphan, V. J., Amend, J. P., Rappé, M. S., Tyson, G. W. Divergent methyl-coenzyme M reductase genes in a deep-subseafloor Archaeoglobi. *ISME J.* **13**, 1269-1279, doi:10.1038/s41396-018-0343-2 (2019).
224. Wongnate, T., Sliwa, D., Ginovska, B., Smith, D., Wolf, M. W., Lehnert, N., Raugei, S., Ragsdale, S. W. The radical mechanism of biological methane synthesis by methyl-coenzyme M reductase. *Science* **352**, 953-8, doi:10.1126/science.aaf0616 (2016).
225. Ermler, U., Grabarse, W., Shima, S., Goubeaud, M., Thauer, R. K. Crystal structure of methyl-coenzyme M reductase: The key enzyme of biological methane formation. *Science* **278**, 1457-62, doi:10.1126/science.278.5342.1457 (1997).
226. Grabarse, W., Mahlert, F., Shima, S., Thauer, R. K., Ermler, U. Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: unusual amino acid modification, conservation and adaptation. *J. Mol. Biol.* **303**, 329-44, doi:10.1006/jmbi.2000.4136 (2000).
227. Wagner, T., Kahnt, J., Ermler, U., Shima, S. Didehydroaspartate modification in methyl-coenzyme M reductase catalyzing methane formation. *Angew. Chem. Int. Ed Engl.* **55**, 10630-3, doi:10.1002/anie.201603882 (2016).
228. Wagner, T., Wegner, C. E., Kahnt, J., Ermler, U., Shima, S. Phylogenetic and structural comparisons of the three types of methyl coenzyme M reductase from *Methanococcales* and *Methanobacteriales*. *J. Bacteriol.* **199**, doi:10.1128/JB.00197-17 (2017).
229. Chadwick, G. L., Joiner, A. M. N., Ramesh, S., Mitchell, D. A., Nayak, D. D. McrD binds asymmetrically to methyl-coenzyme M reductase improving active-site accessibility during assembly. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2302815120, doi:10.1073/pnas.2302815120 (2023).
230. Wagner, T., Koch, J., Ermler, U., Shima, S. Methanogenic heterodisulfide reductase (HdrABC-MvhAGD) uses two noncubane [4Fe-4S] clusters for reduction. *Science* **357**, 699-703, doi:10.1126/science.aan0425 (2017).
231. Goodsell, D. *Molecule of the month: Methyl-coenzyme M reductase*, <https://pdb101.rcsb.org/motm/179> (accessed 14.11.2024).
232. Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W., Hedderich, R. Methanogenic archaea: Ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **6**, 579-91, doi:10.1038/nrmicro1931 (2008).
233. Wagner, T., Ermler, U., Shima, S. The methanogenic CO_2 reducing-and-fixing enzyme is bifunctional and contains 46 [4Fe-4S] clusters. *Science* **354**, 114-117, doi:10.1126/science.aaf9284 (2016).
234. Arcondéguy, T., Jack, R., Merrick, M. P_{II} signal transduction proteins, pivotal players in microbial nitrogen control. *Microbiol. Mol. Biol. Rev.* **65**, 80-105, doi:10.1128/MMBR.65.1.80-105.2001 (2001).
235. Huergo, L. F., Dixon, R. The emergence of 2-oxoglutarate as a master regulator metabolite. *Microbiol. Mol. Biol. Rev.* **79**, 419-35, doi:10.1128/MMBR.00038-15 (2015).
236. Conroy, M. J., Durand, A., Lupo, D., Li, X. D., Bullough, P. A., Winkler, F. K., Merrick, M. The crystal structure of the *Escherichia coli* AmtB-GlnK complex reveals how GlnK regulates the ammonia channel. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 1213-8, doi:10.1073/pnas.0610348104 (2007).

237. Gruswitz, F., O'Connell, J., 3rd, Stroud, R. M. Inhibitory complex of the transmembrane ammonia channel, AmtB, and the cytosolic regulatory protein, GlnK, at 1.96 Å. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 42-7, doi:10.1073/pnas.0609796104 (2007).
238. Pedro-Roig, L., Lange, C., Bonete, M. J., Soppa, J., Maupin-Furlow, J. Nitrogen regulation of protein-protein interactions and transcript levels of GlnK PII regulator and AmtB ammonium transporter homologs in archaea. *MicrobiologyOpen* **2**, 826-40, doi:10.1002/mbo3.120 (2013).
239. Maier, S., Schleberger, P., Lü, W., Wacker, T., Pflüger, T., Litz, C., Andrade, S. L. Mechanism of disruption of the Amt-GlnK complex by P_{II}-mediated sensing of 2-oxoglutarate. *PLoS One* **6**, e26327, doi:10.1371/journal.pone.0026327 (2011).
240. Müller, M. C., Wagner, T. The oxoglutarate binding site and regulatory mechanism are conserved in ammonium transporter inhibitors GlnKs from *Methanococcales*. *Int. J. Mol. Sci.* **22**, doi:10.3390/ijms22168631 (2021).
241. Eisenberg, D., Gill, H. S., Pfluegl, G. M., Rotstein, S. H. Structure-function relationships of glutamine synthetases. *Biochim. Biophys. Acta* **1477**, 122-45, doi:10.1016/s0167-4838(99)00270-8 (2000).
242. Ehlers, C., Weidenbach, K., Veit, K., Forchhammer, K., Schmitz, R. A. Unique mechanistic features of post-translational regulation of glutamine synthetase activity in *Methanosarcina mazei* strain Gö1 in response to nitrogen availability. *Mol. Microbiol.* **55**, 1841-54, doi:10.1111/j.1365-2958.2005.04511.x (2005).
243. Pedro-Roig, L., Camacho, M., Bonete, M. J. Regulation of ammonium assimilation in *Haloferax mediterranei*: Interaction between glutamine synthetase and two GlnK proteins. *Biochim. Biophys. Acta* **1834**, 16-23, doi:10.1016/j.bbapap.2012.10.006 (2013).
244. Belay, N., Sparling, R., Daniels, L. Dinitrogen fixation by a thermophilic methanogenic bacterium. *Nature* **312**, 286-8, doi:10.1038/312286a0 (1984).
245. Murray, P. A., Zinder, S. H. Nitrogen fixation by a methanogenic archaebacterium. *Nature* **312**, 284-286, doi:10.1038/312284a0 (1984).
246. Harris, D. F., Lukoyanov, D. A., Kallas, H., Trncik, C., Yang, Z. Y., Compton, P., Kelleher, N., Einsle, O., Dean, D. R., Hoffman, B. M., et al. Mo-, V-, and Fe-nitrogenases use a universal eight-electron reductive-elimination mechanism to achieve N_2 reduction. *Biochemistry* **58**, 3293-3301, doi:10.1021/acs.biochem.9b00468 (2019).
247. Dodsworth, J. A., Leigh, J. A. NifI inhibits nitrogenase by competing with Fe protein for binding to the MoFe protein. *Biochem. Biophys. Res. Commun.* **364**, 378-82, doi:10.1016/j.bbrc.2007.10.020 (2007).
248. Maslać, N. *Structural and biochemical characterisation of N_2-assimilation system and its regulation in thermophilic methanogenic archaea* Dissertation. University Bremen, doi:10.26092/elib/2553, (2023).
249. Müller, M.-C., Lemaire, O. N., Kurth, J. M., Welte, C. U., Wagner, T. Differences in regulation mechanisms of glutamine synthetases from methanogenic archaea unveiled by structural investigations. *Commun. Biol.* **7**, 111, doi:10.1038/s42003-023-05726-w (2024).

CHAPTER II.

Crystallomics on *Methanocaldococcus infernus*

Main

The following results are preliminary, and some studies have not been completed. However, insights from the experiments are valuable and therefore presented briefly.

The untargeted "shotgun" simultaneous purification approach was first performed on a pure culture to evaluate whether blind protein purification and pure *de novo* structure solving is feasible. A major hurdle in this process is the phase problem[1].

During the diffraction experiment the intensities of waves scattered from the planes in the crystal are measured. The recorded diffraction patterns are then utilized to calculate the 3D electron density. For this, three main parameters are required: the unit cell, the amplitude and their associated phase. While the wave amplitude can be extracted from the spot intensity and the unit cell from the spot positions, the phase is lost in the experiment. The only relationship between amplitudes and phases exists through the structure itself. Therefore, prior knowledge of the electron density or structure is required for all phasing methods. If a homology model is available approximate phases can be generated during molecular replacement.

However, this approach cannot be applied for solving novel enzyme structures. Alternatively, anomalous scattering can be utilized. For some elements, at a specific absorption edge the X-ray photon energy can promote an electron from the inner shell, this creates an anomalous difference in the scattering behavior that can be used to localize the anomalous scatterers and derive the phase. The method of single-wavelength anomalous diffraction (SAD) includes the data collection at a single wavelength, typically the absorption peak[2]. Experiments are often carried out through the introduction of derivatives of scattering atoms such as selenium in the form of Se-

methionine which is not possible in the case of *de novo* structure determination from microorganisms lacking a genetically accessible system. Alternatively, co-crystallization or soaking with heavy elements[3] or specialized crystallophores[4] can be used to introduce anomalous scatterers. In a more challenging approach intrinsic scatterers like metal atoms in cofactors or prosthetic groups (*e.g.*, [Fe-S] clusters, heme groups) and even intrinsic sulfur atoms can be targeted[2]. However, especially the latter elements are only weak scatterers. Moreover, SAD experiments require a specific data collection strategy which must be performed at discrete wavelengths. For some scattering atoms, long wavelengths must be reached, these experiments can only be performed at specialized beamlines and for an extended time to obtain sufficient redundancy which increases the risk of introducing radiation damage. Here, three different phasing approaches were utilized to solve structures obtained through the crystallomics approach.

Methanocaldococcus infernus was chosen as the target organism for the approach. It was isolated from a deep sea hydrothermal vent at the mid-atlantic ridge and can be grown in pure culture with an optimal temperature of 85 °C[5]. The hyperthermophilic hydrogenotrophic methanogen might exhibit interesting protein adaptations for thriving at such high temperatures and the high rigidity of its hyperthermophilic enzymes could facilitate crystallization[6].

Cultivation was performed in a 7 L fermenter at 75 °C with H_2, CO_2, $NaNO_3$, Na_2SeO_3 and Na_2SO_3. For the purification 23.5 g of cells were lysed (yielding 1.9 g soluble protein) and then processed by multiple chromatography steps in an anaerobic environment (Fig. 1). Proteins were followed during the purification by their absorbance at 280, 415 and 550 nm in the elution fractions and on SDS-PAGE so that major protein populations could be identified and separated into pools (detailed methods are listed in supplementary).

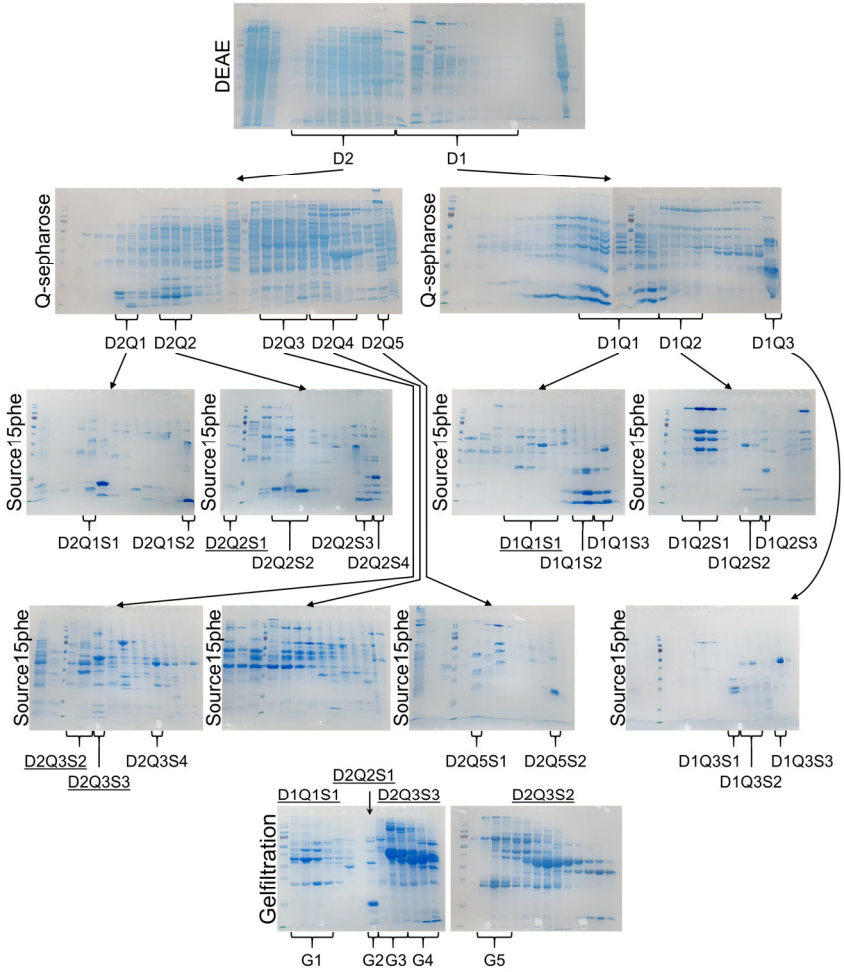


Figure 1. Crystallomics on *Methanocaldococcus infernus.* SDS-PAGE of all chromatography steps. The type of chromatography is detailed on the left of each gel. Fractions of the elution are loaded from left to right, the ladder includes the following sizes from top to bottom: 180, 130, 100, 70 (red), 55, 40, 35, 25, 15 and 10 (green) kDa. Pools of interest are labeled by brackets with their assigned name, arrows highlight purification paths for all pools. Underlined pools were passed on a final gel filtration step shown at the bottom. The used ladder contains the following molecular weight markers (in kDa) from top to bottom: 180, 130, 100, 70 (red), 55,40,35,25,15,10 (green).

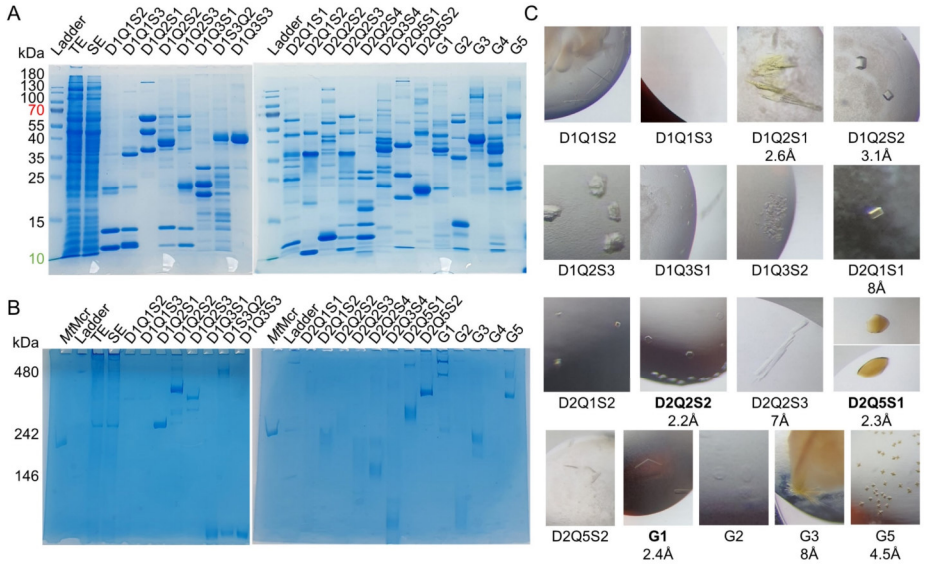


Figure 2. Final fractions and obtained crystals. (**A**) SDS-PAGE and (**B**) hrCN-PAGE of the total extract (TE), soluble extract (SE) final obtained pools (names according to Fig. 1). MCR from *Methanothermococcus thermolithotrophicus* (*Mt*MCR) was loaded as a control. (**C**) Obtained crystals from the final fractions, if X-ray diffraction was measured the obtained resolution is listed below. Bold fractions have yielded structures.

A total of 22 final fractions were separated (Fig. 2A-B, Table S1). Most enriched fractions contained a single, relatively pure, protein population though some were still contaminated with multiple other proteins. All fractions were concentrated to <100 µL and crystallized anaerobically with a crystallization robot to reduce the required sample amounts. From the 22 crystallized fractions 17 yielded crystals or microcrystals after one month, the latter might be exploitable for microED experiments. Due to the unexpectedly high amount of crystal hits (Fig. 2C), focus was put on crystals with a decent size to perform X-ray diffraction experiments. While some did only diffract to low resolution, three exploitable datasets were obtained (Table S2). Additionally, another dataset was obtained for the final fraction D1Q2S1 which was identified as

MCR by its color, size and SDS-PAGE profile. The dataset could potentially reach a resolution of ~2.6 Å, however it was not further processed in order to prioritize the other datasets and due to its lower quality, including twinning, potentially affected by crystallization at a too high concentration.

The first dataset (fraction D2Q2S2) was solved via SAD utilizing the native sulfur atoms of the protein (Fig. 3A). Substructure solving, phasing and backbone tracing was performed with ShelX C/D/E. In the initial backbone model the characteristic fold of a P_{II} regulatory protein was recognized, informed by prior work on this protein family[7]. Four P_{II} sequences were found in the genome of *M. infernus* and one could be immediately excluded, as an insertion caused the total amino acid count to increase beyond the suggested count in the initial model. The correct sequence among the three remaining ones was determined by inspecting the electron density protruding from the initial backbone of the experimental map to distinguish sidechain identities (Fig. 3B). Only partial densities are present at the canonical ligand binding sites for ATP, Mg^{2+} and 2OG. The T-loops, responsible for protein-protein interaction, are positioned similar to the dADP bound state observed in *M. thermolithotrophicus*[7] (Fig. 3C). Further analysis will be required to resolve whether the density reflects a weakly bound

dADP or compounds of the crystallization solution mimicking the binding. Due to time restraints, the structure was not further refined. The closest annotated homolog was found to be $GlnK_1$ from *Methanocaldococcus jannaschii* (WP_010869551.1 Cov:100 % ID:83.93 %). GlnK proteins are involved in the regulation of ammonium import[8, 9] and ammonium assimilation[10-12] to optimize cellular energy expenses for N-acquisition. They are differentially expressed at growth on different nitrogen sources and the production of some is upregulated under nitrogen limitation[13]. Previous work has also demonstrated the protein's strong crystallization properties[7]. The fact that GlnK from *M. infernus* was identified through the crystallomics approach underscores its critical

cellular role and demonstrates that the approach can capture proteins beyond the main metabolic enzymes.

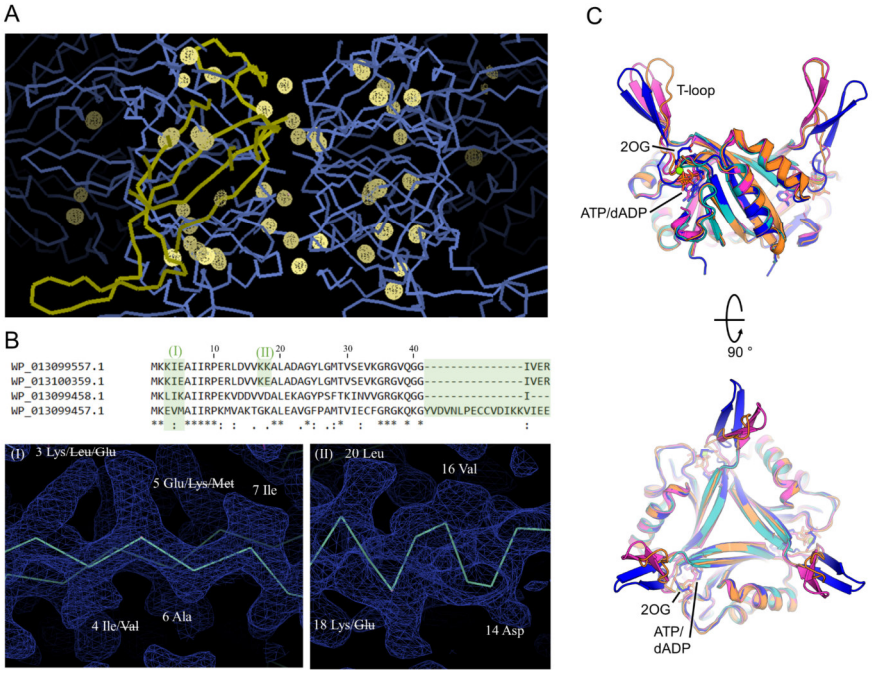


Figure 3. S-SAD, *ab initio* sequencing and preliminary structure of *M. infernus* GlnK. (**A**) Sulfur anomalous signal shown as yellow mesh. The asymmetric unit content is shown in yellow, the symmetry mates in blue ribbons. (**B**) Stretches with deviating sequences were identified in the alignment of the potential candidates (highlighted in green) and inspected in the electron density (blue mesh $2F_o$-F_c at 1σ). Roman numbers indicated corresponding sequence stretches and model sections. All listed sequences are annotated as: P-II family nitrogen regulator. (**C**) Superposition of the initial *M. infernus* GlnK model (orange) with $GlnK_1$ + dADP (PDB: 7P4V, pink), $GlnK_2$ apo (7P4Y, teal) and $GlnK_2$ + Mg^{2+}, ATP and 2OG (7P50, blue) from *M. thermolithotrophicus*. Structures are shown in cartoon and ligands as sticks with the atoms O in red, N in blue, P in orange, Mg in light green.

The second dataset (fraction G1) was solved by soaking the crystals with the terbium (Tb)-containing crystallophore TbXo4[4] followed by SAD at the L_{III}-Tb edge. The substructure was solved by ShelX C/D/E and a poly-alanine model was generated (Fig. 4A). The generated backbone was then used for fold-similarity analysis which identified the target as a phosphoenolpyruvate hydratase / enolase. Only a single archaeal enolase has been deposited in the PDB (PDB: 2PA6 from *M. jannaschii*) and its description remains to be published. The next close structurally characterized homolog is from *B. subtilis* (4A3R). Structural comparison to the initial *M. infernus* enolase model highlights a strong conservation of the homo-octameric complex (root mean square deviation (rmsd) of 0.334 Å for 742 atoms for *M. jannaschii* and 0.581 Å for 699 atoms for *B. subtilis,* Fig. 4B). No ligands could be identified in the electron density and both obtained chains are present in the open conformation as shown by superposition to the open and closed chains captured for the homolog in *E. coli* (1E9I, Fig. 4C).

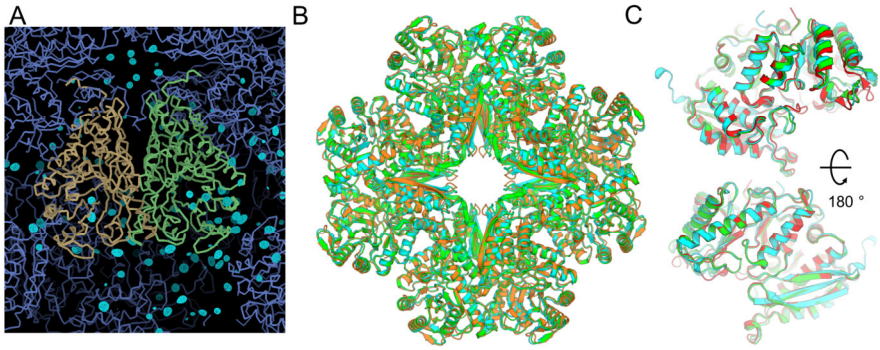


Figure 4. Enolase of *M. infernus*. (**A**) Anomalous map of Tb utilized for phasing shown as cyan mesh at 5σ. (**B**) Superposition of *M. infernus* enolase (cyan) to homologs from *M. jannaschii* (2PA6, green) and *B. subtilis* (4A3R, orange). (**C**) Superposition with the open and closed chain of *E. coli* enolase. Rmsd of 0.675/0.627 for 366/356 atoms to the open chain A for the *M. infernus* enolase chain A and B respectively. Rmsd of 0.859/0.850 for 339/345 atoms to the closed chain D for the *M. infernus* enolase chain A and B respectively.

M. infernus likely utilizes the enolase to convert phosphoenolpyruvate and H_2O to 2-phospho-D-glycerate during the reverse modified archaeal Embden-Meyerhof-Parnas (EMP) pathway operating in the direction of gluconeogenesis[14]. Generated carbohydrates can be utilized for glycogen storage or biosynthesis. Due to time constraints, the model was not refined further.

The 2-oxoglutarate:ferredoxin oxidoreductase of *Methanocaldococcus infernus*

The final target (D2Q5S1) immediately sparked interest due to the brown-colored crystals. An X-ray fluorescence scan confirmed the presence of iron. Therefore, a dataset was collected at the Fe K-edge to perform a SAD experiment using the intrinsic anomalous scatterer as a phasing agent (see below). Again, the backbone was modelled, followed by a fold-similarity analysis. The enzyme was identified as a member of the 2-oxoacid:ferredoxin oxidoreductase (OFOR) superfamily. The operon consists of four genes encoding for an α-, β-, and γ-subunit as well as a ferredoxin family protein that constitutes the δ-subunit (Fig. 5, Table S3). A comparison to homologs shows a diverse set of operon organizations and organisms regularly possess multiple operons likely encoding variants specialized for different substrates. While some operons contain a ferredoxin-like subunit, others do not.

The enzyme complex consists of a dimer of hetero-tetramers (2(αβγδ)). Each half contains an active site and three [4Fe-4S] clusters, the first being coordinated by the β-subunit and the other two by the δ-subunit. Previously characterized homologs include the 2-oxoglutarate: ferredoxin oxidoreductase from *Magnetococcus marinus* (PDB: 6N2O), the oxalate oxidoreductase from *Moorella thermoacetica* (5C4I) and pyruvate:ferredoxin oxidoreductases from *Moorella thermoacetica* (6CIN), *Desulfocurvibacter africanus* (7PLM) and *Methanosarcina acetivorans* (9BT4)[15-19] (Fig. 6). These structures capture a diversity of complex assemblies and some contain fused γ- and δ-subunits. Regarding the complex assembly *Mi*OFOR is most similar to

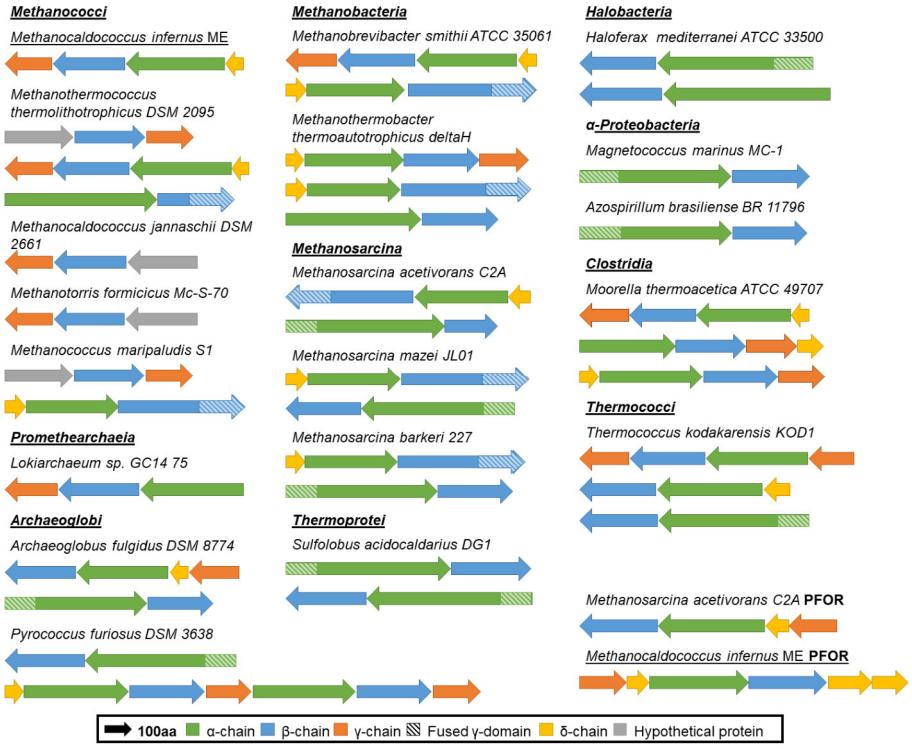


Figure 5. Operon organization of 2-oxoacid:ferredoxin oxidoreductase homologs. Homolog operons were retrieved by manual basic local alignment search tool (BLAST) search against the *Mi*OFOR β-subunit and confirmed by BLAST against the corresponding α/β/γ/δ of *Mi*OFOR. Subunits are colored according to the legend. Hypothetical proteins among the *Methanococci* did not yield a match against the α-subunit. Full accession numbers, annotation, identity and coverage can be found in Table S3.

the pyruvate:ferredoxin oxidoreductase (PFOR) from *Methanosarcina acetivorans* as both complexes contain four separate chains. However, sequence similarity is rather low (23.68, 27.5, 28.09 and 47.37 % identity of the α-, β-, γ- and δ-subunit respectively). Inspection of the active site reveals a more opened accessible conformation in *Mi*OFOR (Fig. 6C). Superposition of both complexes shows that the γ- and δ-subunits are turned

outwards. The [4Fe-4S] clusters are shifted further away from the active site, the proximal cluster is rotated slightly, the medial cluster is shifted by ~2.5 Å and the distal cluster is shifted by ~3 Å. The more opened conformation might be caused by the absence of the cofactor thiamine pyrophosphate (TPP) which is present in all available homologous structures while the obtained structure represents the apo state. TPP is likely contributing to the stabilization of the active site. The cofactor might have been lost during the purification (*e.g.*, due to the exposure to high salt concentrations during the hydrophobic interaction chromatography). Further structural analysis could provide

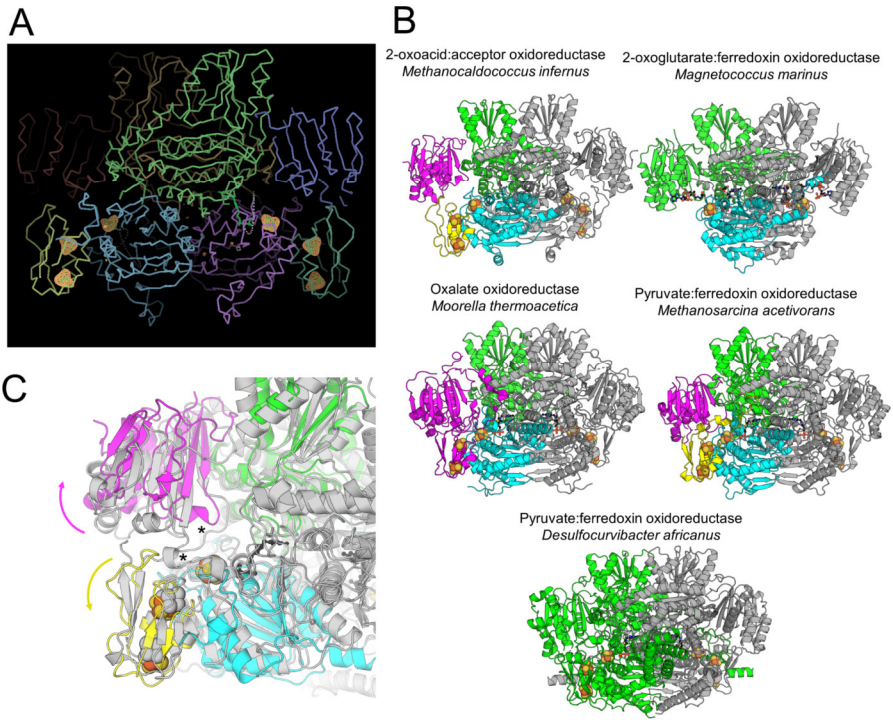


Figure 6. The structure of OFOR from *M. infernus* and homologs. (**A**) Anomalous map of Fe shown as orange mesh at 2.5 σ. (**B**) Homologous OFOR structures. Models are shown as cartoon. The subunits are colored: green for the α-, blue for the β-, pink for the γ- and yellow for the δ-chain. The dimeric equivalent is shown in grey.

Figure 6. cont.[4Fe-4S] clusters are shown as spheres with yellow for S and brown for Fe. Ligands including TPP are shown as black balls-sticks. PDB codes are: 6N2O, 5C4I, 9BT4, 7PLM (from top left to bottom right). (**C**) Zoom in to the active site superposition of *Mi*OFOR to PFOR from *M. acetivorans* (shown in grey). Arrows illustrate the shifted subunits and stars mark the active site loops absent in *Mi*OFOR.

insights into the mechanism and structural remodeling during the cofactor binding. For this it would be beneficial to study intermediate states. To maintain the cofactor during purification, a targeted purification could be carried out to decrease the total purification time and TPP could be added during the last purification step, concentration or co-crystallization. Besides the lack of TPP *Mi*OFOR deviates from the *M. acetivorans* PFOR through the absence of some loops. In the *Ma*PFOR an extended N-terminus of the δ-subunit and loop 37-46 of the γ-subunit are located near the active site and might further contribute to subunit binding. Their absence in *Mi*OFOR might further contribute to the observed apo state.

The enzyme was further investigated with the goal of determining the function as the annotation of a 2-oxoacid:acceptor oxidoreductase was ambiguous. A streamlined purification protocol was performed to purify the enzyme again, and the same complex could be obtained, highlighting the reproducibility of the purification protocol.

On the basis of the annotation the following enzyme reaction was expected:

$$\textbf{2-oxoacid}_{n} + \textbf{CoA} + \textbf{Acceptor}_{ox} \rightleftharpoons \textbf{CO}_2 + \textbf{Acceptor}_{red} + \textbf{Acyl}_{n-1}\textbf{-CoA}$$

To determine the substrate specificity, activity assays were performed. Preliminary results indicate substrate specificity for 2-oxoglutarate (2OG). Under the tested conditions no activity could be detected with other 2-oxoacids (oxalate, pyruvate, 2-oxobutyrate, 2-oxoadipic acid, phenyl pyruvate, 4-methyl-2-oxovalerate, 3-methyl-2-oxobutanoic acid, Fig. 7). Higher rates could be obtained through the addition of TPP,

however activities could also be obtained without TPP addition. This could either indicate that the cofactor could be successfully maintained during the faster second purification or that the purified pool contains a mixture of the apo and holo form of which the apo form was favored in the tested crystallization conditions. Due to large standard deviations, these assays need to be repeated. Nevertheless, these results suggest that the enzyme isolated from *M. infernus* functions as a 2-oxoglutarate:ferredoxin oxidoreductase (*Mi*OGOR). Designing future experiments to follow activity at 80 °C would be beneficial to track physiological rates. The structure-function relationship corroborates the enzyme's specificity for 2-oxoglutarate as the three-dimensional arrangements of the residues Ile30, Thr31, Arg107, Thr112 and Arg56 near the active site are conserved compared to OGOR from *M. marinus* (*Mm*OGOR)[15]. These residues differ compared to canonical pyruvate:ferredoxin oxidoreductases[18, 20] or the *M. thermoacetica* homolog utilizing oxalate[16]. In-depth analysis of OFOR sequences and the conservation of active site residues could provide insights into substrate specificity for biochemically uncharacterized variants, adding context to generalized annotations.

OGOR from *M. marinus* has been reported to function as a CO_2-fixing 2OG synthase *in vivo*[15], while PFOR from *M. acetivorans* has been proposed to switch directionality based on the metabolic needs reflected by the available pools of free coenzyme-A and acetyl-CoA[20]. Autotrophic methanogens utilize the enzyme for pyruvate production for anabolic reactions. Interestingly, OGORs have been described to often lack the δ-subunit[21] but *Mi*OGOR includes this domain. Follow-up protein phylogenetic analyses would be beneficial to resolve whether *Mi*OGOR represents an exception or if other OGORs with ferredoxin domains are common within archaea. Such analyses could also trace possible deletion or horizontal gene transfer events.

This enzyme is a noteworthy discovery among the extracted proteins prompting further investigations. Its complex assembly and different cofactors underscore the methods

capability to capture sophisticated enzymes. Moreover, it emphasizes the potential of crystallomics to not only capture enzymes of the main catabolism but also important players of the anabolism. In summary, this test of the approach on a pure culture highlighted the feasibility of blind purification and *de novo* structure solving was successfully performed by the utilization of three different SAD approaches.

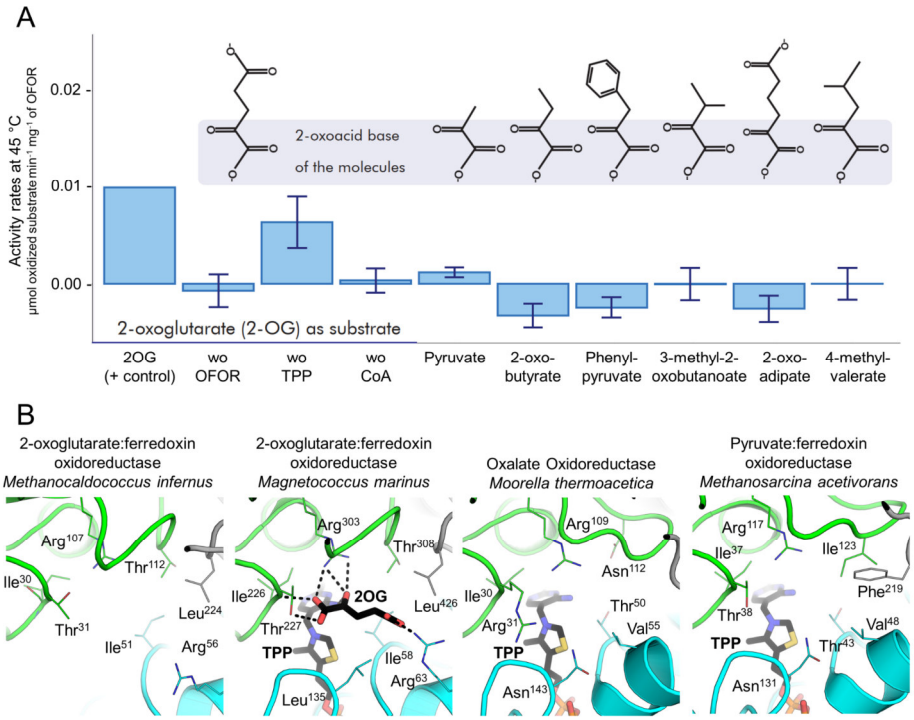


Figure 7. Substrate specificity of *Mi*OFOR. (**A**) Preliminary results: activity followed through methyl-viologen reduction. 2-oxoacid structures are shown above the corresponding activity assay. Standard deviations are indicated by error bars, n=3. The first condition exhibited very large deviations and must be repeated. Figure by Kinga Sántha who performed the activity assays. (**B**) Active site comparison between *Mi*OFOR and OFOR members. PDB codes 6N2O, 5C4I, 6CIN.The protein backbone is shown as ribbon, ligands as sticks and adjacent residues as lines.

Supplementary material

Fermenter cultivation medium for *M. infernus* (DSM 11812)

4.1 mM KH_2PO_4, 13.4 mM KCl, 430 mM NaCl, 10 mM $NaHCO_3$, 2.5 mM $CaCl_2$ x 2 H_2O, 38 mM $MgCl_2$ x 6 H_2O, 10 mM $NaNO_3$, 10 mM Na_2SO_4, 0.32 mM nitrilotriacetic acid, 0.031 mM $FeCl_2$ x 4 H_2O, 0.01 mM Na_2Wo_4 x 2 H_2O, 0.01 mM Na_2MoO_4 x 2 H_2O, 50 mM 2-(N-morpholino)ethanesulfonic acid (MES) were dissolved in demineralized water (dH_2O), then 0.015 mM resazurin and 10 mL trace element solution were added. The pH was set with NaOH. The volume was filled up with dH_2O. 5 µM Na_2SeO_3 x 5 H_2O, 2 mM Na_2SO_3 were added freshly before inoculation

Trace element solution (100x concentrated): 7.1 mM Nitrilotriacetic acid (pH set to 6.5 with KOH), 0.45 mM $MnCl_2$ x 6 H_2O, 0.68 mM $FeCl_3$ x 6 H_2O, 0.76 mM $CaCl_2$ x 2 H_2O, 0.76 mM $CoCl_2$ x 6 H_2O, 0.66 mM $ZnSO_4$ x 7 H_2O, 0.28 mM $CuSO_4$ x 5 H_2O, 0.19 mM Na_2MoO_4 x 2 H_2O, 0.38 mM $NiCl_2$ x 6 H_2O. The pH was set to 6.3 and the volume was filled up with dH_2O. The solution was stored at 4 °C in the dark.

Cell lysis and protein purification

The cell lysis was performed in a N_2/CO_2 atmosphere (90:10 %) and the purification was carried out in a N_2/H_2 atmosphere (97:3 %) with yellow light. The pH of all buffers was either set with HCl or NaOH. 23.5 g cells were resuspended in IEC A buffer (50 mM tricine pH 8, 2 mM dithiothreitol (DTT)) and lysed by sonication (4 x 10 sec at 70 % with a MS73 probe (Bandelin, Germany)) followed by centrifugation at 45,000 × *g* for 45 min at 18 °C. The supernatant was filtered through a 0.2 µm filter (Sartorius, Germany) and loaded on a 4 x 5 mL HiTrap™ DEAE Sepharose FF column (Cytiva, Sweden). Elution was performed at 0-60 % IEC B (50 mM tricine pH 8, 1 M NaCl, 2 mM DTT) at 2 $mL.min^{-1}$. Fractions were pooled into two pools: D1 300-460 mM and D2 80-260 mM NaCl, both pools were diluted 1:3 with IEC A. Pool D1 was filtered and loaded on a 2 x 5 mL HiTrap Q HP™ column (Cytiva, Sweden). Elution was performed

at 15-55 % IEC B at 1.5 mL.min^{-1}. Three pools were obtained: D1Q1 370-410 mM, D1Q2 420-450 mM and D1Q3 550-10000 mM NaCl. Pool D2 was also filtered and loaded on a Q-sepharose column. Elution was performed at 0-40 % IEC B at 1 mL.min^{-1}. Fractions were pooled into: D2Q1 120-140 mM, D2Q2 155-170 mM, D2Q3 235-270 mM, D2Q4 285-325 mM and D2Q5 400-1000 mM NaCl. All pools obtained from the Q-sepharose runs were diluted 1:3 with HIC B buffer (25mM tris pH 7.6, 2 M $(NH_4)_2SO_4$, 2 mM DTT), filtered and subsequently separately loaded on a Source^{TM}15 Phe 4.6/100 PE column (Cytiva, Sweden), elution was always performed at 75-0 % HIC B (against buffer without $(NH_4)_2SO_4$) at 1 mL.min^{-1}. Following fractions were obtained: D1Q1S1 740-460 mM, D1Q1S2 290-160 mM, D1Q1S3 140-12 mM, D1Q2S1 960-740 mM, D1Q2S2 430-280 mM, D1Q2S3 200-80 mM, D1Q3S1 700-580 mM, D1Q3S2 540-420 mM, D1Q5S1 980-860 mM, D2Q1S1 1000-900 mM, D2Q1S2 60-0 mM, D2Q2S1 2000 mM, D2Q2S2 900-670 mM, D2Q2S3 100-0 mM, D2Q2S4 0 mM, D2Q3S1 1500 mM, D2Q3S2 1500-1040 mM, D2Q3S3 1020-700 mM, D2Q3S4 160-40 mM and D2Q4S1 2000-1040 mM $(NH_4)_2SO_4$. Pools D1Q1S1, D2Q2S1 and D2Q3S3 were concentrated to ~ 200 µL, filtered and loaded on a Superdex™ 200 Increase 10/300 GL (Cytiva, Sweden). Elution was performed at 0.4 mL.min^{-1} GF buffer (25 mM tris pH 7.6, 10 % glycerol, 2 mM DTT) and four pools were obtained: (D1Q1S1)G1 8.5-10.1 mL, (D2Q2S1)G2 11.7-12.9 mL, (D2Q3S3)G3 10.3-11.5 mL and G4 11.9-13.5, (D2Q3S2)G5 9.7-10.1 mL. All final pools were washed 1:1000 in GF buffer (apart from the pools of the gel filtration step) and concentrated to ~ 60 µL. Protein concentrations were estimated via Bradford assay using bovine serum albumin as a standard (Table S1). Finally, all fractions were loaded on SDS-PAGE, 20 µg total and soluble extract were loaded and 5 µg for all other fractions. Additionally, 10 µg of the total and soluble extract and 2 µg of all other fractions were loaded on a hrCN-PAGE.

Crystallization setup

Crystallization was performed anaerobically followed by storage in a N_2/H_2 atmosphere (97:3 %). 0.4 µL protein was spotted with 0.4 µL crystallization solution on 96-Well MRC 2-Drop Crystallization Plates in polystyrene (SWISSCI, UK) with an OryxNano (Douglas instruments, UK) with 90 µL in the reservoir. The following crystallization kits were used: Wizard 1&2, Wizard 3&4, JCSG++, PEP (Jena Bioscience, Germany). Crystals yielding datasets were obtained in 30 % v/v Polyethylene glycol 400, 100 mM Tris pH 8.6, 200 mM MgCl for D2Q2S2 (GlnK), in 45 % w/v Pentaerythritol propoxylate, 100 mM Tris pH 8.5 for G1 (enolase) and 25.0 % w/v polyethylene glycol 1,500, 100 mM 2:7:7 Succinic acid:Sodium dihydrogen phosphate:Glycine (SPG) buffer pH 8.5 for D2Q5S1 (OGOR).

Data collection and processing

X-ray crystallography datasets were collected at: Deutsches Elektronen-Synchrotron (DESY) beamline P11 for OGOR, Swiss light source (SLS) beamline PXIII for GlnK and synchrotron SOLEIL beamline PROXIMA-1 for enolase. In all cases, the data processing, experimental phasing and poly-alanine backbone model building was conducted with aimless (CCP4[22]), ShelX C/D/E[23], and Crank-2 (CCP4[22]). Fold similarity analysis for identification was performed with the DALI server[24]. Molecular replacement was performed for the enolase dataset with the homolog form *M. jannaschii* (2PA6) while the sequence for GlnK was adjusted manually. For OGOR an AlphaFold 2 model[25] was generated for each subunit and superposed on the backbone model to generate the model of the whole complex. Preliminary refinement was done with phenix.refine[26].

Condensed purification protocol for the 2-oxoacid:ferredoxin oxidoreductase from *M. infernus*

20 g cells were resuspended in 120 mL IEC A. Lysis was performed by sonication at 70 % for 10 sec of 4 cycles with a KE76 probe (Bandelin, Germany) and repeated three times with short breaks to let se sample cool down between rounds. The total extract was centrifuged at 45,000 × *g* for 45 min at 18 °C. The resulting supernatant was filtered through a 0.2 µm filter (Sartorius, Germany) and passed on 4 x 5 mL HiTrap™ DEAE Sepharose FF column (Cytiva, Sweden) and eluted with a gradient from 0-60 % IEC B at 2 mL.min^{-1} for 60 min. The fractions eluted at 80-260 mM NaCl are pooled and diluted 1:3 in IEC A. Then the pool was passed on 2 x 5 mL HiTrap™ Q HP column (Cytiva, Sweden) and eluted in two steps at 40 % and 100 % IEC B. The peak fractions from 400-1000 mM NaCl were pooled and diluted 1:3 in HIC B. The pool was loaded on a Source™15 Phe 4.6/100 PE column (Cytiva, Sweden) and eluted with 75-0 % HIC B at 1 mL.min^{-1} for 60 min. The fractions eluting from 980-860 mM $(NH_4)_2SO_4$ displaying the four-band pattern (~48 kDa, 30 kDa, 22 kDa and 10 kDa) were pooled, concentrated and washed 1:1000 with GF buffer. The enzyme was stored at 4° for short periods or frozen at −80 °C.

Activity assay (performed by Kinga Sántha during her lab rotation)

Activities were measured with freshly purified enzyme stored at 4 °C in an anaerobic environment. Assays were carried out in a 96 well plate, with a reaction volume of 200 µL at 45 °C following the color change of methyl viologen at 600 nm with a FLUOstar omega multimode microplate reader (BMG Labtech, Germany). Final concentrations were 5 mM methyl viologen, 0.1 mM TPP, 1 mM $MgCl_2$ x $6H_2O$, 0.2 mM CoA, 10 mM oxoacid, 2.44 µg enzyme in 50 mM Tris-HCl (pH 7.5). Reactions were initiated by addition of the enzyme. All measurements were performed in triplicate.

Table S1. Estimated final protein amounts of the purified fractions.

Fraction	Estimated final protein amount (mg)	Fraction	Estimated final protein amount (mg)
TE	1,647 (some loss through aggregation)	D2Q1S2	0.2
		D2Q2S2	2.3
SE	1,875	D1Q2S3	0.4
D1Q1S2	3.3	D2Q2S4	0.5
D1Q1S3	2.2	S2Q3S4	0.8
D1Q2S1	8.4	D2Q5S1	0.4
D1Q2S2	0.7	D2Q5S2	0.4
D1Q2S3	0.5	(D1Q1S1)G1	0.4
D1Q3S1	0.3	(D2Q2S1)G2	0.3
D1Q3S2	0.3	(D2Q3S3)G3	2.0
D1Q3S3	0.3	(D2Q3S3)G4	1.5
D2Q1S1	0.4	(D2Q3S2)G5	0.4

Table S2. X-ray analysis statistics

	D2Q2D2	G1	D2Q5S1
Data collection			
Wavelength (Å)	2.07504	1.00000	1.73892
Space group	$P\,6_3\,2\,2$	$I\,4$	$P\,2_1\,2_1\,2_1$
Resolution (Å)	42.86-2.10 (2.16-2.10)	105.00-2.02 (2.16-2.02)	49.12-2.30 (2.34-2.30)
Cell dimensions			
a, b, c (Å)	82.01 82.01 107.50	148.497 148.497 90.965	77.32 154.06 168.51
α, β, γ (°)	90, 90, 120	90,90, 90	90, 90, 90
R_{merge}(%)[a]	13.2 (218.6)	17.7 (118.2)	14.9 (375.5)
R_{pim} (%)[a]	1.3 (27.3)	5.5 (44.8)	2.9 (75.2)
$CC_{1/2}$ [a]	1 (0.944)	0.994 (0.667)	1.0 (0.737)
$I/\sigma I$[a]	41.1 (3)	8.1 (1.5)	21.9 (1.6)
Spherical completeness[a]	99.7 (99.8)	61.1 (16.4)	99.8 (98.8)
Ellipsoidal completeness[a]	-	92.4 (71.6)	-
Redundancy[a]	171.2 (114.3)	11.5 (7.7)	52.9 (47.3)
Nr. unique reflections[a]	13,132 (1047)	39,692 (1987)	90,079 (4371)
Preliminary Refinement			
Resolution (Å)	42.86-2.10	46.96-2.02	49.12-2.30
Number of reflections	13,037	39,680	89,842
R_{work}/R_{free}[b] (%)	25.64/26.40	18.60/21.72	21.69/25.64
Number of atoms			
Protein	867	6,652	13,175
Ligands/ions	0	2	48
Solvent	10	206	54
Mean B-value (Å^2)	60.48	44.93	85.05
Molprobity clash score	3.34	4.19	12
Ramachandran plot			
Favored regions (%)	97.27	95.54	91.12
Outlier regions (%)	3.19	0	6.22
rmsd[c] bond lengths (Å)	0.008	0.013	0.004
rmsd[c] bond angles (°)	0.879	1.638	0.799

[a] Values relative to the highest resolution shell are within parentheses. [b] Rfree was calculated as the Rwork for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

Table S3. BLAST hits against *M. infernus* 2-oxoglutarate:ferredoxin oxidoreductase. Putative OFORs were retrieved by manual BLAST search against the *Mi*OFOR β-subunit as query followed by BLAST against the corresponding α/β/γ/δ of *Mi*OFOR (colored accordingly). PFOR for *M. acetivorans* and *M. infernus* have been searched for separately and are shown at the bottom of the list. Search was conducted manually and is preliminary, more detailed analysis will follow.

Organism	Accession number	Length (AA)	Annotation	% ID	% Cov
Methanocaldococcus infernus ME	WP_013099506.1	177	2-oxoacid:acceptor oxidoreductase family protein gamma	100	100
	WP_013099507.1	269	2-oxoacid:acceptor oxidoreductase family proteinbeta	100	100
	WP_013099508.1	371	2-oxoacid:acceptor oxidoreductase family protein alpha	100	100
	WP_013099509.1	67	ferredoxin family protein	100	100
Methanothermococcus thermolithotrophicus DSM 2095	WP_018154440.1	257	hypothetical protein	na	na
	WP_018154439.1	265	2-oxoacid:ferredoxin oxidoreductase subunit beta	68.44	97
	WP_018154438.1	178	2-oxoacid:ferredoxin oxidoreductase subunit gamma	54.6	96
	WP_018153336.1	180	2-oxoacid:acceptor oxidoreductase family protein	74.29	98
	WP_018153335.1	287	2-oxoacid:ferredoxin oxidoreductase subunit beta	72.39	98
	WP_018153334.1	381	2-oxoacid:acceptor oxidoreductase subunit alpha	72.5	96
	WP_018153333.1	65	ferredoxin family protein	64.06	95
	WP_018154292.1	577	2-oxoacid:acceptor oxidoreductase subunit alpha	31.4	93
				28.09	98
	WP_018154291.1	289	thiamine pyrophosphate-dependent enzyme	36.54	94
Methanococcus maripaludis S1	WP_181503784.1	258	hypothetical protein	na	na
	WP_181487424.1	265	2-oxoacid:ferredoxin oxidoreductase subunit beta	65.65	97
	WP_181503785.1	177	2-oxoacid:ferredoxin oxidoreductase subunit gamma	52.02	96
	WP_181487460.1	79	ferredoxin family protein	33.33	85
	WP_181503811.1	351	3-methyl-2-oxobutanoate dehydrogenase subunit VorB	34.18	95
	WP_181503812.1	479	2-oxoacid:acceptor oxidoreductase family	35.18	69
				35.77	76
Methanotorris formicicus Mc-S-70	WP_007043677.1	180	2-oxoacid:ferredoxin oxidoreductase subunit gamma	53.45	97
	WP_007043678.1	265	2-oxoacid:ferredoxin oxidoreductase subunit beta	70.45	98
	WP_007043679.1	267	hypothetical protein	na	na
Methanocaldococcus jannaschii DSM 2661	WP_064496537.1	180	2-oxoacid:ferredoxin oxidoreductase subunit gamma	52.54	98
	WP_010870041.1	270	2-oxoacid:ferredoxin oxidoreductase subunit beta	71.54	99
	WP_010870042.1	260	hypothetical protein	na	na
Methanobrevibacter smithii ATCC 35061	WP_011954178.1	180	2-oxoacid:ferredoxin oxidoreductase subunit gamma	49.16	97
	WP_011954179.1	290	2-oxoacid:ferredoxin oxidoreductase subunit beta	59.92	97
	WP_004033007.1	375	2-oxoacid:acceptor oxidoreductase subunit alpha	59.89	98
	WP_011954180.1	69	4Fe-4S binding protein	46.77	92
	WP_004034532.1	79	ferredoxin family protein	38.18	82
	WP_004034534.1	370	3-methyl-2-oxobutanoate dehydrogenase subunit VorB	35.23	94
	WP_004034536.1	486	2-oxoacid:acceptor oxidoreductase family protein	36	69
				32.18	97
	WP_011953859.1	107	hypothetical protein	na	na
	WP_004032133.1	92	hypothetical protein	na	na
	WP_011953860.1	194	indolepyruvate oxidoreductase subunit beta	28.57	98
	WP_004032135.1	632	indolepyruvate ferredoxin oxidoreductase subunit alpha	26.05	97
				23.91	87
	WP_011954046.1	512	thiamine pyrophosphate-binding protein	26.04	35
Haloferax mediterranei ATCC 33500	WP_004057369.1	287	2-oxoacid:ferredoxin oxidoreductase subunit beta	35.56	83
	WP_004057368.1	586	2-oxoacid:acceptor oxidoreductase subunit alpha	34.03	98
				28.77	79
	WP_004057148.1	312	thiamine pyrophosphate-dependent enzyme	33.46	91
	WP_004057146.1	631	2-oxoacid:acceptor oxidoreductase subunit alpha	32.87	89
Methanothermobacter thermoautotrophicus Delta H	WP_048060956.1	67	ferredoxin family protein	54.84	92
	WP_191216118.1	376	2-oxoacid:acceptor oxidoreductase subunit alpha	63.78	98
	WP_010876665.1	286	2-oxoglutarate synthase subunit KorB	61.28	98
	WP_048060957.1	186	2-oxoacid:ferredoxin oxidoreductase subunit gamma	53.63	97
	WP_010876342.1	79	3-methyl-2-oxobutanoate dehydrogenase subunit VorC	38.18	82

	WP_048060884.1	352	3-methyl-2-oxobutanoate dehydrogenase subunit VorB	33.7	98
	WP_048060885.1	490	2-oxoacid:acceptor oxidoreductase family protein	35.18	69
				34.3	95
	WP_010876175.1	512	2-oxoacid:acceptor oxidoreductase subunit alpha	28.61	88
	WP_010876176.1	288	thiamine pyrophosphate-dependent enzyme	36.84	77
		618	indolepyruvate ferredoxin oxidoreductase subunit alpha	23.26	87
	WP_010877454.1			22.99	67
	WP_010877455.1	196	indolepyruvate ferredoxin oxidoreductase subunit beta	23.44	94
Lokiarchaeum sp. GC14_75	KKK43733.1	198	2-oxoacid:ferredoxin oxidoreductase, gamma subunit	34.97	97
	KKK43734.1	303	2-oxoacid:ferredoxin oxidoreductase, beta subunit	45.76	94
	KKK43735.1	390	2-oxoacid:ferredoxin oxidoreductase, alpha subunit	41.6	97
Pyrococcus furiosus DSM 3638	WP_011012908.1	86	2-oxoglutarate ferredoxin oxidoreductase subunit delta	51.56	95
	WP_011012909.1	396	2-oxoacid:acceptor oxidoreductase subunit alpha	56.45	98
	WP_011012910.1	282	2-oxoacid:ferredoxin oxidoreductase subunit beta	58.49	98
	WP_011012911.1	171	2-oxoacid:ferredoxin oxidoreductase subunit gamma	36.42	97
	WP_011012912.1	387	2-oxoacid:acceptor oxidoreductase subunit alpha	61.93	98
	WP_011012913.1	281	2-oxoacid:ferredoxin oxidoreductase subunit beta	58.94	97
	WP_011012914.1	180	2-oxoacid:ferredoxin oxidoreductase subunit gamma	48.88	98
	WP_011011881.1	149	peroxiredoxin	na	na
	WP_011011882.1	288	thiamine pyrophosphate-dependent enzyme	39.2	92
	WP_011011883.1	570	2-oxoacid:acceptor oxidoreductase subunit alpha	31.37	98
				26.14	97
Thermococcus kodakarensis KOD1	WP_011250074.1	186	2-oxoacid:ferredoxin oxidoreductase subunit gamma	48.62	98
	WP_011250075.1	281	2-oxoacid:ferredoxin oxidoreductase subunit beta	55.89	97
	WP_011250076.1	382	2-oxoacid:acceptor oxidoreductase subunit alpha	58.87	98
	WP_011250077.1	170	2-oxoacid:ferredoxin oxidoreductase subunit gamma	37.93	97
	WP_011250080.1	288	2-oxoacid:ferredoxin oxidoreductase subunit beta	56.6	98
	WP_011250081.1	396	2-oxoacid:acceptor oxidoreductase subunit alpha	53.49	98
	WP_011250082.1	99	2-oxoglutarate ferredoxin oxidoreductase subunit delta	53.12	95
	WP_011249767.1	294	thiamine pyrophosphate-dependent enzyme	38.52	90
	WP_011249768.1	570	2-oxoacid:acceptor oxidoreductase subunit alpha	31.78	96
				26.52	74
Archaeoglobus fulgidus DSM 8774	WP_010877975.1	267	2-oxoacid:ferredoxin oxidoreductase subunit beta	51.36	95
	WP_048064242.1	344	hypothetical protein	47.81	98
	WP_010877977.1	67	4Fe-4S binding protein	36.92	95
	WP_010877978.1	187	2-oxoacid:acceptor oxidoreductase family protein	35.63	97
	WP_048095276.1	532	2-oxoacid:acceptor oxidoreductase subunit alpha	29.52	97
				25.42	93
	WP_010878253.1	245	thiamine pyrophosphate-dependent enzyme	34.55	91
	WP_048064723.1	187	pyruvate ferredoxin oxidoreductase subunit gamma	29.67	96
	WP_048096050.1	97	4Fe-4S binding protein	27.59	89
	WP_048096052.1	391	transketolase C-terminal domain-containing protein	25.94	80
	WP_010879198.1	294	thiamine pyrophosphate-dependent enzyme	28.78	94
Sulfolobus acidocaldarius DG1	WP_011279094.1	621	2-oxoacid:acceptor oxidoreductase subunit alpha	31.19	97
				25.26	51
	WP_011279095.1	303	2-oxoacid:ferredoxin oxidoreductase subunit beta	39.48	84
	WP_011277128.1	304	2-oxoacid:ferredoxin oxidoreductase subunit beta	35.93	84
	WP_015385384.1	627	2-oxoacid:ferredoxin oxidoreductase subunit alpha	30.67	97
				21.92	98
Methanosarcina acetivorans C2A	WP_048066425.1	480	2-oxoacid:acceptor oxidoreductase family protein	35.18	69
				33.72	95
	WP_011022859.1	351	3-methyl-2-oxobutanoate dehydrogenase subunit VorB	31.97	98
	WP_011022860.1	82	ferredoxin family protein	35.09	85
	WP_011023014.1	589	2-oxoacid:acceptor oxidoreductase subunit alpha	44.19	46

				26.01	94
	WP_011023015.1	283	2-oxoacid:ferredoxin oxidoreductase subunit beta	38.78	72
Methanosarcina mazei Go1	WP_011035076.1	82	ferredoxin family protein	35.09	85
	WP_011035077.1	351	3-methyl-2-oxobutanoate dehydrogenase subunit VorB	32.6	98
	WP_011035078.1	481	2-oxoacid:acceptor oxidoreductase family protein	31.66	91
				33.72	95
	WP_011032372.1	283	2-oxoacid:ferredoxin oxidoreductase subunit beta	33.73	93
	WP_011032373.1	579	2-oxoacid:acceptor oxidoreductase subunit alpha	30.29	98
				26.12	74
	WP_015411087.1	49	hypothetical protein	na	na
Methanosarcina barkeri 227	WP_048116514.1	82	ferredoxin family protein	35.09	85
	WP_048116516.1	351	3-methyl-2-oxobutanoate dehydrogenase subunit VorB	33.7	98
	WP_048116519.1	481	2-oxoacid:acceptor oxidoreductase family protein	34.25	76
				35.06	95
	WP_230629075.1	574	2-oxoacid:acceptor oxidoreductase subunit alpha	43.6	46
				24.44	74
	WP_048116621.1	283	2-oxoacid:ferredoxin oxidoreductase subunit beta	37.76	72
Azospirillum brasiliense MTCC4039	WP_137139500.1	573	2-oxoacid:acceptor oxidoreductase subunit alpha	34.5	88
				32.48	66
	WP_051658017.1	280	2-oxoacid:ferredoxin oxidoreductase subunit beta	43.48	76
Magnetococcus marinus MC-1	WP_011713405.1	573	2-oxoacid:acceptor oxidoreductase subunit alpha	35.62	99
				26.74	94
	WP_011713406.1	292	2-oxoacid:ferredoxin oxidoreductase subunit beta	35.91	95
Moorella thermoacetica DSM 103132	WP_148871516.1	187	2-oxoacid:acceptor oxidoreductase family protein	36.21	96
	WP_187422226.1	250	thiamine pyrophosphate-dependent enzyme	41.41	79
	WP_069591021.1	356	3-methyl-2-oxobutanoate dehydrogenase subunit VorB	35.56	99
	WP_069591023.1	67	4Fe-4S binding protein	38.81	98
Moorella thermoacetica DSM 103132	WP_069587673.1	363	transketolase C-terminal domain-containing protein	34.38	92
	WP_011391583.1	264	thiamine pyrophosphate-dependent enzyme	40.48	74
	WP_011391584.1	190	2-oxoacid:acceptor oxidoreductase family protein	31.72	97
	WP_011391585.1	97	ferredoxin family protein	33.9	88
Moorella thermoacetica DSM 103132	WP_011393481.1	72	ferredoxin family protein	37.1	92
	WP_069590833.1	390	2-oxoglutarate ferredoxin oxidoreductase subunit alpha	41.26	95
	WP_069590836.1	281	thiamine pyrophosphate-dependent enzyme	32.38	100
	WP_069590838.1	175	2-oxoacid:acceptor oxidoreductase family protein	36.57	97
Methanocaldococcus infernus ME	WP_013100732.1	178	pyruvate ferredoxin oxidoreductase subunit gamma	35.26	96
	WP_013100733.1	85	pyruvate synthase subunit PorD	42.86	91
	WP_013100734.1	377	pyruvate synthase subunit PorA	25.53	99
	WP_013100735.1	295	pyruvate synthase subunit PorB	28.97	71
	WP_013100736.1	166	4Fe-4S dicluster domain-containing protein	28.57	77
	WP_013100737.1	136	4Fe-4S dicluster domain-containing protein	31.25	95
Methanosarcina acetivorans C2A	WP_011020090.1	296	pyruvate synthase subunit PorB	27.49	71
	WP_011020091.1	403	pyruvate synthase subunit PorA	23.68	89
	WP_011020092.1	86	pyruvate synthase subunit PorD	47.37	58
	WP_011020093.1	182	pyruvate ferredoxin oxidoreductase subunit gamma	28.09	96

References

1. Taylor, G. The phase problem. *Acta Crystallogr. D Struct. Biol.* **59**, 1881-1890, doi:10.1107/S0907444903017815 (2003).
2. Doutch, J., Hough, M. A., Hasnain, S. S., Strange, R. W. Challenges of sulfur SAD phasing as a routine method in macromolecular crystallography. *J. Synchrotron Radiat.* **19**, 19-29, doi:10.1107/S0909049511049004 (2012).
3. Dauter, M., Dauter, Z. Many ways to derivatize macromolecules and their crystals for phasing. In *Protein Crystallography: Methods and Protocols* (eds Wlodawer, A.; Dauter, Z.; Jaskolski, M.) 349-356 (Springer New York, 2017).
4. Engilberge, S., Wagner, T., Santoni, G., Breyton, C., Shima, S., Franzetti, B., Riobé, F., Maury, O., Girard, E. Protein crystal structure determination with the crystallophore, a nucleating and phasing agent. *J. Appl. Crystallogr.* **52**, 722-731, doi:10.1107/S1600576719006381 (2019).
5. Jeanthon, C., L'Haridon, S., Reysenbach, A. L., Vernet, M., Messner, P., Sleytr, U. B., Prieur, D. *Methanococcus infernus* sp. nov., a novel hyperthermophilic lithotrophic methanogen isolated from a deep-sea hydrothermal vent. *Int. J. Syst. Bacteriol.* **48 Pt 3**, 913-9, doi:10.1099/00207713-48-3-913 (1998).
6. Fontana, A., De Filippis, V., de Laureto, P. P., Scaramella, E., Zambonin, M. Rigidity of thermophilic enzymes. In *Progr Biotechnol* Vol. 15 (eds Ballesteros, A.; Plou, F. J.; Iborra, J. L.; Halling, P. J.) 277-294 (Elsevier, 1998).
7. Müller, M. C., Wagner, T. The oxoglutarate binding site and regulatory mechanism are conserved in ammonium transporter inhibitors GlnKs from *Methanococcales*. *Int. J. Mol. Sci.* **22**, doi:10.3390/ijms22168631 (2021).
8. Gruswitz, F., O'Connell, J., 3rd, Stroud, R. M. Inhibitory complex of the transmembrane ammonia channel, AmtB, and the cytosolic regulatory protein, GlnK, at 1.96 Å. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 42-7, doi:10.1073/pnas.0609796104 (2007).
9. Conroy, M. J., Durand, A., Lupo, D., Li, X. D., Bullough, P. A., Winkler, F. K., Merrick, M. The crystal structure of the *Escherichia coli* AmtB-GlnK complex reveals how GlnK regulates the ammonia channel. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 1213-8, doi:10.1073/pnas.0610348104 (2007).
10. Pedro-Roig, L., Lange, C., Bonete, M. J., Soppa, J., Maupin-Furlow, J. Nitrogen regulation of protein-protein interactions and transcript levels of GlnK PII regulator and AmtB ammonium transporter homologs in archaea. *MicrobiologyOpen* **2**, 826-40, doi:10.1002/mbo3.120 (2013).
11. Ehlers, C., Weidenbach, K., Veit, K., Forchhammer, K., Schmitz, R. A. Unique mechanistic features of post-translational regulation of glutamine synthetase activity in *Methanosarcina mazei* strain Gö1 in response to nitrogen availability. *Mol. Microbiol.* **55**, 1841-54, doi:10.1111/j.1365-2958.2005.04511.x (2005).
12. Ehlers, C., Grabbe, R., Veit, K., Schmitz, R. A. Characterization of $GlnK_1$ from *Methanosarcina mazei* strain Gö1: Complementation of an *Escherichia coli glnK* mutant strain by $GlnK_1$. *J. Bacteriol.* **184**, 1028-40, doi:10.1128/jb.184.4.1028-1040.2002 (2002).
13. Maslać, N., Sidhu, C., Teeling, H., Wagner, T. Comparative transcriptomics sheds light on remodeling of gene expression during diazotrophy in the thermophilic methanogen *Methanothermococcus thermolithotrophicus*. *mBio* **13**, e0244322, doi:10.1128/mbio.02443-22 (2022).
14. Bräsen, C., Esser, D., Rauch, B., Siebers, B. Carbohydrate metabolism in *Archaea*: Current insights into unusual enzymes and pathways and their regulation. *Microbiol. Mol. Biol. Rev.* **78**, 89-175, doi:10.1128/Mmbr.00041-13 (2014).
15. Chen, P. Y., Li, B., Drennan, C. L., Elliott, S. J. A reverse TCA cycle 2-oxoacid:ferredoxin oxidoreductase that makes C-C bonds from CO_2. *Joule* **3**, 595-611, doi:10.1016/j.joule.2018.12.006 (2019).
16. Gibson, M. I., Brignole, E. J., Pierce, E., Can, M., Ragsdale, S. W., Drennan, C. L. The structure of an oxalate oxidoreductase provides insight into microbial 2-oxoacid metabolism. *Biochemistry* **54**, 4112-20, doi:10.1021/acs.biochem.5b00521 (2015).

17. Chen, P. Y., Aman, H., Can, M., Ragsdale, S. W., Drennan, C. L. Binding site for coenzyme A revealed in the structure of pyruvate:ferredoxin oxidoreductase from *Moorella thermoacetica*. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3846-3851, doi:10.1073/pnas.1722329115 (2018).
18. Cherrier, M. V., Vernede, X., Fenel, D., Martin, L., Arragain, B., Neumann, E., Fontecilla-Camps, J. C., Schoehn, G., Nicolet, Y. Oxygen-sensitive metalloprotein structure determination by cryo-electron microscopy. *Biomolecules* **12**, doi:10.3390/biom12030441 (2022).
19. Yan, Z., Maruyama, A., Arakawa, T., Fushinobu, S., Wakagi, T. Crystal structures of archaeal 2-oxoacid:ferredoxin oxidoreductases from *Sulfolobus tokodaii*. *Sci. Rep.* **6**, 33061, doi:10.1038/srep33061 (2016).
20. Cossu, M., Catlin, D., Elliott, S. J., Metcalf, W. W., Nair, S. K. Structural organization of pyruvate: ferredoxin oxidoreductase from the methanogenic archaeon *Methanosarcina acetivorans*. *Structure*, doi:10.1016/j.str.2024.08.011 (2024).
21. Gibson, M. I., Chen, P. Y. T., Drennan, C. L. A structural phylogeny for understanding 2-oxoacid oxidoreductase function. *Curr. Opin. Struct. Biol.* **41**, 54-61, doi:10.1016/j.sbi.2016.05.011 (2016).
22. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., et al. Overview of the *CCP4* suite and current developments. *Acta Crystallogr. D Struc. Biol.* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).
23. Pape, T., Schneider, T. R. *HKL2MAP*: a graphical user interface for macromolecular phasing with *SHELX* programs. *J. Appl. Crystallogr.* **37**, 843-844, doi:10.1107/S0021889804018047 (2004).
24. Holm, L., Laiho, A., Toronen, P., Salgado, M. DALI shines a light on remote homologs: one hundred discoveries. *Protein. Sci.* **32**, e4519, doi:10.1002/pro.4519 (2023).
25. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zídek, A., Potapenko, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
26. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. *Acta Crystallogr. D Struct. Biol.* **75**, 861-877, doi:10.1107/S2059798319011471 (2019).

CHAPTER III.

Differences in regulation mechanisms of glutamine synthetases from methanogenic archaea unveiled by structural investigations

Marie-Caroline Müller[1], Olivier N. Lemaire[1], Julia M. Kurth[2,3], Cornelia U. Welte[2], Tristan Wagner[1]*

[1] Microbial Metabolism Research Group, Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany.

[2] Department of Microbiology, Institute for Water and Wetland Research, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands.

[3] Present address: Microcosm Earth Center, Philipps-University Marburg and Max Planck Institute for Terrestrial Microbiology, Hans-Meerwein-Str. 4, 35032 Marburg, Germany

***Corresponding author. Email:** twagner@mpi-bremen.de

Published in COMMUNICATIONS BIOLOGY

Müller, MC., Lemaire, O.N., Kurth, J.M. et al. Differences in regulation mechanisms of glutamine synthetases from methanogenic archaea unveiled by structural investigations. *Commun. Biol.* **7**, 111 (2024). https://doi.org/10.1038/s42003-023-05726-w (Open Access, CC BY 4.0)

The text and figures have been reformatted to match the thesis style.

Received June 2023, accepted December 2023

Published January 2024

Abstract

Glutamine synthetases (GS) catalyze the ATP-dependent ammonium assimilation, the initial step of nitrogen acquisition that must be under tight control to fit cellular needs. While their catalytic mechanisms and regulations are well-characterized in bacteria and eukaryotes, only limited knowledge exists in archaea. Here, we solved two archaeal GS structures and unveiled unexpected differences in their regulatory mechanisms. GS from *Methanothermococcus thermolithotrophicus* is inactive in its resting state and switched on by 2-oxoglutarate, a sensor of cellular nitrogen deficiency. The enzyme activation overlays remarkably well with the reported cellular concentration for 2-oxoglutarate. Its binding to an allosteric pocket reconfigures the active site through long-range conformational changes. The homolog from *Methermicoccus shengliensis* does not harbor the 2-oxoglutarate binding motif and, consequently, is 2-oxoglutarate insensitive. Instead, it is directly feedback-inhibited through glutamine recognition by the catalytic Asp50′-loop, a mechanism common to bacterial homologs, but absent in *M. thermolithotrophicus* due to residue substitution. Analyses of residue conservation in archaeal GS suggest that both regulations are widespread and not mutually exclusive. While the effectors and their binding sites are surprisingly different, the molecular mechanisms underlying their mode of action on GS activity operate on the same molecular determinants in the active site.

Keywords: Nitrogen-assimilation, allosteric regulation, methanogenic archaea, structural biology, enzyme mechanism.

Introduction

Nitrogen is an essential component of all living cells. Its most reduced state, ammonia (here representative for NH_3 and NH_4^+ in equilibrium), is one of the most common nitrogen sources assimilated by the microbial world[1-4]. Ammonia enters the central nitrogen metabolism via two systems: the glutamine synthetase – glutamate synthase (GS-GOGAT) couple and/or the glutamate dehydrogenase (GDH). While the GDH carries out the reversible reaction of reductive amination of 2-oxoglutarate (2OG) using NAD(P)H, the couple GS/GOGAT performs an ATP-dependent process[5-8]. Because of its higher affinity for ammonia, the GS-GOGAT couple is notably known to be more effective than the GDH in ammonia-limited environments[1]. The nitrogen assimilation by GS-GOGAT is operated in two steps (Fig. S1a). The GS initially produces glutamine via the condensation of ammonia on glutamate, a reaction coupled to ATP-hydrolysis that requires Mg^{2+} or Mn^{2+} [1]. Then, GOGAT performs the deamination of the synthesized glutamine and the transfer of ammonia to 2OG to ultimately form two molecules of glutamate, with the concomitant oxidation of an electron donor (*e.g.*, NADH)[1, 7].

GS can be categorized into three types[1, 9]. Type I is present in prokaryotes, and some homologs were found in eukaryotes. It organizes as a homo-dodecamer of ~55 kDa-large subunits[10]. Type II is composed of a homo-decameric assembly of ~40 kDa-large subunits and is common in bacteria and eukaryotes[11, 12]. Type III is found in bacteria, archaea and eukaryotes and is composed of a homo-dodecamer of ~75 kDa-large subunits[1, 13]. The GS type I (GSI) is further subdivided into three classes: (i) the GSI-α found in *Archaea, Actinobacteria, Desulfobacterota* and *Bacillota* (formerly *Firmicutes*), (ii) the GSI-β, present in many bacteria and a few archaea and (iii) the GSI-γ mostly found in bacteria[9]. Studies on the evolution of this ancient enzyme led to different scenarios, explaining the puzzling phylogeny of GS that should be the result of gene deletion and horizontal gene transfer, including that GSI-α and GSI-β was

already separated in the Last Universal Common Ancestor and a loss of GSI-α and/or GSI-β occurred in different lineages[9, 14-17]. The GSI-α in *Bacillota* was obtained through lateral gene transfer from an archaeal ancestor[9, 14]. The structural characterization of some GSI-α and GSI-β unveiled a similar organization[6, 18, 19]. The dodecamer is organized as two stacked hexameric rings, and each polypeptide is organized into two parts, a shorter N-terminal domain (β-Grasp fold) and the remaining C-terminal domain (Fig. S1b-g). The N-terminal domain is mainly responsible for the ring association, while the C-terminal domain composes the main part of the ring and is responsible for inter-ring interactions. GSI-β differs structurally from GSI-α by a 25 amino acid-long extension involved in inter-ring stabilization (Fig. S1b-d). The active sites are located at the interface of the N- and C- terminal domains of the adjacent subunit in the hexameric ring. Each active site is structured as a "bifunnel" with ATP and glutamate binding on opposite sides. The ATP binding site is usually referred to as the top of the funnel since the opening lies towards the external ring surface, and the glutamate-binding site as the bottom of the funnel. Divalent metal cations (Mn^{2+} or Mg^{2+}) are coordinated on the C-terminal domain and positioned at the center of the bifunnel. Two loops in the active center are critical for catalysis: the Glu-flap, involved in shielding the active site during catalysis and deprotonating the intermediate product, and the Asp-50′ loop, which binds and deprotonates ammonium. In this concerted orchestra, the glutamate is first phosphorylated on its γ-carboxyl group by the ATP donor, and secondly, the ammonia is incorporated, releasing the products glutamine, ADP and inorganic phosphate[6, 20]. GSI-γ are comprised of only the catalytic domain lacking the N-terminus. A majority of biochemically characterized GSI-γ members lack GS biosynthetic activity and instead function as γ-glutamyl-polyamine synthetases[9].

The GSI-α and GSI-β also differ regarding their regulation. The GSI-β have a complex multi-level regulation by adenylylation, a feature not conserved in GSI-α[8]. The regulation of GSI-α in *Bacillus subtilis* was shown to depend on feedback inhibition by

the enzyme's product glutamine[18]. The glutamine-inhibited enzyme also binds the regulator GlnR, triggering a switch from a dodecameric GS to an inactive tetradecameric form[19]. The GS-GlnR complex finally acts as a transcription repressor for genes involved in nitrogen assimilation, including the GS.

The archaeal GSI-α has never been structurally characterized, and only sporadic studies exist[21-25]. While the activity of the GSI-α of *Haloferax mediterranei* is inhibited by glutamine[26], other radically different regulations have been discovered. In the same archaeon, the molecule 2OG was shown to be a more efficient regulator with a 12-fold activity stimulation, reaching 18-fold when P_{II}-family regulatory proteins $GlnK_1$ or $GlnK_2$ were added in addition to 2OG[26]. 2OG and GlnK also enhance the activity of the enzyme from the methanogen *Methanosarcina mazei*, as well as the 23 residue-large sP26 protein[27, 28], although an inhibitory effect of GlnK (under different buffer conditions) has also been shown[28]. The regulatory proteins GlnK and sP26, whose expression is triggered under nitrogen starvation, form a complex and make a tight and specific interaction with the GS[27, 29]. The direct control by 2OG is particularly elegant as the metabolite is a cellular sensor for nitrogen deprivation and has been shown to act on several key nodes of nitrogen acquisition in methanogens (Fig. S1a)[8, 26-28, 30].

To shed light on the regulatory mechanisms of archaeal GSI-α, we performed biochemical and structural characterizations on the enzymes from two thermophilic methanogens belonging to different phyla. Our results illustrate the different regulation systems at the molecular level, including an unforeseen switch-on controlled by 2OG.

Results

*Mt*GS activity is strictly dependent on 2OG in contrast to *Ms*GS.

Methanothermococcus thermolithotrophicus belonging to the *Methanococcales* order is a strictly hydrogenotrophic methanogen which possesses a single gene coding for a GSI-α (the gene product is referred to as *Mt*GS) while lacking any genes coding for GDH. The purification of *Mt*GS was previously published[31]. We improved the purification protocol to obtain an anaerobically purified *Mt*GS allowing its detailed enzymatic characterization. Denaturing and high-resolution clear native polyacrylamide gel electrophoresis (hrCN PAGE) are coherent with a dodecameric organization of the ~50 kDa peptide (complex size estimated at 540.3 kDa, Fig. 1a and b). The activity of *Mt*GS, measured via a coupled enzyme assay (Fig. S2), could not be detected despite the addition of Mg^{2+} or Mn^{2+}, an increase of the protein concentration (up to 0.75 mg.mL^{-1} final concentration) or an extension of the incubation time.

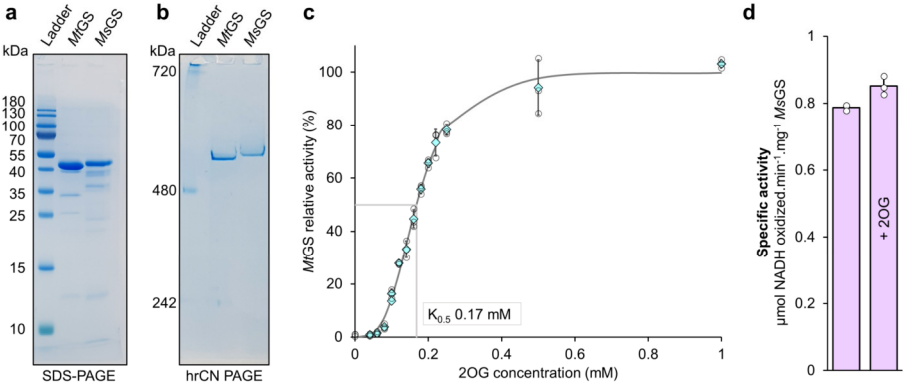


Fig. 1 Purification and characterization of GS from methanogenic archaea. a SDS-PAGE and **b** Native PAGE of 2 µg purified *Mt*GS and *Ms*GS. **c** Relative activity of *Mt*GS at different 2OG concentrations. The presented fit (grey plain line) is extrapolated from the determined Michaelis-Menten parameters. **d** The specific activity of *Ms*GS without or with 2 mM 2OG at 40 mM glutamate and 20 mM NH_4Cl. For **c** and **d** data is represented as mean (by cyan diamonds on panel (**c**)) ± standard deviation (s.d) and individual values are shown as white circles (n=3).

As 2OG acts as an activity stimulator in archaeal GS[26-28], 2OG was added to the assay and revealed to be essential for *Mt*GS activity under these experimental conditions (Fig. 1c). No activity was detected below 0.06 mM 2OG, 50 % activity was reached at 0.17 mM and saturation occurred above ~ 0.6 mM. Noteworthy, exposure to oxygen severely reduced the enzyme activity (1 h of exposure to O_2 decreased the activity to 6 %). Specificity for 2OG was tested by substitution with malate or succinate, which are structurally similar to 2OG. At 15 mM, both surrogates could not activate *Mt*GS. Malate or succinate addition to the 2OG-activated enzyme slightly impaired the activity, an effect that might be due to a competition for the binding site (Fig. S3).

To investigate whether the 2OG dependency is a feature unique to *M. thermolithotrophicus* or also distributed among other methanogens, we purified the GS from the phylogenetically distant methanogen *Methermicoccus shengliensis*. This archaeon belongs to the *Methanosarcinales* order and uses methylotrophic methanogenesis for energy and carbon acquisition[32]. Its genome codes for a single GSI-α (*Ms*GS) and no GDH. *Ms*GS was anaerobically purified from *M. shengliensis* leading to a major band at ~50 kDa on SDS PAGE (Fig. 1a). A dodecameric assembly was suggested by hrCN PAGE (estimated complex size at 547.5 kDa, Fig. 1b). Both the SDS and hrCN PAGE indicate a molecular weight for *Ms*GS (WP_042685700.1, predicted molecular weight: 49.53 kDa) superior to that of *Mt*GS (WP_018154487.1, predicted molecular weight: 50.24 kDa), a difference that we attributed to an artifact from the electrophoretic migration. When subjected to size exclusion chromatography, both proteins exhibited an elution volume in the range of the proposed dodecameric assembly in which *Ms*GS is smaller than *Mt*GS (Fig. S4). Native electrophoresis and size exclusion chromatography indicate a single oligomeric state for the purified GS. Unlike for *Mt*GS, 2OG was not required for *Ms*GS activity, and its addition (2 mM) did not affect its activity (Fig. 1d). Kinetic parameters of both enzymes were determined and are summarized in Table 1 (kinetic curves are provided in Fig. S5). Notably, *Mt*GS

Table 1: Kinetic parameters of *Mt*GS and *Ms*GS for NH_4Cl and glutamate. All activities were measured in technical triplicates (n=3). *Mt*GS was measured in the presence of 2 mM 2OG. The curves used for kinetics parameter determination are provided in Fig. S5.

Enzyme	Substrate	AppK_m (mM)	AppV_{max} (U.mg^{-1})	Hill coefficient
*Mt*GS	NH_4Cl	0.14 ± 0.01	3.268 ± 0.115	1.48 ± 0.21
	Glutamate	8.76 ± 0.87	3.419 ± 0.134	1.69 ± 0.23
*Ms*GS	NH_4Cl	0.07 ± 0.01	0.98 ± 0.02	0.84 ± 0.12
	Glutamate	32.57 ± 3.13	1.40 ± 0.07	2.35 ± 0.41

reaches a 30-fold higher apparent V_{max} than *Ms*GS once fully activated with 2OG. Both enzymes exhibit the same order of magnitude for the apparent K_m for glutamate and ammonium, and binding of both substrates exhibits a marginal positive cooperativity in *Mt*GS, while in *Ms*GS glutamate binding showed a positive cooperativity and almost no cooperativity for ammonium. Such observed differences in cooperativity between both enzymes might hide a more sophisticated discrepancy in their mode of regulation and sensing the intracellular metabolite balance. Structural studies were undertaken to investigate the activation mechanism and the difference between both GS systems.

Structural snapshots of archaeal GS

The structure of *Mt*GS in its resting state was solved by using the diffraction data obtained by Engilberge *et al.*[31]. The data was collected on a crystal obtained by co-crystallization with a crystallophore (TbXo4) that corrected the twinning issue inherent to this crystalline form. The structure (*Mt*GS-apo-TbXo4) was solved by molecular replacement using the closest structural homolog from *Bacillus subtilis* (*Bs*GS resting state, PDB 4LNN[18]) and refined to a resolution of 1.65 Å (Table 2). The *Mt*GS structure presents the typical homo-dodecameric architecture as seen in other structural homologs, with a monomeric unit divided into Nter- (1-111) and Cter- (112-448)

domains (Fig. 2a-c, Fig. S6). A root mean square deviation (rmsd) of 0.847 Å (on 338 Cα aligned from one monomer) exists between this structure and the apo structure of *Bs*GS (Fig. S7). The TbXo4, contributing to the crystal packing, might have artificially provoked this deviation. Moreover, terbium atoms are located in the active site, coordinated by the residues involved in the Mg^{2+}/Mn^{2+} recognition (Fig. S8). To exclude any artefactual effect of TbXo4, we performed crystallization in the absence of the compound. The resulting crystal belonged to the same space group as the TbXo4-containing form, but the crystallographic data exhibited a pseudomerohedral twinning with a fraction of 0.12. The protein structure (*Mt*GS-apo-without TbXo4) was refined to 2.43 Å and is almost identical to the TbXo4-containing form (rmsd of 0.206 Å, 410 atoms aligned Fig. S8, residues 67-69 could not be modeled in the TbXo4-lacking structure), confirming that the observed conformation is not the artefactual result of the TbXo4 binding. For this reason, the *Mt*GS TbXo4-containing structure obtained at the higher resolution of 1.65 Å was used for further analyses.

The enzyme from *M. shengliensis* was crystallized in the absence of ligands, and two different crystalline forms, referred to as *Ms*GS-apo 1 and *Ms*GS-apo 2, were analyzed (Table 3). *Mt*GS-apo-TbXo4 was selected as a template for molecular replacement, and the structures were refined to 2.64 Å and 3.09 Å, respectively (Fig. S9). As *Mt*GS, the homolog from *M. shengliensis* has a dodecameric organization with the N-ter (1-106) and C-ter (107-442) domains forming the protomeric unit (Fig. 2d-f, Fig. S10). *Ms*GS-apo 1 was used for the following structural analyses due to the better resolution.

A more detailed view of the monomeric structures highlights the key catalytic elements surrounding the active site in both GS (Fig. 2c and f, Fig. S11,[18]): the Glu-flap (*Mt*GS 306-311, *Ms*GS 299-304), the Tyr-loop (*Mt*GS 370-378, *Ms*GS 363-371), the Asn-loop (*Mt*GS 236-247, *Ms*GS 229-240), Tyr179-loop (*Mt*GS 153-163, *Ms*GS 146-156) and Asp50′-loop (*Mt*GS 57-71, *Ms*GS 51-65). A structural comparison of the monomeric

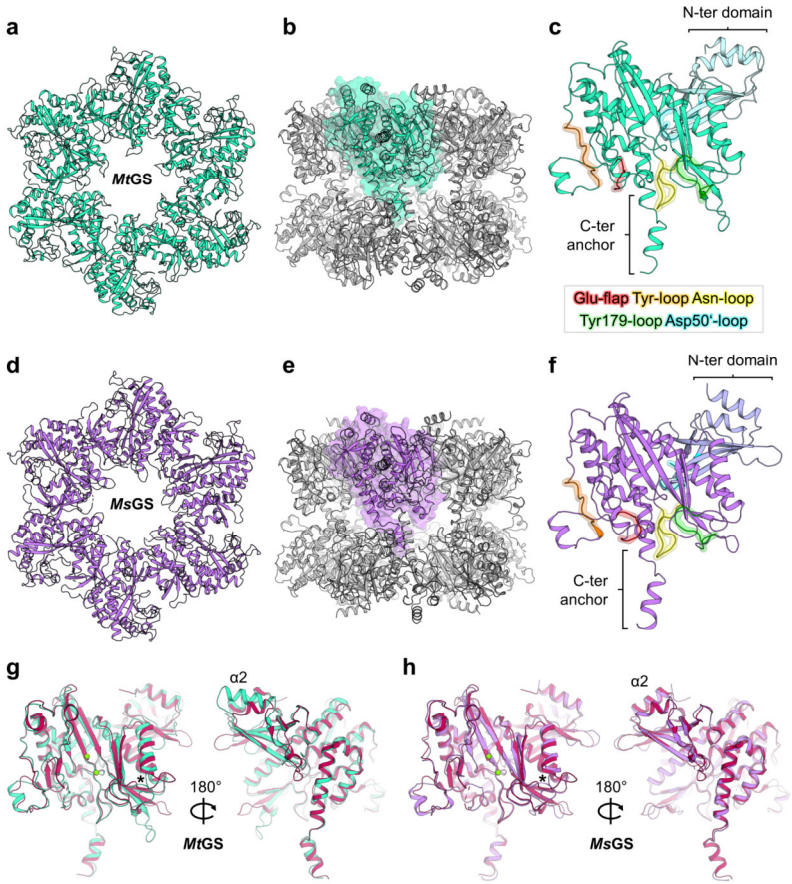


Fig. 2 Structural organization of archaeal GS. All models are represented in cartoons. **a** Top view of one *Mt*GS hexamer. **b** Side view of the *Mt*GS dodecamer with one subunit shown in a transparent surface (cyan). **c** *Mt*GS monomer with the main loops highlighted. The color coding of the loops is indicated in the box. The N- and C-terminal domains are colored light blue and cyan, respectively. **d** Top view of one *Ms*GS hexamer. **e** The side view of the *Ms*GS dodecamer with one subunit shown as a transparent surface (purple). **f** *Ms*GS monomer with the main loops highlighted with the same color coding as in panel (**c**). The N-and C- terminal domains are colored lavender and purple, respectively. **g** Overlay of *Mt*GS (cyan) and *Bs*GS apo state (red, PDB 4LNN). **h** Overlay of *Ms*GS (purple) and *Bs*GS apo state (red). For panels (**g**, **h**), the star indicates the position of the loop deviating in both archaeal GS compared to *Bs*GS. Mg atoms are displayed as green spheres.

unit reveals that *Ms*GS has a closer fit with *Bs*GS compared to *Mt*GS (rmsd *Ms*GS-*Bs*GS: 0.749 Å, 341 atoms aligned, *Mt*GS-*Bs*GS: 0.847 Å, 338 atoms aligned), which exhibits larger deviations such as an extended helix α2 resulting from a seven-residue insertion (Fig. 2g, Fig. S7). While some deviations of local loops might be attributed to the increase in flexibility (Fig. S7a), both archaeal GS deviate from *Bs*GS at the loop juxtaposed to the Tyr179-loop (*Mt*GS 164-171, *Ms*GS 157-164) before helix α6 (highlighted by a star on Fig. 2g and h) at the subunit interface.

The GS are well known to be dynamic, and structural rearrangement can occur upon ligand binding[6, 18, 19, 33, 34]. For instance, *Bs*GS undergoes dramatic intersubunit conformational movements between the apo and transition state. Therefore, to investigate why *Mt*GS apo resting state would be catalytically inactive, co-crystallization with 2OG was undergone.

A 2OG allosteric site localized at the inter-subunit junction

The complex with 2OG/Mg^{2+} was obtained with and without ATP (*Mt*GS-2OG/Mg^{2+}/ATP, *Mt*GS-2OG/Mg^{2+}) in a new crystalline form and both were refined to 2.15 Å and 2.91 Å resolution, respectively (Figs. 3 and 4, Table 2). Both 2OG-containing structures show high tridimensional conservation (rmsd 0.279 Å, 441 atoms aligned, Fig. S12) while deviating from the apo resting state (2OG/Mg^{2+}/ATP: rmsd 0.694 Å, 374 atoms aligned, 2OG/Mg^{2+}: 0.631 Å, 371 atoms aligned, Fig. S13 and S14). Since the 2OG/Mg^{2+}/ATP complex is at a higher resolution, this structure was preferred for deeper structural analyses (Fig. S15). The allosteric

2OG binding site, ~15 Å distant from the catalytic cleft, is localized at the interface between the C-terminal domain and the adjacent N-terminal domain. 2OG is coordinated via ionic bonds by Arg20′ Arg88′, Arg174, Arg175, and a hydrogen bond via Ser191 (primed numbers indicate the adjacent unit, Fig. 3a). Phe18′ stabilizes the 2OG via stacking with its phenyl ring. In the apo structure, water molecules replace

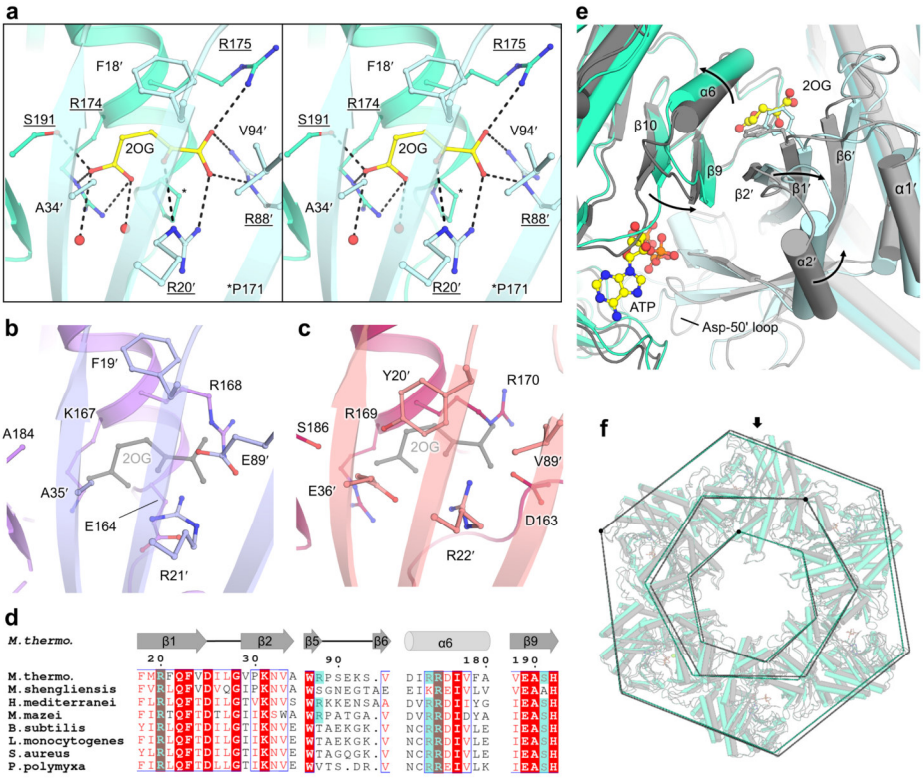


Fig. 3. 2OG binding site and structural rearrangement in *Mt*GS. a Close-up of the 2OG binding site in *Mt*GS (cyan cartoon, the adjacent monomer in light blue) shown as a stereo view. 2OG and the residues in its vicinity are shown as balls and sticks with contacts in black dashes. **b, c** Same view as in (**a**) showing *Ms*GS apo (**b**), purple cartoon, with the adjacent subunit in light purple) and *Bs*GS apo (**c**), red cartoon, with the adjacent subunit in light red, PDB 4LNN). 2OG from *Mt*GS (grey) was superposed to visualize the clash with E89′ for *Ms*GS and E36′/Y20′ for *Bs*GS. **d** Sequence alignment of different GSI-α in which 2OG-binding residues observed in *Mt*GS are highlighted with a cyan box (see Fig. S11 for the entire alignment). **e** Structural rearrangements between the apo (grey cartoon) and 2OG/Mg^{2+}/ATP bound state (cyan cartoon). The adjacent monomer is colored lighter. Phe18′ is shown as sticks. Arrows highlight the movements caused by 2OG binding. **f** *Mt*GS apo (grey cartoon) superposed to *Mt*GS-2OG/Mg^{2+}/ATP (cyan cartoon). The superposition was done on one monomer (indicated by an arrow), and a dashed line was drawn on the Cα position of Val4, Gly198, and Asn267 to illustrate the overall movements. For all, oxygen,

nitrogen, and phosphorus are colored in red, blue, and orange, respectively. Carbons are colored depending on the chain and in yellow for ligands.

2OG in a pocket with a similar volume and slightly open to the solvent (Fig. S15c). In the 2OG-insensitive *Ms*GS structure, the side chain of Glu89′ substituting the *Mt*GS Val94′ would collide with the 2OG position, and its negative charge would additionally lead to the ionic repulsion of 2OG (Fig. 3b). Moreover, in *Ms*GS, most of the residues relevant for 2OG coordination (Arg88′, Arg174 and Ser191) are substituted. Similarly, the structures of the bacterial homologs show a tyrosine and glutamate at the equivalent positions 18′ and 34′ in *Mt*GS that would collide with the modeled 2OG and would contribute to ionic repulsion (Fig. 3c). Additionally, the 2OG-coordinating Arg88′ is absent (Fig. 3d). The Arg88′ might be the essential component for 2OG-based activation as it is positioned on the loop that undergoes the largest conformational change during 2OG binding. Therefore, its movement provoked by 2OG binding would trigger further motions throughout the polypeptide chain (see supplementary movie "*Mt*GS_conformational_changes").

2OG-dependent motions and active site remodeling in *Mt*GS.

2OG presence at the interface of monomers reorients the bulky side chain of Phe18′, and remodels salt bridge networks in its surrounding. Consequently, the N-terminal domain is pushed away from the neighboring C-terminal domain, and the α-helix 6 and β-hairpin 9-10 (residues 167-206, in the vicinity of the active site) are shifted (Fig. 3e). The structural alignment on the C-terminal part from one monomer exemplifies a 4.8° oscillation of the N-terminal domain, leading to a movement of 3 Å (Fig. S13b). The repositioning of the N-terminal domain would clash with the next monomer and therefore, a domino effect occurs in the hexamer ring upon 2OG binding. This probably explains the cooperative behavior observed in the presented kinetic data (Hill coefficient 3.4 for 2OG, Fig. 1c). Overall, the transition between *Mt*GS resting and

2OG-bound states provokes a compaction of the outer ring and an opening in the center and middle parts (Fig. 3f), leading to a more "opened" conformation. The observed switch and the 2OG-bound active state are similar to the one observed between apo and transition state complex of *B. subtilis* (Fig. S16). The most dramatic change occurs in the Asp50′-loop (critical for glutamate and ammonia coordination) interacting with the loop 318-325, which reorganize to reach a position similar to that observed in the bacterial homologs co-crystallized with the L-methionine-S-sulfoximine-phosphate (abbreviated as SOX, a mimic of a reaction transition state)[18, 19] (Fig. 5a-c and S17).

In the *Mt*GS-2OG/Mg^{2+}/ATP structure, the ATP is located in a similar position compared to bacterial homologs (Fig. 4a). The adenosine part is stacked in between Phe206 and Arg336 and coordinated via hydrogen bonds with Lys333 main chain and Ser254 side chain. The ribose moiety is coordinated by the Phe204 and Phe206 main chains and the Glu189 side chain. His252, Arg321, Arg326 and Arg336 bind the triphosphate backbone with ionic bonds. In the resting state, the ATP would clash with the Phe206 and would not bind without displacing the β-sheet 10 (residues 199-205) and the following loop (206-209) (Fig. 4c and S18).

We also obtained and refined to 2.70 Å a *Ms*GS structure in complex with ATP and Mg^{2+} (Table 3). Only minor differences exist compared to the apo state (Fig. S14d). While the adenine part of the ATP is bound in a similar way as described before, the ribose backbone is tilted by 90°, pointing the triphosphate away from the active site (Fig. 4d). This results in a non-catalytic state, similar to the one observed in the *Bs*GS structure loaded with the non-hydrolysable ATP analogue AMPPCP (PDB 4LNK[18]). Nevertheless, a sequence alignment of *Bs*GS, *Mt*GS and *Ms*GS revealed a complete conservation of the residues involved in nucleotide binding (Fig. 4e). Therefore, we concluded that the three GS share the same ATP coordination throughout the catalytic steps.

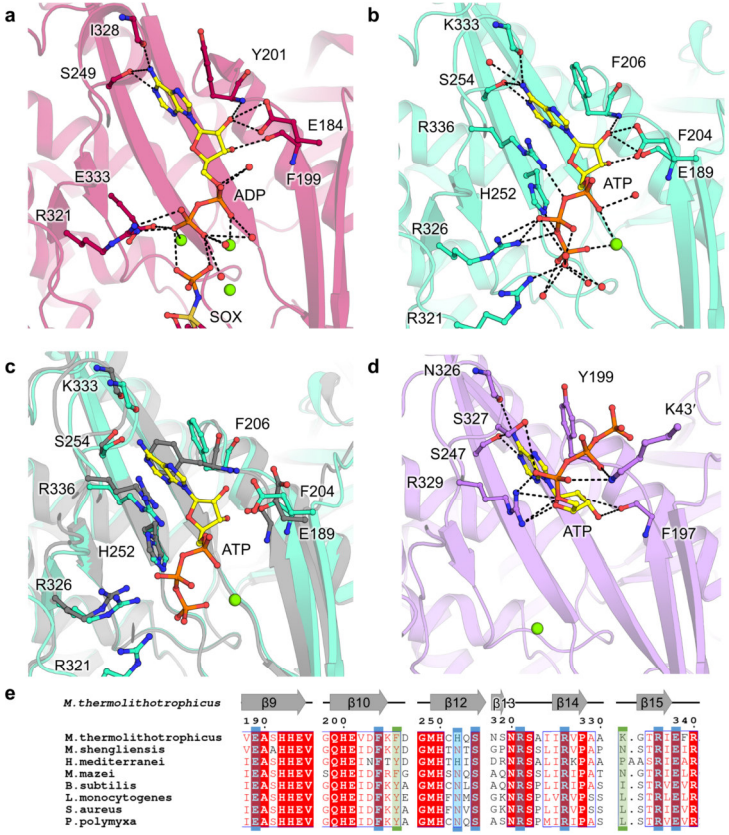


Fig. 4 ATP-binding site comparison between different GSI-α. a ATP binding site in *Bs*GS transition state (containing SOX/Mg^{2+}/ADP, PDB 4LNI) and **b** *Mt*GS-2OG/Mg^{2+}/ATP. **c** Superposition of the C-terminal domain of *Mt*GS apo (grey) on *Mt*GS-2OG/Mg^{2+}/ATP (cyan), with an overlay of the ATP-binding residues. **d** ATP binding site in *Ms*GS-Mg^{2+}/ATP. Models are represented in transparent cartoons with the ligands (yellow) and interacting residues shown as balls and sticks. Oxygen, nitrogen, sulfur, phosphorus, and magnesium are colored red, blue, dark yellow, orange, and green, respectively. Carbons are colored by chain and ATP carbons in yellow. Hydrogen bonds are visualized as black dashes. **e** Sequence alignment of the ATP binding residues. Residues coordinating the nucleotide via side chain and main chain hydrogen bonds are highlighted by a blue and green box, respectively (based on *Mt*GS). *Ms*GS K43′/S327 were omitted from the analysis due to the artefactual γ-phosphate position.

2OG binding in *Mt*GS also impacts the glutamate binding site architecture. The residues involved in glutamate binding (or the mimic SOX molecule) in *B. subtilis* are perfectly conserved in *Mt*GS, *Ms*GS and other archaeal GS (Fig. 5). The conservation between the archaeal and bacterial domains reinforces the importance of these residues to orchestrate the catalysis. While the predicted residues involved in Mg^{2+}/glutamate binding share the same position between the *Mt*GS resting and 2OG-bound states (Glu137, Glu139, Glu194, Gly246, His250, Glu309, Arg303, Arg340), the catalytic Arg321 (homologous to Arg316 in *Bs*GS) shows a drastic difference of position (Fig. 5b, c). In the resting state, the loop 318-325 is constrained by a network of hydrogen bonds and salt bridges from the Asp50′-loop, the side chain of the Arg321 itself being sequestered by Glu70′. Strikingly, in the 2OG-bound state of *Mt*GS, the 318-325 and Asp50′-loops adopt a position similar to all described active transition states in bacterial homologs (Fig. S17). This suggests that the binding of 2OG triggers conformational changes leading to a catalytically competent conformation and explaining the 2OG dependency for activity. Moreover, the Glu309 from the Glu-flap is disengaged and in a more relaxed state in the 2OG-bound model, allowing its assistance for catalysis. In *Ms*GS, the β-hairpin β3-β4 carrying the Asp50′-loop appears to adopt a similar position with or without ATP. However, the Asp50′ loop is flexible and cannot be modeled. The flexibility might directly affect the stabilization of the adjacent loop 313-319 (318-325 in *Mt*GS) that exhibits a different position than that found in *Mt*GS and bacterial GS with the catalytically important Arg314 (Arg321 in *Mt*GS) retracted on the Tyr-loop (Fig. 5d).

Fig. 5 Glutamate-binding site comparison and glutamine feedback inhibition (see next page). a-d Glutamate binding site in **a** *Bs*GS-SOX/Mg^{2+}/ADP (red, 4LNI), **b** *Mt*GS-2OG/Mg^{2+}/ATP (cyan), **c** *Mt*GS apo (cyan), and **d** *Ms*GS-Mg^{2+}/ATP (purple). Models are in cartoons with ligands and equivalent residues binding SOX as balls and sticks. Oxygen, nitrogen, sulfur, phosphorus, and magnesium are colored red, blue, dark yellow, orange, and green, respectively. Carbons are colored by chain and ATP carbons in yellow. Hydrogen bonds are visualized as black dashes. The Asp-50′ loop region is

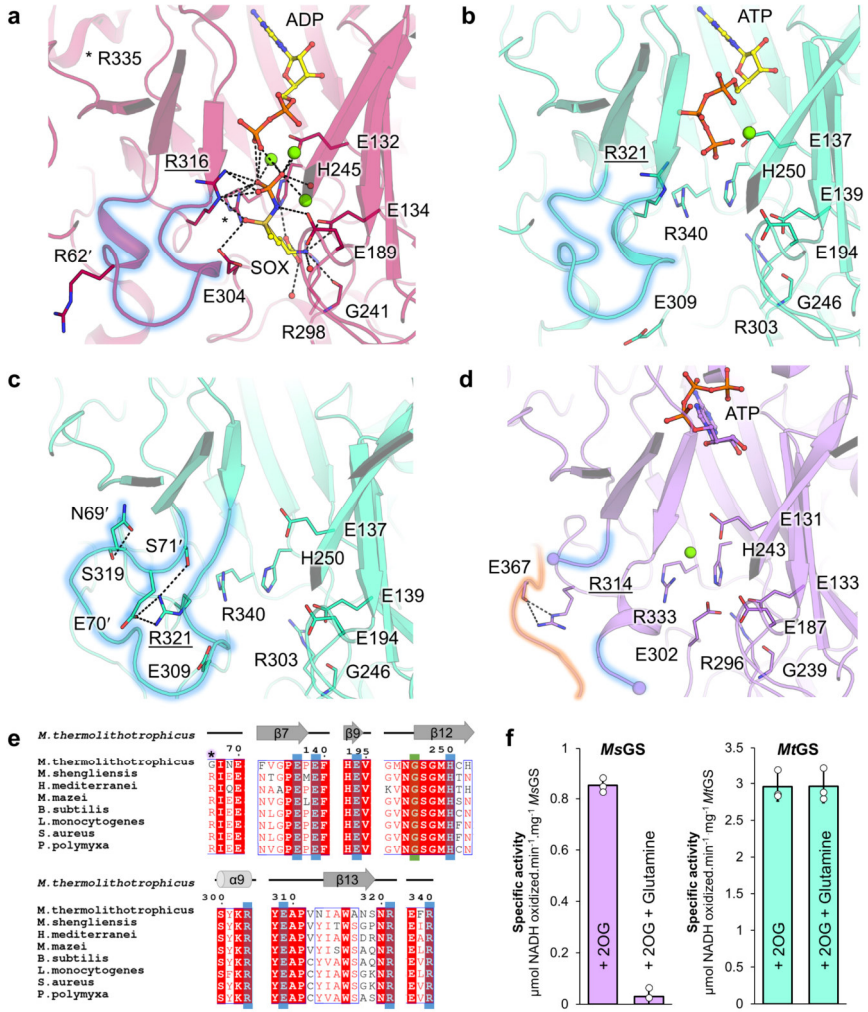


Fig. 5 cont. highlighted by a blue glow.For *Ms*GS, the Tyr-loop is highlighted with an orange glow. An underlined label highlights the catalytic arginine (*e.g.*, R321 in *Mt*GS). **e** Alignment of the residues involved in glutamate binding (based on *Bs*GS). Side chain and main chain interactions are highlighted by a blue and green box, respectively. A star highlights the arginine responsible for glutamine feedback inhibition in *Bs*GS. **f** Specific activity in the absence and presence of glutamine in both archaeal GS. Data is represented as mean ± s.d and individual values are shown as spheres (n=3).

_Ms_GS is feedback inhibited by glutamine

*Bs*GS is inhibited by the reaction product glutamine[18, 19], a feedback regulation under the control of the Asp50′ loop. When glutamine occupies the substrate-binding site, it interacts with Glu304 (*Bs*GS numbering), required for catalysis. The Arg62′ from the Asp50′ loop stabilizes this interaction, resulting in a lock of the Glu-flap, closing the active site. The Arg62′ is not conserved in *Mt*GS (substituted by Gly67′) but is present in *Ms*GS (Arg61′, Fig. 5e). To investigate this feedback inhibition mechanism, we performed activity assays with glutamine addition for *Mt*GS and *Ms*GS. As suspected from the sequences, the addition of glutamine (20 mM) inhibited the activity of *Ms*GS by 97 %, while no difference was observed for *Mt*GS (Fig. 5f). The substitution of the arginine appears to completely abolish glutamine inhibition.

Direct regulation by 2OG and glutamine within the domain *Archaea*

In addition to the previous studies of archaeal GSI-α[26-28], the characterization of the enzyme from *M. thermolithotrophicus* and *M. shengliensis* highlighted that the mechanisms of post-translational regulation are not conserved among archaea. Our study points out that the direct binding of 2OG and glutamine depends on a restricted number of residues, suggesting that the potential regulation of GS activity by both 2OG and glutamine might be predictable from the amino acid sequences. The conservation of residues involved in regulator-binding was analyzed through 500 GS sequences regrouping *Mt*GS and its closest archaeal homologs in the RefSeq database (Fig. 6 and Fig. S19). The GSI-α sequences used to build the tree are, in most cases, forming monophyletic branches that separate archaeal orders (sometimes different groups of GS within orders) and that apparently share a regulation mechanism (see residue conservation in Fig. S19), except for some sequences including *Ms*GS which do not clearly branch with any other GS (Fig. 6). Around half of the GS groups isolated from this tree (8 over the 17 archaeal GS groups) appear to harbor the residues necessary for binding either 2OG or glutamine. In comparison, a third of the GS groups can be

theoretically regulated by both (6 over the 17 archaeal GS groups, Fig. 6), of which the enzymes from *H. mediterranei* and *M. mazei* are representatives. In certain archaeal orders (*e.g.*, *Methanosarcinales*), the genome encodes two GS isoforms that appear to have different sensitivity toward 2OG and glutamine. The enzymes found in *Thermococcales*, *Thermoplasmatales* or *Methanomassiliicoccales* present substitutions that would hinder their capacities to bind both 2OG and glutamine. Yet, binding at another position in the protein structure cannot be excluded to allow potential interactions with regulatory partners. This analysis, yet restricted to a small number of archaeal orders and therefore far from describing the overall archaeal GS family, suggests variability in regulation strategies among these enzymes pointed out by the enzymatic and structural analyses.

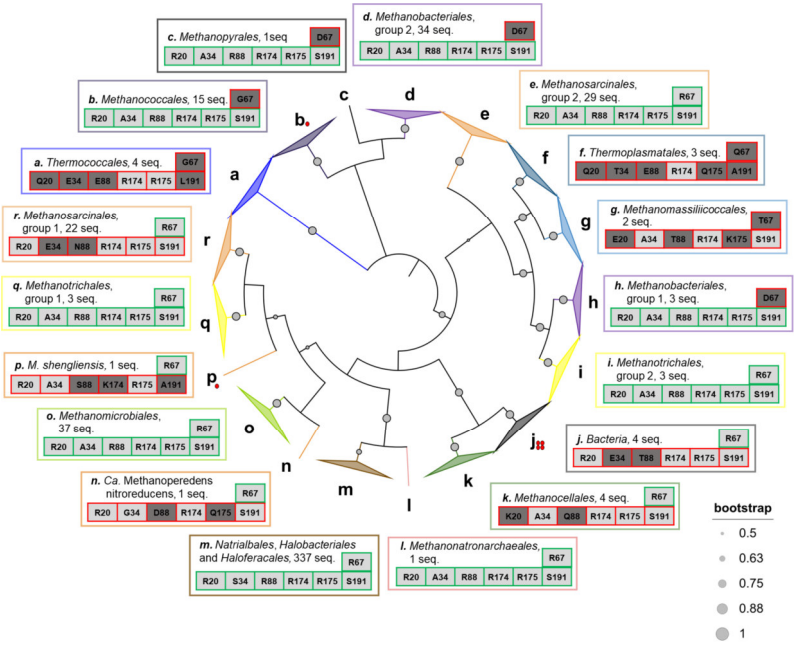


Fig. 6 (see next page)

Fig. 6 Conservation of residues binding 2OG and glutamine in archaeal GS. The presented phylogenic tree (maximum likelihood) was constructed with the 500 closest sequences to *Mt*GS in the RefSeq database, restricted to the domain *Archaea*, as well as the sequences of the bacterial GSI-α from *B. subtilis*, *S. aureus*, *L. monocytogenes* and *P. polymyxa*. The tree was colored by orders (except for *Bacteria*), and grey dots with different radii represent the bootstrap support of each branch. Branches containing monophyletic groups are collapsed. Sequences forming monophyletic branches are gathered and labeled by a letter. The branch containing 337 sequences belonging to *Natrialbales*, *Halobacteriales* and *Haloferales* orders is condensed as no clear monophyletic groups could have been extracted. The outer panels display the most common residues at equivalent positions involved in 2OG (Arg20, Arg88, Arg174, Arg175 and Ser191, *Mt*GS numbering, bottom line) and glutamine (position 67 in *Mt*GS, upper line) binding. Panels are framed and labeled with the color and letter used in the inner tree. Ala34 in *Mt*GS is not involved in 2OG coordination, but a substitution by a bulky residue (*e.g.,* glutamate) would hinder its fixation. The residues are colored in light or dark grey depending on whether they allow metabolite binding or not, respectively. The residue distribution for each position in each group is presented in Fig. S19, and the sequences can be found in Supplementary Data 2. The GS that are predicted to be able or unable to bind 2OG and glutamine are framed in green and red, respectively. Red dots indicate the GSI-α structurally characterized previously or in the present work.

Table 2 X-ray analysis statistics for *Mt*GS.

	***Mt*GS-apo-TbXo4**	***Mt*GS-apo without TbXo4**	***Mt*GS-2OG/Mg^{2+}/ATP**	***Mt*GS-2OG/Mg^{2+}**
Data collection				
Wavelength (Å)	0.97625	1.00004	1.30511	0.99999
Space group	$C222_1$	$C222_1$	$P1$	$P1$
Resolution (Å)	49.78 – 1.65 (1.74 – 1.65)	76.67 – 2.43 (2.57 – 2.43)	110.03 – 2.15 (2.30 – 2.15)	203.54 – 2.91 (3.12 – 2.91)
Cell dimensions				
a, b, c (Å)	131.43, 228.45, 204.80	132.65, 230.24, 205.59	112.34, 131.77, 131.51	131.81, 131.93, 203.54
α, β, γ (°)	90, 90, 90	90, 90, 90	60.04, 87.72, 67.34	89.95, 89.86, 60.05
R_{merge}(%)[a]	6.3 (91.8)	26.9 (179.5)	11.0 (95.0)	12.2 (50.1)
R_{pim} (%)[a]	3.9 (56.6)	7.3 (46.4)	6.8 (59.0)	7.6 (34.4)
$CC_{1/2}$ [a]	0.999 (0.782)	0.998 (0.782)	0.995 (0.569)	0.993 (0.548)
I/σ$_I$[a]	11.9 (1.6)	7.4 (1.5)	9.3 (1.3)	5.9 (1.6)
Spherical completeness[a]	98.9 (97.9)	86.6 (28.7)	76.9 (21.1)	64.8 (16.8)
Ellipsoidal completeness[a]	/	93.3 (49.5)	92.2 (58.8)	91.7 (69.3)
Redundancy[a]	6.9 (6.9)	14.5 (15.6)	3.6 (3.4)	3.5 (3.0)
Nr. unique reflections[a]	361,584 (51,999)	102,291 (5,117)	248,419 (12,420)	169,224 (8,461)
Refinement				
Resolution (Å)	49.78 – 1.65	58.89 – 2.43	42.88-2.15	43.74 – 2.91
Number of reflections	361148	102,286	248,376	169,113
R_{work}/R_{free}[b] (%)	16.30/18.50	22.94/26.82	17.10/20.30[d]	25.41/27.80
Number of atoms				
Protein	21,321	21,082	42,408	84,792
Ligands/ions	81	154	710	371
Solvent	2,932	293	3,041	0
Mean B-value (Å^2)	33.52	52.65	38.36	63.17
Molprobity clash score	2.32	4.01	1.99	1.09
Ramachandran plot				
Favored regions (%)	98.20	96.51	97.92	96.26
Outlier regions (%)	0	0.04	0.22	0.24
rmsd[c] bond lengths (Å)	0.012	0.004	0.009	0.010
rmsd[c] bond angles (°)	1.485	0.654	0.93	1.274
PDB ID code	8OOL	8OON	8OOO	8OOQ

[a] Values relative to the highest resolution shell are within parentheses. [b] Rfree was calculated as the Rwork for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation. [d] Rfactor and Rfree are from the PDB validation report.

Table 3 X-ray analysis statistics for *Ms*GS.

	***Ms*GS-apo 1**	***Ms*GS-apo 2**	***Ms*GS- Mg^{2+}/ATP**
Data collection			
Wavelength (Å)	1.00004	1.00002	0.97949
Space group	$P2_1$	$P4_332$	$P2_1$
Resolution (Å)	109.99 - 2.64 (2.93 - 2.64)	130.75 - 3.09 (3.17 - 3.09)	78.85 - 2.70 (2.88 - 2.70)
Cell dimensions			
a, b, c (Å)	130.92, 195.65, 133.44	226.46, 226.46, 226.46	131.55, 197.40, 135.17
α, β, γ (°)	90, 94.71, 90	90, 90, 90	90, 94.89, 90
R_{merge}(%)[a]	7.6 (87.0)	34.5 (309.9)	6.2 (68.8)
R_{pim} (%)[a]	3.5 (46.7)	6.8 (60.7)	3.3 (37.0)
$CC_{1/2}$ [a]	0.999 (0.600)	0.998 (0.486)	0.999 (0.701)
I/σ_I[a]	15.5 (1.6)	12.2 (1.4)	17.4 (2.0)
Spherical completeness[a]	62.8 (11.5)	96.6 (63.9)	68.5 (19.8)
Ellipsoidal completeness[a]	90.3 (78.7)	96.6 (63.8)	88.1 (89.8)
Redundancy[a]	5.7 (4.4)	26.6 (26.8)	4.3 (4.4)
Nr. unique reflections[a]	123,574 (6,180)	35,711 (1,793)	128,572 (6,429)
Refinement			
Resolution (Å)	62.96 - 2.64	65.37 - 3.09	49.10 - 2.70
Number of reflections	123,529	35,696	128,533
R_{work}/R_{free}[b] (%)	19.14/22.52	18.63/21.62	19.61/21.72
Number of atoms			
Protein	40,821	6,920	40,815
Ligands/ions	76	64	431
Solvent	66	0	33
Mean B-value (Å^2)	72.24	80.11	70.78
Molprobity clash score	2.19	2.30	2.60
Ramachandran plot			
Favored regions (%)	98.14	96.31	97.95
Outlier regions (%)	0	0	0.02
rmsd[c] bond lengths (Å)	0.011	0.004	0.011
rmsd[c] bond angles (°)	1.35	0.678	1.348
PDB ID code	8OOW	8OOX	8OOZ

[a] Values relative to the highest resolution shell are within parentheses. [b] Rfree was calculated as the Rwork for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

Discussion

In the absence of a GDH system, GS represents the main entry point for ammonia assimilation, feeding the cellular metabolism with nitrogen. The ATP-dependent reaction must be under tight control to fit cellular needs. The GS, as does their regulation, come in different flavors in eukaryotes and prokaryotes[8, 18, 19, 26-28]. Many studies have been performed on bacterial and eukaryotic enzymes, but only a few are focused on the GS from the domain *Archaea*. This work unveiled the first archaeal GSI-α crystal structures. Their homo-dodecameric organization is highly similar to their bacterial counterparts. This is in accordance with published phylogenetic analyses in which archaeal and bacterial GSI-α are part of a monophyletic group (Fig. 6 [9]). Our results expand the knowledge gathered on this enzyme and aim to dissect the regulation mechanism at the molecular level.

2OG is a sensor for cellular nitrogen availability and has recently been recognized as a master regulator of multiple biochemical pathways[30], especially other elements of nitrogen assimilation. For instance, most of the P_{II}-family regulatory proteins, operating the signaling for nitrogen fluxes (including the regulation of GS), are regulated by 2OG[5, 30, 35-37]. Previous studies highlighted 2OG as a direct regulator of GSI-α in *Archaea*[26-28], such as *Haloferax mediteraneii* and *Methanosarcina mazei*, in which 2OG activates the enzyme by 12-fold[26] and 16-fold[28], respectively. Our study found that *Mt*GS activity is strictly dependent on this metabolite under the described experimental conditions. *Mt*GS activation is saturated at 0.6 mM 2OG and is inactive below 0.06 mM. This range remarkably fits the measured 2OG cellular concentration in the methanogen *Methanococcus maripaludis*, which is around 0.8 mM under N_2-fixing conditions (nitrogen-limited) and 0.08 mM after ammonium addition[38]. Relying on 2OG concentration as a sensor of cellular nitrogen availability could benefit an energy-limited organism such as *M. thermolithotrophicus*, and might represent a “primitive” regulation mechanism, which evolved before more elaborate regulatory networks

emerged. In addition, regulation at the transcriptional level also occurs as *M. thermolithotrophicus* upregulates *glnA* expression when cells switch to N_2 fixation[39], and it is known that small RNAs are involved in nitrogen metabolism regulation in prokaryotes, including methanogens[40].

The structure of *Mt*GS reveals an allosteric pocket perfectly suited to accommodate 2OG via specific salt bridges, hydrogen bonds, and Van der Waals contacts. Binding of 2OG leads to a succession of conformational rearrangements resulting in a catalytically competent state. A conservation of all five residues of the 2OG binding site can be found in other orders of archaea (Fig. 6 and S19). However, while the motif would allow 2OG binding, the influence on the activity remains to be verified as several other structural features are involved in the accurate positioning of catalytic residues. A 2OG binding motif is far from being a feature shared by all archaea, as exemplified by the characterization of the 2OG-insensitive *Ms*GS. Previous works on the enzymes from *H. mediterranei* and *M. mazei* pointed towards a unified 2OG activation mechanism as main regulation in archaeal GS. However, spreading the analysis to other species rather highlights a variety of different regulation mechanisms depending on archaeal groups and enzyme isoforms (Fig. 6). The glutamine inhibition for *Ms*GS, previously described in *Bs*GS[18], is dependent on one essential arginine (R62′) responsible for the Glu-flap sequestration preventing the release of the product from the active site, which is not happening in *Mt*GS due to a substitution (Gly67′).

Even if both, 2OG and glutamine, act through different mechanisms, *e.g.,* competitive inhibition *versus* allosteric activation, it is interesting that their modulation acts on the same key determinants. In the resting state of *Mt*GS without 2OG or the glutamine feedback-inhibited *Bs*GS, the Glu-flap and the catalytic arginine are locked in an unproductive conformation by the Asp50′-loop. In *Mt*GS the 2OG binding provokes a major displacement of the N-terminal domain towards the adjacent C-terminal domain

and the overall conformational changes promote a restructuration of the Asp50′-loop, leading to the repositioning of Arg321 and of the Glu-flap, yielding an active conformation. Additionally, the 2OG-dependent motions provide enough flexibility to allow the correct structuration of the Mg^{2+}/ATP binding site, similar to *Ms*GS in its resting state.

Kinetic measurements revealed similar apparent K_m for ammonia and glutamate between both GS (Table 1) and are comparable to previous studies[18, 41]. The determined V_{max} also lies within the previously reported range of specific activity for archaeal and bacterial GS (up to around 6 μmol.min^{-1}.mg^{-1} in both *H. mediterranei* and *M. mazei*[26, 28] and up to 23 μmol.min^{-1}.mg^{-1} for *B. subtilis*[42]). P_{II}-family proteins might play a role in *Ms*GS and *Mt*GS activation that could mediate 2OG sensitivity, a hypothesis that future exploratory studies will hopefully confirm. In contrast, the sP26 protein described in *M. mazeii* is not encoded in the genomes of *M. thermolithotrophicus* and *M. shengliensis*, and the same applies to the repressor GlnR forming higher oligomer species in the bacterial systems. However, non-homologous functional equivalents may exist. Similarly, if our experiments do not suggest any modification of the oligomeric state, such a regulatory mechanism could occur in the cell in the presence of other regulatory partners, which might become separated during the purification process.

The works gathered on GS illustrate that despite the universal requirement of assimilating ammonia to fuel nitrogen metabolism and a conserved reaction mechanism, microbes rely on different ways to modulate this ATP-dependent activity. Such regulatory networks might have been elaborated over time to adapt to a particular physiology, environment, or catabolism. Opening scientific investigation to a broader group of organisms will hopefully contribute to extending our knowledge of this crucial and ancient enzyme.

Material and Methods

Cultivation

M. thermolithotrophicus (DSM 2095) cells were obtained from the Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). Cells were grown in a minimal medium as described in Jespersen *et al.*[43]. Anaerobic cultivation of *M. thermolithotrophicus* was performed in a fermenter with NH_4Cl and sulfate as described in Jespersen & Wagner[44] with slight modifications, or in a fermenter with N_2 and sulfate as described in Maslać *et al.*[39]. We did not observe dramatic changes in the expression of GS nor its properties between both cultivation processes. *M. shengliensis* ZC-1 (DSM 18856) was also obtained from the DSMZ (Braunschweig, Germany) and was grown anaerobically on methanol as previously described in Kurth & Müller *et al.*[45].

Purification

Cell lysis and extracts were prepared in an anaerobic chamber at room temperature filled with an N_2/CO_2 atmosphere (90:10 %). Enzyme purification was carried out under anaerobic conditions in a Coy tent filled with an N_2/H_2 atmosphere (97:3 %), at 20 °C and under yellow light. For each step, chromatography columns were washed with at least three column volumes (CV) of the corresponding loading buffer, and samples were filtrated on 0.2 µm filters (Sartorius, Germany) prior to loading. During purification, the enzyme was followed by sodium dodecyl sulfate polyacrylamide-gel electrophoresis (SDS-PAGE) and absorbance monitoring at 280 nm. Purifications were performed with adapting protocols, and the optimal one is described below for *Mt*GS and *Ms*GS.

*Mt*GS purification

25 g (wet weight) cells were thawed and diluted in 200 mL lysis buffer (50 mM Tricine/NaOH pH 8, 2 mM dithiothreitol (DTT)), sonicated (~5 × 10 s at 60 %, probe

KE76 Bandelin SONOPULS, Germany) and centrifuged at 45,000 × *g* for 45 min at 18 °C. The supernatant was passed twice on a 4 × 5 mL HiTrap™ DEAE Sepharose FF column (Cytiva, Sweden). Elution was performed with a NaCl gradient ranging from 150 to 500 mM in the same buffer. The gradient was applied for 90 min at 2 mL.min^{-1}, and *Mt*GS was eluted at 170 to 260 mM NaCl. The pool was diluted 1:4 in 20 mM NaH_2PO_4 pH 7.6, 2 mM DTT. The sample was passed twice on 2 x Mini CHT ™ type I column (Bio-Rad, United States). Elution was performed with a gradient of 20 to 200 mM NaH_2PO_4 in 10 min, followed by a second gradient from 200 to 500 mM NaH_2PO_4 in 10 min. The gradient was run at 1.5 mL.min^{-1}. Under these conditions, *Mt*GS was eluted at 20 to 200 mM NaH_2PO_4. The resulting pool was concentrated to 900 µL with a 30-kDa cutoff concentrator (Sartorius, Germany). The sample was injected thrice on a Superose™ 6 Increase 10/300 GL column (Cytiva, Sweden). The protein was eluted at 0.4 mL.min^{-1} in 25 mM Tris/HCl pH 7.6, 10 % glycerol, 150 mM NaCl, 2 mM DTT. The resulting pool was diluted 1:1 in 25 mM Tris/HCl pH 7.6, 2 M $(NH_4)_2SO_4$, 2 mM DTT, and injected on a Source™15 Phe 4.6/100 PE column (Cytiva, Sweden). Elution was performed with a gradient ranging from 1 to 0 M $(NH_4)_2SO_4$ for 60 min at 0.5 mL.min^{-1}. *Mt*GS was eluted between 0.86 to 0.74 M $(NH_4)_2SO_4$. The final pool was diluted 1:100 in 50 mM Tricine/NaOH pH 8.0, 2 mM DTT, and loaded on a MonoQ™ 5/50 GL column (Cytiva, Sweden). *Mt*GS was eluted with a NaCl gradient ranging from 0 to 200 mM NaCl in 10 min followed by 200 to 500 mM NaCl in 60 min at 0.5 mL.min^{-1} (eluted at 375-395 mM NaCl). The final pool was washed 1:1000 in 25 mM Tris/HCl pH 7.6, 10 % glycerol, 150 mM NaCl, 2 mM DTT and concentrated to 200 µL with a 30-kDa cutoff concentrator (Sartorius, Germany). The protein concentration was estimated via Bradford assay, and the sample was flash-frozen and stored anaerobically at −80 °C.

*Ms*GS purification

The pellet (4 g) was suspended in 50 mM Tris/HCl pH 8.0 and 2 mM DTT (lysis buffer). Cell lysis and preparation of extracts were performed similarly, except that the pellet after centrifugation was resuspended in the lysis buffer, sonicated and centrifuged a second time to extract additional proteins. Soluble fractions were pooled and diluted with the lysis buffer to obtain a final 15-fold dilution (final volume 60 mL). The filtered sample was loaded on a 2 × 5 mL HiTrap™ DEAE Sepharose FF (Cytiva, Sweden) equilibrated with the same buffer. The protein was eluted with a 0 to 400 mM NaCl linear gradient for 150 min at a 2 mL.min^{-1} flow rate. *Ms*GS eluted between 240 mM and 310 mM NaCl. The pooled sample was diluted with three volumes of lysis buffer and was loaded on a 5 mL HiTrap Q HP™ column (Cytiva, Sweden). The protein was eluted with a 150 to 500 mM NaCl linear gradient for 70 min at a 1 mL.min^{-1} flow rate. *Ms*GS eluted between 0.38 and 0.41 M NaCl under these conditions. The pooled sample was diluted with 1 volume of 25 mM Tris/HCl pH 7.6, 2 M $(NH_4)_2SO_4$, 2 mM DTT before loading on a 5 mL HiTrap™ Phenyl Sepharose HP column (Cytiva, Sweden) equilibrated with the same buffer. *Ms*GS was eluted with a 1.1 to 0 M $(NH_4)_2SO_4$ linear gradient for 70 min at a 1 mL.min^{-1} flow rate. The protein eluted between 0.91 M and 0.81 M $(NH_4)_2SO_4$. Pooled fractions were concentrated on a 10-kDa cutoff centrifugal concentrator (Sartorius, Germany), and the buffer was exchanged for 25 mM Tris/HCl pH 7.6, 10 % glycerol, 2 mM DTT.

Size estimation of the GS

The protocol of hrCN PAGE was adapted as described in Lemaire *et al.* [46] (originally described in Lemaire *et al.*[47]) and run at 40 mA for 1 h using 5-12 % gradient gels. The size determination of *Mt*GS (540.3 kDa) and *Ms*GS (547.5 kDa) was obtained using a fit derived from the migration distances and sizes of the standard proteins.

Size exclusion chromatography was carried out under anaerobic conditions in a Coy tent filled with an N_2/H_2 atmosphere (97:3 %), at 20 °C and under yellow light, using a Superose^{TM} 6 Increase 10/300 GL column (Cytiva, Sweden). Chromatography was performed in 25 mM Tris/HCl buffer, pH 7.6, 10 % glycerol, 2 mM DTT, at a flow rate of 0.4 mL.min^{-1}. 76 and 67.5 µg of *Mt*GS and *Ms*GS, respectively, were used. Size determination was performed using a fit derived from the elution volumes and sizes of the standard proteins.

Activity assays

GS activities were measured using the pyruvate kinase/ lactate dehydrogenase (PK/LDH) coupled enzymes from rabbit muscle ordered from Sigma-Aldrich (containing 600-1000 units.mL^{-1} pyruvate kinase and 900-1400 units.mL^{-1} lactate dehydrogenase). The activity was measured by following NADH oxidation resulting in the change of absorbance at 340 nm (Fig. S2).

Absorbance was measured aerobically in a 96-well plate with a SPEKTROstarNano (BMG Labtech, Germany) at 45 °C in 100 µL final reaction volume. GS was added immediately after opening the anaerobic storage flask to minimize the negative effects of oxygen and low temperature. Activities were performed with an enzyme coming from a single preparation and measured in technical triplicates.

All reagents were prepared in the reaction buffer (200 mM KH_2PO_4 pH 7, 10 mM KCl) apart from $MgCl_2 \times 6\ H_2O$ which was dissolved in water. The standard reaction mix contained as a final concentration: 1 mM NADH (freshly prepared), 2 mM phosphoenolpyruvate (freshly prepared), 3 mM ATP (freshly prepared), 50 or 80 mM sodium glutamate (for *Mt*GS and *Ms*GS, respectively), 20 mM NH_4Cl, 25 mM $MgCl_2 \times 6H_2O$, 2 mM sodium 2OG (freshly prepared), 0.02 mg.mL^{-1} GS and 5 µL PK/LDH from a 1:10 dilution of the stock. Kinetic parameters were determined by varying the glutamate concentrations from 0 to 100 mM for *Mt*GS and 0 to 200 mM for *Ms*GS, and

NH_4Cl from 0 to 5 mM for *Mt*GS and 0 to 20 mM for *Ms*GS. The determination of the $K_{0.5}$ and Hill coefficient for *Mt*GS activation was performed by varying the 2OG concentration from 0 to 5 mM. Glutamine feedback inhibition was measured at 20 mM glutamine. Unspecific interactions with malate and succinate for *Mt*GS were measured at 15 mM of the compounds by preparing the reaction mix without 2OG, measuring for 20 min and then adding 2OG (2 mM final). O_2 sensitivity was measured in anaerobic cuvettes with 0.01 $mg.mL^{-1}$ *Mt*GS, starting the reaction with glutamate addition. All solutions were prepared anaerobically, and then half of the master mix, GS and glutamate were separately incubated for 1 h at ambient air. The anaerobic or aerobic compounds were combined, and activity was measured immediately. It has been previously observed that certain buffer conditions contribute to a faster activity loss, including compounds like β-mercaptoethanol or dithiothreitol[41]. Therefore, it should be considered when comparing specific activities from different studies, as it might also affect the O_2-effect on enzymes.

Rates were measured by exploiting the linear regression for each dataset through the points with the steepest, most linear slope. The control slope without the enzyme was subtracted from the slope obtained with the enzyme in the same timeframe. Activities are presented in µmol of NADH oxidized per minute per mg of added GS, using a molar extinction coefficient of ε_{340nm}= 6220 $M^{-1}.cm^{-1}$ for NADH. The apparent K_m and V_{max} were calculated with the K_M V_{max} Tool Kit (ic50tk/kmvmax.html).

Crystallization

All proteins were crystallized fresh without any freezing step, and all crystals were obtained through the sitting drop method on a 96-Well MRC 2-Drop Crystallization Plates in polystyrene (SWISSCI, United Kingdom) at 20 °C under anaerobic conditions (N_2/H_2, gas ratio of 97:3).

*Mt*GS apo without TbXo4 was crystallized at a concentration of 15 mg.mL^{-1}. The crystallization reservoir contained 90 µL of mother liquor (20 % w/v polyethylene glycol 3,350 and 100 mM potassium sodium tartrate), the crystallization drop contained a mixture of 0.55 µL protein and 0.55 µL precipitant. Hexagonal plates appeared after few weeks and were soaked in the mother liquor supplemented with 25 % v/v ethylene glycol prior to freezing in liquid nitrogen.

*Mt*GS with 2OG/Mg^{2+}/ATP was crystallized at 3.7 mg.mL^{-1} with a final concentration of 2 mM 2OG, 2 mM ATP and 2 mM MgCl. The crystallization reservoir contained 90 µL of mother liquor (20 % w/v polyethylene glycol 3,350 and 200 mM sodium fluoride). The crystallization drop contained 0.6 µL GS with ligands and 0.6 µL precipitant. Short thick hexagonal rods appeared after few weeks and were soaked in the mother liquor supplemented with 20 % v/v glycerol prior to freezing in liquid nitrogen.

*Mt*GS with 2OG/Mg^{2+} was crystallized at 3.4 mg.mL^{-1} with a final concentration of 2 mM 2OG and 2 mM MgCl. The crystallization reservoir contained 90 µL of mother liquor (25 % w/v polyethylene glycol 1,500 and 100 mM SPG (succinic acid, sodium dihydrogen phosphate, and glycine) buffer pH 5.0), and the crystallization drop contained 0.6 µL GS with ligands and 0.6 µL precipitant. Hexagonal plates obtained after few weeks were soaked in the mother liquor supplemented with 15 % v/v glycerol prior to freezing in liquid nitrogen.

All *Ms*GS crystals were obtained by mixing 0.55 µL protein at 9 mg.mL^{-1} with 0.55 µL precipitant.
For *Ms*GS apo 1 the crystallization reservoir contained 90 µL of 200 mM ammonium formate and 20 % (w/v) polyethylene glycol 3,350. Rectangular rods appeared within weeks, and crystals were soaked in the mother liquor supplemented with 25 % v/v ethylene glycol prior to freezing in liquid nitrogen.

*Ms*GS-Mg^{2+}/ATP was obtained from the same condition. Here, crystals were soaked for 4 min in the crystallization solution containing ATP, $MgCl_2$ and sodium glutamate, each at a final concentration of 10 mM, then back soaked in the crystallization solution supplemented with 25 % glycerol prior to freezing in liquid nitrogen. It is worth noting that no additional electron density could be attributed despite soaking the *Ms*GS crystal with 10 mM glutamate. Therefore, the structure was named *Ms*GS-Mg^{2+}/ATP bound complex. The absence of binding might come from a packing artifact, a low occupancy of magnesium in the active site, or the aberrant position of the ATP phosphate backbone.

For *Ms*GS apo 2 the crystallization reservoir contained 90 µL of 1.6 M sodium citrate tribasic dihydrate. Flat squares appeared within weeks, and crystals were soaked in the mother liquor supplemented with 30 % v/v glycerol prior to freezing in liquid nitrogen.

Data collection, processing and structure refinement

All datasets were collected at 100 K at the Swiss Light Source (SLS), beamline PXIII (X06DA). Data processing was performed with *autoPROC*[48] combined with STARANISO[49] except for *Mt*GS apo co-crystallized with the TbXo4 (obtained from Engilberge *et al.*[31]), which was processed by XDS[50] and scaled with SCALA from the CCP4 package. The structure of *Mt*GS-apo-TbXo4 was solved by molecular replacement with MOLREP (CCP4)[51] by using the closest structural homolog from *Bacillus subtilis* (*Bs*GS resting state, PDB 4LNN). Other *Mt*GS structures were solved by molecular replacement with PHASER from the PHENIX package with the *Mt*GS-apo-TbXo4 model. *Mt*GS-2OG/Mg^{2+}/ATP was used as template to solve *Mt*GS-2OG/Mg^{2+} and *Ms*GS-apo 1 structures with PHASER. The two other *Ms*GS structures were solved with PHASER by using *Ms*GS-apo 1 as template.

Refinement was performed with *PHENIX*[52] and *BUSTER*[53] in combination with fast automatic visual model building in *COOT*[54]. All models were systematically validated by using Molprobity[55]. *Mt*GS apo co-crystallized with TbXo4 was refined with

considering all atoms except water anisotropic. Additionally, riding hydrogens were added during the refinement. *Mt*GS apo without TbXo4 was refined by applying translation libration screw model (TLS) and adding riding hydrogens. Hydrogens were omitted in these final deposited models. In addition, this dataset was refined by applying the following twin operator: 1/2h+1/2k,3/2h-1/2k,-l (twin fraction of 0.12). *Mt*GS-2OG/Mg^{2+}/ATP was refined in BUSTER by applying non crystallography symmetry (NCS) and TLS, while *Mt*GS-2OG/Mg^{2+} was refined in PHENIX with TLS and without NCS and with riding hydrogens. All *Ms*GS models were refined by applying TLS, without NCS and without generating hydrogens.

The PDB ID codes, data collection and refinement statistics for the deposited models are listed in Tables 2 and 3. The structures of *Mt*GS-apo without TbXo4 (PDB 8OON) and *Mt*GS-2OG/Mg^{2+} (PDB 8OOQ) presented R-factors slightly higher than that averaged from structures of similar resolution, due to imperfections in the crystals. The refinement and electron density quality of the *Mt*GS-apo without TbXo4 structure was hampered by pseudomerohedral twinning. Numerous crystals were analyzed to remove or reduce the twinning, which was only obtained by adding the TbXo4 molecule[31]. In the present work, this structure is mainly used as a control to exclude an effect of TbXo4 on the protein. A total of 114 different crystals were analyzed in order to obtain the structure of *Mt*GS-2OG/Mg^{2+}, and the presented structure was obtained from the best dataset from this intensive screening and collection.

Structural analyses.

All structure visualization was performed with PyMol (2.2.0 Schrödinger, LLC, New York, USA). Rmsd putty graphs were generated by generating rmsd values with SUPERPOSE from the CCP4 package[51], substituting the B-values in the according PDB file with the new rmsd values and displaying it in the Putty preset in PyMol. Coloring

was set to "spectrum b, yellow_orange_magenta, minimum=0, maximum=5". Putty scaling was set to absolute linear scaling from the B-factor column.

When chain selection was necessary for structural alignment, the chain with the lowest average B-factor was used. The structures do not exhibit any major differences of B-factors between chains, and therefore similar conclusions would have been drawn with the selection of other chains for analysis.

The movie visualizing the conformational changes upon 2OG binding was generated with UCSF Chimera (Version 1.16, University of California, USA). A morph between *Mt*GS-apo-TbXo4 to *Mt*GS-2OG/Mg^{2+} to *Mt*GS-2OG/Mg^{2+}/ATP is shown. The enzyme is displayed as cartoon with the ligands as spheres. 2OG carbons are colored yellow and ATP carbons light blue, while the other atoms are colored according to element (O in red, N in blue, P in orange). In the second half a zoom-in of the active site is shown with the Asp-50′ loop in a darker blue and Arg321 is displayed as sticks. The two chains composing the active site are colored in different shades of cyan.

Sequence alignments were generated with Clustal Omega[56] and superposed to the secondary structure with the Espript 3.0 server (https://espript.ibcp.fr)[57].

The oligomerization state of all obtained structures was predicted via PDBePISA (Proteins, Interfaces, Structures and Assemblies, https://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver). To facilitate processing, 8OOQ was processed as two separate dodecamers, and ligands were removed from 8OOO. The two best results for each structure are listed in Table S1.

Phylogeny and residues conservation analysis.

The *Mt*GS sequence was used as query for a BLAST P[58] search in the RefSeq database, restricted to the *Archaea* domain. This database was used to limit the redundancy of sequences from similar species. The research limit was set to 500 sequences. The

sequences of the GS from *B. subtilis*, *S. aureus*, *P. polymyxa* and *L. monocytogenes* were added, leading to a total of 504 sequences, all listed in Supplementary Data 2. The phylogenic tree was constructed by the MEGA X program[59]. The alignment was done by MUSCLE with default parameters and the tree was constructed using the Maximum Likelihood method and JTT matrix-based model [60]. The tree with the highest log likelihood (-191831.27) is shown. The node scores were calculated with 200 replications. There were a total of 656 positions in the final dataset. The branches were manually colored according to the taxonomy from the NCBI database[61]. The tree was visualized and imaged using iTol v6[62].

The sequences were manually separated in different monophyletic groups. The different sequences forming the groups were aligned by Clustal Omega[56] and the residue conservation images were constructed using Weblogo 3 (version 3.7.12)[63] based on the residues position in *Mt*GS and multiple sequences alignments.

Acknowledgments, author contributions and data availability

We would like to thank the Max Planck Institute for Marine Microbiology and the Max-Planck-Society for their continuous support. We are also grateful for the beam time allocation at SLS and the support of the PXIII beamline staff. We would also like to thank Christina Probian and Ramona Appel from the Microbial Metabolism laboratory for their continuous support, as well as Marion Jespersen and Nevena Maslać for providing *M. thermolithotrophicus* cells. We acknowledge Sylvain Engilberge, Éric Girard, Olivier Maury and François Riobé for their contribution to the *Ms*GS-TbXo4 structure. We are grateful to Stian Torset and Tomás Alarcón Schumacher for their help with data formatting for the graph generation. C.W. and J.M.K. were supported by the SIAM gravitation program (grant #024002002) granted by the Netherlands Organization for Scientific Research and the Ministry of Education, Culture and Science. J.M.K was furthermore supported by the Deutsche Forschungs Gesellschafts (DFG) Grant KU 3768/1-1.

Methanothermococcus thermolithotrophicus cultivation was mainly performed by M-C.M. *M. shengliensis* was cultivated by J.M.K. *Mt*GS and *Ms*GS were purified and crystallized by M-C.M and O.N.L., respectively. Biochemical characterization and activity assays were performed by M-C.M and O.N.L. X-ray data collection was performed by M-C.M., O.N.L and T.W. Data processing, model building, structure refinement, validation and deposition were performed by M-C.M and T.W. Structures were analyzed by M-C.M. and T.W. O.N.L. performed phylogenetic analyses. C.U.W. and T.W. acquired funding to realize the project. The paper was written by M-C.M, O.N.L. and T.W. with contributions and final approval of all co-authors.

All protein models were deposited in the Protein data bank under the following PDB codes: 8OOL (*Mt*GS-apo-TbXo4), 8OON (*Mt*GS-apo without TbXo4), 8OOO *Mt*GS-2OG/Mg^{2+}/ATP), 8OOQ (*Mt*GS-2OG/Mg^{2+}), 8OOW (*Ms*GS-apo 1), 8OOX (*Ms*GS-apo 2), 8OOZ (*Ms*GS-Mg^{2+}/ATP). Source data for the graphs in Fig. 1 and Fig. 5 can

be found in the Supplementary Data 1. Biochemical and crystallographic raw data will be made available on request.

Authors declare no competing interests.

Supplementary information

Table S1. Oligomeric state prediction by PDBePISA. Predictions were made using the asymmetric units (except for 8OOQ, see below), the two top results being listed in the table 8OOQ was analyzed as two separate dodecamers, and ligands were removed from 8OOO to facilitate processing.

Structure	PDB code	Monomer / Asymmetric unit	Multimeric state for calculations	ΔG^{int} kcal/mol	ΔG^{diss} kcal/mol
*Mt*GS-apo-TbXo4	8OOL	6	12	-811.42	219.48
			6	-323.76	41.76
*Mt*GS-apo without TbXo4	8OON	6	12	-30.34	149.81
			6	27.22	10.64
*Mt*GS-2OG/Mg^{2+} (A)	8OOQ	2x12	12	-237.25	195.65
			6	-56.28	86.29
*Mt*GS-2OG/Mg^{2+} (B)	8OOQ	2x12	12	-190.32	175.67
			6	-47.76	86.65
*Mt*GS- 2OG/Mg^{2+}/ATP	8OOO	12	12	-262.52	189.38
			6	-68.01	79.33
*Ms*GS-apo 1	8OOW	12	12	-348.95	200.48
			6	-103.83	35.98
*Ms*GS-apo 2	8OOX	2	12	-380.81	182.89
			6	-121.07	53.67
*Ms*GS-Mg^{2+}/ATP	8OOZ	12	12	-413.23	219.03
			6	-118.09	44.24

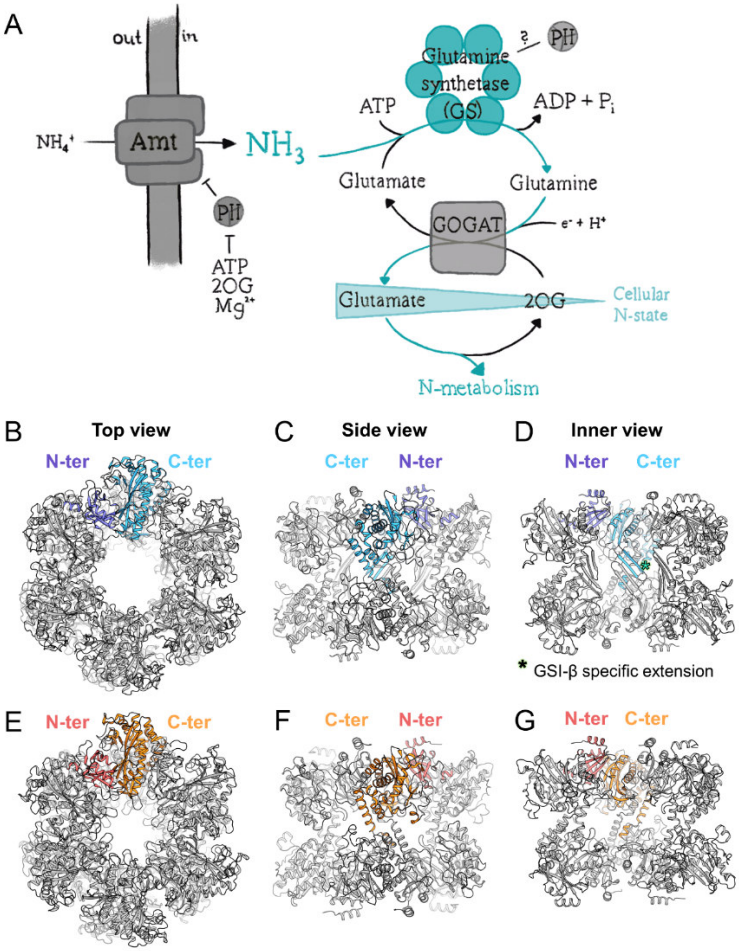


Figure S1: Physiological role of GS-GOGAT in methanogens and structural organization of GSI-α and GSI-β. (**A**) GS-GOGAT reaction scheme. 2OG concentration increases under nitrogen starvation, while it decreases when cellular nitrogen is available. (**B-G**) Comparison of homododecameric GSI-β from *Escherichia coli* (B-D, PDB 7W85) and GSI-α from *B. subtilis* (E-G, PDB 4LNN). The proteins are in cartoons with the highlighted N-terminal (dark blue and red) and C-terminal (light blue and orange). (B and E) Top view of the dodecameric GS. (C and F) outer-side view. (D and G) Inner-side view. An asterisk marks the extension characterizing the GSI-β.

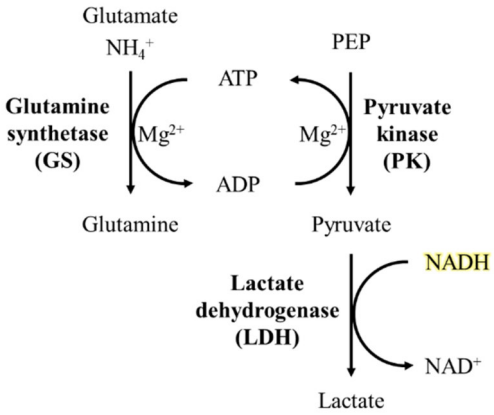


Figure S2: Illustration of the coupled enzyme assay used in this study. The reaction was tracked by following NADH oxidation at 340 nm. Enzyme names are bold. PEP stands for phosphoenolpyruvate. The phosphate liberated by the GS reaction is not shown.

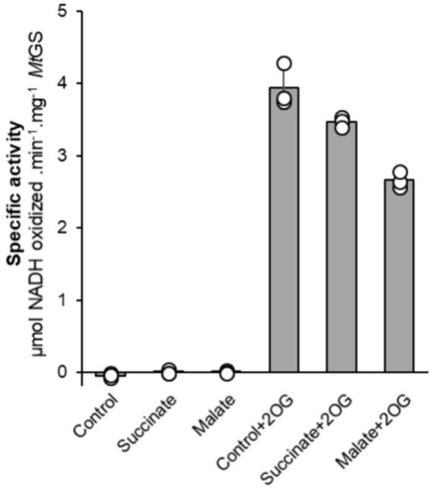


Figure S3: 2OG specificity of *Mt*GS. Specific activity of *Mt*GS after incubation with succinate or malate (15 mM final, first three columns) and after 20 min incubation with succinate and malate (15 mM final) and subsequent 2OG addition (2 mM 2OG, last three columns). Data is represented as mean ± s.d and individual values are shown as white circles (n=3).

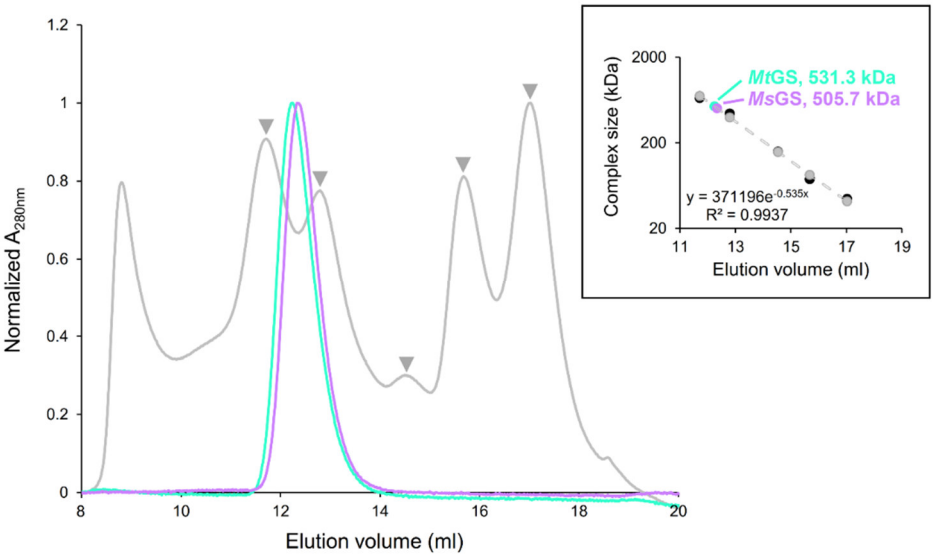


Figure S4. Size exclusion chromatography profile of the purified archaeal GS. Chromatograms of *Mt*GS and *Ms*GS are shown as cyan and purple lines, respectively. The High Molecular Weight (HMW) calibration kit (GE Healthcare) was used as a standard, and its chromatogram is shown as a grey line. Grey arrows indicate the peaks corresponding to the different proteins of the standard used for size estimation: from left to right, thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa) and ovalbumin (44 kDa). All chromatograms were normalized for clarity. Inset: determination of the size of the GS complex. The elution volumes of the standard proteins (black dots) were used to establish a fit (dashed grey line) to estimate the molecular weight of GS (cyan and purple dots). Grey dots are experimental calculations of the size of the standard proteins.

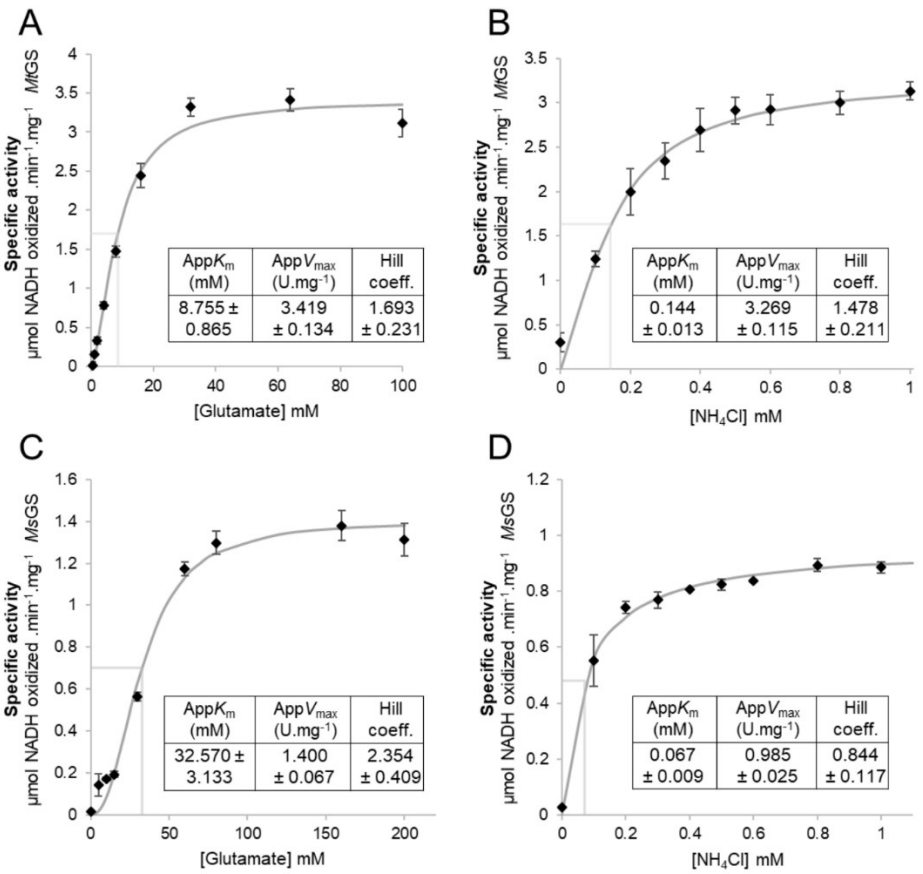


Figure S5: Kinetic parameters of archaeal GS for glutamate and ammonium. Determination of the kinetic parameters of *Mt*GS for glutamate (**A**) and ammonium (**B**). Determination of the kinetic parameters of *Ms*GS for glutamate (**C**) and ammonium (**D**). The following data points are not shown in the figure for clarity reasons: panel (B) NH_4Cl: 5 mM, specific activity of 3.20 ± 0.15 µmol.min^{-1}.mg^{-1} *Mt*GS and panel (D) NH_4Cl: 2.5 mM, specific activity 0.91 ± 0.02 µmol.min^{-1}.mg^{-1} *Ms*GS and 20 mM, 1.00 ± 0.01 µmol.min^{-1}.mg^{-1} *Ms*GS. Data is represented as mean ± s.d (*n*=3).

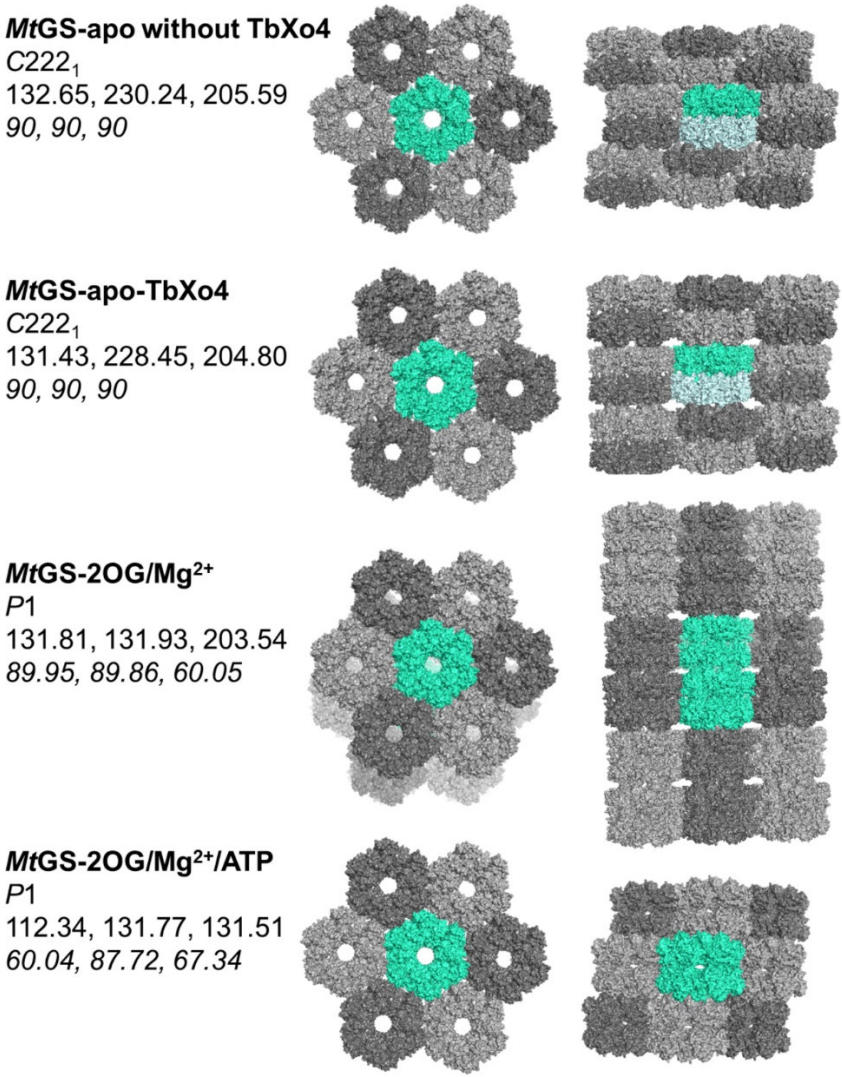


Figure S6: Crystal packing of the *Mt*GS structures. Top and side view of the asymmetric unit in the crystal packing. The asymmetric unit is colored in cyan, symmetry mates are colored in different grey shades, and if the biological dodecamer is generated through a symmetry operator, the additional chains are colored in light blue. The unit cell parameters are given in a (Å), b (Å), c (Å), α (°), β (°), γ (°) format.

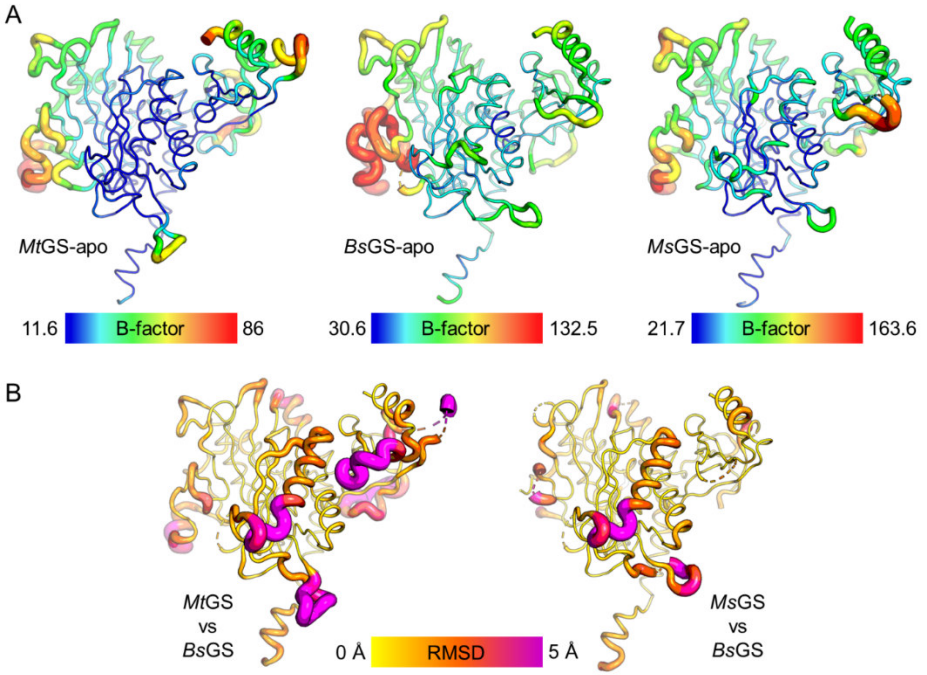


Figure S7: Structural variation between *Mt*GS/*Ms*GS apo and *Bs*GS apo. (**A**) Cartoon putty of Cα B-factors of the chain with the lowest average B-factor. (**B**) Root mean square deviation (rmsd) values between *Mt*GS or *Ms*GS with *Bs*GS. Average rmsd values were generated from aligning all chains of *Mt*- or *Ms*GS on chain A of *Bs*GS apo (PDB 4LNN).

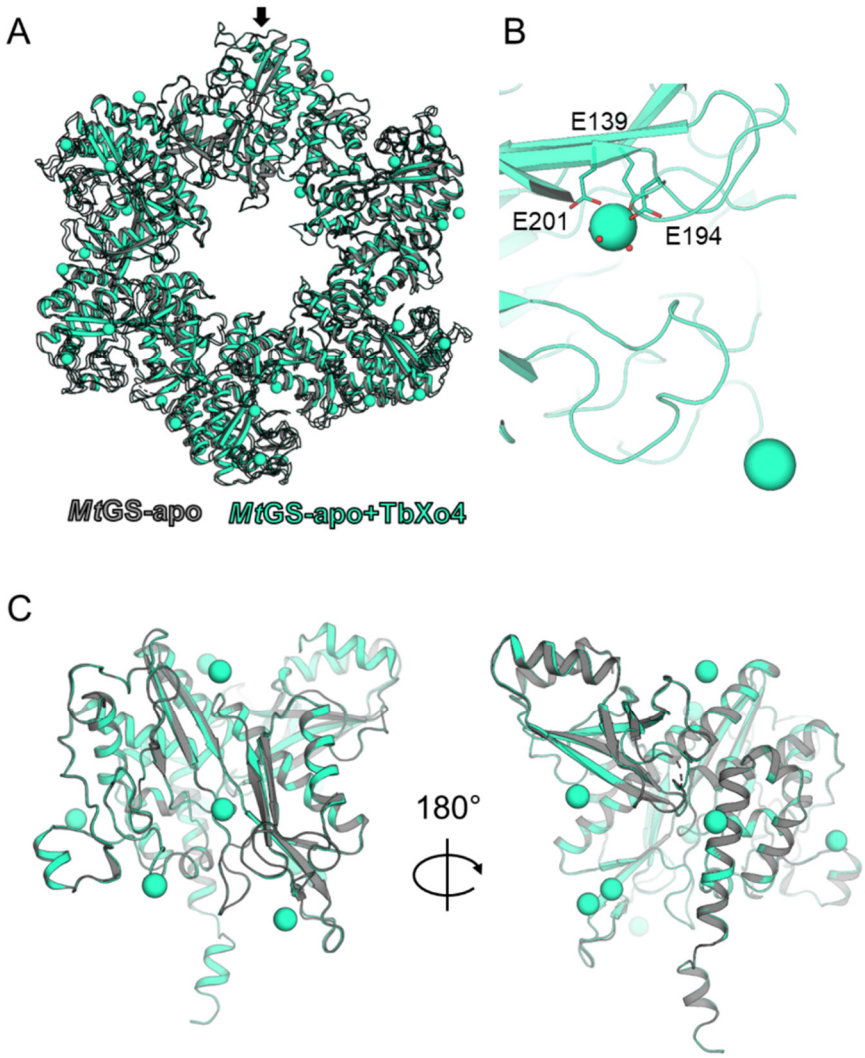


Figure S8: Differences between *Mt*GS apo state with or without TbXo4. (**A**) Overview of the hexamer from the top. An arrow indicates the subunit used for superposition. (**B**) Binding of Tb in the active site. (**C**) Overview of the superposed monomers with the same pose as in Fig. 2c. For A-C, models are represented in cartoon with Tb (cyan) and H_2O as spheres and the coordinating residues as sticks with oxygen atoms in red.

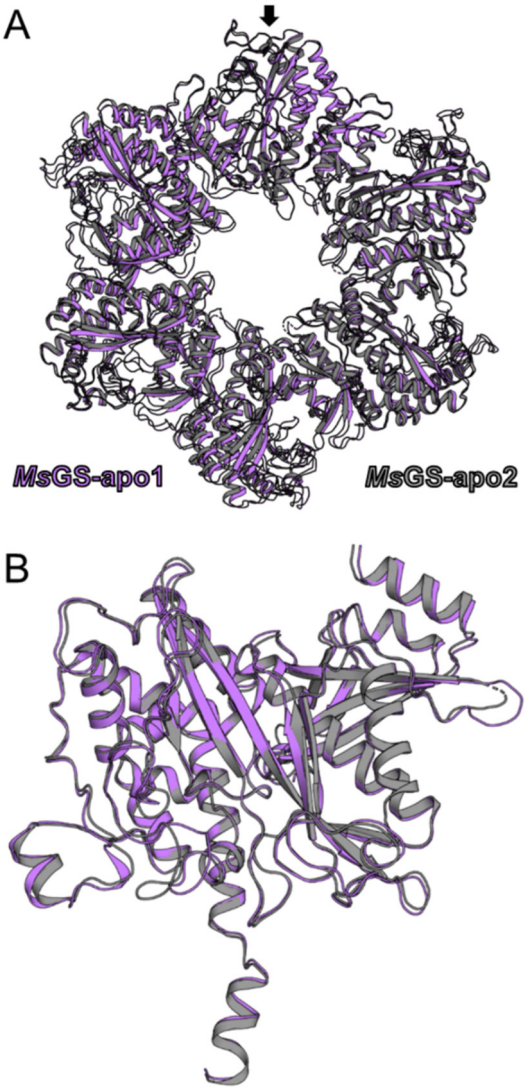


Figure S9: Differences between the two apo forms of *Ms*GS. (**A**) Superposition of the *Ms*GS apo1 and apo2 hexamers. An arrow indicates the subunit used for superposition. (**B**) Superposition of the monomers. All models are represented as cartoons with the same pose as in Fig. 2f.

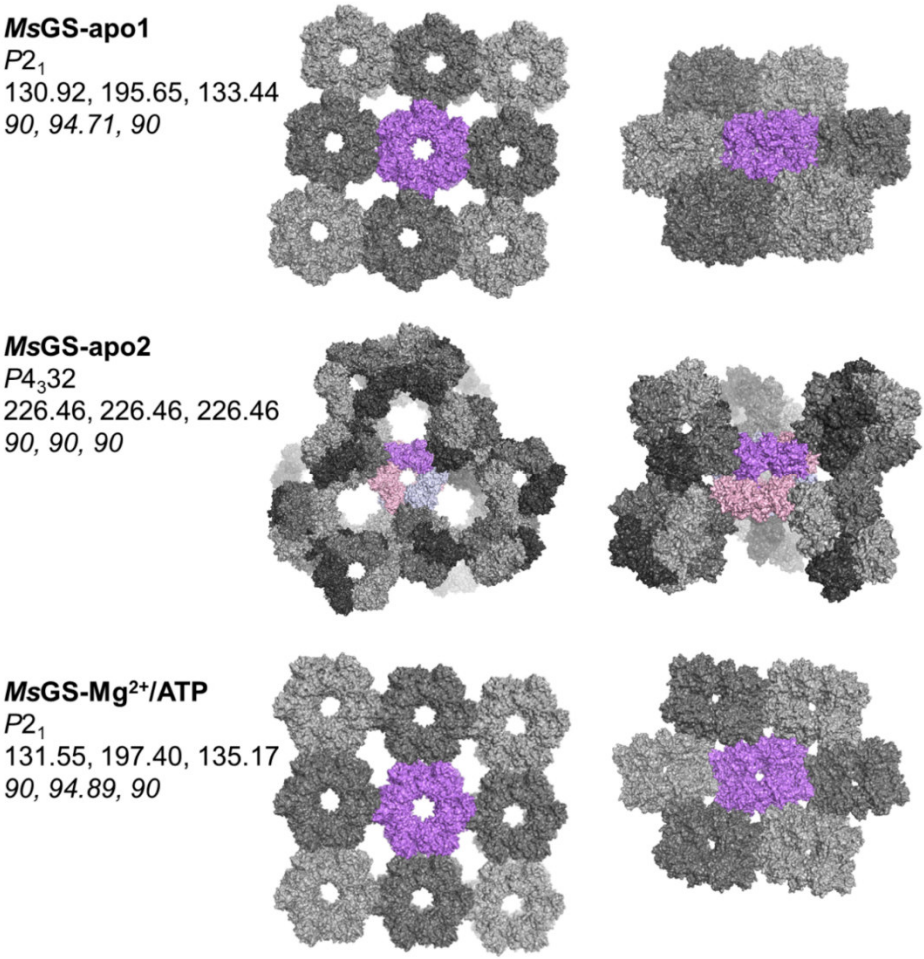


Figure S10: Crystal packing of the obtained *Ms*GS structures. Top and side view of the asymmetric unit in the crystal packing. The asymmetric unit is colored in purple, symmetry mates are colored in different grey shades, and if the biological dodecamer is generated through a symmetry operator, the additional chains are colored in lighter shades of pink. The unit cell parameters are given in a (Å), b (Å), c (Å), α (°), β (°), γ (°) format.

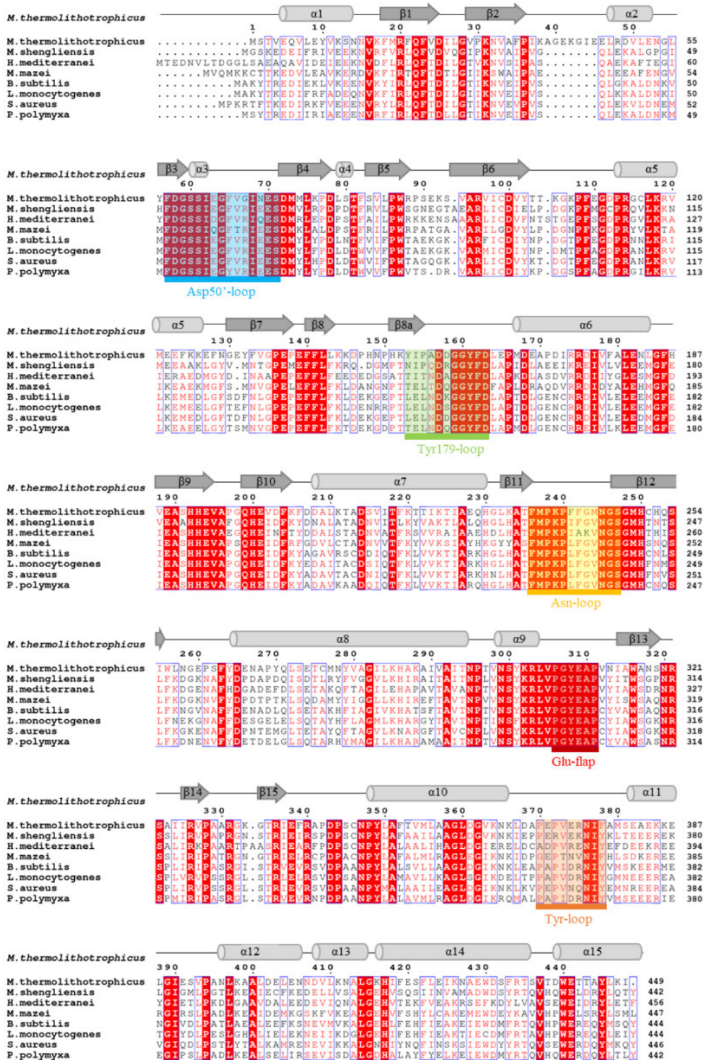


Figure S11: Sequence conservation superposed to the secondary structure. The secondary structure is annotated based on Travis *et al.*[19] . Relevant loops are labeled. WP numbers from top to bottom: WP_018154487.1, WP_042685700.1, CAR67815.1 (*H. mediterranei* GS3), WP_011032914.1 (*M. mazei* GS1), NP_389628.1, EAC9051058.1 (*Listeria monocytogenes*), WP_086038154.1 (*Staphylococcus aureus*), WP_016822091.1 (*Paenibacillus polymyxa*).

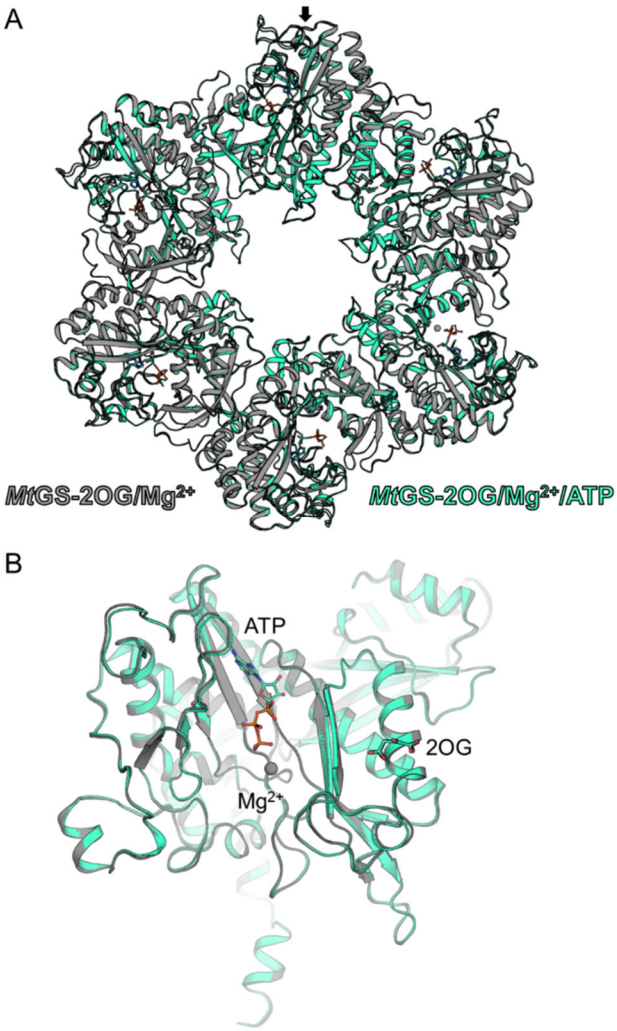


Figure S12: Differences between 2OG/Mg^{2+}-containing and the 2OG/Mg^{2+}/ATP-containing *Mt*GS. (**A**) Top view of the hexamer aligned on the C-terminal domain. An arrow indicates the subunit used for superposition. (**B**) Side view of the monomer with the same pose as in Fig. 2c. Models are represented as cartoons with Mg^{2+}, 2OG, and ATP as balls and sticks. Carbon, oxygen, nitrogen, and phosphorus are colored in cyan, red, blue, and orange, respectively.

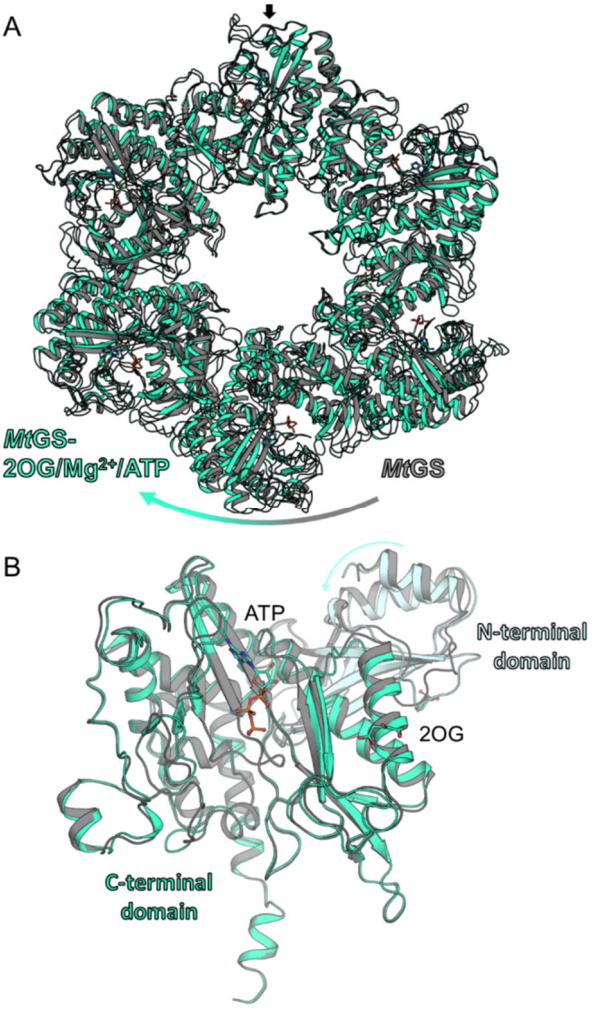


Figure S13: Differences between apo and the 2OG/Mg^{2+}/ATP-containing *Mt*GS. (**A**) Top view of the superposed *Mt*GS apo and *Mt*GS with 2OG/Mg^{2+}/ATP hexamers. An arrow indicates the subunit used for superposition. (**B**) Side view of the monomer, aligned on the C-terminal domain with the same pose as in Fig. 2c. Models are represented as cartoons with 2OG, Mg^{2+}, and ATP as balls and sticks. Carbon, oxygen, nitrogen, and phosphorus are colored in cyan, red, blue, and orange, respectively. Structural rearrangements are indicated with arrows.

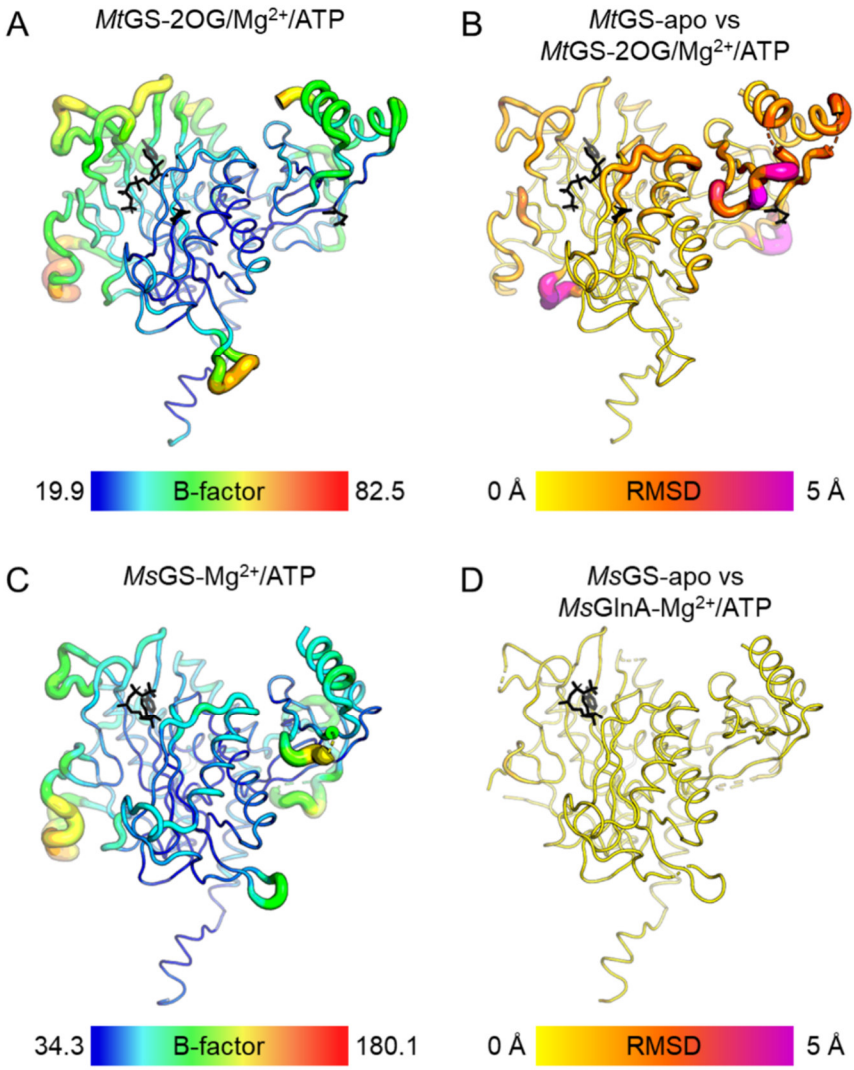


Figure S14: Structural variation between *Mt*GS/*Ms*GS apo and ligand-bound state. (**A, C**) Cartoon putty of Cα B-factors of the chain with the lowest average B-factor. (**B, D**) rmsd values with the average of all chains is displayed. *Ms*GS-Mg^{2+}/ATP and *Mt*GS-2OG/Mg^{2+}/ATP *versus* their respective apo state. Ligands are shown as black sticks.

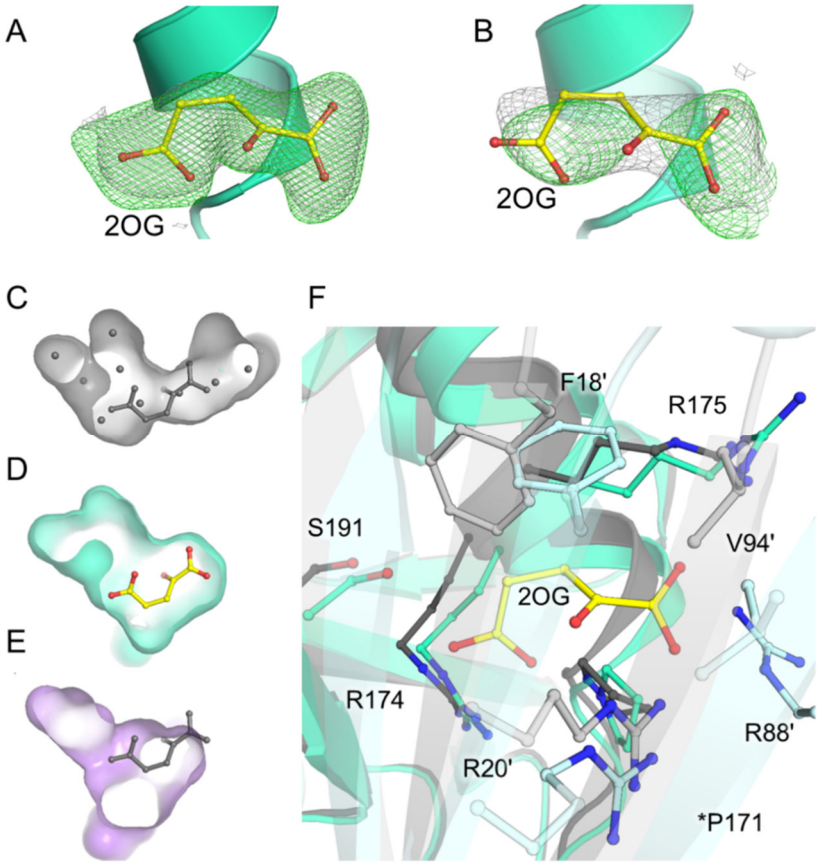


Figure S15: 2OG binding site in *Mt*GS and *Ms*GS. (**A and B**), Omit maps ($2F_o$-F_c map contoured at 1 sigma as grey mesh and F_o-F_c map contoured at 3 sigma as green mesh) of 2OG for *Mt*GS-2OG/Mg^{2+}/ATP (**A**) and *Mt*GS-2OG/Mg^{2+} (**B**). 2OG are modelled as balls and sticks. (**C-E**) 2OG pocket surface in *Mt*GS apo (grey), with waters as grey spheres (**C**), *Mt*GS-2OG/Mg^{2+}/ATP (cyan) (**D**), and the equivalent position in *Ms*GS-Mg^{2+}/ATP (purple) (**E**). 2OG was superposed on the C-terminal domain from *Mt*GS-2OG/Mg^{2+}/ATP (grey balls and sticks) for panels C and E. (**F**) Overlay of *Mt*GS apo (grey) and *Mt*GS-2OG/Mg^{2+}/ATP (cyan). Residues in the close vicinity of 2OG and 2OG itself are shown as balls and sticks. The adjacent subunit is colored in a lighter shade. A, B, and F are represented in cartoons. Oxygen and nitrogen are colored red and blue, respectively. Carbons are colored by chains, and 2OG carbons are yellow. Pro171 is located behind Arg20′.

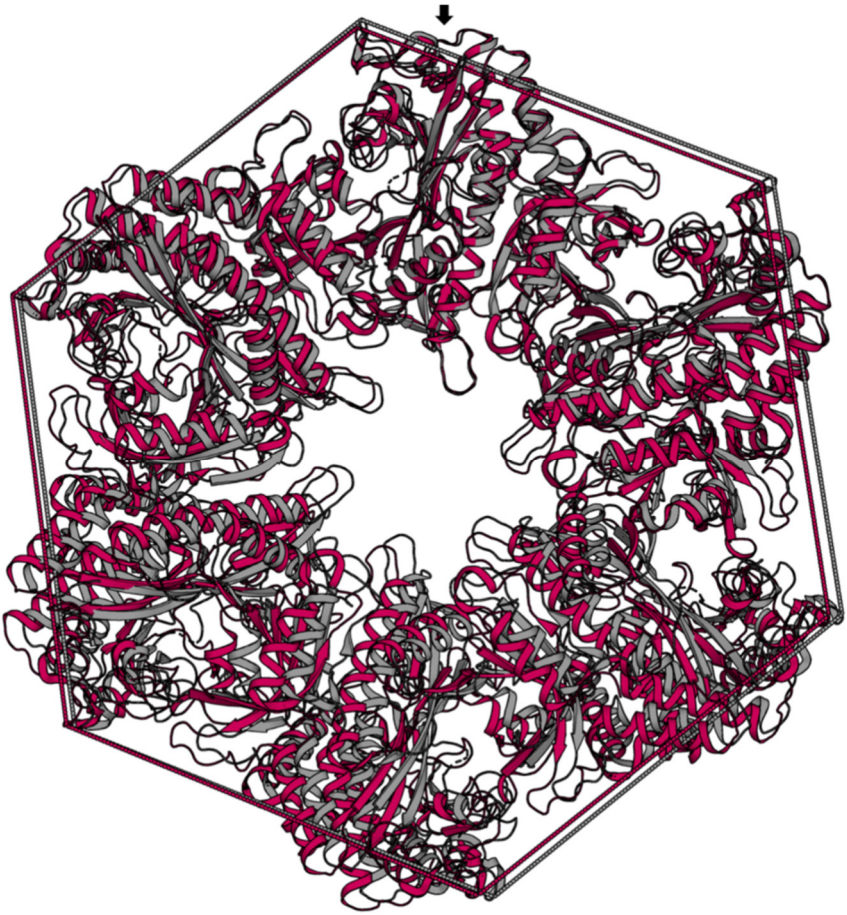


Figure S16: Conformational change in *Bs*GS. Superposition of the *Bs*GS apo (grey, PDB 4LNN) and *Bs*GS containing L-methionine-S-sulfoximine phosphate (SOX, which mimics the reaction intermediate, colored red, PDB 4LNI). Hexamers are represented as cartoons. The superposition was done on one monomer (indicated by an arrow), and a dashed line was drawn on the Cα position of Asn262 to illustrate the overall movements.

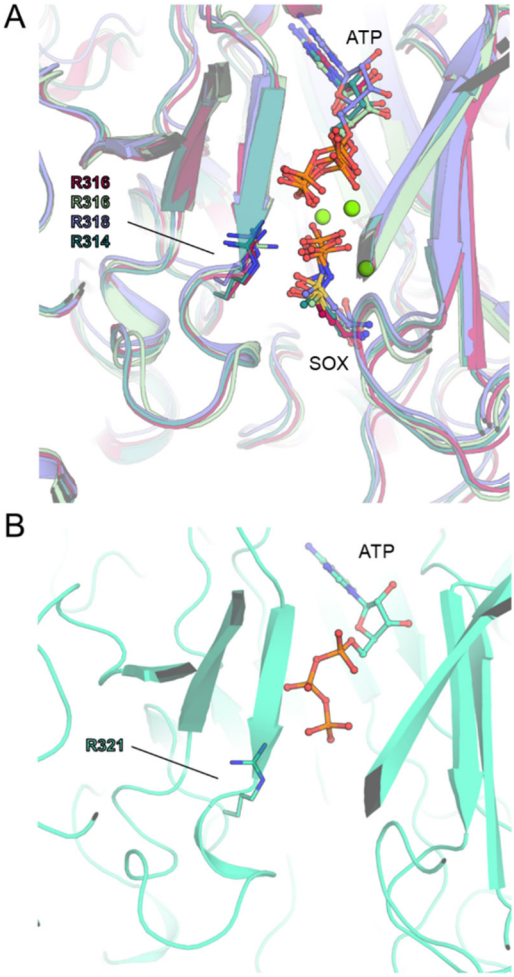


Figure S17: Glutamate binding site view in all structurally characterized bacterial GSI-α and *Mt*GS. (**A**) *Bs*GS (red, 4LNI), *Staphylococcus aureus* (blue, 7TDV), *Paenibacillus polymyxa* (dark green, 7TDP), *Listeria monocytogenes* (light green, 7TEN), all in the transition state with SOX/Mg^{2+}/ADP bound. (**B**) *Mt*GS-2OG/Mg^{2+}/ATP. All models are in cartoons with ligands in balls and sticks. Oxygen, nitrogen, sulfur, phosphorus, and magnesium are colored red, blue, yellow, orange and green, respectively. Carbons are accordingly colored to the model. The catalytic arginine is represented as sticks.

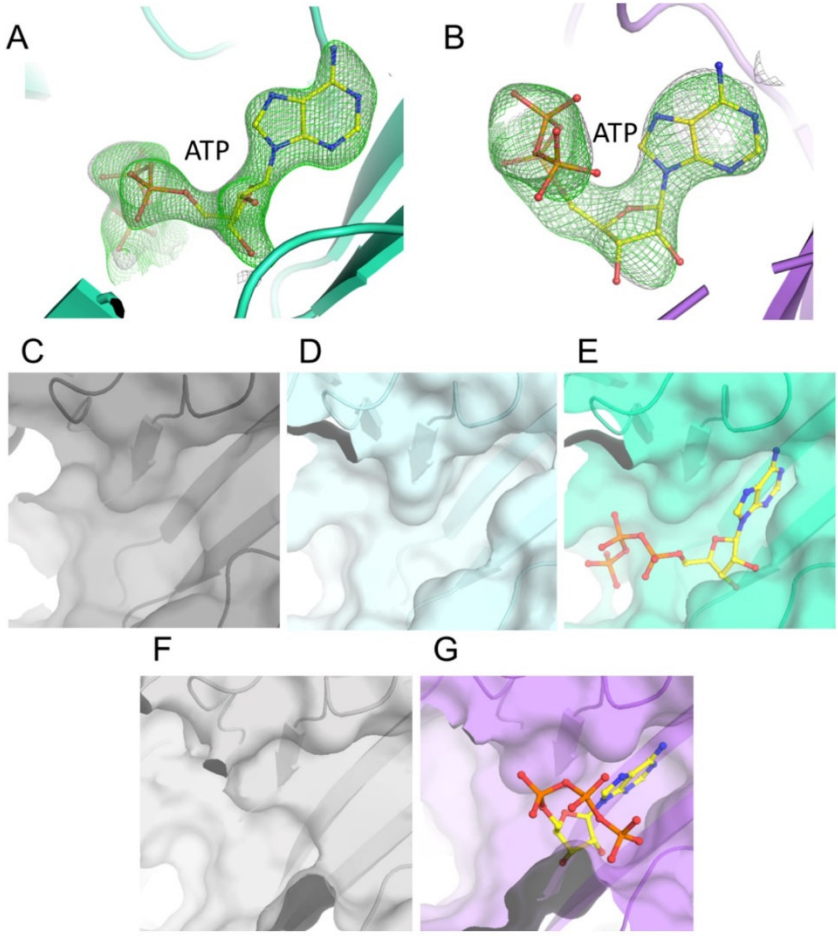


Figure S18: ATP binding pocket in *Mt*GS and *Ms*GS. (**A-B**) Omit maps ($2F_o$-F_c contoured at 1 sigma, colored in grey mesh and F_o-F_c map contoured at 3 sigma as green mesh) of ATP for *Mt*GS-2OG/Mg^{2+}/ATP (**A**) and *Ms*GS-Mg^{2+}/ATP (**B**). (**C-E**) ATP binding pockets are displayed as transparent surfaces in *Mt*GS apo (**C**), *Mt*GS-2OG/Mg^{2+} (**D**), and *Mt*GS-2OG/Mg^{2+}/ATP (**E**). (**F-G**) ATP binding pocket in *Ms*GS apo (**F**) and *Ms*GS-Mg^{2+}/ATP (**G**). ATP is highlighted as balls and sticks. Carbon, oxygen, nitrogen, and phosphorus are colored yellow, red, blue, and orange, respectively. All models are represented as cartoons.

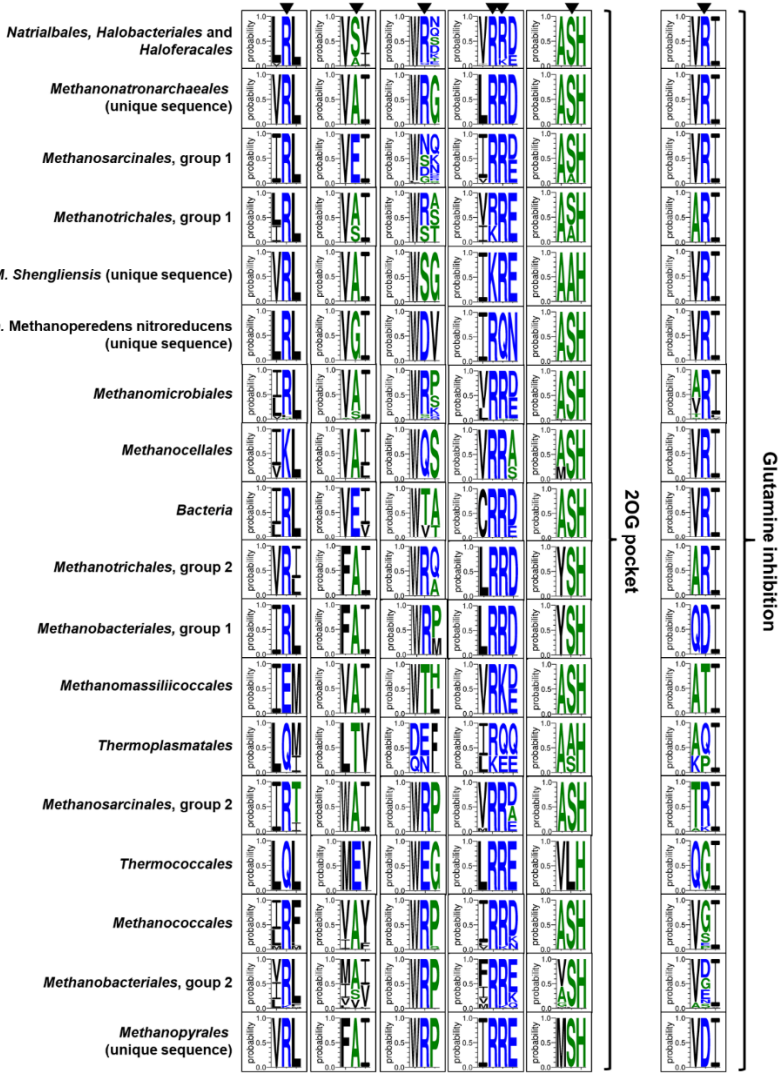


Figure S19: Conservation of residues binding 2OG and glutamine in GSI-α. The sequence conservation at the residues allowing or preventing binding of 2OG and glutamine (according to *Mt*GS, *Ms*GS, and *Bs*GS and shown by black triangles) are illustrated by the Weblogo3 figures constructed on the sequences forming the GS monophyletic groups selected on the tree from Fig. 6.

Supplementary data

Supplementary Movie: Movie illustrating the impact of 2OG binding on *Mt*GS. Accessible under https://www.nature.com/articles/s42003-023-05726-w#Sec22

Supplementary Data 1: Source data for the graphs in Fig. 1 and Fig. 5. Vcalc was calculated by Michaelis-Menten-fit utilizing the substrate concentration, V_{max}, $K_{0.5}$ for 2OG, and Hill coefficient to plot the grey curve in the graph.

mM	MtGS relative activity	Vcalc in %
0	0.03025	0
0	-0.4235	
0	-0.1513	
0.01	-0.8773	0.00625
0.01	-0.5143	
0.01	-0.7865	
0.04	1.17982	0.72336
0.04	0.45378	
0.04	0.8168	
0.05	-0.4235	1.54331
0.05	-0.4235	
0.05	-0.7865	
0.06	1.11932	2.84774
0.06	1.30083	
0.06	2.02687	
0.08	3.4487	7.29587
0.08	2.99493	
0.08	4.99154	
0.1	16.6385	14.4795
0.1	14.5511	
0.1	17.8183	
0.12	28.0736	24.0472
0.12	27.4384	
0.12	27.9829	
0.14	30.1005	34.9582
0.14	33.1862	
0.14	36.4534	

mM	MtGS relative activity	Vcalc in %
0.16	41.9592	45.947769
0.16	43.5928	
0.16	48.4936	
0.18	53.9692	56.018626
0.18	56.5103	
0.18	57.1456	
0.2	63.922	64.648717
0.2	66.0094	
0.2	66.9169	
0.22	67.5825	71.72467
0.22	76.295	
0.22	76.1135	
0.25	76.658	79.733373
0.25	80.6512	
0.25	77.384	
0.5	84.5537	97.700852
0.5	93.0847	
0.5	105.155	
1	104.762	99.782604
1	103.038	
1	101.585	
2	95.7469	99.979833
2	99.1956	
2	107.273	
5	95.4141	99.999132
5	102.221	
5	100.043	

mM	MtGS relative activity	rel stddev
0	-0.1815	0.2284
0.01	-0.726	0.18892
0.04	0.8168	0.36302
0.05	-0.5445	0.20959
0.06	1.48234	0.48023
0.08	3.81172	1.04664
0.1	16.336	1.65447
0.1	13.583	1.17982
0.12	27.8316	0.34359
0.14	33.2467	3.17687
0.16	44.6819	3.4006
0.18	55.875	1.68081
0.2	65.6161	1.53571
0.22	73.3303	4.9786
0.25	78.2311	2.12711
0.5	94.2645	10.3513
1	103.128	1.59016
2	100.738	5.91583
5	99.2258	3.47606

Parameter	Best-fit value	Standard deviation
K0.5	0.16775	0.00212
Vmax	3.01279	0.03229
Vmax %	100	1.0717
Hill Coeff.	3.43301	0.13381

Supplementary Data 2: The accession numbers of the sequences used for construction of Figure 6. Accessible under https://www.nature.com/articles/s42003-023-05726-w#Sec22

References

1. Cabello, P., Roldán, M. D., Moreno-Vivián, C. Nitrate reduction and the nitrogen cycle in archaea. *Microbiology* **150**, 3527-3546, doi:10.1099/mic.0.27303-0 (2004).
2. Kuypers, M. M. M., Marchant, H. K., Kartal, B. The microbial nitrogen-cycling network. *Nat. Rev. Microbiol.* **16**, 263-276, doi:10.1038/nrmicro.2018.9 (2018).
3. Zhang, X., Ward, B. B., Sigman, D. M. Global nitrogen cycle: critical enzymes, organisms, and processes for nitrogen budgets and dynamics. *Chem. Rev.* **120**, 5308-5351, doi:10.1021/acs.chemrev.9b00613 (2020).
4. Martínez-Espinosa, R. M. Microorganisms and their metabolic capabilities in the context of the biogeochemical nitrogen cycle at extreme environments. *Int. J. Mol. Sci.* **21**, doi:10.3390/ijms21124228 (2020).
5. Arcondéguy, T., Jack, R., Merrick, M. P_{II} signal transduction proteins, pivotal players in microbial nitrogen control. *Microbiol. Mol. Biol. Rev.* **65**, 80-105, doi:10.1128/MMBR.65.1.80-105.2001 (2001).
6. Eisenberg, D., Gill, H. S., Pfluegl, G. M., Rotstein, S. H. Structure-function relationships of glutamine synthetases. *Biochim. Biophys. Acta* **1477**, 122-45, doi:10.1016/s0167-4838(99)00270-8 (2000).
7. van den Heuvel, R. H. H., Curti, B., Vanoni, M. A., Mattevi, A. Glutamate synthase: A fascinating pathway from L-glutamine to L-glutamate. *Cell. Mol. Life Sci.* **61**, 669-681, doi:10.1007/s00018-003-3316-0 (2004).
8. Leigh, J. A., Dodsworth, J. A. Nitrogen regulation in bacteria and archaea. *Annu. Rev. Microbiol.* **61**, 349-77, doi:10.1146/annurev.micro.61.080706.093409 (2007).
9. de Carvalho Fernandes, G., Turchetto-Zolet, A. C., Pereira Passaglia, L. M. Glutamine synthetase evolutionary history revisited: Tracing back beyond the Last Universal Common Ancestor. *Evolution* **76**, 605-622, doi:10.1111/evo.14434 (2022).
10. Almassy, R. J., Janson, C. A., Hamlin, R., Xuong, N. H., Eisenberg, D. Novel subunit-subunit interactions in the structure of glutamine-synthetase. *Nature* **323**, 304-309, doi:10.1038/323304a0 (1986).
11. Unno, H., Uchida, T., Sugawara, H., Kurisu, G., Sugiyama, T., Yamaya, T., Sakakibara, H., Hase, T., Kusunoki, M. Atomic structure of plant glutamine synthetase: A key enzyme for plant productivity. *J. Biol. Chem.* **281**, 29287-96, doi:10.1074/jbc.M601497200 (2006).
12. He, Y. X., Gui, L., Liu, Y. Z., Du, Y., Zhou, Y. Y., Li, P., Zhou, C. Z. Crystal structure of *Saccharomyces cerevisiae* glutamine synthetase Gln1 suggests a nanotube-like supramolecular assembly. *Protein: Struct. Funct. Bioinf.* **76**, 249-254, doi:10.1002/prot.22403 (2009).
13. van Rooyen, J. M., Abratt, V. R., Belrhali, H., Sewell, T. Crystal structure of type III glutamine synthetase: Surprising reversal of the inter-ring interface. *Structure* **19**, 471-483, doi:10.1016/j.str.2011.02.001 (2011).
14. Pesole, G., Gissi, C., Lanave, C., Saccone, C. Glutamine synthetase gene evolution in bacteria. *Mol. Biol. Evol.* **12**, 189-97, doi:10.1093/oxfordjournals.molbev.a040197 (1995).
15. Brown, J. R., Masuchi, Y., Robb, F. T., Doolittle, W. F. Evolutionary relationships of bacterial and archaeal glutamine synthetase genes. *J. Mol. Evol.* **38**, 566-76, doi:10.1007/BF00175876 (1994).
16. Kumada, Y., Benson, D. R., Hillemann, D., Hosted, T. J., Rochefort, D. A., Thompson, C. J., Wohlleben, W., Tateno, Y. Evolution of the glutamine-synthetase gene, one of the oldest existing and functioning genes. *P. Natl. Acad. Sci. U. S. A.* **90**, 3009-3013, doi:10.1073/pnas.90.7.3009 (1993).
17. Pesole, G., Bozzetti, M. P., Lanave, C., Preparata, G., Saccone, C. Glutamine synthetase gene evolution: A good molecular clock. *P. Natl. Acad. Sci. U. S. A.* **88**, 522-526, doi:10.1073/pnas.88.2.522 (1991).
18. Murray, D. S., Chinnam, N., Tonthat, N. K., Whitfill, T., Wray, L. V., Jr., Fisher, S. H., Schumacher, M. A. Structures of the *Bacillus subtilis* glutamine synthetase dodecamer reveal large intersubunit catalytic conformational changes linked to a unique feedback inhibition mechanism. *J. Biol. Chem.* **288**, 35801-11, doi:10.1074/jbc.M113.519496 (2013).
19. Travis, B. A., Peck, J. V., Salinas, R., Dopkins, B., Lent, N., Nguyen, V. D., Borgnia, M. J., Brennan, R. G., Schumacher, M. A. Molecular dissection of the glutamine synthetase-GlnR nitrogen regulatory circuitry in Gram-positive bacteria. *Nat. Commun.* **13**, 3793, doi:10.1038/s41467-022-31573-0 (2022).

20. Liaw, S. H., Kuo, I. C., Eisenberg, D. Discovery of the ammonium substrate site on glutamine synthetase, a third cation binding site. *Prot. Sci.* **4**, 2358-2365, doi:10.1002/pro.5560041114 (1995).
21. Martínez-Espinosa, R. M., Esclapez, J., Bautista, V., Bonete, M. J. An octameric prokaryotic glutamine synthetase from the haloarchaeon *Haloferax mediterranei*. *FEMS Microbiol. Lett.* **264**, 110-6, doi:10.1111/j.1574-6968.2006.00434.x (2006).
22. Adul Rahman, R. N., Jongsareejit, B., Fujiwara, S., Imanaka, T. Characterization of recombinant glutamine synthetase from the hyperthermophilic archaeon *Pyrococcus* sp. strain KOD1. *Appl. Environ. Microbiol.* **63**, 2472-6, doi:10.1128/aem.63.6.2472-2476.1997 (1997).
23. Neelon, K., Schreier, H. J., Meekins, H., Robinson, P. M., Roberts, M. F. Compatible solute effects on thermostability of glutamine synthetase and aspartate transcarbamoylase from *Methanococcus jannaschii*. *Biochim. Biophys. Acta - Proteins Proteom.* **1753**, 164-173, doi:10.1016/j.bbapap.2005.08.009 (2005).
24. Robinson, P., Neelon, K., Schreier, H. J., Roberts, M. F. β-Glutamate as a substrate for glutamine synthetase. *Appl. Environ. Microbiol.* **67**, 4458-63, doi:10.1128/AEM.67.10.4458-4463.2001 (2001).
25. Yin, Z. M., Purschke, W. G., Schäfer, G., Schmidt, C. L. The glutamine synthetase from the hyperthermoacidophilic crenarcheon *Sulfolobus acidocaldarius*: isolation, characterization and sequencing of the gene. *Biol. Chem.* **379**, 1349-1354, doi:10.1515/bchm.1998.379.11.1349 (1998).
26. Pedro-Roig, L., Camacho, M., Bonete, M. J. Regulation of ammonium assimilation in *Haloferax mediterranei*: Interaction between glutamine synthetase and two GlnK proteins. *Biochim Biophys Acta* **1834**, 16-23, doi:10.1016/j.bbapap.2012.10.006 (2013).
27. Gutt, M., Jordan, B., Weidenbach, K., Gudzuhn, M., Kiessling, C., Cassidy, L., Helbig, A., Tholey, A., Pyper, D. J., Kubatova, N., et al. High complexity of glutamine synthetase regulation in *Methanosarcina mazei*: Small protein 26 interacts and enhances glutamine synthetase activity. *FEBS J.* **288**, 5350-5373, doi:10.1111/febs.15799 (2021).
28. Ehlers, C., Weidenbach, K., Veit, K., Forchhammer, K., Schmitz, R. A. Unique mechanistic features of post-translational regulation of glutamine synthetase activity in *Methanosarcina mazei* strain Gö1 in response to nitrogen availability. *Mol Microbiol* **55**, 1841-54, doi:10.1111/j.1365-2958.2005.04511.x (2005).
29. Ehlers, C., Grabbe, R., Veit, K., Schmitz, R. A. Characterization of $GlnK_1$ from *Methanosarcina mazei* strain Gö1: Complementation of an *Escherichia coli glnK* mutant strain by $GlnK_1$. *J. Bacteriol.* **184**, 2331-2331, doi:10.1128/jb.184.4.1028-1040.2002 (2002).
30. Huergo, L. F., Dixon, R. The emergence of 2-oxoglutarate as a master regulator metabolite. *Microbiol. Mol. Biol. Rev.* **79**, 419-35, doi:10.1128/MMBR.00038-15 (2015).
31. Engilberge, S., Wagner, T., Santoni, G., Breyton, C., Shima, S., Franzetti, B., Riobé, F., Maury, O., Girard, E. Protein crystal structure determination with the crystallophore, a nucleating and phasing agent. *J. Appl. Crystallogr.* **52**, 722-731, doi:10.1107/S1600576719006381 (2019).
32. Cheng, L., Qiu, T. L., Yin, X. B., Wu, X. L., Hu, G. Q., Deng, Y., Zhang, H. *Methermicoccus shengliensis* gen. nov., sp. nov., a thermophilic, methylotrophic methanogen isolated from oil-production water, and proposal of *Methermicoccaceae* fam. nov. *Int. J. Syst. Evol. Microbiol.* **57**, 2964-2969, doi:10.1099/ijs.0.65049-0 (2007).
33. Liaw, S. H., Eisenberg, D. Structural model for the reaction mechanism of glutamine synthetase, based on five crystal structures of enzyme-substrate complexes. *Biochemistry* **33**, 675-81, doi:10.1021/bi00169a007 (1994).
34. Krajewski, W. W., Collins, R., Holmberg-Schiavone, L., Jones, T. A., Karlberg, T., Mowbray, S. L. Crystal structures of mammalian glutamine synthetases illustrate substrate-induced conformational changes and provide opportunities for drug and herbicide design. *J. Mol. Biol.* **375**, 217-228, doi:10.1016/j.jmb.2007.10.029 (2008).
35. Forchhammer, K., Selim, K. A., Huergo, L. F. New views on PII signaling: From nitrogen sensing to global metabolic control. *Trends Microbiol.* **30**, 722-735, doi:10.1016/j.tim.2021.12.014 (2022).
36. L. F. Huergo, G. Chandra, M. Merrick. P_{II} signal transduction proteins: nitrogen regulation and beyond. *Fems Microbiology Reviews* **37**, 251-283, doi:10.1111/j.1574-6976.2012.00351.x (2013).

37. Müller, M.-C., Wagner, T. X The oxoglutarate binding site and regulatory mechanism are conserved in ammonium transporter inhibitors GlnKs from *Methanococcales*. *Int. J. Mol. Sci.* **22**, doi:10.3390/ijms22168631 (2021).
38. Dodsworth, J. A., Cady, N. C., Leigh, J. A. 2-Oxoglutarate and the PII homologues $NifI_1$ and $NifI_2$ regulate nitrogenase activity in cell extracts of *Methanococcus maripaludis*. *Mol. Microbiol.* **56**, 1527-38, doi:10.1111/j.1365-2958.2005.04621.x (2005).
39. Maslać, N., Sidhu, C., Teeling, H., Wagner, T. Comparative transcriptomics sheds light on remodeling of gene expression during diazotrophy in the thermophilic methanogen *Methanothermococcus thermolithotrophicus*. *mBio* **13**, doi:10.1128/mbio.02443-22 (2022).
40. Prasse, D., Schmitz, R. A. Small RNAs involved in regulation of nitrogen metabolism. *Microbiol. Spectr.* **6**, doi:10.1128/microbiolspec.RWR-0018-2018 (2018).
41. Wedler, F. C., Shreve, D. S., Kenney, R. M., Ashour, A. E., Carfi, J., Rhee, S. G. Two glutamine synthetases from *Bacillus caldolyticus*, an extreme thermophile. Isolation, physicochemical and kinetic properties. *J. Biol. Chem.* **255**, 9507-16, doi:10.1016/S0021-9258(19)70592-4 (1980).
42. Wang, Q., Min, C., Yan, T. T., Pu, H. F., Xin, Y. Q., Zhang, S. Q., Luo, L., Yin, Z. M. Production of glutamine synthetase in *Escherichia coli* using SUMO fusion partner and application to L-glutamine synthesis. *World J. Microbiol. Biotechnol.* **27**, 2603-2610, doi:10.1007/s11274-011-0733-3 (2011).
43. Jespersen, M., Pierik, A. J., Wagner, T. Structures of the sulfite detoxifying F_{420}-dependent enzyme from *Methanococcales*. *Nat. Chem. Biol.*, doi:10.1038/s41589-022-01232-y (2023).
44. Jespersen, M., Wagner, T. How a methanogen assimilates sulfate: Structural and functional elucidation of the complete sulfate-reduction pathway. *bioRxiv*, doi:10.1101/2022.10.18.512691 (2022).
45. Kurth, J. M., Müller, M.-C., Welte, C. U., Wagner, T. Structural insights into the methane-generating enzyme from a methoxydotrophic methanogen reveal a restrained gallery of post-translational modifications. *Microorganisms* **9**, doi:10.3390/microorganisms9040837 (2021).
46. Lemaire, O. N., Müller, M. C., Kahnt, J., Wagner, T. Structural Rearrangements of a Dodecameric Ketol-Acid Reductoisomerase Isolated from a Marine Thermophilic Methanogen. *Biomolecules* **11**, doi:10.3390/biom11111679 (2021).
47. Lemaire, O. N., Infossi, P., Ali Chaouche, A., Espinosa, L., Leimkuhler, S., Giudici-Orticoni, M. T., Méjean, V., Iobbi-Nivol, C. Small membranous proteins of the TorE/NapE family, crutches for cognate respiratory systems in Proteobacteria. *Sci. Rep.* **8**, 13576, doi:10.1038/s41598-018-31851-2 (2018).
48. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., Bricogne, G. Data processing and analysis with the *autoPROC* toolbox. *Acta Crystallogr. D Struct. Biol.* **67**, 293-302, doi:10.1107/S0907444911007773 (2011).
49. Tickle, I. J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., Bricogne, G. STARANISO. *Global Phasing Ltd., Cambridge, United Kingdom* (2020).
50. Kabsch, W. *XDS*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).
51. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., et al. Overview of the *CCP4* suite and current developments. *Acta Crystallogr. D Struc. Biol.* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).
52. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. *Acta Crystallogr. D Struct. Biol.* **75**, 861-877, doi:10.1107/S2059798319011471 (2019).
53. Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C. Buster Version 2.10.4. *Cambr. U. K. Glob. Phasing Ltd.* (2017).
54. Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
55. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., Richardson, D. C. *MolProbity*: All-atom structure validation for

macromolecular crystallography. *Acta Crystallogr. D Struct. Biol.* **66**, 12-21, doi:10.1107/S0907444909042073 (2010).
56. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W. Z., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Sys. Biol.* **7**, doi:10.1038/msb.2011.75 (2011).
57. Robert, X., Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-W324, doi:10.1093/nar/gku316 (2014).
58. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., Lipman, D. J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402, doi:10.1093/nar/25.17.3389 (1997).
59. Kumar, S., Stecher, G., Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870-1874, doi:10.1093/molbev/msw054 (2016).
60. Jones, D. T., Taylor, W. R., Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275-282, doi:10.1093/bioinformatics/8.3.275 (1992).
61. Schoch, C. L., Ciufo, S., Domrachev, M., Hotton, C. L., Kannan, S., Khovanskaya, R., Leipe, D., Mcveigh, R., O'Neill, K., Robbertse, B., et al. NCBI Taxonomy: A comprehensive update on curation, resources and tools. *Database: J. Biol. Databases Curation*, doi:10.1093/database/baaa062 (2020).
62. Letunic, I., Bork, P. Interactive tree of life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127-128, doi:10.1093/bioinformatics/btl529 (2007).
63. Crooks, G. E., Hon, G., Chandonia, J. M., Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188-1190, doi:10.1101/gr.849004 (2004).

CHAPTER IV.

Structural insights into the methane-generating enzyme from a methoxydotrophic methanogen reveal a restrained gallery of post-translational modifications

Julia Maria Kurth[1*], Marie-Caroline Müller[2*], Cornelia Ulrike Welte[1] and Tristan Wagner[2#]

[1] Department of Microbiology, Institute for Water and Wetland Research, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

[2] Microbial Metabolism Research Group, Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen

[#]Correspondence: twagner@mpi-bremen.de

*These authors contributed equally to this work.

Published in *Microorganisms*

Kurth, J.M.; Müller, M.-C.; Welte, C.U.; Wagner, T. Structural Insights into the Methane-Generating Enzyme from a Methoxydotrophic Methanogen Reveal a Restrained Gallery of Post-Translational Modifications. *Microorganisms* **2021**, *9*, 837. https://doi.org/10.3390/ microorganisms9040837 (Open Access, CC BY 4.0)

The text and figures have been reformatted to match the thesis style.

Received February 2021, accepted April 2021

Published April 2021

Abstract

Methanogenic archaea operate an ancient, if not primordial, metabolic pathway that releases methane as an end-product. This last step is orchestrated by the methyl-coenzyme M reductase (MCR), which uses a nickel-containing F_{430}-cofactor as the catalyst. MCR astounds the scientific world by its unique reaction chemistry, its numerous post-translational modifications, and its importance in biotechnology not only for production but also for capturing the greenhouse gas methane. In this report, we investigated MCR natively isolated from *Methermicoccus shengliensis*. This methanogen was isolated from a high-temperature oil reservoir and has recently been shown to convert lignin and coal derivatives into methane through a process called methoxydotrophic methanogenesis. A methoxydotrophic culture was obtained by growing *M. shengliensis* with 3,4,5- trimethoxybenzoate as the main carbon and energy source. Under these conditions, MCR represents more than 12 % of the total protein content. The native MCR structure refined at a resolution of 1.6-Å precisely depicts the organization of a dimer of heterotrimers. Despite subtle surface remodeling and complete conservation of its active site with other homologs, MCR from the thermophile *M. shengliensis* contains the most limited number of post-translational modifications reported so far, questioning their physiological relevance in other relatives.

Keywords: methyl-coenzyme M reductase; post-translational modifications; methoxydotrophic methanogenesis; X-ray crystallography; F_{430}-cofactor; thermophilic archaeon

Introduction

Methanogenesis is a primitive energy metabolic pathway found only in the archaeal domain that evolved more than 3.46 Gyr ago[1, 2]. During evolution, different types of methanogenesis arose, all of them sharing the common trait of releasing methane. Hydrogenotrophic methanogenesis reduces CO_2 by using H_2 or alternatively formate; acetoclastic methanogenesis disproportionates acetate in CO_2 and CH_4 and finally, methylotrophic methanogenesis uses methylated molecules such as methanol, methylamine(s), or methylsulfides[3, 4]. In 2016, Mayumi and co-workers discovered a new methylotrophic pathway named methoxydotrophic methanogenesis in which the substrates are methoxylated aromatic compounds derived from lignin, oil, and coal[5]. The organism exhibiting this novel pathway is *Methermicoccus shengliensis*, a thermophilic archaeon that has been isolated from oil production water (75–80 °C)[6]. Methoxydotrophic methanogens such as *M. shengliensis* might play an important role in the carbon cycle of coal- and lignin-rich subsurface sediments as well as of oil reservoirs. These specialized methoxydotrophic methanogens are able to metabolize methoxy compounds intracellularly and transfer the methyl group on a carrier by a so far unknown mechanism (Figure 1A). As for all other methanogenic pathways, the methyl group must be transferred onto coenzyme M (HS-CoM) to be released as methane[4] by the Methyl-coenzyme M reductase (MCR).

MCR is a three-subunit complex harboring the cofactor F_{430}, a nickel-containing corrinoid that gives its yellow color to the enzyme[7]. The chemical reaction catalyzed by the cofactor is a thiyl-radical mechanism in which the Ni(I)-active state will attack the thiol group of the methyl-S-CoM forcing the generation of methane and the formation of the heterodisulfide made of the HS-CoM and Coenzyme B (CoB-SH)[8]. Structural studies[9-17] revealed how MCR precisely coordinates the cofactor and coenzymes, and they also depicted a gallery of post-translational modifications that vary depending on the species[18]. Recent studies based on genetic manipulation of *Methanosarcina*

acetivorans confirmed that thioglycine, S-methylcysteine and 5(S)-methylarginine are not required for catalysis[17, 19, 20] while the role of N^1-methylhistidine is yet unknown[21]. Nevertheless, combinatorial interactions between modified residues were shown to alter the thermal stability of MCR as well as the growth fitness on different carbon sources[17]. It is assumed that these modifications might have a tuning-up function to improve the enzyme's robustness under stress conditions and stimulate its turn-over[15, 16]. Unfortunately, because of the high instability of the active Ni(I) state, enzymatic studies which characterize the impact of the loss of each modification are still a challenging task.

The overall reaction of methane generation by MCR is highly exergonic with a $\Delta G^{0'} = -30$ kJ.mol^{-1} of methane formed[7]. Surprisingly, despite its thermodynamic difficulty, anaerobic methane oxidizers are using the reverse reaction to capture methane[14, 22, 23]. The methane activation by MCR-homologs highlights biotechnological potentials to mitigate the concentration of atmospheric methane[24], of which 50 % worldwide is released by methanogens[1, 25]. By domesticating the enzyme, it would be possible to trap and transform methane[26] or alternatively to block the methane release by inhibiting the enzyme[27]. Characterization of MCR from various methanogens is yielding an overview of the enzyme's variability and provides templates for targeted mutagenesis. The MCR from *M. shengliensis* (abbreviated as *Ms*MCR) is offering a new variation of the enzyme that might contain typical adaptations for methoxydotrophic growth in a high-temperature ecological niche. The structural features of *Ms*MCR presented in this report highlight a conserved active site with the lowest post-translational modification content reported so far.

Results

Purification and crystallization of *Ms*MCR obtained under methoxydotrophic methanogenesis

The cell extracts of *M. shengliensis* grown with 3,4,5-trimethoxybenzoate or methanol as the main carbon and energy source, were first compared and showed a similar profile for the three subunits constituting MCR (Supplemental Figure S2A). McrA identification was confirmed by MALDI-TOF MS with a molecular weight search (MOWSE) score of 97 and an amino acid sequence coverage of 43 %. MCR was anaerobically purified to homogeneity by anionic exchange and hydrophobic interaction chromatography (Figure 1B and Supplemental Figure S2B-C), yielding 43.6 mg of purified protein (see Materials and Methods) that corresponds to 12 % of the total protein extract. It is generally assumed that MCR is catalyzing the rate-limiting step of methanogenesis and methanogens maintain their high-flux metabolism by expressing enormous amounts of the enzyme[7].

The purified MCR, containing an equal stoichiometry of the three subunits, has a characteristic yellow color coming from its F_{430}-cofactor. The UV/Visible spectra (Supplemental Figure S2D) is typical of the Ni(II) red1-silent state with an absorption peak at 424 nm[28]. O_2-incubation for one hour did not modify the spectra and therefore the sample was crystallized aerobically.

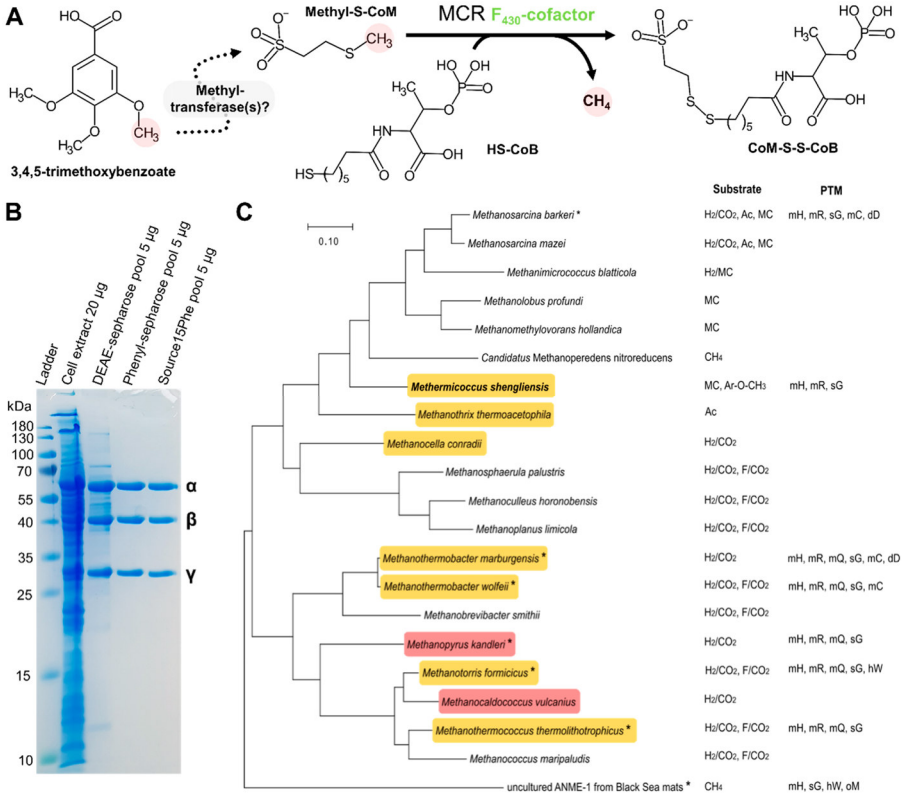


Figure 1. MCR metabolic function, purification, and phylogeny. (**A**) During methoxydotrophic growth, 3,4,5- trimethoxybenzoate (TMB) methyl-groups are transferred by an unknown mechanism to the central carbon metabolism of the methanogen. The methyl-group will be transferred onto HS-CoM and MCR will branch methyl-S-CoM to CoB-SH by a thiyl-radical based reaction catalyzed by its F_{430} cofactor. The end products of the reaction are methane and the heterodisulfide made of HS-CoM and CoB-SH. (**B**) Purification profile on SDS-PAGE of MCR α, β and γ subunits from TMB-grown cells. (**C**) Phylogenetic tree of concatenated MCR generated with MegaX using the Maximum Likelihood method and JTT matrix-based model (see Materials and Methods). Orange and red backgrounds indicate thermophiles and hyperthermophiles, respectively. Structural information exists for the species with asterisks. Post-translational modifications (PTM) observed in the structures are shown: mH, N1-methylhistidine; mR, 5(S)-methylarginine; mQ, 2(S)-methylglutamine; sG, thioglycine; mC, S-methylcysteine; dD, didehydroaspartate; hW, 6-hydroxytryptophan

Fig. 1 cont. in *M. formicicus* and 7-hydroxytryptophan in ANME-1; oM, oxidized methionine. Growth substrates are also indicated (Ac, acetate; F, formate; MC, methylated compounds; Ar-O-CH3, methoxylated compounds).

A conserved overall structure and active site

X-ray diffraction measurements were performed on *Ms*MCR crystals and revealed a primitive orthorhombic crystalline form. The structure of MCR from *Methanosarcina barkeri* (*Mb*MCR) was used for molecular replacement based on a phylogenetic analysis (Figure 1C) and the *Ms*MCR structure was refined to 1.6-Å resolution. The asymmetric unit contains two dimers with the typical $(\alpha\beta\gamma)_2$ organization (Figure 2A). Interestingly, while the first dimer shows an excellent fit in the electron density (average B-factor = 22.6 Å^2), the second dimer has a very high average B-factor (51.3 Å^2), which made its accurate modelling challenging. All following analyses were therefore performed on the first stable dimer.

The model was compared with three homologs: the terrestrial mesophile *M. barkeri* (optimal growth temperature of 35 °C, *Mb*MCR PDB code 1E6Y), the terrestrial thermophile *Methanothermobacter marburgensis* (optimal growth temperature of 65 °C, *Mm*MCR type I PDB code 5A0Y), and the marine thermophile *Methanothermococcus thermolithotrophicus* (optimal growth temperature of 65 °C, *Mt*MCR PDB code 5N1Q). It is worth noting that *M. marburgensis* and *M. thermolithotrophicus* are hydrogenotrophic methanogens growing at the same temperature as *M. shengliensis*. *Ms*MCR and its structural homologs aligned very well with a root mean square deviations below 1 Å for the three different chains (Supplemental Table S1). This is not surprising considering the high sequence identity between the four MCRs (Supplemental Table S1). The extended loop following the N-terminal helix of the α-subunit found in *Mb*MCR (residues 18–29) is also conserved in *Ms*MCR (residues 19–33) and might have a stabilizing role. Only one discrepancy was noticeable, the loop 53–66 of the β-subunit is shifted in one of the monomers

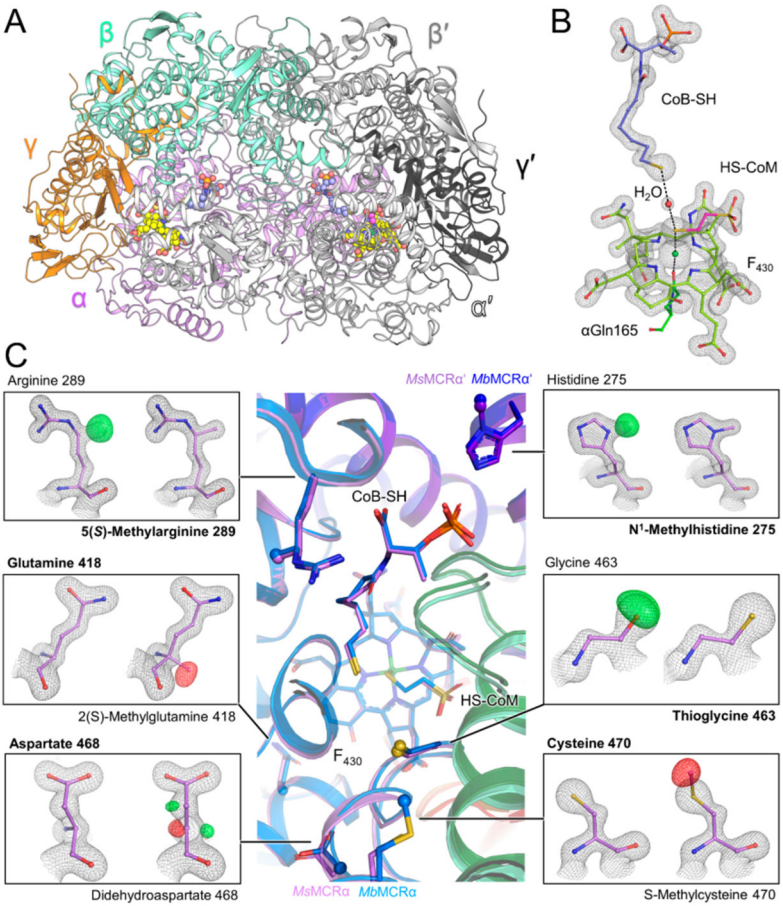


Figure 2. *Ms*MCR structure and its post-translational modifications. (**A**) *Ms*MCR $(\alpha\beta\gamma)_2$ organization with each chain colored differently. F_{430}, HS-CoM and CoB-SH are in balls and sticks and colored in yellow, pink, and light blue, respectively. (**B**) Close up of the active site. $2F_o$-F_c electron density map for the F_{430} and coenzymes is contoured at 2-σ. (**C**) Superposition of *Ms*MCR (same color code as panel A) on *Mb*MCR (α, β, γ in blue, dark green and red respectively). The ligands and modified residues are in balls and sticks with the modifications as spheres. Each panel presents the $2F_o$-F_c map contoured at 2-σ (black mesh) and the F_o-F_c map contoured at 4-σ (green, positive, and red, negative) after refinement for a classic (left) or modified (right) residue. Final modelled residue is highlighted in bold.

(Supplemental Figure S3A-B). Since the loop is on the surface, distant from the active site and involved in a crystallographic contact, this shift was most probably due to a packing artefact rather than a typical trait for this family. An inspection of the electrostatic charge profile on the proteins surface reflects the classic positively charged entrance of the CoB-SH channel. The electrostatic charge repartition of *Ms*MCR fits very well with the one from *Mb*MCR rather than the two other thermophiles (Supplemental Figure S3C–F) showing that thermophilic and high salt adaptations of *M. shengliensis* have not drastically modified the enzyme surface.

The active site is identical compared to the three other structural homologs with the same coordination of the coenzymes and F_{430}-cofactor (Figure 2B). Both coenzymes are bound at very high occupancy with a distance of 6.2 Å separating their sulfur groups as previously seen in other Ni(II) red1-silent structures[10, 11, 13, 15, 16], the density between the thiol groups was interpreted as a water molecule. The well-defined electron density for the F_{430} perfectly fits the classic cofactor observed in structural homologs (Figure 2B). The HS-CoM has an average B-value of 21.6 $Å^2$ that is 4.1 $Å^2$ higher than CoB-SH in the most defined *Ms*MCR dimer This could come from a higher vibration or slightly lower occupancy of the HS-CoM already seen in MCR structures[15].

The smallest post-translational modification gallery observed in methanogens

The high quality of the obtained electron density map confirmed the presence of three modified residues: N^1-methylhistidine275, 5(*S*)-methylarginine289, and the thioglycine463. The calculation of an omit map for the three modifications unambiguously confirmed this result (Figure 2C). Surprisingly, the S-methylcysteine and didehydroaspartate found in the close relatives *M. barkeri*[13, 18] and *Methanosarcina acetivorans*[17] are not detected in *M. shengliensis*. After forcing the modelling of a didehydroaspartate at position 468 and S-methylcysteine 470 in *M. shengliensis* the resulting $2F_o$-F_c and F_o-F_c maps post-refinement confirmed the absence of both

modifications (Figure 2C). With help of MALDI-TOF MS analysis we were able to detect the additional methylation in the peptide containing the His275 (1452.9 Da vs predicted mass of 1452.7 Da). We also observed the peptide containing the thioglycine463, however with a low signal/noise ratio of 1.9 (see Materials and Methods). The tryptic digestion and MALDI-TOF MS analysis could not detect the peptide containing the methylated Arg289 due to its small size. Asp468 and Cys470 in *Ms*MCR present a similar position and coordination compared to the modified versions in *Mb*MCR and *Mm*MCR and no mutation in the direct surrounding appeared to counterbalance the absence of modifications (Supplemental Figure S4A). Rather subtle readjustments take place, such as shorter hydrogen bond distances, which might ultimately affect the loops coordinating the coenzymes as previously hypothesized for *Mm*MCR and MCR from *Methanothermobacter wolfeii*[15, 16].

Methanobacteriales, Methanococcales, *Methanopyrus kandleri,* and *Methanoculleus thermophilus*[18] contain a 2(S)-methylglutamine close to the F_{430}. This modification is not present in *Mb*MCR and in *M. acetivorans* and the presented structural data shows a classic glutamine at this position in *Ms*MCR (Figure 2C). A water molecule "fills" the absence of the methylation that might indirectly stabilize the F_{430} position via the αTyr350 (Supplemental Figure S4B). The peptide containing a classic Gln418 in *Ms*MCR was also detected by mass spectrometry (see Materials and Methods). Surprisingly, these results reveal that despite the high similarity in sequence and structure, *Ms*MCR operates with a reduced gallery of post-translational modifications compared to *M. barkeri* or *M. acetivorans*.

Discussion

Anaerobic archaea have already been thriving on earth for billions of years and developed a variety of metabolic pathways to utilize a broad range of substrates. Methanogens living in deeper sediment layers managed to utilize coal, oil, and lignin derivatives as methyl donors, which provide methanogens with an abundant source of carbon and energy. This process could be an inspiration to transform methoxy-compounds to methane and use it as biofuel. *M. shengliensis* represents an excellent model organism to study this new pathway in depth, as its enzymes, involved in the methoxydotrophic metabolism are accessible for biochemical characterization. In this work, we isolated and structurally characterized the enzyme involved in the methane release, one of the last reactions of methoxydotrophy. Under methoxydotrophic growth conditions, *M. shengliensis* contains a tremendous amount of MCR as also found in methanol-grown cultures (Supplemental Figure S2A).

M. shengliensis contains the different required machineries to feed on a broad variety of substrates such as methanol, methylamine(s), and different methoxylated aromatic compounds[5, 6]. It was suggested that the post-translational modifications of the MCR from *M. acetivorans* might affect its growth robustness when grown on different substrates[17]. The systematic or complete deletion of the three genes involved in the arginine and cysteine methylation, as well as thioglycine formation, have indeed some impact on the growth when the methanogen uses different carbon sources[17]. Based on these results, it can be assumed that the acquisition of the methoxydotrophic pathway will favor the apparition of additional modifications or adaptive traits. However, instead of harboring new modified residues, *Ms*MCR shows a reduced set of modifications, which was unexpected. Only the core modifications methyl-histidine, methyl-arginine, and thioglycine are present.

The absence of S-methylcysteine in *Ms*MCR is explained by the fact that the gene coding for the methyl-transferase responsible of its installation (MA_RS23695 in *M. acetivorans*) is absent in the current genome of *M. shengliensis* strain DSM 18856 (Assembly number GCA_000711905.1). No conclusions can be drawn for the didehydroaspartate and methylglutamine since the biosynthetic machineries are still unknown. However, the genes coding for the enzyme responsible of the thioamidation (YcaO coding gene: BP07_RS07665 and TfuA coding gene: BP07_RS07670) of αGly463 and methylation of the αArg289 (coding gene: *mmp10*) are present in the *M. shengliensis* genome, which allow the installation of these modifications as observed in the *Ms*MCR structure. Despite their different ecological niches, *Ms*MCR and *Mb*MCR share a remarkably similar organization and electrostatic surface, illustrating the close relationship between the two organisms in accordance with the phylogenetic studies (Figure 1C). The active site is identical to already described MCRs. Such perfect conservation, even the one from anaerobic methane oxidizers, such as marine ANME-1 clade archaea[14], depicts how challenging the chemical mechanism of methane generation/capture is. As always, in the absence of an active structure of MCR Ni(I) state, it is difficult to derive final conclusions on the possible structural roles of modifications during the catalysis.

Although the active site of different MCRs is quite conserved, the overall amino acid sequence of MCR enzymes from various archaea differs (Figure 1C) and notable differences exist between the structurally characterized MCRs and *Ms*MCR. Interestingly, *Ms*MCR is closely related to MCR from *Candidatus* Methanoperedens nitroreducens, which is an anaerobic methane oxidizer of the ANME-2d clade. This indicates that the MCR enzymes of some methane producing, and methane consuming, archaea might not only be very similar regarding their active site, but most likely also regarding overall structural features. The structural information we gained on *Ms*MCR might therefore be useful to understand MCR enzymes from ANME-2 archaea better.

To conclude, accumulating structural information from metabolically and ecologically diverse MCRs is broadening our scope on their natural diversity, as well as their post-translational modification repertoire. The latter is not following phylogenetic relationships or consistency regarding the growth conditions (*e.g.*, temperature, and salt concentration) and it is still a mystery why such energy-extremophiles sacrifice cellular energy to install MCR post-translational modifications. The synergistic effort of genetic modification by using *M. acetivorans* as chassis and the exploration of the broad natural MCR landscape will hopefully provide more clues to further investigate the function of the modifications and could ultimately improve the robustness of the biotechnological application of MCRs.

Material and Methods

Phylogenetic analyses

Protein sequences used for phylogenetic analyses are, organized by organism (α,β,γ): *Methermicoccus shengliensis* DSM 18856 (WP_084174107.1, WP_042686194.1, WP_042686201.1), *Methanosarcina barkeri* Fusaro (WP_011305916.1, WP_011305920.1, WP_011305917.1), *Methanosarcina mazei* Gö1 (WP_011033189.1, WP_011033193.1, WP_048045871.1), *Methanothrix thermoacetophila* PT (WP_011695757.1, WP_011695760.1, WP_011695758.1), *Candidatus* Methanoperedens nitroreducens ANME-2d (WP_048089615.1, WP_048089608.1, WP_048089613.1), *Methanolobus profundi* Mob M (WP_091936029.1, WP_091936035.1, WP_091936030.1), *Methanomethylovorans hollandica* DSM 15978 (WP_015325028.1, WP_015325024.1, WP_015325027.1), *Methanimicrococcus blatticola* DSM 13328 (WP_133517056.1, WP_133517053.1, WP_133517055.1), *Methanotorris formicicus* Mc-S-70 (WP_007043982.1, WP_007043986.1, WP_007043983.1), *Methanothermococcus ther- molithotrophicus* DSM 2095 (WP_018153522.1, WP_018153526.1, WP_018153523.1), *Methan- othermobacter marburgensis* strain Marburg type I (WP_013296337.1, WP_013296341.1, WP_013296338.1), *Methanopyrus kandleri* AV19 (WP_011019025.1, WP_011019021.1, WP_011019024.1), *Methanothermobacter wolfeii* isolate SIV6 (SCM58307.1, SCM58314.1, SCM58308.1), *Methanosphaerula palustris* E1-9c (WP_012618913.1, WP_012618909.1, WP_012618912.1), *Methanoculleus horonobensis* JCM 15517 (WP_067078350.1, WP_067078343.1, WP_067078348.1), *Methanoplanus limicola* DSM 2279 (WP_004079635.1, WP_004079639.1, WP_004079636.1), *Methanocella conradii* isolate 1 (WP_174590719.1, WP_174590722.1, WP_014405505.1), *Methanococcus maripaludis* C7 (WP_011977191.1, WP_011977187.1, WP_011977190.1), *Methanocaldococcus vulcanius* M7 (WP_012819563.1, WP_012819559.1, WP_012819562.1),

Methanobrevibacter smithii DSM 2375 (WP_019262578.1, WP_004035807.1, WP_004035804.1), ANME-1 from Black Sea mats, Uncultured archaeon ANME-1 (D1JBK4, D1JBK2, D1JBK3). The sequences were aligned by using the Clustal W tool in MegaX[29] followed by the evolutionary analyses conducted with the same software. The evolutionary history was inferred using the Maximum Likelihood method and JTT matrix-based model[30]. The tree with the highest log likelihood (−27,171.46) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 21 amino acid sequences. There were a total of 1,328 positions in the final dataset. The references regarding the substrate utilization for each methanogen and MCR post-translational modifications, presented in the phylogenetic tree, can be found in the supplemental information.

Cultivation of Methermicoccus shengliensis

Methermicoccus shengliensis ZC-1 (DSM 18856)[6] was obtained from the DSMZ (Braunschweig, Germany) and cultivated in modified DSM medium 1084. Sludge fluid was replaced by trace element solution (100 × trace element solution: 1.5 $g.L^{-1}$ nitrilotriacetic acid, 3 $g.L^{-1}$ $MgSO_4 \cdot 7\ H_2O$, 0.45 $g.L^{-1}$ $MnSO_4 \cdot 2\ H_2O$, 1 $g.L^{-1}$ NaCl, 0.1 $g.L^{-1}$ $FeSO_4 \cdot 7\ H_2O$, 0.18 $g.L^{-1}$ $CoSO_4 \cdot 6\ H_2O$, 0.1 $g.L^{-1}$ $CaCl_2 \cdot 2\ H_2O$, 0.18 $g.L^{-1}$ $ZnSO_4 \cdot 7\ H_2O$, 0.01 $g.L^{-1}$ $CuSO_4 \cdot 5\ H_2O$, 0.02 $g.L^{-1}$ $KAl(SO_4)_2 \cdot 12\ H_2O$, 0.01 $g.L^{-1}$ H_3BO_3, 0.01 $g.L^{-1}$ $Na_2WO_4 \cdot 2\ H_2O$, 0.01 $g.L^{-1}$ $Na_2MoO_4 \cdot 2\ H_2O$, 0.025 $g.L^{-1}$ $NiCl_2 \cdot 6\ H_2O$, 0.01 $g.L^{-1}$ Na_2SeO_3) and vitamin solution (1,000 × vitamin solution: 20 $mg.L^{-1}$ biotin, 20 $mg.L^{-1}$ folic acid, 100 $mg.L^{-1}$ pyridoxine-HCl, 50 $mg.L^{-1}$ thiamin-HCl·2 H_2O, 50 $mg.L^{-1}$ riboflavin, 50 $mg.L^{-1}$ nicotinic acid, 50 $mg.L^{-1}$ D-Capantothenate, 2 $mg.L^{-1}$ vitamin B_{12}, 50 $mg.L^{-1}$ p-aminobenzoic acid, 50 $mg.L^{-1}$ lipoic acid). The

amount of supplied coenzyme M was reduced 20-fold (0.13 $g.L^{-1}$) and 2.5 $g.L^{-1}$ $NaHCO_3$ instead of 1 $g.L^{-1}$ Na_2CO_3 was used. The medium was sparged with N_2:CO_2 in an 80:20 ratio before autoclaving. 10 mM 3,4,5-trimethoxybenzoate was used (0.5 M stock solution was prepared by adjusting the pH to 8). Methanol-grown cultures were provided with 100 mM methanol instead of TMB. The cultures were incubated at 65 °C. *M. shengliensis* cells were harvested anaerobically (10,000× *g*, 25 min and 4 °C) after reaching the late exponential phase and cells were frozen anaerobically at −80 °C.

Native purification of *Ms*MCR

About 6 g of cells were defrosted while gassing for 10 min with N_2 gas. Afterwards, cells were resuspended in 15 mL anaerobic IEC buffer (50 mM Tris/HCl pH 8, 2 mM dithiothreitol (DTT)), sonicated (6 × 75 % amplitude for 10 s with 20 s break, Bandelin sonopuls, Berlin, Germany), centrifuged (16,250× *g*, 30 min at room temperature) and the supernatant was collected. The pellets were resuspended in 15 mL anaerobic IEC buffer, sonicated (5 × 75 % amplitude for 10 s with 20 s break), centrifuged (16,250× *g*, 30 min) and the supernatant was combined with the supernatant from the previous step. The supernatant was then diluted 5-fold with IEC buffer, filtered through a 0.2 µm filter (Sartorius, Göttingen, Germany) and loaded on a 15 mL DEAE column (GE healthcare, Chicago, IL, USA). Proteins were eluted by applying a 0 to 0.45 M NaCl gradient, over 120 min with a flow rate of 2 $mL.min^{-1}$. Under these conditions, MCR eluted between 0.28 and 0.33 M NaCl. The fractions containing MCR were pooled, diluted with 4 volumes of HIC buffer (25 mM Tris/HCl pH 7.6, 2 mM DTT, 2 M $(NH_4)_2SO_4$), filtered through a 0.2 µm filter and loaded on a 5 mL phenyl sepharose column (GE healthcare). Proteins were eluted by applying a 1.7 to 0 M gradient of $(NH_4)_2SO_4$ over 60 min with a flow rate of 1 $mL.min^{-1}$. MCR was eluting between 1.25 and 1 M of $(NH_4)_2SO_4$. Pooled MCR fractions were diluted with 4 volumes of HIC buffer, filtered through a 0.2 µm filter and loaded on a Source15Phe 4.6/100 PE column

(GE healthcare). Proteins were eluted by applying a gradient of 1.6 to 0 M $(NH_4)_2SO_4$, over 60 min with a flow rate of 1 mL.min^{-1}. Fractions of apparently pure MCR were eluting between 1.45 and 1.2 M $(NH_4)_2SO_4$. Pooled fractions were concentrated with 15 mL Millipore Ultra-10 centrifugal filter units (Merck, Darmstadt, Germany) and the buffer was exchanged for storage buffer (25 mM Tris pH 7.6, 10 % v/v glycerol, 2 mM DTT). MCR was concentrated to 47 g.L^{-1} and immediately used for crystallization and spectrophotometry. Protein concentration was evaluated by the Bradford method according to manufacturer (Bio-Rad, Hercules, CA, USA) recommendations. MCR from methanol-grown cells was purified following a similar protocol consisting of DEAE and Phenyl sepharose.

To compare the Stokes radius of *Ms*MCR purified from methanol and TMB-grown cells, both proteins (0.55 mg of purified *Ms*MCR) were injected on a Superdex 200 10/300 Increase GL (GE Healthcare) at a flow rate of 0.4 mL.min^{-1} at 20 °C. Both *Ms*MCRs showed an elution volume of 10.55 mL.

High-resolution clear native (hrCN) polyacrylamide gel electrophoresis (PAGE)

The hrCN-PAGE protocol was adapted from Lemaire *et al.*[31]. Glycerol (20 % v/v final) was added to samples and 0.001 % w/v Ponceau S was used as a protein migration marker. The electrophoresis cathode buffer contained a buffer mixture of 50 mM Tricine; 150 mM Bis-Tris pH 7 supplemented with 0.05 % w/v sodium deoxycholate; 0.01 % w/v dodecyl maltoside. The anode buffer contained 150 mM Bis-Tris buffer, pH 7. The NativeMark™ unstained protein standard from Thermo Fisher Scientific (Darmstadt, Germany) was used as a ladder. hrCN-PAGE were carried out using an 8 to 15 % linear polyacrylamide gradient, gels were run with a constant 20 mA current using a PowerPac™ Basic Power Supply (Bio- Rad). After electrophoresis, the protein bands were stained with Instant Blue™ (Expedeon, Heidelberg, Germany).

Mass spectrometry

MCR α-subunit was identified with help of matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) by the following protocol. Protein bands were cut into small pieces (about 3 × 3 mm) and destained by adding the following solvents/buffers successively: 20 µL acetonitrile (ACN), 20 µL 50 mM ammonium bicarbonate (ABC) buffer, 50 % v/v ACN in ABC buffer and 20 µL ACN. After each addition, samples were swirled and incubated for 10 min at room temperature (RT) followed by removing the liquid from the sample. Those steps were repeated until the gel pieces were destained. For reduction and alkylation, samples were incubated in 20 µL 10 mM DTT at 56 °C for 30 min, the liquid removed, and the following solvents/buffers successively added: 20 µL ACN, 20 µL 50 mM 2-chloroacetamide in 50 mM ABC buffer, 20 µL ACN, 20 µL ABC buffer, 20 µL ACN, and 20 µL ABC buffer. After each addition, samples were incubated for 10 min at RT followed by removing the liquid from the sample. For trypsin digestion, 10 µL of 5 $ng.\mu L^{-1}$ trypsin (V5518, Promega, Madison, WI, USA) in 50 mM ABC buffer were added to the gel pieces followed by 30 min incubation at RT. Afterwards 20 µL ABC buffer were added and the samples were incubated overnight at 37 °C. The samples were sonicated for 20 s in a sonication bath (Branson 2510, Brookfield, CT, USA) and 20 µL 0.1 % v/v trifluoroacetic acid were added. The samples were incubated for 20 min at RT before the extract liquid was transferred to a new tube. 20 µL ACN were added to the remaining trypsin digests, the samples were incubated for 30 min at RT before the extract liquid was combined with the extract liquid from before. The samples were then dried in a Sanvant ISS110 speedVac (Thermo Scientific, Waltham, MA, USA) until ~5 µL remained. Then, 0.5 µL of the extracted peptides was pipetted on a MALDI-TOF sample plate and directly mixed with an equal volume of matrix solution containing 10 $mg.mL^{-1}$ α-cyano-4-hydroxy-α-cyanocinnamic acid in 50 % v/v ACN/0.05 % v/v trifluoroacetic acid. After drying of the sample this process was repeated once more. A spectrum in the

range of 600 to 4,000 *m/z* was recorded using a Microflex LRF MALDI-TOF (Bruker). The Biotools software (Bruker Life Sciences) was used to perform a MASCOT search (Matrix Science Ltd., London, UK) by using the *M. shengliensis* protein database (GenBank accession number NZ_JONQ00000000.1). Search parameters allowed a mass deviation of 0.3 Da, one miscleavage, a variable modification of oxidized methionines and a fixed modification of carbamidomethylated cysteines. The N^1-methylhistidine275 containing peptide (1,452.9 Da vs predicted mass of 1,452.7 Da) and the Gln418 containing peptide (3,468.8 Da vs predicted mass of 3,468.6 Da) were detected (see Supplemental Figure S1A). The mass of the 5(*S*)-methylarginine289, containing peptide is below the 600 m/z threshold and was therefore not detected. The alkylated peptide containing the thioglycine463, aspartate468 and cysteine470 (LGFF**G**YDLQ**D**Q**C**GAANVFSYQSDEGLPLELR) was observable (3,524.7 Da vs predicted mass of 3,525.6 Da; Supplemental Figure S1A-B) at a signal over noise threshold of 1.9 and the mass fits in the 1 Dalton range.

Crystallization

MCR crystals were obtained aerobically by initial screening at 18 °C using the sitting drop method on 96-Well MRC 2-Drop Crystallization Plates in polystyrene (SWISSCI). The crystallization reservoir contained 90 µL of the following crystallization condition: 25 % w/v polyethylene glycol 3,350, 100 mM Bis-Tris pH 5.5, and 200 mM lithium sulfate. The crystallization drop contained a mixture of 0.6 µL *Ms*MCR at a concentration of 47 $mg.mL^{-1}$ and 0.6 µL of the crystallization condition. Thick yellow brick-shaped crystals appeared within two weeks.

X-ray data collection and model refinement/validation

All X-ray crystallographic data and refinement statistics are presented in Table 1. MCR crystals were soaked in the crystallization solution supplemented with 20 % v/v glycerol for 6 s before being transferred to liquid nitrogen. All diffraction experiments were

performed at 100 K on Proxima-1 beamline, SOLEIL synchrotron, Saclay, France. The data were processed with xdsme and scaled with SCALA from the CCP4 package[32]. *Ms*MCR structure was solved by molecular replacement with Phenix[33] using MCR from *Methanosarcina barkeri* (PDB 1E6Y[13]) as a template. The model was manually built via Coot[34] and refined with BUSTER[35] by using the non-crystallographic symmetry and translational-libration screw (TLS). The last refinement steps were performed with hydrogens in riding position. The model was ultimately validated by the MolProbity server[36] (http://molprobity.biochem.duke.edu, accessed on 15 of February 2021). Hydrogens were omitted in the final deposited model (PDB code 7NKG). All figures were generated and rendered with PyMOL (V. 1.8, Schrödinger, LLC).

Table 1. X-ray crystallographic data and refinement statistics.

	MCR from *M. shengliensis*
Data collection	
Wavelength (Å)	0.97856
Space group	$P2_12_12_1$
Resolution (Å)	49.41-1.60 (1.69-1.60)
Cell dimensions: a, b, c (Å)	132.62 148.18 235.41
R_{merge} (%)[a]	9.1 (121.6)
R_{pim} (%)[a]	5.1 (66.1)
$CC_{1/2}$[a]	0.997 (0.356)
I/σ_I[a]	8.3 (1.0)
Completeness[a]	99.7 (99.3)
Redundancy[a]	4.2 (4.3)
Number of unique reflections[a]	602614 (87124)
Refinement	
Resolution (Å)	48.36-1.60
Number of reflections	602442
R_{work}/R_{free}[b] (%)	0.1725/0.1904
Number of atoms	
Protein	38087
Ligands/ions	405
Solvent	4298
Mean B-value (Å^2)	35.0
Molprobity clash score, all atoms	0.67
Ramachandran plot	
Favored regions (%)	97.71
Outlier regions (%)	0.16
rmsd[c] bond lengths (Å)	0.007
rmsd[c] bond angles (°)	0.95
PDB ID code	**7NKG**

[a] Values relative to the highest resolution shell are within parentheses. [b] R_{free} was calculated as the R_{work} for 5 % of the reflections that were not included in the refinement. Refined models contained hydrogens. [c] rmsd, root mean square deviation.

Acknowledgments, author contributions, data availability and funding

We thank the Max Planck Institute for Marine Microbiology and the Max-Planck-Society for continuous support. We acknowledge the SOLEIL synchrotron for beam time allocation and the beamline staff of Proxima-1 for assistance with data collection, with specific regards to Pierre Legrand for his precious advices. We are also thankful to Christina Probian and Ramona Appel for their continuous support in the Microbial Metabolism laboratory. We further thank Huub Op den Camp for his assistance with MALDI-TOF MS.

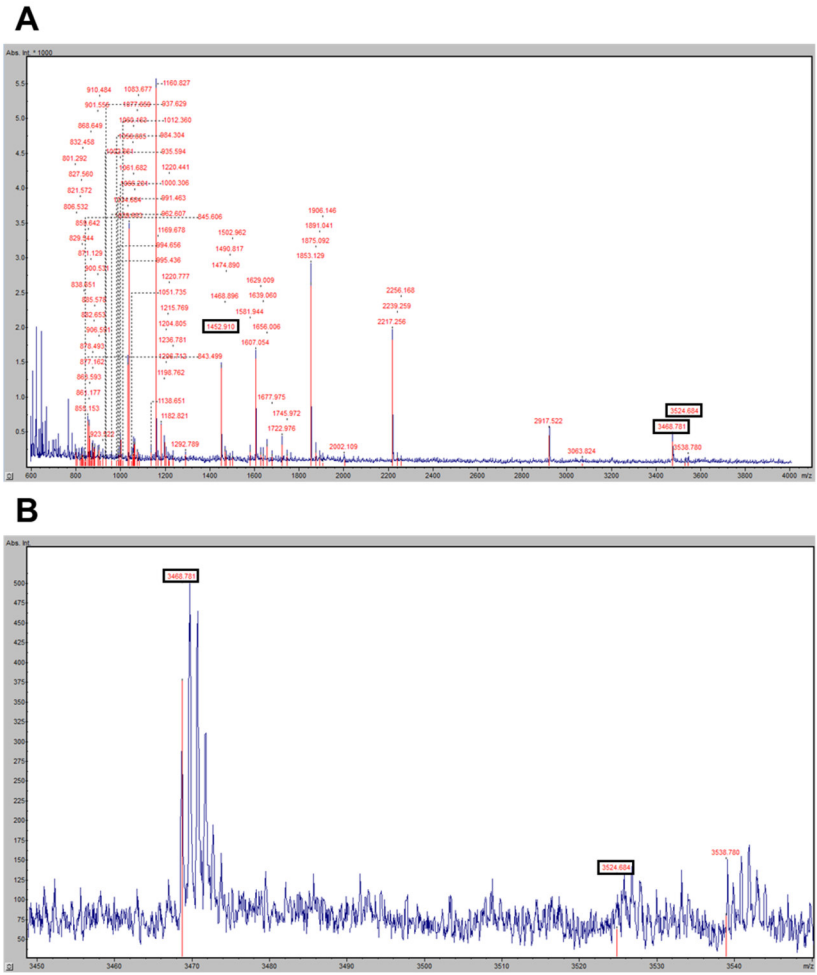
This study was designed by J.M.K. and T.W. Cultivation was performed by J.M.K. Protein extraction and purification was performed by J.M.K., M.-C.M. and T.W., J.M.K. performed phylogenetic analyses and mass spectrometry. Crystallization, X-ray data collection, processing and model building were performed by M.-C.M. and T.W. Crystal structure analyses were performed by M.-C.M. and T.W. The data was interpreted, and the paper was written by J.M.K., M.-C.M., C.U.W. and T.W. with contributions and final approval of all co-authors. All authors have read and agreed to the published version of the manuscript.

The structure was deposited in the protein data bank under the ID: 7NKG.

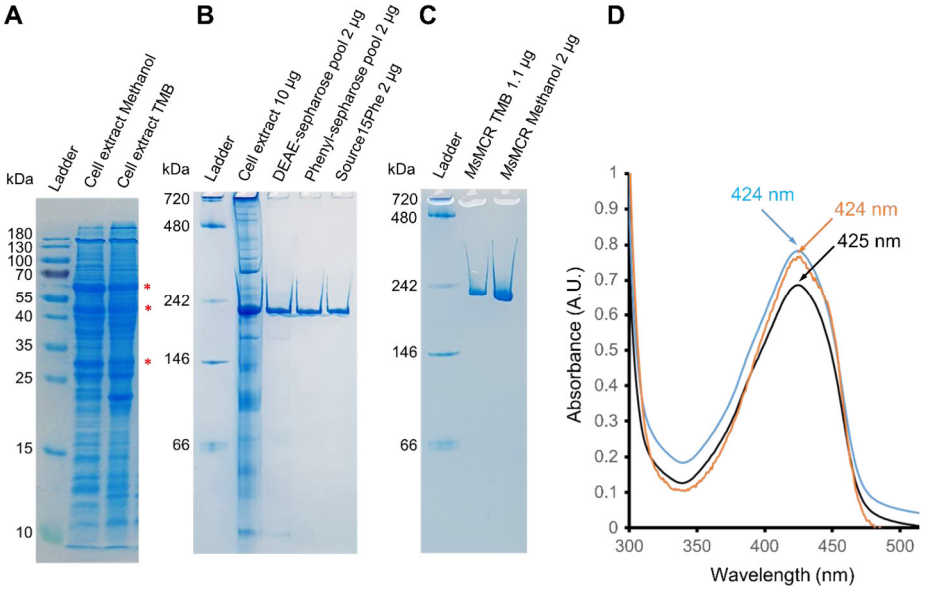
The research was funded by the Max-Planck Gesellschaft for M.-C.M. and T.W., C.U.W. and J.M.K. were supported by the Dutch Research Council through the Soehngen Institute for Anaerobic Microbiology (SIAM) Gravitation Grant 024.002.002. J.M.K. was further supported by the Deutsche Forschungsgemeinschaft (DFG) Grant KU 3768/1-1.

The authors declare no conflict of interest.

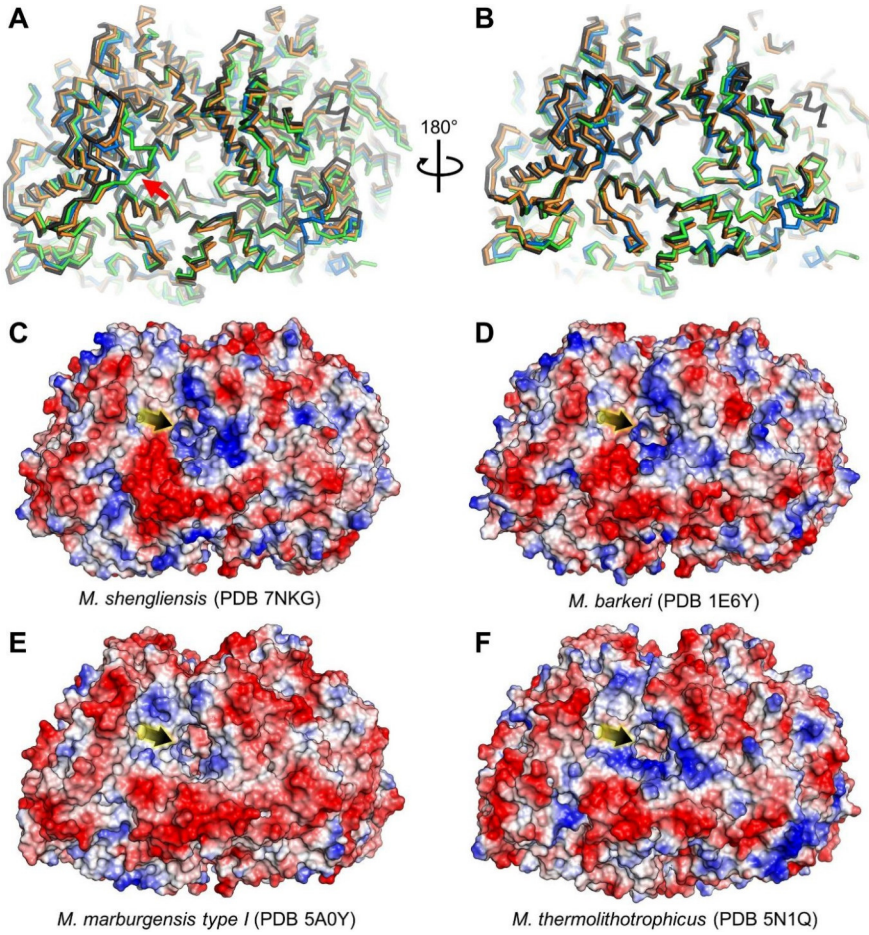
Supplementary materials



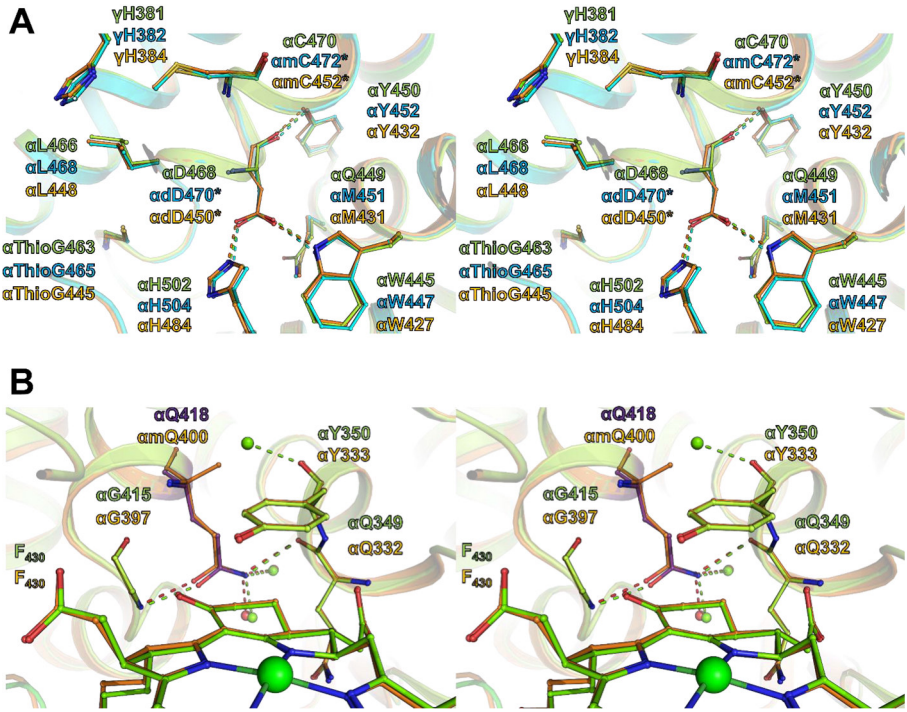
Supplemental Fig. 1. Mass spectrometry data obtained on MCRα peptides. **(A)** Full spectrum **(B)** Close up of the 3,450-3,550 m/z region. Peaks of interest are highlighted with boxes: N^1-methylhistidine275 containing peptide (1,452.9 Da vs predicted mass of 1,452.7 Da), Gln418 containing peptide (3,468.8 Da vs predicted mass of 3,468.6 Da). The alkylated peptide containing the thioglycine463 (3,524.7 Da vs predicted mass of 3,525.6 Da) is detected at a signal over noise threshold of 1.9. Low signal strength might be explained by increased instability of larger peptides.



Supplemental Fig. 2. Native PAGE and UV/visible spectra profile of purified *Ms*MCRs. (**A**) SDS-PAGE of cell extract obtained from methanol and 3,4,5-trimethoxybenzoate (TMB) grown cells. The red stars show the position of the expected three MCR subunits. (**B**) High-resolution clear native PAGE of *Ms*MCR from TMB-grown cells for each purification step. (**C**) High- resolution clear native PAGE of purified *Ms*MCR from TMB and methanol-grown cells. (**D**) UV/visible spectra of purified *Ms*MCR from TMB-grown cells at a concentration of 5.6 mg.mL^{-1} (in 50 mM Tris/HCl pH 7.6, 10 % v/v glycerol and 2 mM dithiothreitol) under anaerobic atmosphere (black line) and after one hour O2-exposition (blue line). The orange line corresponds to the UV/visible spectra of aerobically purified *Ms*MCR from methanol-grown cells at a concentration of 5.6 mg.mL^{-1} (in 25 mM Tris/HCl pH 7.6, 10 % v/v glycerol and 2 mM dithiothreitol). The peak at 424-425 nm is typical for a Ni(II) red1-silent inactive state.



Supplemental Fig. 3. Structural and electrostatic charge differences between *Ms*MCR and its homologs. (**A-B**) Superposition of the α-subunit of *Ms*MCR with MCR homologs. The Cα-trace is represented in ribbon with a color code corresponding to the different homologs: *Ms*MCR (PDB 7NKG) in green, MCR from *M. barkeri* (*Mb*MCR, PDB 1E6Y) in blue, MCR type I from *M. marburgensis* (*Mm*MCR, PDB 5A0Y) in orange and MCR from *M. thermolithotrophicus* (*Mt*MCR, PDB 5N1Q) in black. The deviating loop is highlighted with a red arrow. The difference between panel (**A**) and (**B**) is a 180° rotation along the y-axis. (**C-F**), Electrostatic charge surface representation. The charges from negative to positive are reflected by a color gradient from red to blue, respectively. An arrow points towards the active site channel entrance.



Supplemental Fig. 4. Close up of the environment at the expected modified residues didehydroaspartate, S-methylcysteine and 2(S)-methylglutamine in *Ms*MCR. (**A**) Stereo-view of the close surrounding at the didehydroaspartate (αdDXXX*) and S-methylcysteine (αmCXXX*) positions. Both residues and their close surrounding are shown as balls and sticks and secondary structures are shown in transparent cartoon. *Ms*MCR, *Mb*MCR (PDB 1E6Y) and *Mm*MCR (PDB 5A0Y) are colored in green, cyan and orange, respectively. The only observable mutation in the close surrounding of the expected didehydroaspartate (*i.e.,* αD468 in *Ms*MCR) is the methionine that turns to a glutamine (*i.e.,* αM451 in *Mb*MCR, αM431 in *Mm*MCR, αQ449 in *Ms*MCR). (**B**) Stereo-view of the close surrounding at the 2(S)-methylglutamine (αmQ400 in *Mm*MCR) position. 2(S)-methylglutamine and its close surrounding is shown as balls and sticks and secondary structures are shown in transparent cartoon. *Ms*MCR, and *Mm*MCR are colored in green and orange, respectively. αQ418 in *Ms*MCR is highlighted in deep purple. The nickel from the F_{430} cofactor is shown as a large green ball.

Supplemental Table 1. Sequence identity between the subunits of different MCRs and root mean square deviation (r.m.s.d) of structurally characterized MCRs.

Organism (PDB)	Sequence identity (%)			R.m.s.d (aligned Cα)		
	α-subunit	β-subunit	γ-subunit	α-subunit	β-subunit	γ-subunit
M. shengliensis (7NKG)	100	100	100	0	0	0
M. barkeri (1E6Y)	72.97	70.9	81.38	0.371 (477)	0.411 (385)	0.276 (205)
M. marburgensis type I (5A0Y)	67.73	61.63	58.37	0.410 (424)	0.655 (382)	0.546 (192)
M. thermolithotrophicus (5N1Q)	62.98	57.01	51.59	0.522 (448)	0.716 (399)	0.805 (217)

References regarding the substrate utilization of methanogens and MCR post-translational modifications used in the phylogenetic tree (Fig. 1C).

Methanosarcina barkeri[13, 37]

Methanosarcina mazei[38]

Methanimicrococcus blatticola[39]

Methanolobus profundi[40]

Methanomethylovorans hollandica[41]

Candidatus Methanoperedens nitroreducens[42]

Methermicoccus shengliensis[5, 6]

Methanothrix thermoacetophila[43]

Methanocella conradii[44]

Methanosphaerula palustris[45]

Methanoculleus horonobensis[46]

Methanoplanus limicola[47]

Methanothermobacter marburgensis[15, 48]

Methanothermobacter wolfeii[15, 49, 50]

Methanobrevibacter smithii[51]

Methanopyrus kandleri[13, 52]

Methanotorris formicicus[16, 53]

Methanocaldococcus vulcanius[54]

Methanothermococcus thermolithotrophicus[16, 55]

Methanococcus maripaludis[56]

ANME-1 from Black Sea mats,

Uncultured archaeon ANME-1[14]

References

1. Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W., Hedderich, R. Methanogenic archaea: Ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **6**, 579-91, doi:10.1038/nrmicro1931 (2008).
2. Ueno, Y., Yamada, K., Yoshida, N., Maruyama, S., Isozaki, Y. Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era. *Nature* **440**, 516-9, doi:10.1038/nature04584 (2006).
3. Wagner, T., Watanabe, T., Shima, S. Hydrogenotrophic methanogenesis. In *Biogenesis of Hydrocarbons Handbook of Hydrocarbon and Lipid Microbiology* (eds Stams, A.; Sousa, D.) 1-29 (Springer, 2018).
4. Kurth, J. M., Op den Camp, H. J. M., Welte, C. U. Several ways one goal—methanogenesis from unconventional substrates. *Appl. Microbiol. Biotechnol.* **104**, 6839-6854, doi:10.1007/s00253-020-10724-7 (2020).
5. Mayumi, D., Mochimaru, H., Tamaki, H., Yamamoto, K., Yoshioka, H., Suzuki, Y., Kamagata, Y., Sakata, S. Methane production from coal by a single methanogen. *Science* **354**, 222-225, doi:10.1126/science.aaf8821 (2016).
6. Cheng, L., Qiu, T. L., Yin, X. B., Wu, X. L., Hu, G. Q., Deng, Y., Zhang, H. *Methermicoccus shengliensis* gen. nov., sp. nov., a thermophilic, methylotrophic methanogen isolated from oil-production water, and proposal of *Methermicoccaceae* fam. nov. *Int. J. Syst. Evol. Microbiol.* **57**, 2964-2969, doi:10.1099/ijs.0.65049-0 (2007).
7. Thauer, R. K. Methyl (alkyl)-coenzyme M reductases: Nickel F-430-containing enzymes involved in anaerobic methane formation and in anaerobic oxidation of methane or of short chain alkanes. *Biochemistry* **58**, 5198-5220, doi:10.1021/acs.biochem.9b00164 (2019).
8. Wongnate, T., Sliwa, D., Ginovska, B., Smith, D., Wolf, M. W., Lehnert, N., Raugei, S., Ragsdale, S. W. The radical mechanism of biological methane synthesis by methyl-coenzyme M reductase. *Science* **352**, 953-8, doi:10.1126/science.aaf0616 (2016).
9. Cedervall, P. E., Dey, M., Li, X., Sarangi, R., Hedman, B., Ragsdale, S. W., Wilmot, C. M. Structural analysis of a Ni-methyl species in methyl-coenzyme M reductase from *Methanothermobacter marburgensis*. *J. Am. Chem. Soc.* **133**, 5626-8, doi:10.1021/ja110492p (2011).
10. Cedervall, P. E., Dey, M., Pearson, A. R., Ragsdale, S. W., Wilmot, C. M. Structural insight into methyl-coenzyme M reductase chemistry using coenzyme B analogues. *Biochemistry* **49**, 7683-93, doi:10.1021/bi100458d (2010).
11. Ermler, U., Grabarse, W., Shima, S., Goubeaud, M., Thauer, R. K. Crystal structure of methyl-coenzyme M reductase: The key enzyme of biological methane formation. *Science* **278**, 1457-62, doi:10.1126/science.278.5342.1457 (1997).
12. Grabarse, W., Mahlert, F., Duin, E. C., Goubeaud, M., Shima, S., Thauer, R. K., Lamzin, V., Ermler, U. On the mechanism of biological methane formation: structural evidence for conformational changes in methyl-coenzyme M reductase upon substrate binding. *J. Mol. Biol.* **309**, 315-30, doi:10.1006/jmbi.2001.4647 (2001).
13. Grabarse, W., Mahlert, F., Shima, S., Thauer, R. K., Ermler, U. Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: unusual amino acid modification, conservation and adaptation. *J. Mol. Biol.* **303**, 329-44, doi:10.1006/jmbi.2000.4136 (2000).
14. Shima, S., Krueger, M., Weinert, T., Demmer, U., Kahnt, J., Thauer, R. K., Ermler, U. Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature* **481**, 98-101, doi:10.1038/nature10663 (2012).
15. Wagner, T., Kahnt, J., Ermler, U., Shima, S. Didehydroaspartate modification in methyl-coenzyme M reductase catalyzing methane formation. *Angew. Chem. Int. Ed Engl.* **55**, 10630-3, doi:10.1002/anie.201603882 (2016).
16. Wagner, T., Wegner, C. E., Kahnt, J., Ermler, U., Shima, S. Phylogenetic and structural comparisons of the three types of methyl coenzyme M reductase from *Methanococcales* and *Methanobacteriales*. *J. Bacteriol.* **199**, doi:10.1128/JB.00197-17 (2017).

17. Nayak, D. D., Liu, A. D., Agrawal, N., Rodriguez-Carerro, R., Dong, S. H., Mitchell, D. A., Nair, S. K., Metcalf, W. W. Functional interactions between posttranslationally modified amino acids of methyl-coenzyme M reductase in *Methanosarcina acetivorans*. *PLoS Biol.* **18**, doi:10.1371/journal.pbio.3000507 (2020).
18. Kahnt, J., Buchenau, B., Mahlert, F., Krüger, M., Shima, S., Thauer, R. K. Post-translational modifications in the active site region of methyl-coenzyme M reductase from methanogenic and methanotrophic archaea. *FEBS J.* **274**, 4913-21, doi:10.1111/j.1742-4658.2007.06016.x (2007).
19. Deobald, D., Adrian, L., Schöne, C., Rother, M., Layers, G. Identification of a unique radical SAM methyltransferase required for the sp^3-C-methylation of an arginine residue of methyl-coenzyme M reductase. *Sci. Rep.* **8**, doi:10.1038/s41598-018-25716-x (2018).
20. Nayak, D. D., Mahanta, N., Mitchell, D. A., Metcalf, W. W. Post-translational thioamidation of methyl-coenzyme M reductase, a key enzyme in methanogenic and methanotrophic Archaea. *eLife* **6**, doi:10.7554/eLife.29218 (2017).
21. Chen, H., Gan, Q. L., Fan, C. G. Methyl-coenzyme M reductase and its post-translational modifications. *Front. Microbiol.* **11**, doi:10.3389/fmicb.2020.578356 (2020).
22. Scheller, S., Goenrich, M., Boecher, R., Thauer, R. K., Jaun, B. The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature* **465**, 606-U97, doi:10.1038/nature09015 (2010).
23. Krüger, M., Meyerdierks, A., Glöckner, F. O., Amann, R., Widdel, F., Kube, M., Reinhardt, R., Kahnt, R., Böcher, R., Thauer, R. K., et al. A conspicuous nickel protein in microbial mats that oxidize methane anaerobically. *Nature* **426**, 878-881, doi:10.1038/nature02207 (2003).
24. Schaefer, H., Mikaloff Fletcher, S. E., Veidt, C., Lassey, K. R., Brailsford, G. W., Bromley, T. M., Dlugokencky, E. J., Michel, S. E., Miller, J. B., Levin, I., et al. A 21st-century shift from fossil-fuel to biogenic methane emissions indicated by $^{13}CH_4$. *Science* **352**, 80-4, doi:10.1126/science.aad2705 (2016).
25. Conrad, R. Microbial Ecology of Methanogens and Methanotrophs. In *Advances in Agronomy* Vol. 96 1-63 (Academic Press, 2007).
26. Scheller, S., Ermler, U., Shima, S. Catabolic pathways and enzymes involved in anaerobic methane oxidation. In *Anaerobic Utilization of Hydrocarbons, Oils, and Lipids* (ed Boll, M.) 31-59 (Springer International Publishing, 2020).
27. Duin, E. C., Wagner, T., Shima, S., Prakash, D., Cronin, B., Yáñez-Ruiz, D. R., Duval, S., Rümbeli, R., Stemmler, R. T., Thauer, R. K., et al. Mode of action uncovered for the specific reduction of methane emissions from ruminants by the small molecule 3-nitrooxypropanol. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 6172-7, doi:10.1073/pnas.1600298113 (2016).
28. Ragsdale, S. W. Biochemistry of methyl-coenzyme M reductase: The nickel metalloenzyme that catalyzes the final step in synthesis and the first step in anaerobic oxidation of the greenhouse gas methane. In *The Metal-Driven Biogeochemistry of Gaseous Compounds in the Environment* Vol. 14 (eds Kroneck, P.; Torres, M.) 125-145 (Springer, 2014).
29. Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547-1549, doi:10.1093/molbev/msy096 (2018).
30. Jones, D. T., Taylor, W. R., Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275-82, doi:10.1093/bioinformatics/8.3.275 (1992).
31. Lemaire, O. N., Infossi, P., Ali Chaouche, A., Espinosa, L., Leimkühler, S., Giudici-Orticoni, M. T., Méjean, V., Iobbi-Nivol, C. Small membranous proteins of the TorE/NapE family, crutches for cognate respiratory systems in Proteobacteria. *Sci. Rep.* **8**, 13576, doi:10.1038/s41598-018-31851-2 (2018).
32. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., et al. Overview of the *CCP4* suite and current developments. *Acta Crystallogr. D Struc. Biol.* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).
33. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., et al. Macromolecular structure determination using X-rays,

neutrons and electrons: recent developments in *Phenix*. *Acta Crystallogr. D Struct. Biol.* **75**, 861-877, doi:10.1107/S2059798319011471 (2019).
34. Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
35. Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C. Buster Version 2.10.4. *Cambr. U. K. Glob. Phasing Ltd.* (2017).
36. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., Richardson, D. C. *MolProbity*: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Struct. Biol.* **66**, 12-21, doi:10.1107/S0907444909042073 (2010).
37. Bryant, M. P., Boone, D. R. Emended description of strain MS^T (DSM 800^T), the type strain of *Methanosarcina barkeri*. *Int. J. Syst. Evol. Microbiol.* **37**, 169-170, doi:10.1099/00207713-37-2-169 (1987).
38. Mah, R. A. Isolation and characterization of *Methanococcus mazei*. *Curr. Microbiol.* **3**, 321-326, doi:10.1007/BF02601895 (1980).
39. Sprenger, W. W., van Belzen, M. C., Rosenberg, J., Hackstein, J. H., Keltjens, J. T. *Methanomicrococcus blatticola* gen. nov., sp. nov., a methanol- and methylamine-reducing methanogen from the hindgut of the cockroach *Periplaneta americana*. *Int. J. Syst. Evol. Microbiol.* **50 Pt 6**, 1989-1999, doi:10.1099/00207713-50-6-1989 (2000).
40. Mochimaru, H., Tamaki, H., Hanada, S., Imachi, H., Nakamura, K., Sakata, S., Kamagata, Y. *Methanolobus profundi* sp. nov., a methylotrophic methanogen isolated from deep subsurface sediments in a natural gas field. *Int. J. Syst. Evol. Microbiol.* **59**, 714-8, doi:10.1099/ijs.0.001677-0 (2009).
41. Lomans, B. P., Maas, R., Luderer, R., Op den Camp, H. J., Pol, A., van der Drift, C., Vogels, G. D. Isolation and characterization of *Methanomethylovorans hollandica* gen. nov., sp. nov., isolated from freshwater sediment, a methylotrophic methanogen able to grow on dimethyl sulfide and methanethiol. *Appl. Environ. Microbiol.* **65**, 3641-50, doi:10.1128/AEM.65.8.3641-3650.1999 (1999).
42. Raghoebarsing, A. A., Pol, A., van de Pas-Schoonen, K. T., Smolders, A. J. P., Ettwig, K. F., Rijpstra, W. I. C., Schouten, S., Damsté, J. S. S., Op den Camp, H. J. M., Jetten, M. S. M., et al. A microbial consortium couples anaerobic methane oxidation to denitrification. *Nature* **440**, 918-921, doi:10.1038/nature04617 (2006).
43. Kamataga, Y., Mikami, E. Isolation and characterization of a novel thermophilic *Methanosaeta* strain. *Int. J. Sys. Evol. Microbiol.* **41**, 191-196, doi:10.1099/00207713-41-2-191 (1991).
44. Lu, Z., Lu, Y. *Methanocella conradii* sp. nov., a thermophilic, obligate hydrogenotrophic methanogen, isolated from Chinese rice field soil. *PLoS One* **7**, e35279, doi:10.1371/journal.pone.0035279 (2012).
45. Cadillo-Quiroz, H., Yavitt, J. B., Zinder, S. H. *Methanosphaerula palustris* gen. nov., sp. nov., a hydrogenotrophic methanogen isolated from a minerotrophic fen peatland. *Int. J. Syst. Evol. Microbiol.* **59**, 928-35, doi:10.1099/ijs.0.006890-0 (2009).
46. Shimizu, S., Ueno, A., Tamamura, S., Naganuma, T., Kaneko, K. *Methanoculleus horonobensis* sp. nov., a methanogenic archaeon isolated from a deep diatomaceous shale formation. *Int. J. Syst. Evol. Microbiol.* **63**, 4320-4323, doi:10.1099/ijs.0.053520-0 (2013).
47. Wildgruber, G., Thomm, M., König, H., Ober, K., Richiuto, T., Stetter, K. O. *Methanoplanus limicola*, a plate-shaped methanogen representing a novel family, the methanoplanaceae. *Arch. Microbiol.* **132**, 31-36, doi:10.1007/BF00690813 (1982).
48. Fuchs, G., Stupperich, E., Thauer, R. K. Acetate assimilation and the synthesis of alanine, aspartate and glutamate in *Methanobacterium thermoautotrophicum*. *Arch. Microbiol.* **117**, 61-6, doi:10.1007/BF00689352 (1978).
49. Winter, J., Lerp, C., Zabel, H. P., Wildenauer, F. X., König, H., Schindler, F. *Methanobacterium wolfei*, sp. nov., a new tungsten-requiring, thermophilic, autotrophic methanogen. *Syst. Appl. Microbiol.* **5**, 457-466, doi:10.1016/S0723-2020(84)80003-X (1984).
50. Wasserfallen, A., Nölling, J., Pfister, P., Reeve, J., Conway de Macario, E. Phylogenetic analysis of 18 thermophilic *Methanobacterium* isolates supports the proposals to create a new genus,

Methanothermobacter gen. nov., and to reclassify several isolates in three species, *Methanothermobacter thermautotrophicus* comb. nov., *Methanothermobacter wolfeii* comb. nov., and *Methanothermobacter marburgensis* sp. nov. *Int. J. Syst. Evol. Microbiol.* **50 Pt 1**, 43-53, doi:10.1099/00207713-50-1-43 (2000).

51. Miller, T. L., Wolin, M. J., de Macario, E. C., Macario, A. J. Isolation of *Methanobrevibacter smithii* from human feces. *Appl. Environ. Microbiol.* **43**, 227-232, doi:10.1128/aem.43.1.227-232.1982 (1982).
52. Kurr, M., Huber, R., König, H., Jannasch, H. W., Fricke, H., Trincone, A., Kristjansson, J. K., Stetter, K. O. *Methanopyrus kandleri*, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110°C. *Arch. Microbiol.* **156**, 239-247, doi:10.1007/Bf00262992 (1991).
53. Takai, K., Nealson, K. H., Horikoshi, K. *Methanotorris formicicus* sp. nov., a novel extremely thermophilic, methane-producing archaeon isolated from a black smoker chimney in the Central Indian Ridge. *Int. J. Syst. Evol. Microbiol.* **54**, 1095-1100, doi:10.1099/ijs.0.02887-0 (2004).
54. Jeanthon, C., L'Haridon, S., Reysenbach, A. L., Corre, E., Vernet, M., Messner, P., Sleytr, U. B., Prieur, D. *Methanococcus vulcanius* sp. nov., a novel hyperthermophilic methanogen isolated from East Pacific Rise, and identification of *Methanococcus* sp. DSM 4213^T as *Methanococcus fervens* sp. nov. *Int. J. Syst. Bacteriol.* **49 Pt 2**, 583-9, doi:10.1099/00207713-49-2-583 (1999).
55. Huber, H., Thomm, M., König, H., Thies, G., Stetter, K. O. *Methanococcus thermolithotrophicus*, a novel thermophilic lithotrophic methanogen. *Archives of Microbiology* **132**, 47-50, doi:10.1007/Bf00690816 (1982).
56. Jones, W. J., Paynter, M. J. B., Gupta, R. Characterization of *Methanococcus maripaludis* sp. nov., a new methanogen isolated from salt marsh sediment. *Arch. Microbiol.* **135**, 91-97, doi:10.1007/BF00408015 (1983).

CHAPTER V.

Atomic resolution structures of methane-fixing enzymes from marine and freshwater ANME-2 depict a conserved organization with numerous post-translational modifications.

Marie-C. Müller[1], Martijn Wissink[2], Priya Mukherjee[1], Nicole von Possel[1], Rafa Laso-Pérez[3,4,] Sylvain Engilberge[5,6], Philippe Carpentier[7], Jörg Kahnt[8], Gunter Wegener[3], Cornelia Welte[2], Tristan Wagner[1,6#]

[1] Microbial Metabolism research group, Max-Planck-Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany

[2] Department of Microbiology, Institute for Water and Wetland Research, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands.

[3] HGF MPG Joint Research Group for Deep-Sea Ecology and Technology, Max-Planck-Institute for Marine Microbiology, Celsiusstraße 1, 28359, Bremen, Germany

[4] Systems Biology Department, Centro Nacional de Biotecnología (CNB-CSIC), Madrid, Spain.

[5] European Synchrotron Radiation Facility, Grenoble Cedex 9 38043, France

[6] Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale, Grenoble Cedex 9 38044, France.

[7] French National Centre for Scientific Research, CNRS · Institut de Chimie (INC), 7 Rue du Four Solaire, 61120, Paris, Pyrénées orientales, France

[8] Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Strasse 10, 35043 Marburg, Germany

[#]Correspondence: twagner@mpi-bremen.de

Manuscript in preparation

Abstract

Anaerobic methanotrophic archaea (ANME) play a crucial role in planetary carbon cycling. They oxidize methane in anoxic niches by transferring electrons directly to nitrate or metal oxides and alternatively to sulfate-reducing bacterial partners. Due to their physiological complexity, none of the ANME species have been isolated, hampering the biochemical investigation of the enzymatic processes involved in the anaerobic oxidation of methane. Here, we deciphered the methane capturing enzyme through native purification, circumventing the isolation barrier by utilizing microbial enrichments. Freshwater ANME-2d reducing nitrate, grown in bioreactors, and a marine microbial enrichment containing ANME-2c along with its bacterial symbiont were used as model systems. The crystal structures of the Methyl-Coenzyme M Reductases (MCRs), refined to true atomic resolution, captured its most precise image to date. Despite their physiological differences, these ANMEs have extremely conserved MCR structures, similar to homologs from methanogenic *Methanosarcinales,* rather than the MCR of ANME-1 isolated from Black Sea mats. An exceptional content of seven post-translational modifications was observed *in cristallo* and confirmed by mass spectrometry. Among them was a novel 3(*S*)-methylhistidine on the γ-chain of the ANME-2d MCR. Krypton gassing did not lead to the detection of an internal channel that would facilitate alkane diffusion to the active site, contrary to the ethane-activating homologue. Our results corroborate that MCRs from freshwater ANME-2d and marine ANME-2c should follow the conserved radical reaction mechanism of the methane-generating counterpart. The described pattern of post-translational modifications, different to ANME-1, highlights the importance of native purification as an unbiased approach to discovering intrinsic enzymatic features existing in nature.

Introduction

In light of climate change, studies about planetary methane fluxes have become increasingly important due to the global impact of the greenhouse gas. In anoxic marine sediments, 71 % of biogenic methane is transformed into the less potent greenhouse gas CO_2 through the process of anaerobic oxidation of methane (AOM)[1]. The same is true in terrestrial and freshwater systems where AOM can account for up to 50 % reduction in methane emissions in wetlands[2]. The only microorganisms capable of performing AOM are anaerobic methanotrophic archaea (ANME)[3, 4], a mixed taxonomic group comprising ANME-1 (Family *Methanophagaceae*), ANME-2a/b/c/d (Families *Methanocomedenaceae, Methanogasteraceae, Methanoperedenaceae*) and ANME-3 (Family *Methanophagales*)[5]. ANME-1 has been found in diverse aquatic sediments, ANME-2a/b/c are prevalent in mostly cold or moderate environments, ANME-2d are more abundant in freshwater, and ANME-3 are found in colder environments[6]. Most ANMEs fully oxidize methane to CO_2 by transferring electrons to sulfate-reducing *Deltaproteobacteria*, most likely by direct interspecies electron transfer (DIET) through nanowires[7, 8]. In addition, ANME 1-related, and ANME-2a have been proposed to reduce Fe in marine environments[9]. More versatile, ANME-2d couple AOM to the reduction of nitrate[10], iron oxides or other metal oxides as final electron acceptors[11]. To date, none of the studied ANMEs have been isolated. Nevertheless, some laboratories have succeeded in obtaining enrichments of microbial consortia[12-14], a strenuous process due to the extremely low doubling times of ANME.

Methane activation, the first reaction in AOM, has been proposed to be carried out by the Methyl-coenzyme M reductase (MCR). Accordingly, the reaction would follow the reverse direction involved in methanogenesis[15]. In the proposed mechanism, the heterodisulfide, made of Coenzyme M (HS-CoM) and Coenzyme B (CoB-SH), would react with methane to generate methyl-S-CoM and free CoB-SH through a radical mechanism catalyzed by the Ni-containing tetrapyrrole cofactor F_{430} [16, 17]. Methane

conversion to methyl-S-CoM was previously proved in the methanogen *Methanothermobacter marburgensis*[18] despite its unfavorable thermodynamics (Gibbs energy of +30 kJ.mol^{-1} of transformed methane)[4]. To perform this difficult process, ANMEs produce considerable quantities of MCR, which represents by far the most abundant cellular enzyme[15, 19, 20]. In addition, a thermodynamic pull allowing an immediate conversion of methyl-S-CoM should proceed, leading to subsequent transformation to CO_2.

MCR consists of a dimer of heterotrimers $(\alpha\beta\gamma)_2$ harboring the cofactor F_{430} in each active site[17]. The catalytic chamber and its close surroundings contain many post-translational modifications (PTMs) that vary in different methanogens and are carried on the α-subunit. No correlation has so far been observed between the PTMs and phylogeny. While the installation machinery of some of them has been reported, their physiological functions remain elusive[21-23]. Structures of ten methanogenic MCRs have yielded a large amount of information and allowed for comparisons, highlighting conserved overall features[16, 24]. In contrast, a unique structure has been reported for an MCR from ANME-1 directly purified from Black Sea mats[25]. Biochemical studies, corroborated with the structure refined to 2.1-Å, have revealed unique features, including a modified methylthio-F_{430} cofactor, a cysteine-rich patch, and two novel PTMs (*i.e.*, a C7-hydroxy-tryptophan and a methionine sulfoxide). Recently, a second short-chain alkane-capturing enzyme has been natively purified from a microbial consortium composed of the ethanotroph *Candidatus* Ethanoperedens thermophilum, belonging to a sister group of ANME-2d, and its sulfate reducing partner[20]. The Ethyl-Coenzyme-M Reductase (ECR), proposed to catalyze the first reaction in anaerobic ethane oxidation, exhibits singular features, including a dimethylated-F_{430} cofactor and a widened active site. Moreover, some extensions located on the surface, together with unique PTMs (N^2-methylhistidine and a 3-methylisoleucine), scaffold a gas channel connecting the active site to the solvent to favor ethane diffusion.

The unique traits in these short-chain alkane-activating enzymes hint towards the presence of other unexpected features in further ANME MCRs. Attempts have been made to facilitate structural studies of ANME MCRs, including successful recombinant expression of ANME-1 MCR and *Ca.* E. thermophilum ECR in *M. maripaludis*[26]. However, this approach is biased as the machinery for generating and installing native F_{430} variants, species specific PTMs, and dedicated chaperones required for correct folding are absent in the recombinant host. Moreover, artefactual chimeric assemblies have been observed between the native MCR from the methanogenic host and the recombinant ANME-2 MCR-γ. Similarly, the cultivation-independent approach of computational protein structure modelling exhibits biases as the process is guided by extracted data from known models but cannot predict the physiological cofactor, PTMs, and other unknown traits such as water networks and gas channels. Therefore, to explore the specific features of alkane-capturing enzymes, the investigation of native systems still prevails, which is a challenge in the case of non-isolated microbes such as ANMEs.

In this study, we bypass the isolation barrier by purifying MCRs from microbial enrichments dominated by ANME-2, sampled from different environments. The native structures obtained at (sub)atomic resolution depict the architecture of the methane-capturing enzyme, with the discovery of a novel PTM on the γ-subunit. These findings expand the knowledge about this crucial enzyme in these fascinating organisms, highlighting the importance of studying the natural microbial diversity.

Results

Isolation of the core catabolic enzyme from marine and freshwater ANME-2

Cultures containing ANME are rare due to the laborious enrichment procedures and the long division time of these organisms. Successive breakthroughs in the cultivation of ANME-2 species have allowed us to access marine and freshwater representatives to study differences among their MCRs through native protein purification (Figure 1). In all presented cases, the native protein isolation was based on MCR tracking by following its intrinsic properties: the absorption signature from the F_{430} cofactor and the typical pattern on denaturing gels (SDS-PAGE, Supplementary Fig. 1).

Investigations were started on a relatively simplified system in which ANME-2d grew independent of sulfate-reducing partners. Biological samples were retrieved from Ooijpolder (Netherlands)[27, 28] and Vercelli rice fields (Italy)[14, 29] and cultivated in bioreactors for years, significantly enriching ANME-2d on nitrate. Metagenomic analysis of the Ooijpolder starting samples showed a read-based relative abundance of 20-40 % *Ca.* Methanoperedens[30] with main strain *Ca.* Methanoperedens nitroreducens BLZ2[31] (further referred to as ANME-2d^{O}). In comparison, metagenomic analysis of the Vercelli rice field starting sample showed a read-based relative abundance of 15 % *Ca*. Methanoperedens', with main strain *Ca*. Methanoperedens Vercelli Strain 1[32] (further referred to as ANME-2d^{V}). Both samples were separately processed for protein purification, starting from 17 g and 12.5 g of ANME-2d^{O} and ANME-2d^{v} containing biomass and yielding 114 mg and 14 mg of the major MCR, respectively (Figure 1).

Furthermore, the purification setup was tested on a more complex microbial mixture from a marine system. Here, material was sampled at Amon Mud Volcano in the Mediterranean Sea at 1,120 m with a temperature of 14 °C. The microbial enrichment,

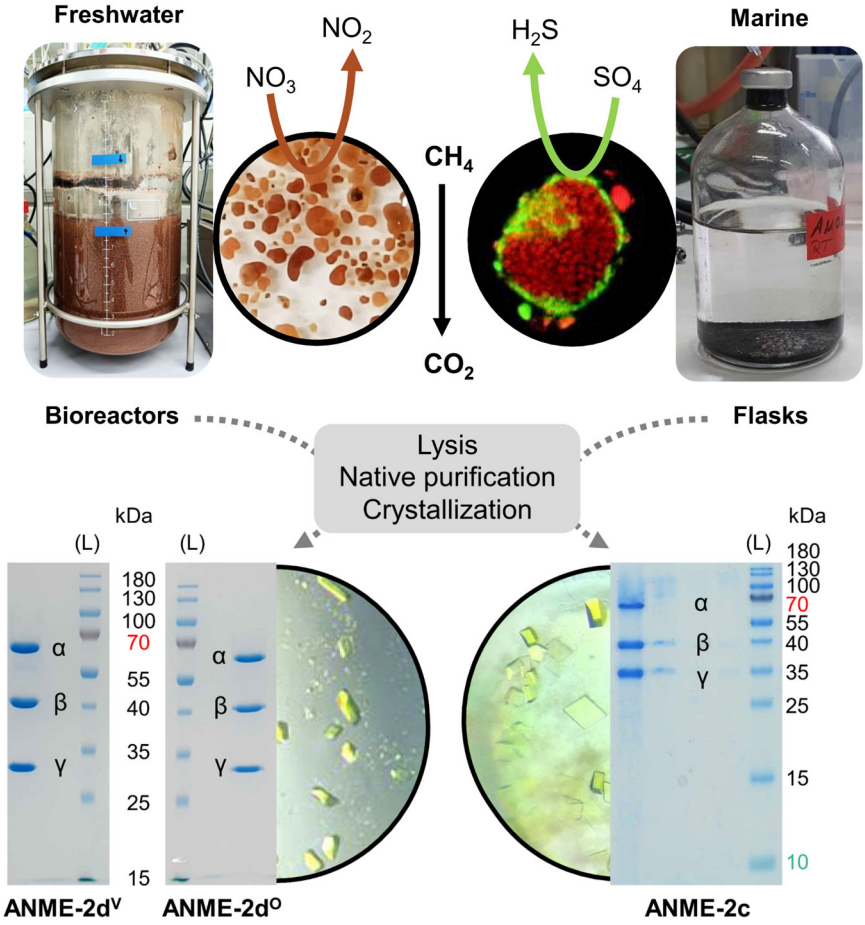


Figure 1. Comparison of the ANME-2 samples investigated in this study and procedure of protein isolation by native purification and crystallisation. Pictures of the bioreactor and flasks containing the respective freshwater and marine ANME-2 are presented. Close-ups of granular biomass of ANME-2d reducing NO_3 and fluorescence *in situ* hybridisation of the marine methanotrophic consortium (red for the archaeon and green for the sulfate-reducing partner) are shown. Biological samples were used to purify the major MCR fraction, as shown on SDS-PAGE at the bottom of the figure. For ANME-2c, the sample passed on SDS-PAGE corresponds to dissolved crystals, and a picture of the purification procedure can be examined in Supplementary Fig. S1. L stands for ladder.

grown at 20 °C in batches since 2003, gathered mostly ANME-2a/b and ANME-2c species based on 16S RNA sequencing[7]. Bacterial partners belonging to the Seep-SRB1a and Seep-SRB2 clades are present together with other microbes that might be important to sustain the consortium. The processed biomass of 3.21 g contained the microbial mixture as well as inorganic particles and a final quantity of 4.9 mg of the major MCR could be extracted. The success of purification was facilitated by the extremely high natural abundance of the enzyme, constituting roughly ~22 % of the soluble extract.

In all cases, the final purified MCR fractions presented a peak at 425 nm due to the F_{430} cofactor in accordance with previous reports for methanogenic MCRs (Supplementary Fig. 1b-c)[20, 33]. While the final SDS-PAGE suggests a high sample homogeneity in solution, we could not exclude a mixture of MCRs from different subspecies. Therefore, we used the ultimate selective crystallization process to isolate and identify a single variant by X-ray crystallography.

MCR identification and a well-conserved overall architecture

All MCR samples crystallized in a relatively short time (*e.g.*, less than a week), which was unexpected based on the plausible heterogeneity. X-ray diffraction data was collected, and models were refined to a final resolution of 0.98 and 0.94 Å (Table 1) for the MCR from ANME-2d^{O} and ANME-2d^{V}, respectively. Both MCRs crystallized in a monoclinic space group with very similar unit cell dimensions. Despite the differences in the organisms and the slight deviations in the crystallization solution (see methods), both MCR models shared the same crystalline packing (Supplementary Fig. S2). This suggests a very close sequence identity that would maintain the contacts on the surface to establish the same crystal lattice. This was confirmed after sequencing the MCR from the initial electron density following the molecular replacement, presenting an average deviation of only 3-5 % in sequence identity (*i.e.*, with an identity of 95.4, 97.2, and

95.6 % for α, β, and γ, respectively), resulting in an almost identical structural superposition (root mean square deviations (rmsd) of 0.09 Å with an overall of 2,186 aligned Cα).

The dataset obtained for the marine system was also of sufficiently high quality to unequivocally determine the sequence from the electron density map of the initial model, refined to a resolution of 1.34 Å. The derived sequence perfectly matched the one of ANME-2c MCR from the metagenomic data (Supplementary Fig. S3). Based on the crystal structure, MCR from ANME-2c seems to form stable dimers (Supplementary Fig. S2), as pointed out by the PISA server analyses (the dimer exhibiting a ΔG^{int} of −617.2 kcal.mol^{-1} and a ΔG^{diss} of 26.8 kcal.mol^{-1}, with analyses performed on the 25th of November 2024). The dimeric interface is exclusively formed by the interaction of the α-subunits presenting many interactions with the C-terminal extension. The dimeric state does not disturb the access to the active site. Due to the low quantities of MCR obtained at the end of the purification and the slow growth of the consortium, we could not test if this dimeric arrangement reflects a crystallization artefact or a stable entity in solution, and more studies will have to be carried out. On the other hand, the predicted trimeric arrangement of the MCR from ANME-2d^{O} observed in blue native PAGE[31] was not detected in the crystal structure. Nevertheless, we cannot exclude that the purification and crystallization process might have disrupted a trimeric organization.

Despite the environmental gaps, the marine and freshwater MCRs share relatively high sequence identity (*i.e.*, with an identity of 63.6, 61.5, and 71.4 % for α, β, and γ, respectively, based on MCR from ANME-2d^{V}, Supplementary Fig. S4). The three MCRs from ANME-2 share a matching overall structure with previously structurally studied methanogenic MCRs as well as the methanotrophic ANME-1 MCR and ECR from *Ca.* E. thermophilum (Figure 2, Supplementary Fig. S4-5), the closest homolog being the MCR from *Methermicoccus* shengliensis[33]. Besides the overall similarity, a

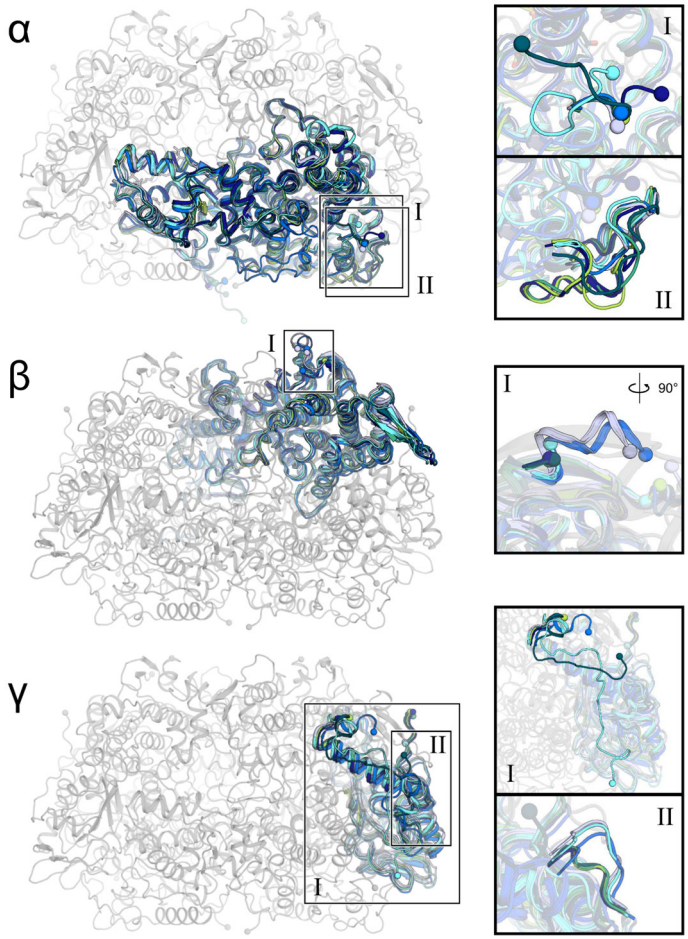


Figure 2. Comparison of ANME-2 MCR structures with previously structurally characterized MCRs. Superposition of the α-, β-, and γ-subunits of MCR. The panels on the right are close-ups of structural deviations. In each panel, the corresponding chain is coloured, while the remaining chains are shown (greyed out) for ANME-2d^{V}. MCR structures are coloured according to the microorganism, with *M. marburgensis* (PDB code: 5A0Y) in light blue, *M. shengliensis* (7NKG) in deep blue, ANME-1 from Black Sea mats (3SQG) in cyan, *Ca.* E. thermophilum ECR (71BS) in marine, ANME-2c in dark green, ANME-2d^{O} in green and ANME-2d^{V} in light green. N- and C-termini are shown as spheres and ligands as sticks.

few structural differences are notable and summarized in Figure 2. The electrostatic profile of the surface is also similar to previously studied MCRs, with both ANME-2d members exhibiting more pronounced positive charges compared to a typical methanogenic MCR (Supplementary Fig. S6) at the entrance leading to the active site.

A sealed, strictly conserved catalytic chamber covered by seven post-translational modifications

The active site of the three ANME-2 MCRs is highly conserved compared to the described MCRs from methanogenic archaea and contains a canonical F_{430} cofactor in accordance with previous studies[19] (Fig. 3, Supplementary Fig. S7). The subatomic resolution of both ANME-2d structures allows the visualization of the majority of hydrogen atoms and depicts the accurate position of each atom of the ligands. Both MCRs from ANME-2c and ANME-2d^{O} contain reduced coenzymes trapped in the active site. However, MCR from ANME-2d^{V} exhibits a mixture of reduced and oxidized coenzymes. The ligand coordinating residues are almost perfectly conserved. The only substitutions are Valine to Isoleucine (position 153/156/159 in ANME-2c, ANME-2d^{O}, ANME-2d^{V} respectively) and Glutamine to Methionine (position 343/346 in ANME-2d^{O} and ANME-2d^{V}, respectively). The former

has no effect on F_{430} coordination since the sidechain does not interact with the ligand, and the latter does abolish a stabilizing hydrogen bond, which seems to have little effect on the F_{430} position (Fig. 3a-b, Supplementary Fig. S8).

The discovery of a gas channel in the ethane-capturing enzyme motivated us to explore whether ANME MCR has a similar accession system for methane. Analysis of the internal cavities in all three models revealed tunnels protruding from the active site with small diameters comparable to the methanogenic counterparts (Supplementary Fig. S9). Their positions and diameters differ from those of the ECR from *Ca.* E. thermophilum. Therefore, we used Krypton instead of Xenon to mimic the sterical hindrance of

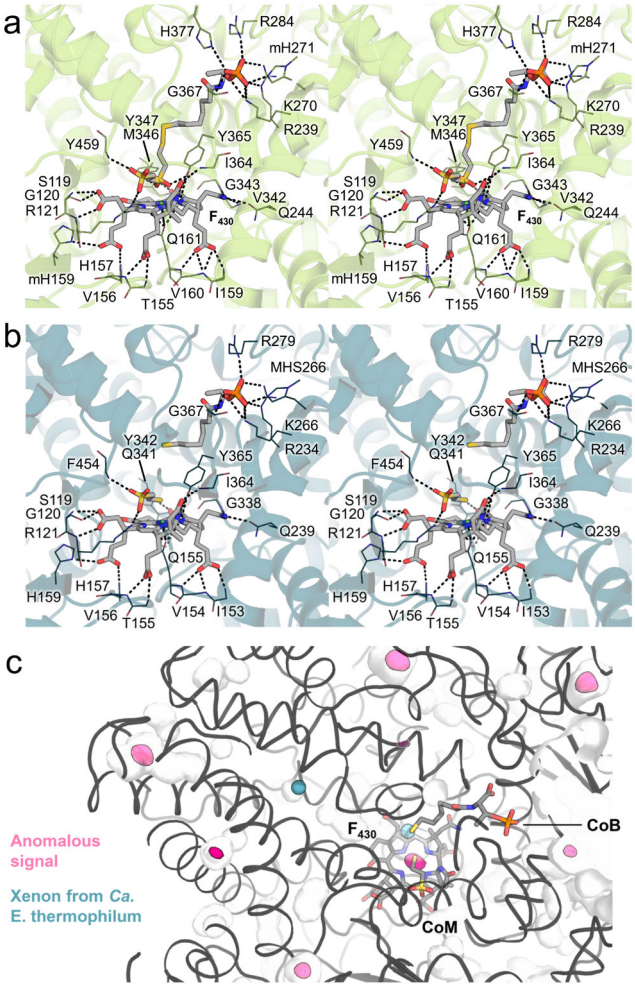


Figure 3. ANME-2 MCR active sites. Stereo-view of the MCR active site in (**a**) ANME-2d^{v} and (**b**) ANME-2c. Coordinating residues are shown as lines, and non-interacting side- or main chains were omitted. Polar contacts are depicted as black dashes. (**c**) Anomalous map for Krypton position shown as pink surface and contoured at 5 σ. Xenon position from *Ca*. E. thermophilum are teal spheres. Proteins are in cartoons, and ligands are shown as sticks coloured according to atom (red for O, blue for N, orange for P, yellow for S).

methane and performed high-pressure gassing for a few minutes on MCR ANME-2c, which presented one of the largest protrusions. Based on the anomalous signal of the dataset collected at the Krypton K-edge, no signals were observed in the active site or tracking in a hydrophobic tunnel (Fig. 3c, Supplementary Fig. S10). Krypton atoms are rather positioned on the outer shell of the protein, trapped in hydrophobic pockets, as previously observed for the methanogenic enzyme from *M. marburgensis*[20, 34]. The absence of a dedicated gas tunnel reinforces the proposed scenario that only ethane-specialized enzymes evolved this strategy through insertions and dedicated PTMs, and that methane should enter before the heterodisulfide in the catalytic chamber[16].

A surprising number of seven PTMs were observed in the electron density from the three MCRs (Fig. 4, Supplementary Fig. S11 and S12), which was confirmed via mass spectrometry. ANME-2c MCR includes the conserved N^1-methylhistidine and thioglycine, as well as the common 5(*S*)-methylarginine, 2(*S*)-methylglutamine, and the rarer S-methylcysteine, didehydro-aspartate and 6-hydroxytryptophane. The same PTMS are present in both ANME-2d MCRs, with the exception of methyl-glutamine. Instead, the histidine 159 from the γ chain harbors a methyl group on its β-carbon. The histidine is highly conserved and positioned at a distance of 3.5Å to the F_{430} cofactor, with atom ND1 forming a hydrogen bond to F_{430}. However, the modified methyl group is distanced at ~5.7 Å and directed away from the cofactor without disturbing the position of surrounding residues (Supplementary Fig. S13).

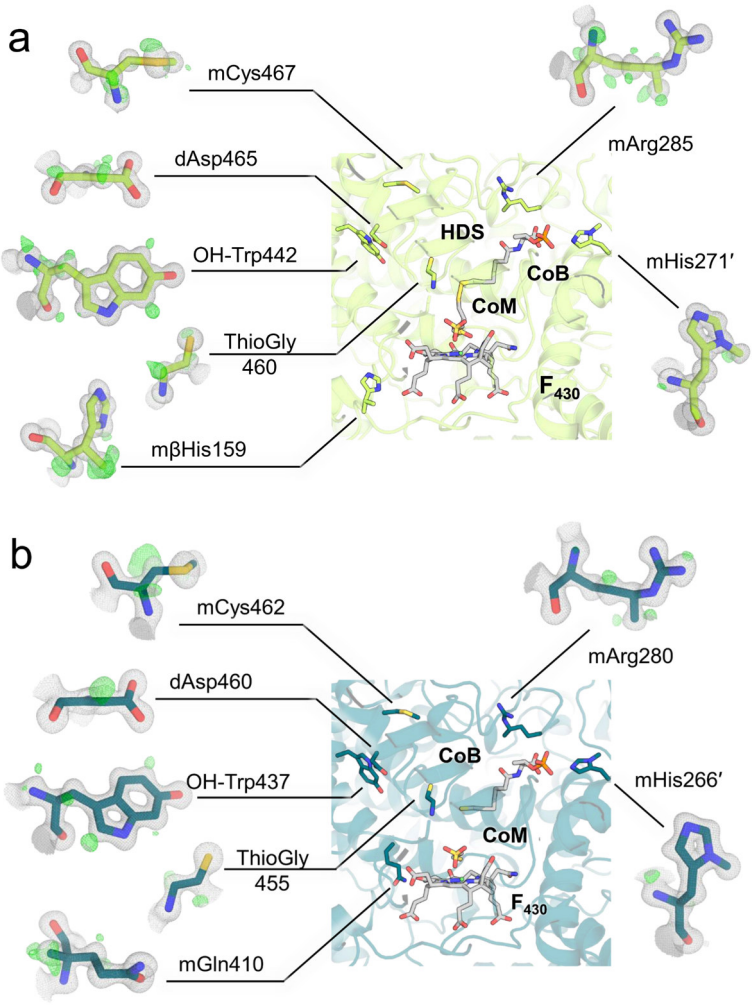


Figure 4. ANME-2 MCR post-translational modifications. PTMs in the active site of (**a**) ANME-2d^{V} and (**b**) ANME-2c. The main chain is shown as a cartoon, with ligands and PTMs shown as sticks coloured according to atom (red for O, blue for N, orange for P, yellow for S). Close-ups of PTMs flank the central picture, highlighting electron densities. The $2F_o-F_c$ map (grey mesh) is contoured at 1 σ, and the F_o-F_c map (green mesh) is contoured at 3 σ. The latter highlights the experimental hydrogen positions.

Table 1: X-ray analysis statistics for ANME-2 MCR

	ANME-2c from microbial enrichment	ANME-2c (Kr pressured)	ANME-2d^{O} from bioreactor	ANME-2d^{V} from bioreactor
Data collection				
Synchrotron and beamline	SOLEIL, PROXIMA-1	ESRF, ID23-1	ESRF, BM07-FIP2	ESRF, BM07-FIP2
Wavelength (Å)	0.97856	0.86100	0.97980	0.97930
Space group	$P2_1$	$P2_12_12_1$	$P2_1$	$P2_1$
Resolution (Å)	127.16 – 1.34 (1.50 – 1.34)	59.40 – 1.80 (1.83 – 1.80)	76.67 – 0.98 (1.04 − 0.98)	96.94 – 0.92 (1.04 − 0.92)
Cell dimensions				
a, b, c (Å)	156.87, 157.46, 215.42	153.48, 153.67, 212.88	81.58, 189.30, 84.11	81.64, 189.89, 83.96
α, β, γ (°)	90, 90.34, 90	90, 90, 90	90, 114.27, 90	90, 114.36, 90
R_{merge} (%)[a]	16.8 (167.1)	12.7 (162.3)	11.9 (116.8)	6.8 (96.2)
R_{pim} (%)[a]	7.3 (73.7)	5.1 (64.6)	4.5 (46.8)	2.6 (38.1)
$CC_{1/2}$ [a]	0.991 (0.560)	0.999 (0.768)	0.996 (0.667)	0.998 (0.684)
I/σ_I[a]	6.6 (1.6)	14.1 (1.8)	9.4 (1.7)	13.0 (1.7)
Spherical completeness[a]	66.2 (12.2)	/	84.4 (24.4)	62.6 (10.4)
Ellipsoidal completeness[a]	96.2 (66.6)	100.0 (100.0)	96.5 (70.9)	84.7 (58.5)
Redundancy[a]	6.3 (6.0)	13.8 (14.2)	7.8 (7.1)	7.6 (7.2)
Nr. unique reflections[a]	1,527,343 (76,367)	461,593 (22,707)	1,130,261 (56,513)	988,467 (49,423)
Refinement				
Resolution (Å)	98.93 – 1.34	/	41.78 – 0.98	41.72 – 0.94
Number of reflections	1,527,209	/	1,130,120	986,474
R_{work}/R_{free}[b] (%)	11.30/14.81	/	10.32/11.90	11.21/13.06
Number of atoms				
Protein	152,083	/	38,640	38,408
Ligands/ions	912	/	283	289
Solvent	9452	/	3203	3,075
Mean B-value (Å^2)	15.86	/	10.40	11.95
Molprobity clash score	1.28	/	1.42	1.18
Ramachandran plot				
Favoured regions (%)	97.31	/	97.16	97.08
Outlier regions (%)	0.07	/	0.08	0.12
rmsd[c] bond lengths (Å)	0.010	/	0.010	0.010
rmsd[c] bond angles (°)	1.235	/	1.298	1.321

[a] Values relative to the highest resolution shell are within parentheses. [b] Rfree was calculated as the Rwork for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

Discussion

Due to their deep implications in the anaerobic biosphere, methanotrophic archaea majorly impact the global carbon cycle. The efficiency of their catabolism could be employed in microbial-driven technology, as illustrated by ANME-2d producing an electrical current from methane oxidation when grown on electrodes[29]. Despite their significance in methane emissions, little is known about their metabolic capabilities as isolation attempts remain unsuccessful. By directly exploiting microbial consortia, we isolated and structurally characterized the catabolic enzymes from ANME representatives, expanding the available pool of methanotrophic MCR structural models from one to four. The enzyme, isolated from only a few grams of material, obtained from a decade-cultivated marine microbial enrichment, was the most challenging. This tour de force was possible due to the extreme abundance of MCR in the biomass, representing almost a fifth of the total soluble extract. Such a high value is in accordance with previous reports in which MCRs constituted 7 % of proteins extracted from Black Sea mats[15] and 10.4 % of peptides in proteomic datasets of seep sediments[35], suggesting that the amount in enriched or isolated cells would be even higher. The excessive intracellular MCR concentration might be necessary to counteract the endergonic methane capture step in AOM[17].

The abundance of MCR is not conserved among anaerobic alkanotrophs, as our previous work reports a lower cellular amount of the ECR in *Ca*. E. thermophilum compared to the microbial enrichment of ANMEs on native PAGE. This is likely due to the thermodynamics of the reaction or the lower activation energy required for the ethane reactivity. While the ethanotroph is a sister group of the ANME-2d clade, its ECR enzyme exhibits consequent deviations compared to the MCR structures obtained in this work that reflect a specialization towards C2 compounds. The larger active site, a dimethylated F_{430}, gas channels supported by extensions and specific PTMs are not conserved in MCRs from ANME-2d/c. In contrast, the C1-specialized enzymes of

ANME-2 are rather highly similar to their methanogenic counterparts with the same active site characteristics and F_{430} cofactor[15]. This is in accordance with the accepted concept that thermodynamics determine the reaction directionality of methane generation/capture as opposed to structural differences between methanogenic and methanotrophic MCRs[26]. It also supports the proposal that methane capture proceeds via the reverse proposed mechanism of methanogenesis, as previously proposed[16].

While assuming the same physiological function, structural differences exist between the MCR structures from ANME-2 from this work and ANME-1 isolated from Black Sea mats, which contains a methylthio-F_{430} [25], a cysteine-rich patch, and two particular PTMs that should compensate for the loss of the 5(*S*)-methylarginine. The differences are more likely linked to taxonomy than to adaptation to ecological niches, as illustrated by the MCRs from freshwater and marine ANME-2d/c.

Together with ECR, the three ANME-2d/c MCRs are the most modified known so far. The seven PTMs are striking compared to the closest structural relative *M. shengliensis* harboring a reduced gallery of only three PTMs. When considering MCR as the rate-limiting step of catabolism and how the reaction directionality differs among C1-metabolisms, then a comparison between ANME-2d/c and *M. shengliensis* would suggest that PTM numbers might be correlated to the overall energy yield of methanogenesis. In other words, due to the endergonic reaction of methane capture, ANMEs gain little energy from methane oxidation in their natural environment where methane might fluctuate to borderline thresholds. On the other hand, *M. shengliensis* could rely on the energy efficiency of the methoxydotrophic pathway and the abundance of natural substrates in their natural environment (oil fields)[36]. If, as postulated by other groups[22, 23], PTMs stabilize the active site and ligands, it might be beneficial for the microbe to invest in equipping its core catabolic machinery to enhance, even slightly, its catalytic properties. Further studies of other ANME MCRs will be necessary to

determine whether an increased PTM amount could be related to the anaerobic oxidation of methane/alkanes or if the presented structures are isolated cases. On the same line, future studies could investigate how common the 3(S)-methylhistidine is and if it exists in other *Methanosarcinales*. To our knowledge, this is the first report of a 3(*S*)-methylhistidine in a natural protein and the first PTM identified on the γ-chain of any characterized MCR. While the methylation occurs on a residue coordinating the F_{430} cofactor, it appears to have no immediate effect on the surrounding active site, raising questions about its role, falling in line with the so far mostly unresolved functions of the majority of PTMs observed in MCRs[16, 24]. However, since all crystallographic structures are mere snapshots of a single state, we cannot exclude a role during catalysis or enzyme assembly.

Being located on a loop at the surface of the γ-subunit, the 3(*S*)-methylhistidine could be installed before the oligomeric assembly by a dedicated methyl-transferase. Due to the low nucleophilicity of the β-carbon, the reaction should occur similarly to the installation of a methylarginine (installed by the Mmp10 methyltransferase[23]) or methylglutamine rather than the N^1-methylhistidine in which an S-adenosylmethionine dependent methyltransferase would substitute a hydrogen of the imidazole ring nitrogen with a methyl group. ANME-2d does possess an Mmp10 homolog upstream of the MCR operon that could potentially fill the role.

The sub-atomic resolution datasets, detailing the active site, cofactor and coenzymes to unprecedented levels, represent an important step for future studies to understand the molecular mechanisms of methane activation. For instance, assigning protonation states of active site residues accessible by Neutron diffraction would be an important asset. Moreover, with its high yield/high purity, rapid micro-crystal formation and excellent diffraction quality, MCR ANME-$2d^{O/V}$ are well suited for time-resolved crystallography. By delivering these accurate models, together with the discovery of

new PTMs in environmentally and biotechnologically relevant anaerobic methanotrophs, this work reinforces the importance of native studies on non-isolated microbes. The unbiased experimental approach, in combination with omics analyses and AI-driven protein modelling could counterbalance possible biases in the future to pave the way for the biochemical investigation of the microbial dark matter in methanogenic and methanotrophic MCRs.

Material and Methods

Biomass acquisition and microbial enrichments.

ANME-2c cultivation: Sediment was sampled by multicore at Amon Mud Volcano, Nile deep-sea fan, Mediterranean Sea at 032° 21.70 N, 031° 23.40 E at a water depth of 1,120 m and 14 °C with the research vessel RV L'Atalante NAUTINIL cruise in 2003. On board samples were transferred into gas-tight bottles. In the home lab, sediment slurries were set up with a synthetic marine sulfate reducer medium at pH 7.4 (Widdel and Bak, 1992[37]) and methane headspace at 20 °C. Incubated slurries showed methane-dependent sulfate reduction within a few days. To remove sediment debris, cultures were continuously incubated with methane as the sole electron donor and sulfate as the electron acceptor. The activity-based doubling time for both cultures, monitored by sulfide measurements, was about 4-6 months. Consecutive dilution and further incubation resulted in debris-free cultures of detached flocks essentially composed of AOM-performing methane-oxidizing archaea and sulfate-reducing bacteria consortia. For more detailed cultivation methods, see Laso-Pérez *et al*., 2018[38].

ANME-2d cultivation, bioreactor cultivation from Ooijpolder sample: Granular biomass was obtained from a highly enriched N-DAMO bioreactor culture[31]. The culture was maintained in an anaerobic 15 L sequencing-fed batch reactor (30 °C, pH 7.3 ± 0.1 controlled with $KHCO_3$, stirred at 200 rpm, with a working volume between 8 and 11 L. The reactor was continuously sparged with 15 mL.min^{-1} CH_4/CO_2 (95:5) and fed with medium (flow rate 2-2.5 L.day^{-1}) - as described in Kurth *et al.* 2019[28] - containing nitrate (15-25 mM) with an adaptation of the KH_2PO_4 concentration to 0.05 g.L^{-1}. Once a day, the biomass settled for 5 min, after which the excess liquid in the reactor was pumped off to a volume of 8 L. Metagenome data from the time of sampling showed a read-based relative abundance of 20 – 40 % '*Ca*. Methanoperedens'[30], with

main strain '*Ca.* Methanoperedens nitroreducens BLZ2'[31] and nearly 50 different bacterial species[28].

ANME-2d cultivation, bioreactor cultivation for Vercelli rice fields sample: Granular biomass was obtained from a freshwater N-DAMO bioreactor enrichment culture first described by Vaksmaa *et al.* 2017[14]. The culture was maintained in an anaerobic 15 L sequencing-fed batch reactor (room temperature, pH 7.3 ± 0.1 controlled with $KHCO_3$, stirred at 200 rpm) with a working volume between 8 and 11 L. The reactor was continuously sparged with 10 mL.min^{-1} CH_4/CO_2 (95:5) and fed with medium (flow rate 1.5-2 L.day^{-1}) containing sodium nitrate (5-8 mM). Once a day, the biomass settled for 5 min, after which the excess liquid in the reactor was pumped off to a volume of 8 L.

The medium contained 0.07 g.L^{-1} $MgSO_4.7H_2O$, 0.1 g.L^{-1} $CaCl_2.2H_2O$, 0.05 g.L^{-1} K_2HPO_4, 0.5 mL.L^{-1} of trace element solution, 0.3 mL iron solution (54 g.L^{-1} $FeCl_3.6H_2O$ and 114.6 g.L^{-1} nitrilotriacetic acid), and 0.1 mL.L^{-1} of Wolin's vitamin solution (DSMZ 141). The trace element solution contained per liter: 0.12 g $CeCl_3.7H_2O$, 0.6 g $CoCl_2.6H_2O$, 4.0 g $CuSO_4$, 0.07 g H_3BO_3, 1.0 g $MnCl_2.4H_2O$, 0.35 g $Na_3MoO_4.2H_2O$, 0.1 g Na2$WO_4.2H_2O$, 0.95 g $NiCl_2.6H_2O$, 0.14 g SeO_2, 1.44 g $ZnSO_4.7H_2O$.

Metagenome sequencing

ANME-2c: The different MCR sequences were obtained from metagenome information processed by Dr. Laso-Pérez. The sequences were retrieved by TIGRFAM motif search in combination with *ab initio* sequencing from the electron density.

ANME-2d: MCR sequences were retrieved from recent metagenomes used in Wissink *et al.*, 2024[30] and Vaksmaa *et al.* 2017[14].

Sample treatment and lysis

ANME-2c: Two anaerobic 150 mL AOM enrichment cultures were used as starting material. Particles were let to settle overnight at RT, followed by the removal of ~130 mL of clear supernatant. H_2S in the culture headspace was replaced with N_2 gas before transfer to an anaerobic chamber. Cell lysis and preparation of extracts were performed in a N_2/CO_2 atmosphere (90:10 %) at 31 °C. Both samples were pooled and centrifuged at 8,000 × *g* for 5 min. The supernatant was carefully removed, resulting in a pellet of ~ 3.21 g (wet weight). The aggregates were resuspended in 30 mL IEC-A buffer (50 mM Tricine pH 8, 2 mM dithiothreitol (DTT)) and lysed via two rounds of sonication interceded with a French Press treatment at 1,000 PSI (6.895 MPa). To ensure anaerobic conditions French press cell was flushed with N_2 gas before use. Unfortunately, 7.5 mL of the lysed sample was lost in the airlock due to an accidental break of the serum flask. The leftover total extract was then diluted to 60 mL with IEC-A followed by centrifugation at 45,000 × *g* for 30 min at 18 °C. 100 µL was withdrawn at each for SDS-PAGE analysis (Supplementary Fig. S1).

ANME-2d^{O}: Biomass highly enriched in *Ca.* Methanoperedens BLZ2 was cultivated in a 10 L batch reactor. The biomass (7.75 g of cells) was transferred in a tent with an N_2/CO_2 atmosphere (90:10 %) and resuspended with 30 mL buffer IEC-A. Cells were first lysed via two rounds of sonication (SONOPULSE Brandelin, 75 % power, 30 sec, 1st round: 4 cycles, 2nd round: 3 cycles) interceded with a French Press treatment at 1,000 PSI (6.895 MPa). The French press cell was previously flushed with N_2 to ensure anaerobic conditions. After the second round of sonication, the sample was centrifuged at 45,000 × *g* for 30 min at 10 °C the supernatant was collected and subjected to ultracentrifugation at 100,000 × *g* for 1.5 hours at 10 °C (rotor used: 70Ti, Beckman Coulter). 100 µL of the total cell extract and soluble fraction was withdrawn for the Bradford assay and SDS-PAGE.

ANME-2d^{V}: 12.47 g of biomass was resuspended to 60 mL buffer IEC-A. Cells were processed in the same way as ANME-2d^{O} except for the following variation. The supernatant from ultracentrifugation was sonicated for 4 extra cycles, and 40 mL of filtered supernatant was obtained before injection on the HiTrap Q HP column.

MCR purification

ANME-2c: Protein purification was carried out under anaerobic conditions in a Coy tent with an N_2/H_2 atmosphere (97:3 %) at 20 °C under yellow light. The sample was filtered using a 0.2 µm nitrocellulose filter (Sartorius, Germany) before every injection on a chromatography column. Proteins were tracked via multi-wavelength absorbance monitoring (λ 280, 415 and 550 nm) in combination with SDS-PAGE. The 60 mL of filtered soluble extract (20.7 mg protein) was loaded on a 1 mL HiTrap^{TM} Q HP column (Cytiva, Sweden) equilibrated with IEC-A. The proteins were eluted at steps of 50 % and 100 % IEC-B (50 mM Tricine pH 8, 1 M NaCl) at 0.7 mL.min^{-1}. MCR eluted at 0.5 M NaCl. The fractions of interest were pooled (~13 mL) and diluted ~1:4 with IEC-A buffer and filtered. The sample was loaded on a 0.982 mL Mono Q^{TM} 5/50 GL column (Cytiva, Sweden). The elution was performed at a gradient of 10-70 % IEC-B at 0.5 mL.min^{-1}. The pooled MCR fractions were diluted ~1:4 (protein sample:buffer) with HIC-B buffer (25 mM Tricine pH 7.6, 2 M $(NH_4)_2SO_4$). The pool was filtered and passed on a 1.622 mL Source15ϕ 4.6/100 PE column (GE healthcare). Proteins were eluted by applying a gradient of 75-0 % HIC-B (against buffer without $(NH_4)_2SO_4$) at 0.6 mL.min^{-1}. The final MCR pool was washed by dilution (1:1000) with storage buffer (25 mM Tris pH 7.6, 10 % v/v glycerol) using a 30 kDa cut-off concentrator to remove salts. Finally, the sample was concentrated to 200 µL. The protein concentration estimated via the Bradford method was 24.6 mg.mL^{-1}. Purification with three chromatography steps yielded a pure MCR fraction displaying the typical three-subunit

pattern on SDS-PAGE with molecular weights of ~65, 39, and 34 kDa (Supplementary Fig. S1).

ANME-2d^{O}: Purification was performed under the previously described conditions, and fractions of interest were followed at λ 280. The soluble extract (90 mL at 7.82 mg.mL^{-1}) could not be filtered due to the large particulate size and was directly injected into the 10 mL HiTrap^{TM} Q HP column (Cytiva, Sweden) pre-equilibrated with IEC-A buffer. After washing the column, the elution was performed with a gradient of 0 to 30 % IEC-B for 10 column volumes (CV) at a 2 mL.min^{-1} flow rate. The pool of interest was diluted 1:1 with HIC-B buffer, filtered, and loaded on the 1.622 mL Source15ϕ 4.6/100 PE column (GE healthcare). A gradient from 80 % to 0 % $(NH_4)_2SO_4$ for 10 CV at a 1.5 mL.min^{-1} flow rate was applied. The final MCR fractions were pooled and concentrated on a Vivaspin centrifugation concentrator with a cut-off of 30 kDa. The buffer was exchanged by ultrafiltration for the storage buffer containing 100 mM NaCl to prevent the spontaneous formation of microcrystals. Around 2 mL of purified MCR were recovered at 57 mg.mL^{-1}, aliquoted into 50 µL, and stored at 4 and −80 °C.

ANME-2d^{V}: Purification was performed under the previously described conditions, and fractions of interest were followed at λ 280. The filtered soluble extract (40 mL) was injected into the 10 mL HiTrap^{TM} Q HP column (Cytiva, Sweden) pre-equilibrated with IEC-A buffer. After washing the column, the elution was performed with a gradient of 0 to 65 % IEC-B for 16 column volumes (CV) at a 2 mL.min^{-1} flow rate. The pool of interest was diluted 1:2 (protein sample:buffer) with HIC-B buffer, filtered, and loaded on a 5 ml HiTrap^{TM} Phenyl HP Sepharose column (Cytiva, Sweden). A gradient from 65 % to 0 % $(NH_4)_2SO_4$ for 12 CV at a 1.0 mL.min^{-1} flow rate was applied. The final MCR fractions were pooled and concentrated on a Vivaspin centrifugation concentrator with a cut-off of 30 kDa. The buffer was exchanged by ultrafiltration for the storage

buffer. Around 0.35 mL of purified MCR were recovered at 40 mg.mL^{-1} and stored at 4 and −80 °C.

UV-visible spectrophotometry

UV-Vis absorption spectra were recorded with a Cary 60 (Agilent Technologies) spectrophotometer on a TrayCell with a 1 mm path.

Phylogeny analyses

The trees presented in Supplementary Fig. S3 have been performed by the BLAST from the NCBI with the Distance tree of results tool. The tree method was by neighbour joining with all parameters set by default. The selected sequences were the three first closest sequence homologues of MCR ANME-2c obtained by blasting each subunit sequence on NCBI BLAST, the two sequences from MCR ANME-2d characterized in this study, and all structural homologues.

Crystallization

ANME-2c: The fresh preparation of MCR was immediately crystallized. Crystals were obtained using the sitting drop method on 96-Well MRC 2-Drop Crystallization Plates in polystyrene (SWISSCI, Switzerland), containing 90 µL of crystallization solution in the reservoir: 0.5 µL protein sample with 0.5 µL crystallization solution were spotted. The crystallization was performed under anaerobic conditions via an Oryx Nano robot at 18 °C in a tent filled with 100 % N_2. And then plates were transferred in a Coy tent (N_2/H_2 atmosphere 97:3 %) at 20 °C. Crystals were formed following the first 24 h.

The most successful condition (30 % v/v 2-Methyl-2,4-pentanediol, 100 mM Tris pH 8, 500 mM NaCl and 8 % w/v Polyethylene glycol 8,000) was replicated on a Junior Clover plate (Jena Bioscience, Germany) under aerobic conditions with 90 µL of the crystallization solution in the reservoir and 2 µL protein sample with 2 µL

crystallization solution in the well. Crystals with yellow square or rectangular morphology appeared over the next two weeks.

ANME-2d^{O}: Following an initial screening 96-Well MRC 2-Drop Crystallization Plates in polystyrene done aerobically at 18 °C, MCR crystals were reproduced using the sitting drop method on the Junior Clover plate. The best crystals were obtained with a reservoir filled with 100 µL of the following crystallization condition: 20 % w/v polyethylene glycol 3,350, 50 mM Tris pH 8.0 and 200 mM potassium nitrate and by spotting 5 µL protein at a concentration of 2.19 mg.mL^{-1} with 2 µL crystallization solution. Thick yellow brick-shaped crystals appeared within 24-48 hours.

ANME-2d^{V}: Following an initial screening 96-Well MRC 2-Drop Crystallization Plates in polystyrene done aerobically at 18 °C, MCR crystals were reproduced using the sitting drop method on the Junior Clover plate. Here, the reservoir was filled with 100 µl of 20 % w/v Polyethylene glycol 3,350 and 200 mM sodium thiocyanate pH 6.9 and the crystallization well contained 3 µL protein at a concentration of 40 mg.mL^{-1} with 2 µL mother liquor.

Mass spectrometry analysis

ANME-2c MCR in solution, on SDS PAGE, and crystallized were used for Mass spectrometry analyses. The best results were obtained with trypsin treatment of the MCR in solution followed by chymotrypsin in combination with vinylpyridine labeling. ANME-2d samples were analyzed from the protein in solution. The whole procedure is described in Hahn *et al*. 2021[20].

Kr pressurization

ANME-2c MCR in solution was shipped to the High-Pressure Freezing Laboratory for Macromolecular Crystallography (HPMX)[39] of the European Synchrotron Radiation Facility (ESRF). The best crystallization hit was repeated, and the crystals obtained

were used for the Krypton-pressurization experiment. Here, crystals were pressurized at 214 bars Krypton for 2 min before flash freeze in liquid nitrogen. X-ray crystallographic data were collected on ID23-1 at 14.4 KeV, close to the Krypton fluorescence K-edge.

X-ray data collection, processing, modelling and validation.

MCR crystals were soaked in the crystallization solution supplemented with 20 % v/v glycerol for ANME-2d^{O} and 20 % v/v ethylene glycol for ANME-2d^{V} for 5-10 seconds before being transferred to liquid nitrogen.

X-ray diffraction for ANME-2c crystals was collected at SOLEIL (Optimized Light Source of Intermediate Energy to LURE) synchrotron, Saint-Aubin, France, at the beamline PROXIMA-1. Crystals for both ANME-2d were collected at the European Synchrotron Radiation Facility (ESRF), Grenoble, France, on the beamline BM07-FIP2 using the PILATUS 6M detector.

The datasets were processed and scaled with *autoPROC* (Version 1.0.5, Global Phasing Limited, Cambridge, UK)[40], except for the Krypton gassed sample that was integrated with XDS and scaled with aimless from the CCP4 package. The structures were solved by molecular replacement with *PHASER* from the *PHENIX* package (Version 1.19.2-4158)[41]. For molecular replacement, we used an AlphaFold 2[42] model based on the metagenome sequence for ANME-2c, then the ANME-2c model for ANME-2d^{O} and the Krypton-pressurized ANME-2c and finally, the ANME-2d^{O} model was used for molecular replacement for ANME-2d^{V}.

All models were then refined with *COOT* (Version 0.8.9.2)[43] and phenix.refine. Because of the high resolution, all atoms were considered anisotropic (except for the Krypton-pressured structure), and hydrogens in riding mode were added. Since only anomalous data were interesting for the Krypton-gassed crystal, a rigid body, followed by three

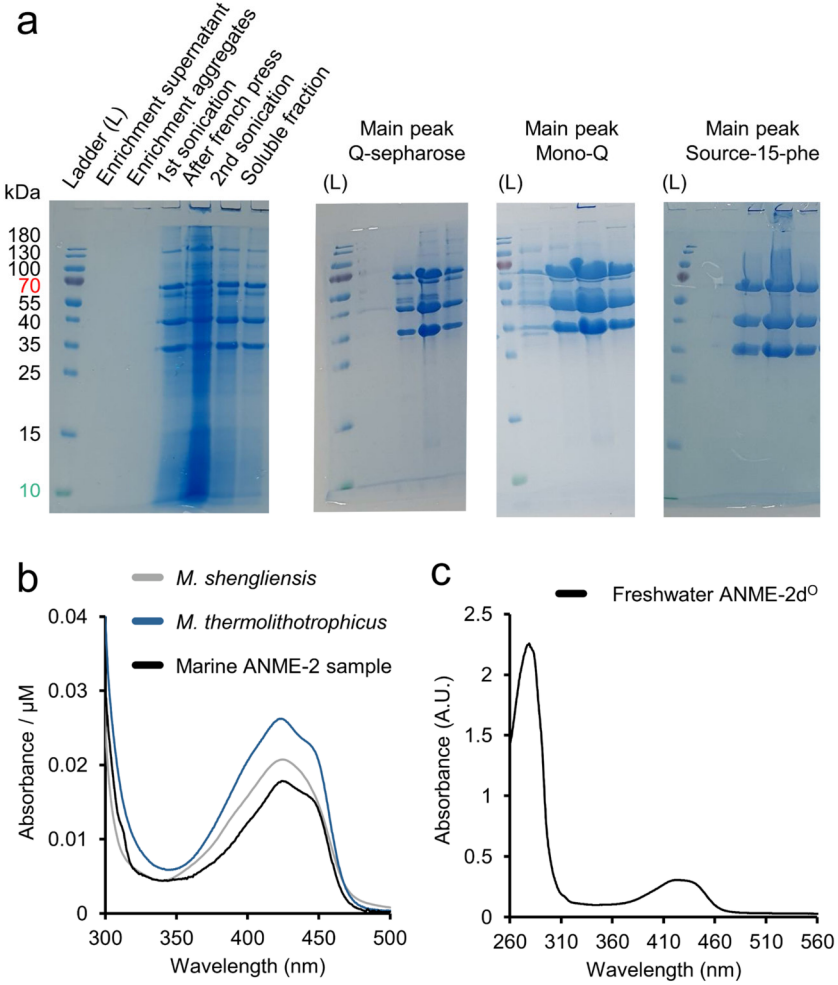
cycles of PHENIX refine, was performed to generate Fig. 4a and Supplementary Fig. S10b. The models were validated by the MolProbity server (http://molprobity.biochem.duke.edu). Figures of MCR structures were generated with PyMOL (V. 2.2.0, Schrödinger, LLC). Alignments were generated with MAFFT 7[44] and visualized with ESPript 3.0[45]. All models from Table 1 have only hydrogens modelled on the protein (*e.g.*, not on ligands and solvents).

Acknowledgements and author contributions

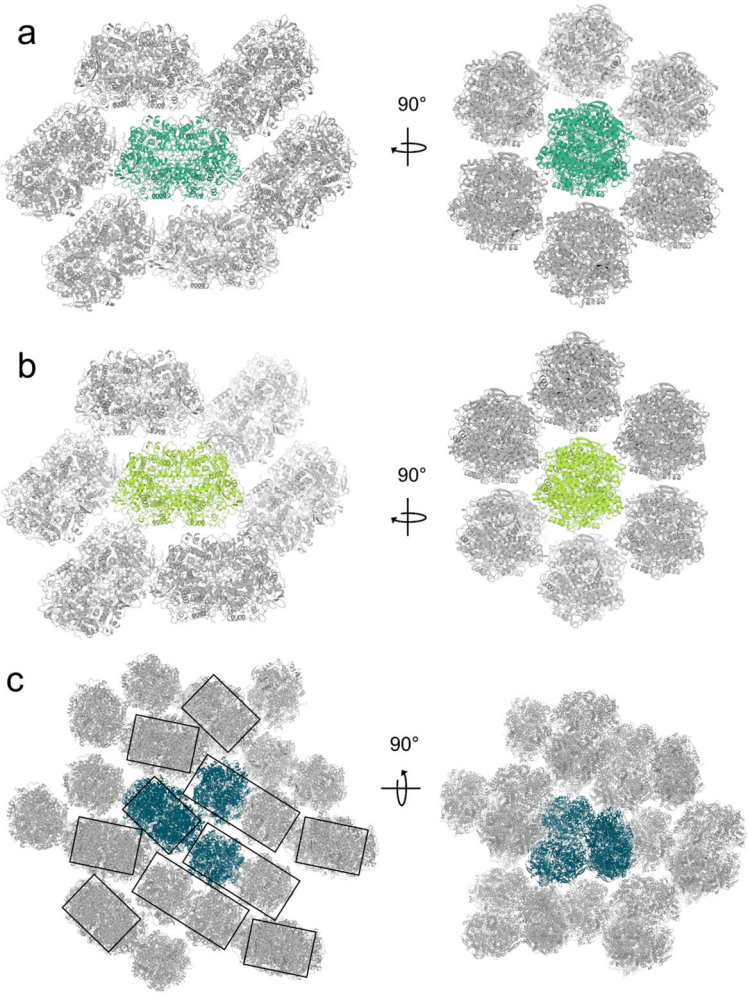
We are thankful to the Max Planck Institute for Marine Microbiology and the Max-Planck-Society for continuous support. We thank Christina Probian and Ramona Appel of the Microbial Metabolism laboratory for their continuous support. We acknowledge the SOLEIL and ESRF synchrotrons for beam time allocation and the beamline staff of Proxima-1, ID23-1 and BM07-FIP-2.

Enrichment of ANME-2c was done by G.W., enrichment of ANME-2d was done by P.M., M.W., C.W., treatment of metagenomic data for ANME-2c was done by R.L., purification and crystallization of ANME-2c MCR was done by M.-C.M. and N.v.P, purification and crystallization of ANME-2d^{O} MCR was done by P.M., purification and crystallization of ANME-2d^{V} was done by M.W. and T.W., data collection was done by S.E, T.W, structural refinement and analysis was done by M.-C.M., *in cristallo* spectra and Kr pressurization were done by S.E. and P.C., the manuscript was written by M.-C.M. and T.W.

Supplementary materials

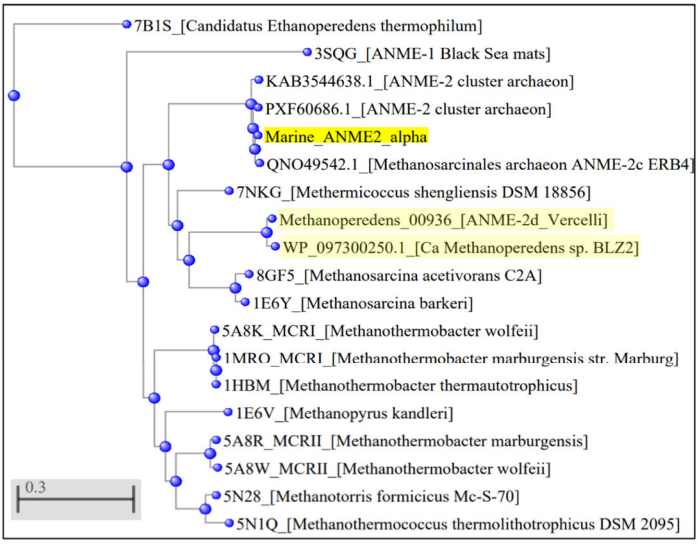


Supplementary figure S1. Purification and crystallization of ANME-2c MCR. (**a**) SDS-PAGE of different purification steps. L stands for Ladder. (**b**) UV-Vis spectrum of ANME-2c MCR compared to methanogenic homologs from *Methermicoccus shengliensis*[33] and *Methanothermococcus thermolithotrophicus*[20]. The measured signal is in absorbance (AU) per µM of measured MCR. (**c**) UV-Vis spectrum of ANME-2d^{O} MCR.

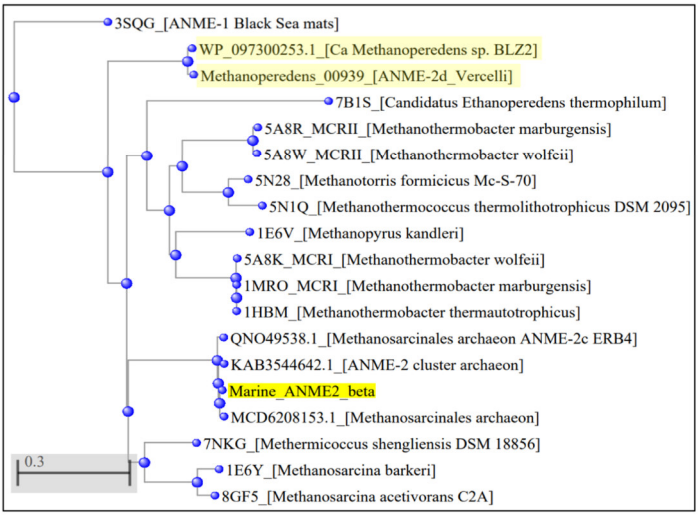


Supplementary figure S2. Presentation of ANME-2 MCR crystalline packing. Packing of MCR from ANME-2d^{O} (**a**), ANME-2d^{V} (**b**), and ANME-2c (**c**). MCR units are shown in cartoons. The asymmetric unit content is coloured, while the symmetry mates are shown in grey. Black boxes indicate the formation of dimeric units for ANME-2c MCR.

α-subunit

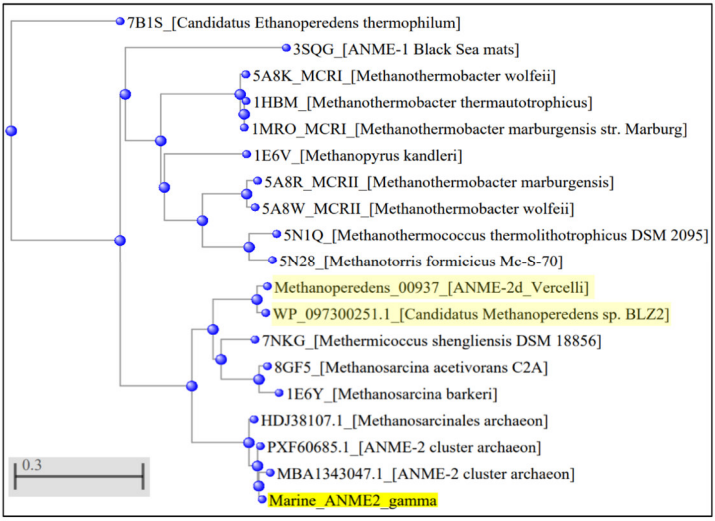


β-subunit

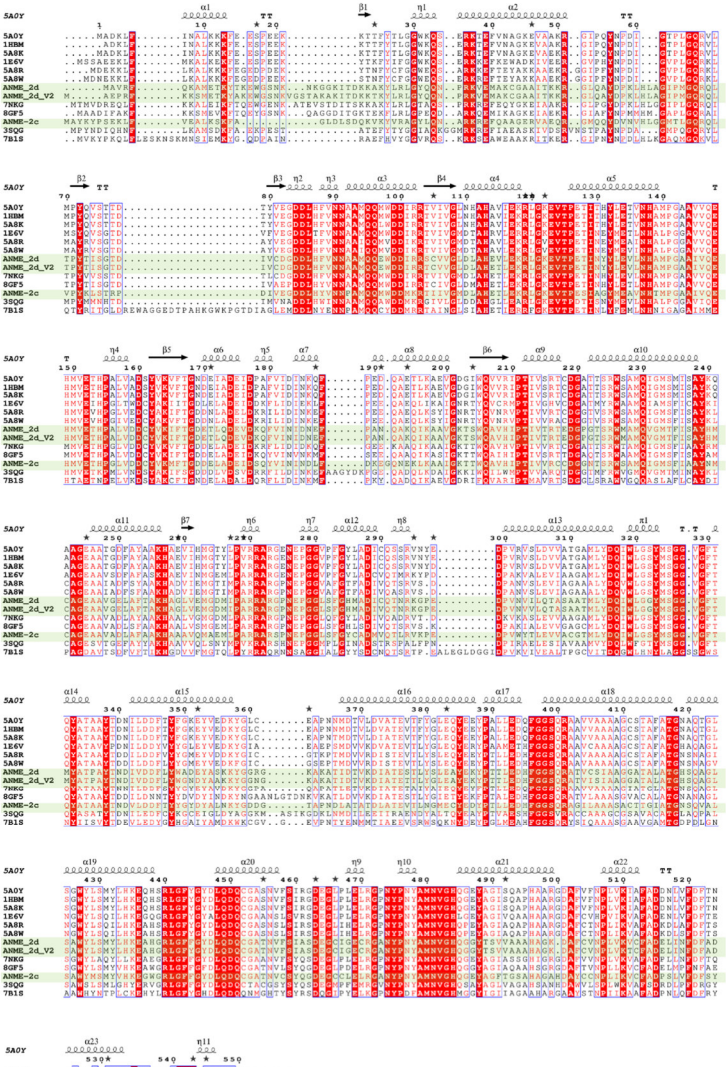


Supplementary figure S3. Continued on next page.

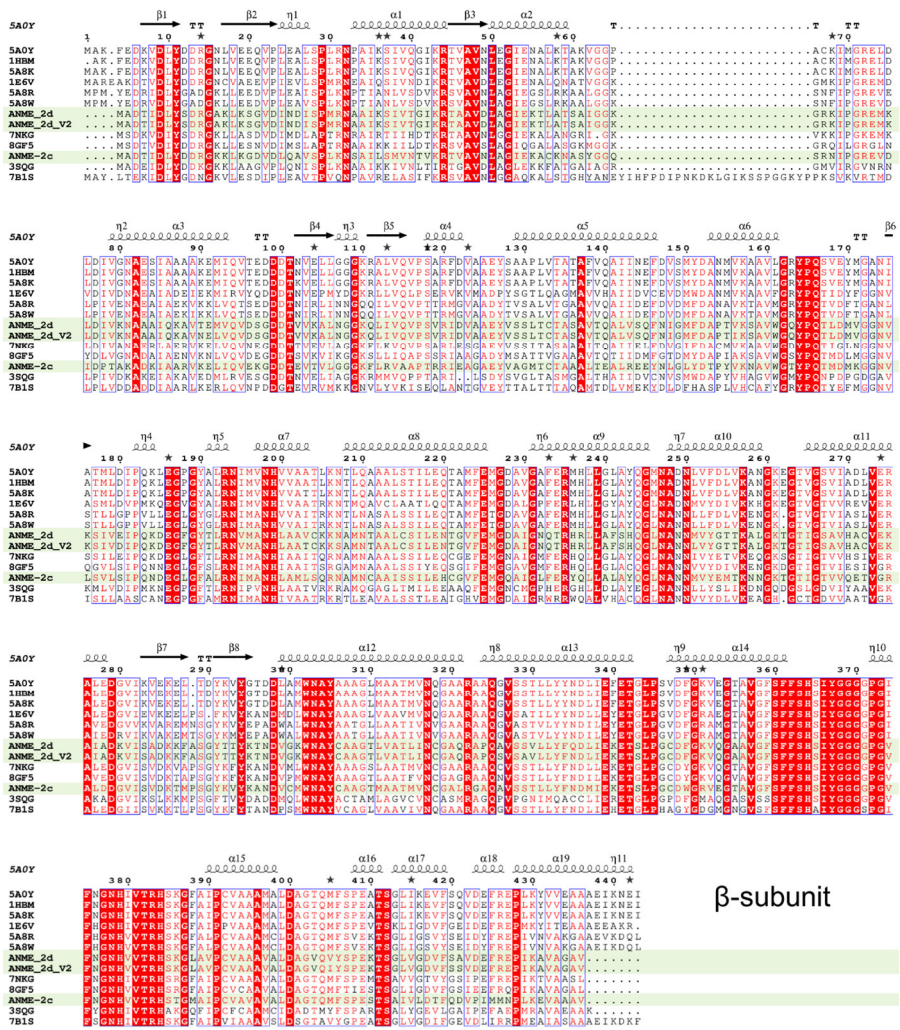
γ-subunit



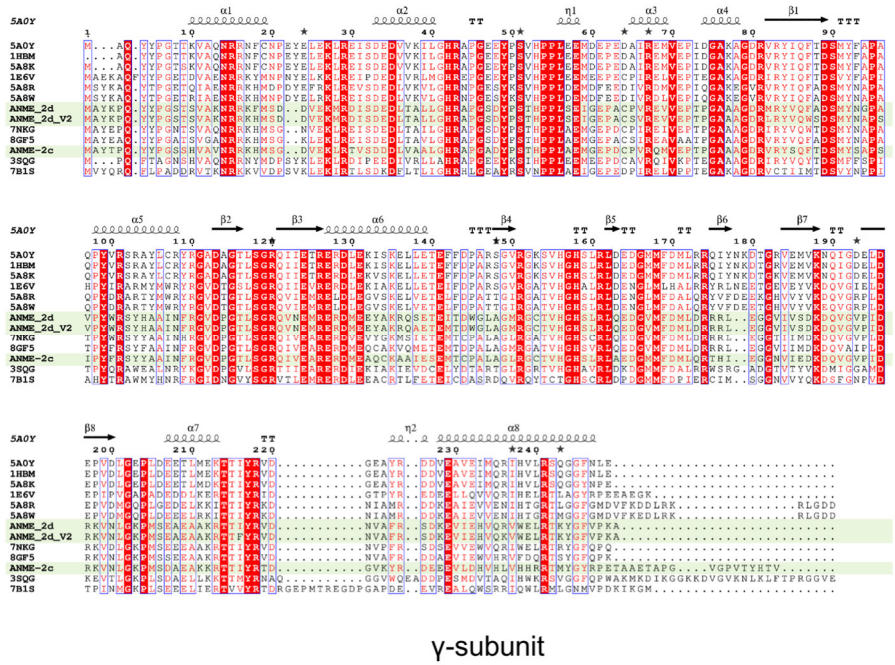
Supplementary figure S3. Phylogeny tree of the structurally characterized MCRs and close sequence homologues from the marine ANME-2 MCR. The models studied in this work are highlighted in yellow.



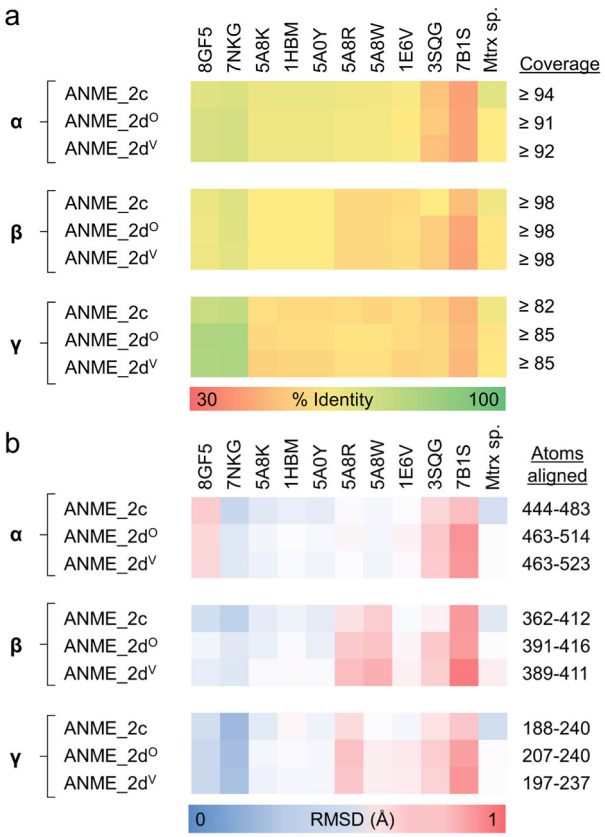
Supplementary figure S4. Continued on next page.



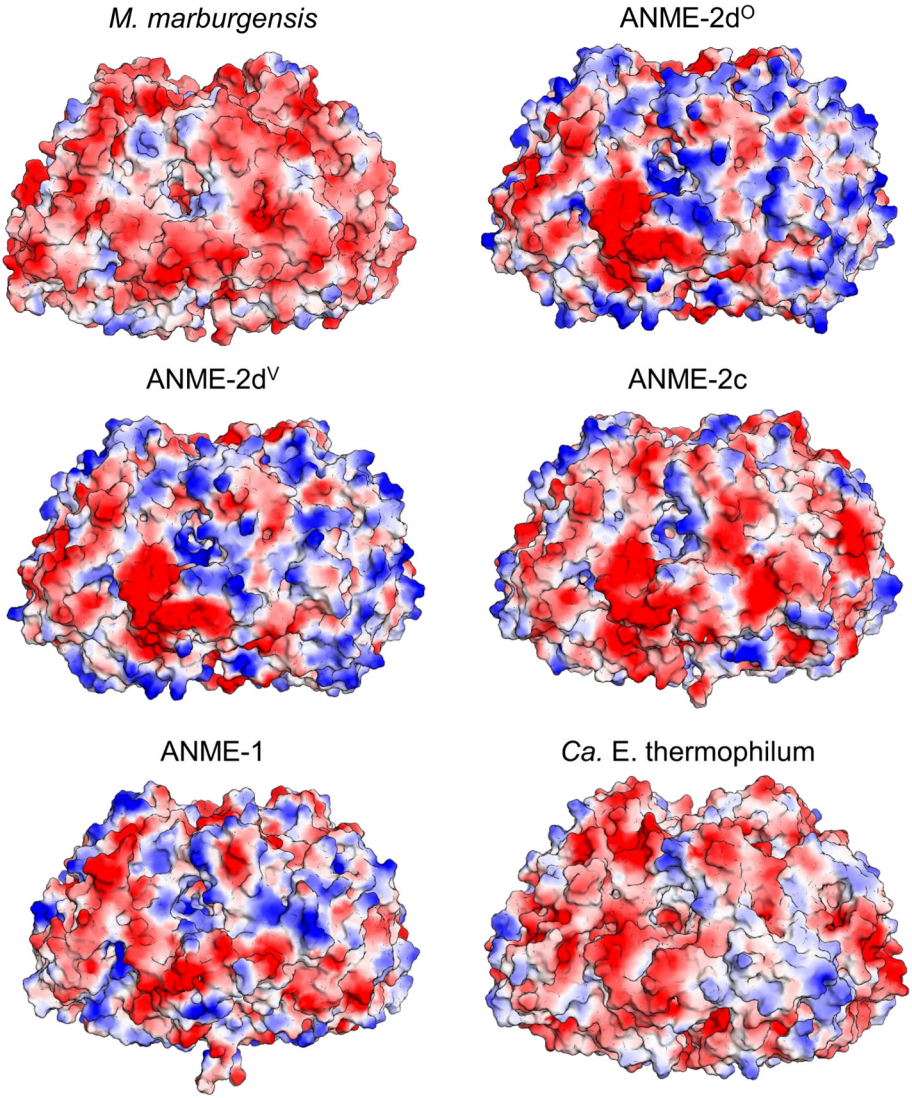
Supplementary figure S4. Continued on next page.



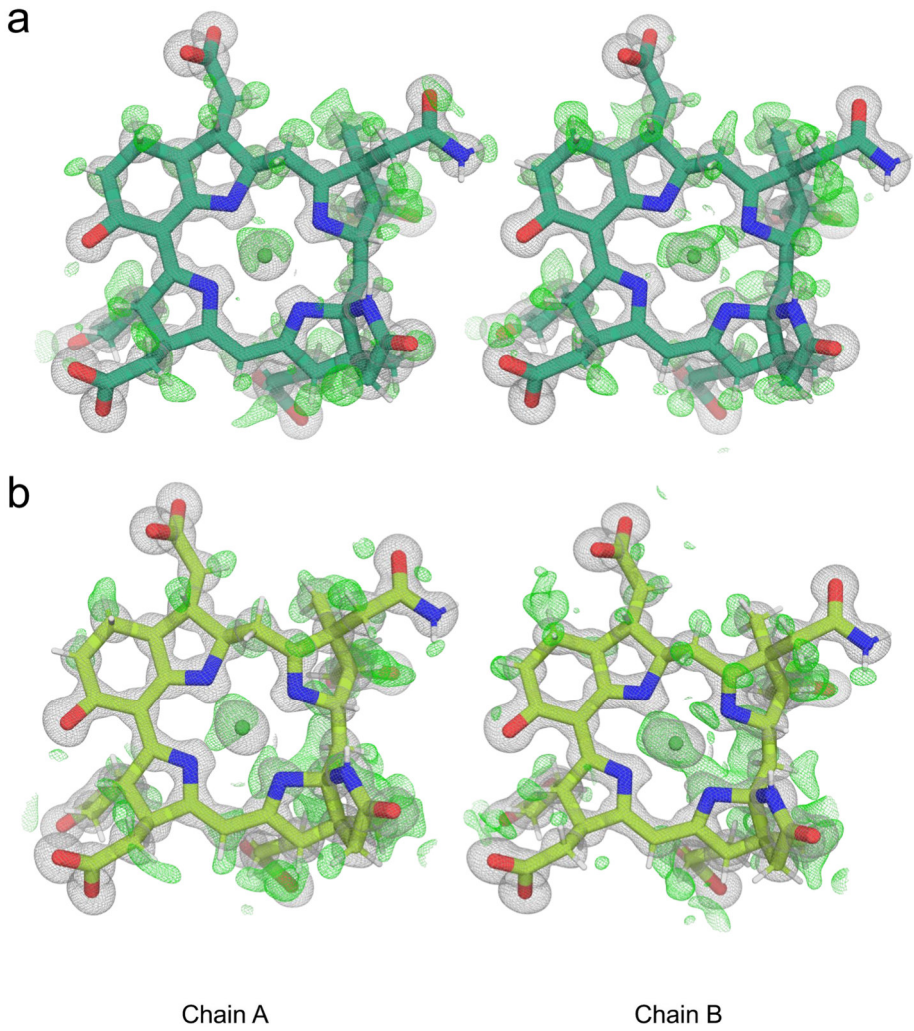
Supplementary figure S4. Sequence alignment of the structurally characterized MCR homologs. ESPript alignment of ANME-2 MCR and structurally characterized homologs from *Methanothermobacter marburgensis* (PDB code: 5A0Y), *Methanothermobacter thermautotrophicus* (1HBM), *Methanothermobacter wolfeii* I (5A8K), *Methanopyrus kandleri* (1E6V), *Methanothermobacter marburgensis* MCR II (5A8R), *Methanothermobacter wolfeii* MCR II (5A8W), *Methermicoccus shengliensis* (7NKG), *Methanosarcina acetivorans* (8GF5), ANME-1 from Black Sea mats (3SQG) and *Candidatus* Ethanoperedens thermophilum (7B1S). ANME-2 MCR sequences are highlighted in green.



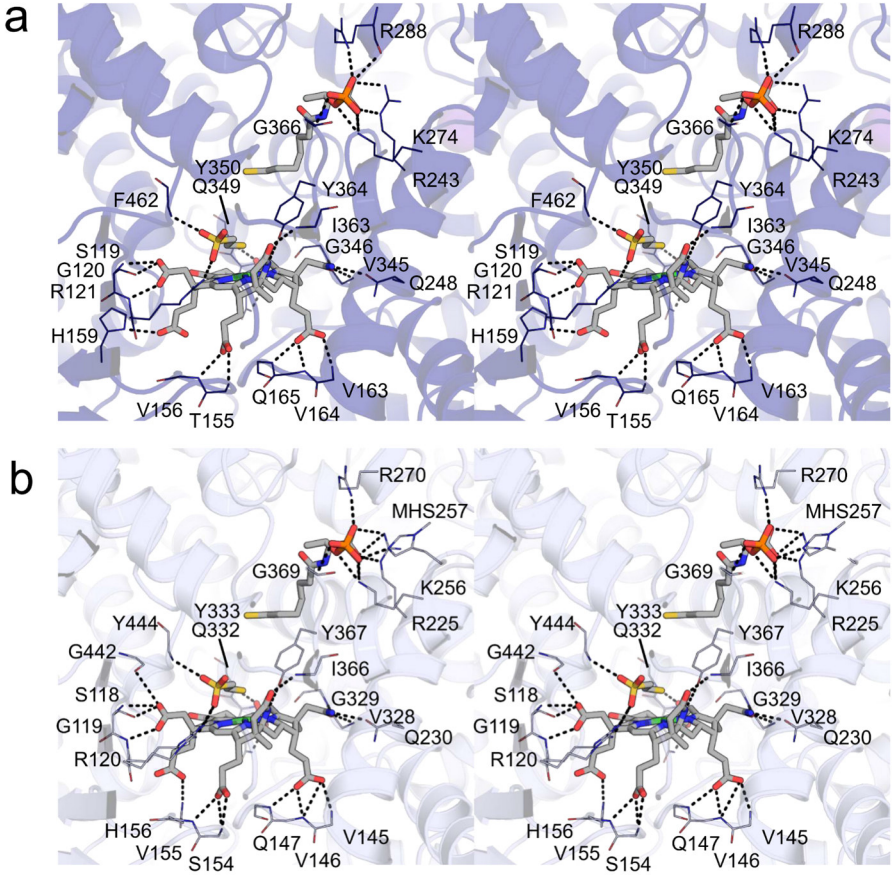
Supplementary figure S5. Sequence and structural similarity of characterized MCR homologs. (**a**) Sequence similarity of the α-, β- and γ-subunits of ANME-2 MCRs to structural homologs. The percentage of identity is displayed as a gradient from red to green for 30 % to 100 % identity. The coverage for all samples is summarised on the right. (**b**) Structural similarity of the α-, β- and γ-subunits of ANME-2 MCRs to structural homologs. The RMSD is displayed as a gradient from blue to red for 0 to 1 Å. The aligned atoms for all samples are summarized on the right. Listed homologs are *Methanothermobacter marburgensis* (PDB code: 5A0Y), *Methanothermobacter thermautotrophicus* (1HBM), *Methanothermobacter wolfeii* I (5A8K), *Methanopyrus kandleri* (1E6V), *Methanothermobacter marburgensis* II (5A8R), *Methanothermobacter wolfeii* II (5A8W), *Methermicoccus shengliensis* (7NKG), *Methanosarcina acetivorans* (8GF5), ANME-1 (3SQG), *Ca.* Ethanoperedens thermophilum (7B1S), and *Methanothrix sp.* (unpublished, see Ch. VI).



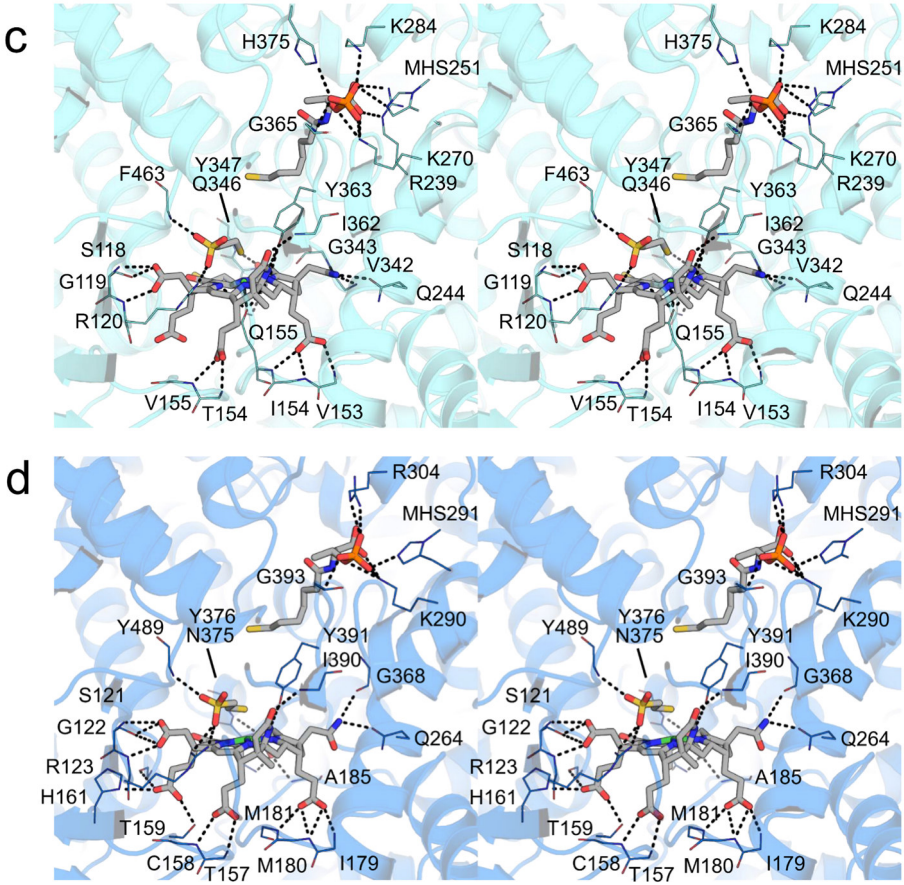
Supplementary figure S6. Electrostatic potential on the surface of MCRs. The vacuum electrostatic potential of MCR is shown as a gradient from red to blue for a more negative to a positive potential, respectively.



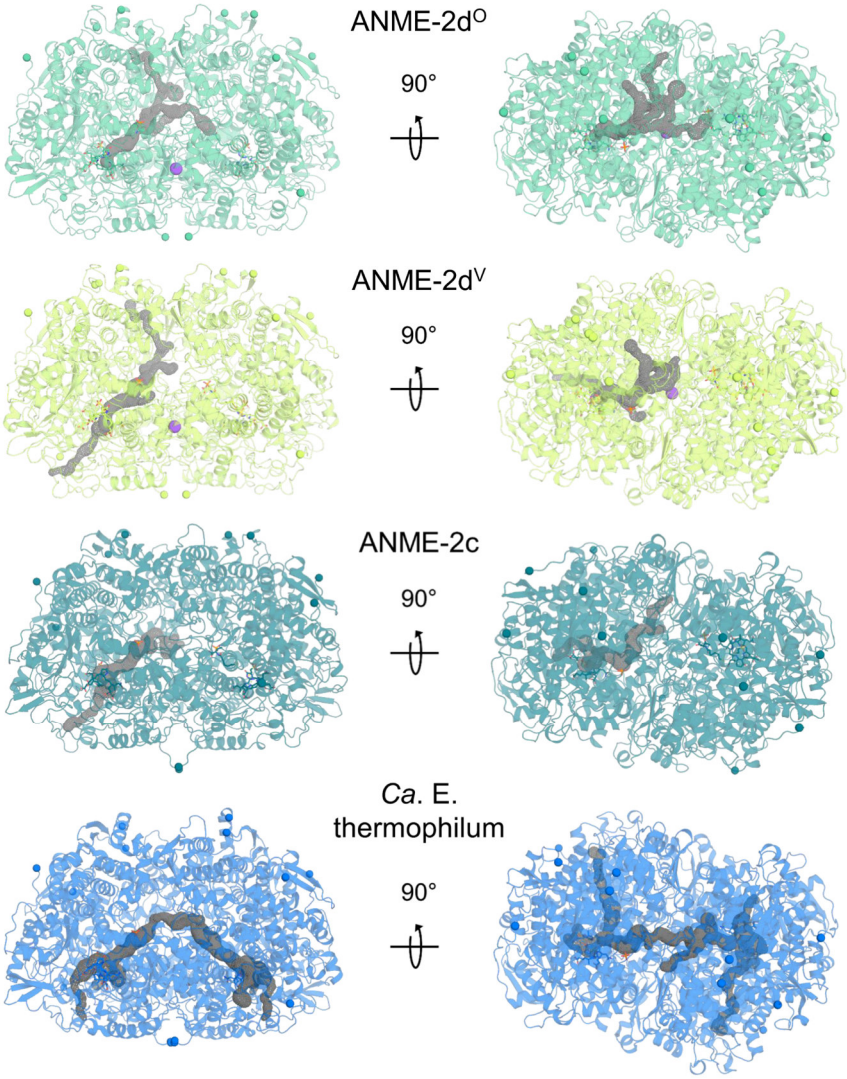
Supplementary figure S7. Electron density map of F_{430}. The electron density is shown for (**a**) ANME-2d^{O} MCR and (**b**) ANME-2d^{V} MCR as grey mesh for the $2F_o{-}F_c$ map at 1 σ and green mesh for the $F_o{-}F_c$ map at 2.8 σ. F_{430} is shown as sticks. Atoms are coloured in red for O, blue for N, and white for riding H. The map was generated with non-hydrogenated F_{430}.



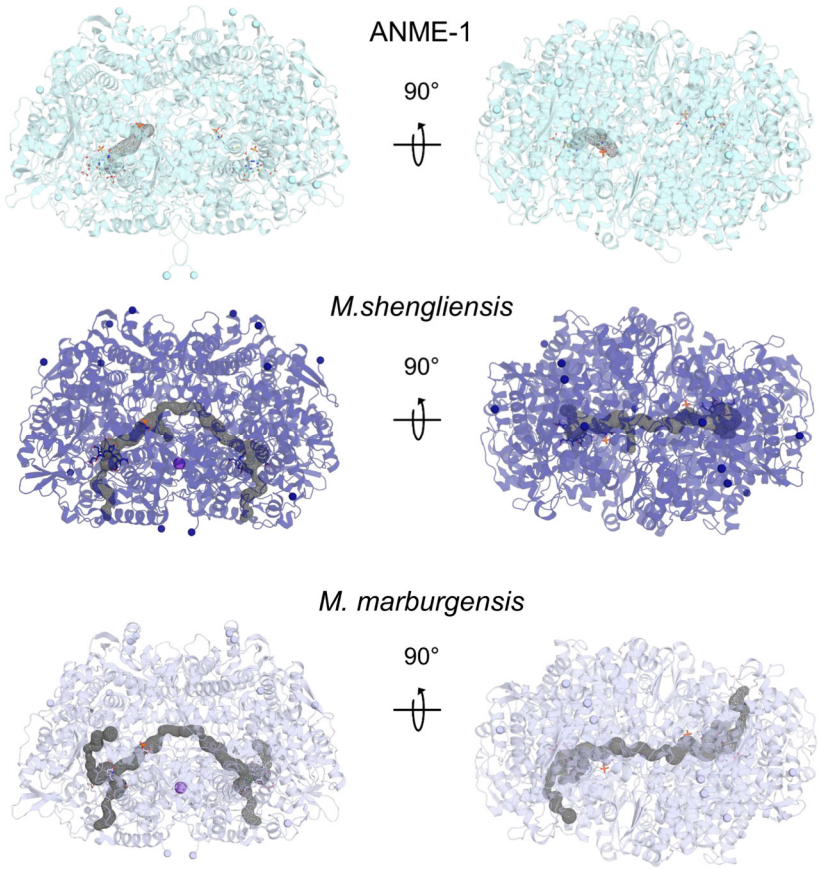
Supplementary figure S8. Continued on the next page.



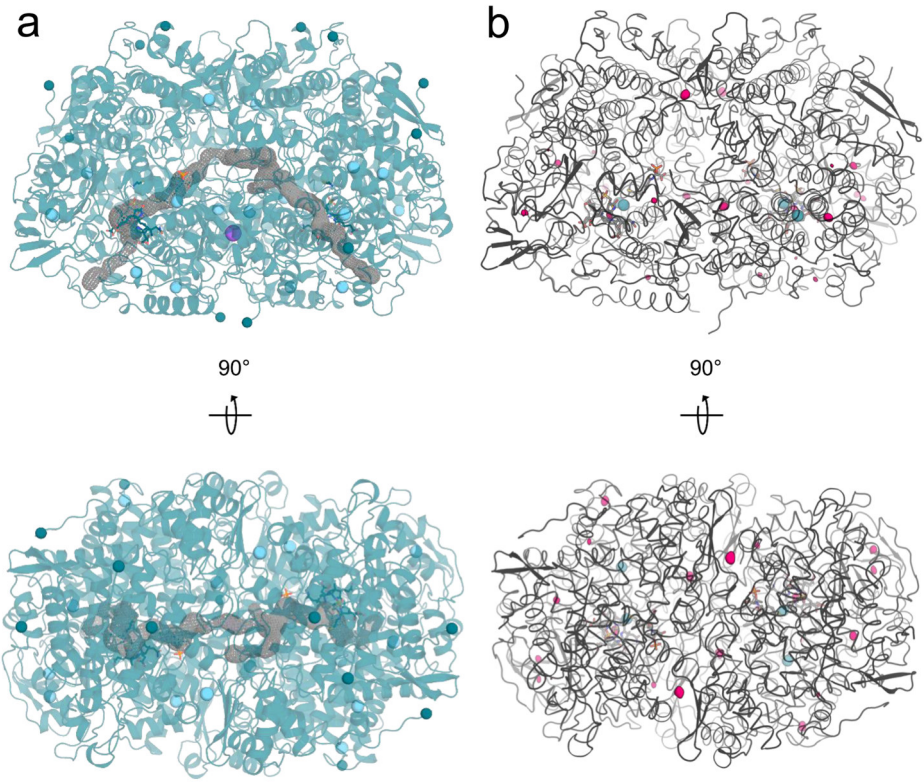
Supplementary figure S8. Active site of MCR homologs. Stereo-view of the MCR active site in (**a**) *M. shengliensis*, (**b**) *M. marburgensis*, (**c**) ANME-1 and (**d**) *Ca.* E. thermophilum. The main chains are shown as cartoons, and ligands are shown as sticks coloured according to the atom with red for O, blue for N, orange for P, and yellow for S. Coordinating residues are shown as lines, non-interacting side- or mainchains were omitted. Polar contacts are depicted as black dashes.



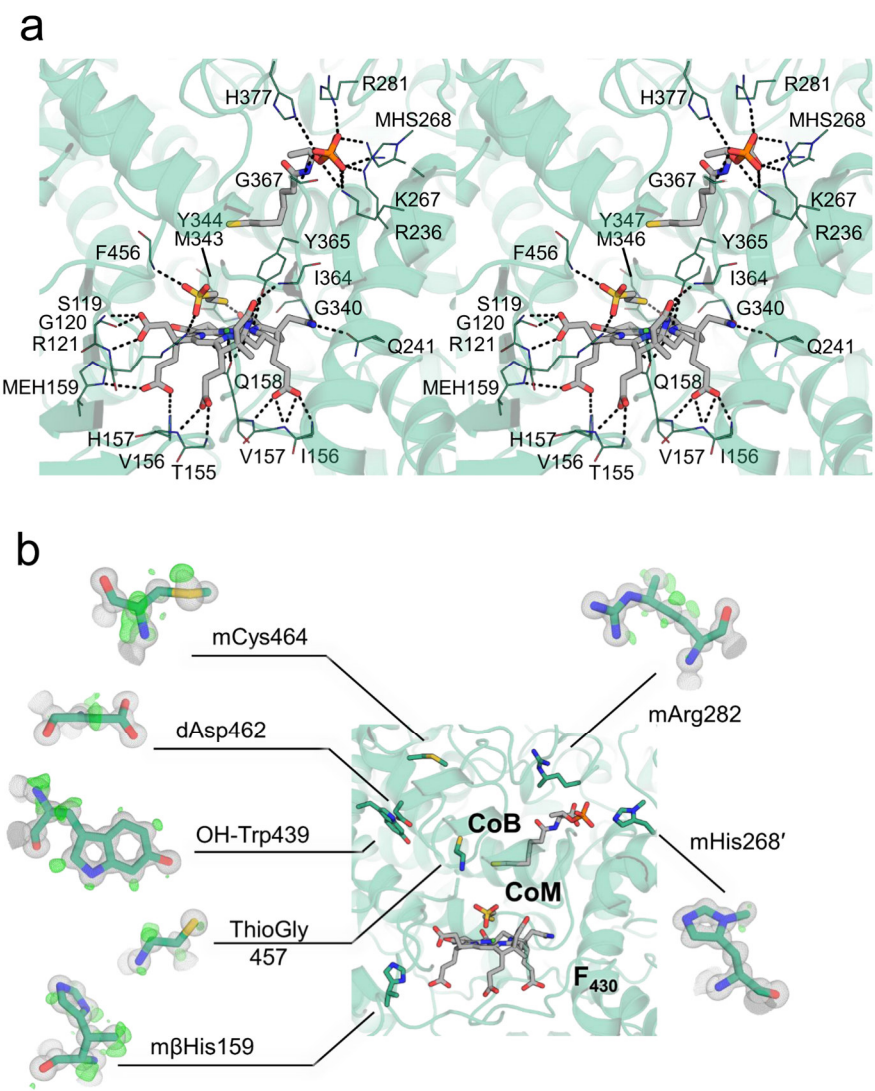
Supplementary figure S9. Continued on the next page.



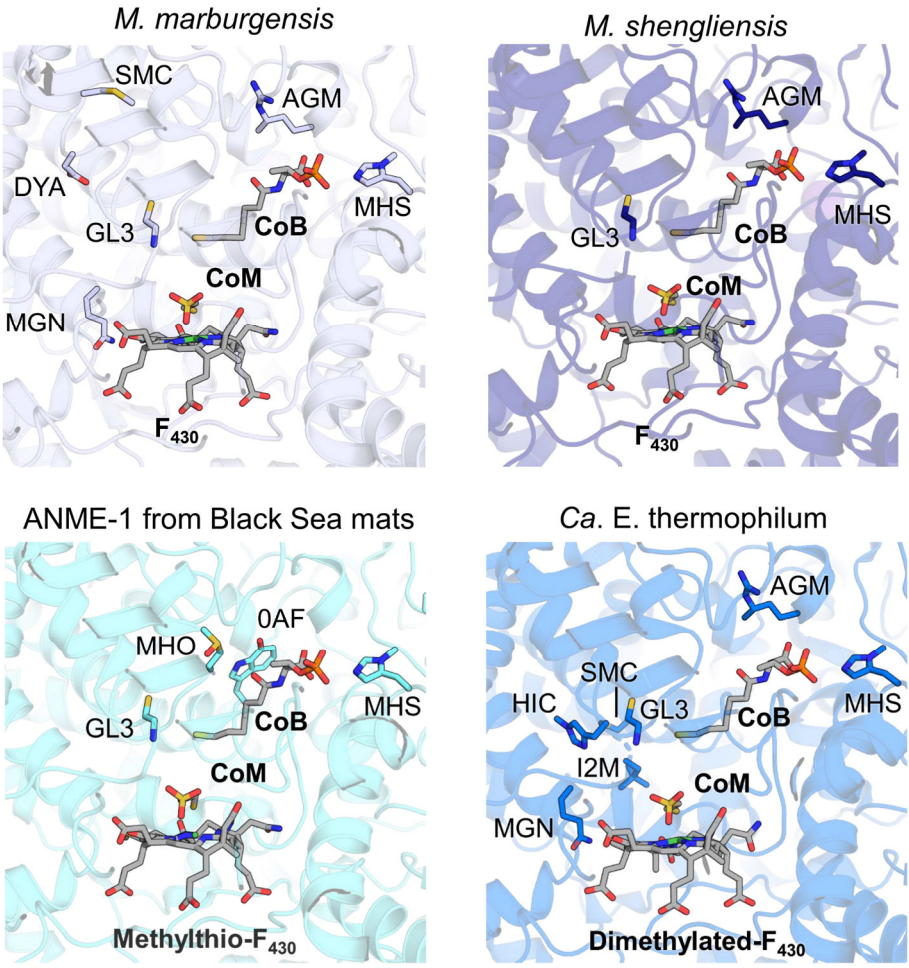
Supplementary figure S9. Internal cavities of MCRs. Prediction of internal cavities by Caver (probe radius = 1Å) for MCR/ECR. Proteins are shown in cartoons with cavities as grey mesh, termini are shown as spheres, and ligands as sticks. Atoms are coloured red for O, blue for N, orange for P, yellow for S, purple for K, and teal for Kr.



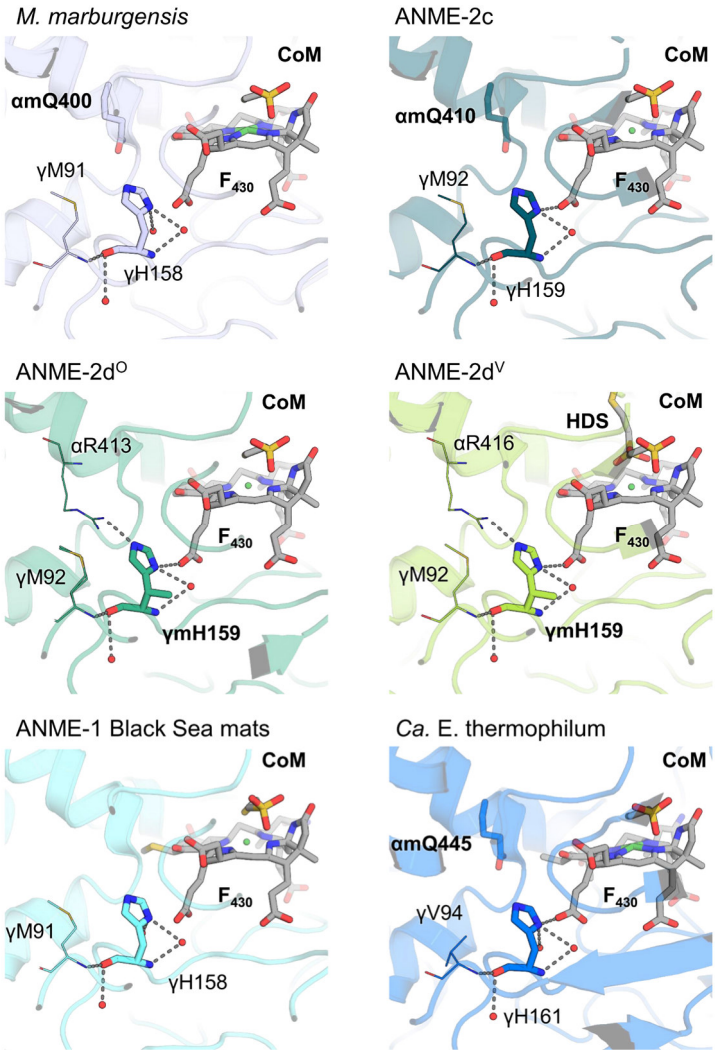
Supplementary figure S10. Internal cavities of MCRs. (**a**) Prediction of internal cavities by Caver (probe radius = 1Å) for MCR ANME-2c gassed with Krypton. The proteins are shown in cartoons with cavities as grey mesh, termini are shown as spheres, ligands as sticks. Atoms are coloured red for O, blue for N, orange for P, yellow for S, purple for K, and teal for Kr. (**b**) Anomalous signal of Kr is depicted by a pink surface contoured to 5 σ. The proteins are shown in cartoons with Xenon position detected in the gassed structure of *Ca.* E. thermophilum is shown in teal spheres.



Supplementary figure S11. Active site and PTMs in MCR-ANME2d^{O}. (**a**) The active site and (**b**) the PTMs localised in the active site surroundings are depicted in the same way as Figure 3b and Figure 4a, respectively.



Supplementary figure S12. PTMs in MCR homologs. The main chain is shown as a cartoon, with ligands and PTMs depicted as sticks coloured according to atom (red for O, blue for N, orange for P, yellow for S). Residues are labelled according to the PDB 3-letter ligand codes, with MHS, GL3, AGM, MGN, SMC, DYA, HIC, MHO, and 0AF standing for N^1-methylhistidine, thioglycine, 5(*S*)-methylarginine, 2(*S*)-methylglutamine, S-methylcysteine, didehydroaspartate, 4-methylhistidine, S-oxymethionine, and 7-hydroxy-L-tryptophan, respectively.



Supplementary figure S13. Structural environment around the PTM 3(*S*)-methylhistidine. The 3(*S*)-methylhistidine or the equivalent unmodified histidine is shown for the different MCR models. The main chain is shown as a cartoon, and ligands are shown as sticks coloured as Supplementary Fig. S12. Polar contacts of histidine/3(*S*)-methylhistidine are depicted as grey dashes. Interacting waters are shown as spheres. Residues are labelled: chain-amino acid initial-residue number.

References

1. Gao, Y. H., Wang, Y., Lee, H. S., Jin, P. K. Significance of anaerobic oxidation of methane (AOM) in mitigating methane emission from major natural and anthropogenic sources: a review of AOM rates in recent publications. *Environ. Sci. Adv.* **1**, 401-425, doi:10.1039/d2va00091a (2022).
2. Segarra, K. E., Schubotz, F., Samarkin, V., Yoshinaga, M. Y., Hinrichs, K. U., Joye, S. B. High rates of anaerobic methane oxidation in freshwater wetlands reduce potential atmospheric methane emissions. *Nat. Commun.* **6**, 7477, doi:10.1038/ncomms8477 (2015).
3. Boetius, A., Ravenschlag, K., Schubert, C. J., Rickert, D., Widdel, F., Gieseke, A., Amann, R., Jorgensen, B. B., Witte, U., Pfannkuche, O. A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* **407**, 623-626, doi:10.1038/35036572 (2000).
4. Timmers, P. H. A., Welte, C. U., Koehorst, J. J., Plugge, C. M., Jetten, M. S. M., Stams, A. J. M. Reverse methanogenesis and respiration in methanotrophic archaea. *Archaea* **2017**, doi:10.1155/2017/1654237 (2017).
5. Chadwick, G. L., Skennerton, C. T., Laso-Pérez, R., Leu, A. O., Speth, D. R., Yu, H., Morgan-Lang, C., Hatzenpichler, R., Goudeau, D., Malmstrom, R., et al. Comparative genomics reveals electron transfer and syntrophic mechanisms differentiating methanotrophic and methanogenic archaea. *PLoS Biol.* **20**, doi:10.1371/journal.pbio.3001508 (2022).
6. Wang, Y., Wegener, G., Ruff, S. E., Wang, F. Methyl/alkyl-coenzyme M reductase-based anaerobic alkane oxidation in archaea. *Environ. Microbiol.* **23**, 530-541, doi:10.1111/1462-2920.15057 (2021).
7. Wegener, G., Krukenberg, V., Ruff, S. E., Kellermann, M. Y., Knittel, K. Metabolic capabilities of microorganisms involved in and associated with the anaerobic oxidation of methane. *Front. Microbiol.* **7**, doi:10.3389/fmicb.2016.00046 (2016).
8. Wegener, G., Krukenberg, V., Riedel, D., Tegetmeyer, H. E., Boetius, A. Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. *Nature* **526**, 587-590, doi:10.1038/nature15733 (2015).
9. Schnakenberg, A., Aromokeye, D. A., Kulkarni, A., Maier, L., Wunder, L. C., Richter-Heitmann, T., Pape, T., Ristova, P. P., Bühring, S. I., Dohrmann, I., et al. Electron acceptor availability shapes anaerobically methane oxidizing archaea (ANME) communities in South Georgia sediments. *Front. Microbiol.* **12**, 617280, doi:10.3389/fmicb.2021.617280 (2021).
10. Raghoebarsing, A. A., Pol, A., van de Pas-Schoonen, K. T., Smolders, A. J. P., Ettwig, K. F., Rijpstra, W. I. C., Schouten, S., Damsté, J. S. S., Op den Camp, H. J. M., Jetten, M. S. M., et al. A microbial consortium couples anaerobic methane oxidation to denitrification. *Nature* **440**, 918-921, doi:10.1038/nature04617 (2006).
11. Ettwig, K. F., Zhu, B. L., Speth, D., Keltjens, J. T., Jetten, M. S. M., Kartal, B. Archaea catalyze iron-dependent anaerobic oxidation of methane. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12792-12796, doi:10.1073/pnas.1609534113 (2016).
12. Nauhaus, K., Albrecht, M., Elvert, M., Boetius, A., Widdel, F. *In vitro* cell growth of marine archaeal-bacterial consortia during anaerobic oxidation of methane with sulfate. *Environ. Microbiol.* **9**, 187-196, doi:10.1111/j.1462-2920.2006.01127.x (2007).
13. Holler, T., Widdel, F., Knittel, K., Amann, R., Kellermann, M. Y., Hinrichs, K. U., Teske, A., Boetius, A., Wegener, G. Thermophilic anaerobic oxidation of methane by marine microbial consortia. *ISME J.* **5**, 1946-1956, doi:10.1038/ismej.2011.77 (2011).
14. Vaksmaa, A., Guerrero-Cruz, S., van Alen, T. A., Cremers, G., Ettwig, K. F., Lüke, C., Jetten, M. S. M. Enrichment of anaerobic nitrate-dependent methanotrophic '*Candidatus* Methanoperedens nitroreducens' archaea from an Italian paddy field soil. *Appl. Microbiol. Biotechnol.* **101**, 7075-7084, doi:10.1007/s00253-017-8416-0 (2017).
15. Krüger, M., Meyerdierks, A., Glöckner, F. O., Amann, R., Widdel, F., Kube, M., Reinhardt, R., Kahnt, R., Böcher, R., Thauer, R. K., et al. A conspicuous nickel protein in microbial mats that oxidize methane anaerobically. *Nature* **426**, 878-881, doi:10.1038/nature02207 (2003).

16. Lemaire, O. N., Wagner, T. A structural view of alkyl-coenzyme M reductases, the first step of alkane anaerobic oxidation catalyzed by archaea. *Biochemistry* **61**, 805-821, doi:10.1021/acs.biochem.2c00135 (2022).
17. Thauer, R. K. Methyl (alkyl)-coenzyme M reductases: Nickel F-430-containing enzymes involved in anaerobic methane formation and in anaerobic oxidation of methane or of short chain alkanes. *Biochemistry* **58**, 5198-5220, doi:10.1021/acs.biochem.9b00164 (2019).
18. Scheller, S., Goenrich, M., Boecher, R., Thauer, R. K., Jaun, B. The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature* **465**, 606-U97, doi:10.1038/nature09015 (2010).
19. Mayr, S., Latkoczy, C., Krüger, M., Günther, D., Shima, S., Thauer, R. K., Widdel, F., Jaun, B. Structure of an F430 variant from archaea associated with anaerobic oxidation of methane. *J. Am. Chem. Soc.* **130**, 10758-10767, doi:10.1021/ja802929z (2008).
20. Hahn, C. J., Lemaire, O. N., Kahnt, J., Engilberge, S., Wegener, G., Wagner, T. Crystal structure of a key enzyme for anaerobic ethane activation. *Science* **373**, 118-121, doi:10.1126/science.abg1765 (2021).
21. Nayak, D. D., Liu, A. D., Agrawal, N., Rodriguez-Carerro, R., Dong, S. H., Mitchell, D. A., Nair, S. K., Metcalf, W. W. Functional interactions between posttranslationally modified amino acids of methyl-coenzyme M reductase in *Methanosarcina acetivorans*. *PLoS Biol.* **18**, doi:10.1371/journal.pbio.3000507 (2020).
22. Nayak, D. D., Mahanta, N., Mitchell, D. A., Metcalf, W. W. Post-translational thioamidation of methyl-coenzyme M reductase, a key enzyme in methanogenic and methanotrophic Archaea. *eLife* **6**, doi:10.7554/eLife.29218 (2017).
23. Deobald, D., Adrian, L., Schöne, C., Rother, M., Layers, G. Identification of a unique radical SAM methyltransferase required for the sp^3-C-methylation of an arginine residue of methyl-coenzyme M reductase. *Sci. Rep.* **8**, doi:10.1038/s41598-018-25716-x (2018).
24. Chen, H., Gan, Q. L., Fan, C. G. Methyl-coenzyme M reductase and its post-translational modifications. *Front. Microbiol.* **11**, doi:10.3389/fmicb.2020.578356 (2020).
25. Shima, S., Krueger, M., Weinert, T., Demmer, U., Kahnt, J., Thauer, R. K., Ermler, U. Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature* **481**, 98-101, doi:10.1038/nature10663 (2012).
26. Shao, N. N., Fan, Y., Chou, C. W., Yavari, S., Williams, R. V., Amster, I. J., Brown, S. M., Drake, I. J., Duin, E. C., Whitman, W. B., et al. Expression of divergent methyl/alkyl coenzyme M reductases from uncultured archaea. *Commun. Biol.* **5**, doi:10.1038/s42003-022-04057-6 (2022).
27. Berger, S., Frank, J., Dalcin Martins, P., Jetten, M. S. M., Welte, C. U. High-quality draft genome sequence of "*Candidatus* Methanoperedens sp." strain BLZ2, a nitrate-reducing anaerobic methane-oxidizing archaeon enriched in an anoxic bioreactor. *Genome Announc.* **5**, doi:10.1128/genomeA.01159-17 (2017).
28. Kurth, J. M., Smit, N. T., Berger, S., Schouten, S., Jetten, M. S. M., Welte, C. U. Anaerobic methanotrophic archaea of the ANME-2d clade feature lipid composition that differs from other ANME archaea. *FEMS Microbiol. Ecol.* **95**, doi:10.1093/femsec/fiz082 (2019).
29. Ouboter, H. T., Mesman, R., Sleutels, T., Postma, J., Wissink, M., Jetten, M. S. M., Ter Heijne, A., Berben, T., Welte, C. U. Mechanisms of extracellular electron transfer in anaerobic methanotrophic archaea. *Nat. Commun.* **15**, 1477, doi:10.1038/s41467-024-45758-2 (2024).
30. Wissink, M., Glodowska, M., van der Kolk, M. R., Jetten, M. S. M., Welte, C. U. Probing denitrifying anaerobic methane oxidation via antimicrobial intervention: implications for innovative wastewater management. *Environ. Sci. Technol.* **58**, 6250-6257, doi:10.1021/acs.est.3c07197 (2024).
31. Berger, S., Cabrera-Orefice, A., Jetten, M. S. M., Brandt, U., Welte, C. U. X Investigation of central energy metabolism-related protein complexes of ANME-2d methanotrophic archaea by complexome profiling. *Biochim. Biophys. Acta. Bioenerg.* **1862**, 148308, doi:10.1016/j.bbabio.2020.148308 (2021).
32. Echeveste Medrano, M. J., Leu, A. O., Pabst, M., Lin, Y., McIlroy, S. J., Tyson, G. W., van Ede, J., Sánchez-Andrea, I., Jetten, M. S. M., Jansen, R., et al. Osmoregulation in freshwater anaerobic methane-oxidizing archaea under salt stress. *ISME J.* **18**, doi:10.1093/ismejo/wrae137 (2024).

33. Kurth, J. M., Müller, M.-C., Welte, C. U., Wagner, T. Structural insights into the methane-generating enzyme from a methoxydotrophic methanogen reveal a restrained gallery of post-translational modifications. *Microorganisms* **9**, doi:10.3390/microorganisms9040837 (2021).
34. Ohmer, C. J., Dasgupta, M., Patwardhan, A., Bogacz, I., Kaminsky, C., Doyle, M. D., Chen, P. Y., Keable, S. M., Makita, H., Simon, P. S., et al. XFEL serial crystallography reveals the room temperature structure of methyl-coenzyme M reductase. *J. Inorg. Biochem.* **230**, 111768, doi:10.1016/j.jinorgbio.2022.111768 (2022).
35. Marlow, J. J., Skennerton, C. T., Li, Z., Chourey, K., Hettich, R. L., Pan, C., Orphan, V. J. Proteomic stable isotope probing reveals biosynthesis dynamics of slow growing methane based microbial communities. *Front. Microbiol.* **7**, doi:10.3389/fmicb.2016.00563 (2016).
36. Kurth, J. M., Nobu, M. K., Tamaki, H., de Jonge, N., Berger, S., Jetten, M. S. M., Yamamoto, K., Mayumi, D., Sakata, S., Bai, L., et al. Methanogenic archaea use a bacteria-like methyltransferase system to demethoxylate aromatic compounds. *ISME J.* **15**, 3549-3565, doi:10.1038/s41396-021-01025-6 (2021).
37. Widdel, F., Bak, F. Gram-negative mesophilic sulfate-reducing bacteria. In *The Prokaryotes: A Handbook on the Biology of Bacteria: Ecophysiology, Isolation, Identification, Applications* (eds Balows, A.; Trüper, H. G.; Dworkin, M.; Harder, W.; Schleifer, K.-H.) 3352-3378 (Springer New York, 1992).
38. Laso-Pérez, R., Krukenberg, V., Musat, F., Wegener, G. Establishing anaerobic hydrocarbon-degrading enrichment cultures of microorganisms under strictly anoxic conditions. *Nat. Protoc.* **13**, 1310-1330, doi:10.1038/nprot.2018.030 (2018).
39. Carpentier, P., van der Linden, P., Mueller-Dieckmann, C. The High-Pressure Freezing Laboratory for Macromolecular Crystallography (HPMX), an ancillary tool for the macromolecular crystallography beamlines at the ESRF. *Acta Crystallog. D* **80**, 80-92, doi:10.1107/S2059798323010707 (2024).
40. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., Bricogne, G. Data processing and analysis with the *autoPROC* toolbox. *Acta Crystallogr. D Struct. Biol.* **67**, 293-302, doi:10.1107/S0907444911007773 (2011).
41. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. *Acta Crystallogr. D Struct. Biol.* **75**, 861-877, doi:10.1107/S2059798319011471 (2019).
42. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zídek, A., Potapenko, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
43. Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
44. Katoh, K., Rozewicki, J., Yamada, K. D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160-1166, doi:10.1093/bib/bbx108 (2019).
45. Robert, X., Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-W324, doi:10.1093/nar/gku316 (2014).

CHAPTER VI. Native isolation of the CH_4-generating enzyme from a mesocosm – steps towards microbial environmental structural biology

Marie-C. Müller[1*], Marion Jespersen[1*], Nevena Maslać[1], Grace D'Angelo[1], Almud Lonsing[1], Sonja Dunemann[2], Bruno Huettel[2], Sylvain Engilberge[3,4], Manuel Liebeke[1], Jens Harder[1], Tristan Wagner[1,4#]

[1] Max-Planck-Institute for Marine Microbiology, Bremen, Germany

[2] Max Planck Genome Centre, Cologne, Germany

[3] European Synchrotron Radiation Facility, Grenoble, France

[4] Univ. Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale, Grenoble, France.

[*]These authors participated equally in the study.

Present addresses: Marion Jespersen - Department of Microbiology, Monash University, Australia. Nevena Maslać - Archaea Biology and Ecogenomics Unit, Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria. Manuel Liebeke - Metabolomics, Christian-Albrechts-University Kiel, Germany. Tristan Wagner - Institut de Biologie Structurale, Grenoble, France.

[#]Correspondence: twagner@mpi-bremen.de

Manuscript in preparation

Abstract

Studying enzymes from non-isolated microorganisms is key to understanding planetary microbial biogeochemistry. Biochemical investigation of these enzymes is feasible through recombinant expression at the price of potential biases due to the absence of specific native factors. Alternatively, enzymatic characterization directly from environmental samples would provide an unbiased snapshot of microbial biochemistry, a task currently considered extremely challenging due to the heterogeneity of environmental biomass. Here, using a modern purification pipeline, we relied on a century-old technique to separate proteins via crystallization. As proof of concept, we isolated two methane-generating enzyme variants from *Methanothrix* species from an active wastewater treatment plant with or without physical cell enrichment. Two structures were obtained from the same crystal diffracting to 1.98-Å, detailing post-translational modifications, cofactors, and active sites. The successful protein isolation from a challenging heterogeneous sample will stimulate future environmental biochemistry studies that overcome the microbial isolation barrier.

Introduction

It is currently estimated that 80 to 99 % of all global microbial species remain uncultivated[1]. However, organisms of this "microbial dark matter" are far from being undescribed. For instance, microscopy techniques using specific fluorescent probes allow their identification[2] while the complementary analysis of genomes[3], transcriptomes[4], proteomes[5], and metabolomes[6] can attribute metabolic roles to the diverse microbial actors in these environmental communities. Since omics techniques rely on sequence homologies with previously characterized enzymes, a bias of metabolic pathway predictions still exists. Therefore, experimental validation remains a necessary means to confirm the physiological function of gene products based on the description of their enzymatic properties, mechanisms, and regulation. Deciphering the reaction mechanisms of the enzymatic treasure trove of microbial dark matter is imperative to inspire chemical processes or biotechnological applications.

For the last decades, biochemists have mostly studied enzymes from non-cultivated microbes by recombinant expression to produce proteins in a host and streamline purification via tagging. While the technique is popular and has led to breakthroughs, pitfalls and biases (*e.g.*, codon bias, misfolding, lacking partners and protein modifications, toxicity and more[7]) persist, resulting in lower production yield, aggregation, and partial or complete absence of activity. To circumvent these barriers, recent work has focused on diversifying expression systems[8-10]. However, as recently shown in the case of methanogenic archaea, other issues arise, like the formation of heterocomplexes through the hybrid assembly of recombinant and native proteins[11].

Alternatively, manipulating native biomass allows direct observation of the natural physiological state of correctly folded proteins fully active and equipped with their required modifications, cofactors, and partners to ensure their biological function. With a few exceptions, native proteins are studied from axenic monoclonal cultures. These

isolated microbes are cultivated in ideal laboratory conditions with unlimited substrates, additional nutrients (*e.g.*, vitamins), controlled physicochemical parameters and the absence of predators and viruses. Similar to domesticated pets, these tamed microbes, evolving without stresses of starvation and harsh habitat conditions, can lose some of their natural traits and metabolic versatility.

The presented work aims to overcome the microbial cultivation barrier and the drawbacks of recombinant expression by directly investigating "environmental" enzymes from a native microbial community. The challenge of this approach is the extreme protein complexity of the biological sample, which we overcame by utilizing crystallization for ultimate protein isolation[12].

As a proof of concept, this study focused on the isolation of enzymes involved in the microbial process of methanogenesis, which contributes significantly to the carbon cycle by contributing to half of the yearly methane release into the atmosphere[13, 14]. Our focus turned to *Methanothrix* species, which have been proposed to be the predominant biogenic methane producers on Earth[15]. By efficiently metabolizing acetate to methane for energy conservation[16-18], *Methanothrix* species outcompete other acetoclastic methanogens (*i.e.*, *Methanosarcina*) in low acetate environments[17]. The catabolic transformation of acetate is a key step in the degradation of organic matter[14, 19], as 70 % of biological methane emissions come from anaerobic fermentation[20, 21]. The process is particularly relevant in wetland soils such as rice paddies[22], where *Methanothrix* is highly active[23]. Despite their significant relevance in the environment[23-25] and in biotechnological applications such as wastewater treatment plants[26], especially in anaerobic digesters, structural studies of *Methanothrix* enzymes have been limited due to slow growth and low biomass yields under laboratory conditions[15]. More importantly, artificial cultivation under laboratory conditions only explores one facet of their methanogenic metabolism, as *Methanothrix* are capable of CO_2-dependent methane

formation when they establish symbiotic interactions with microbial partners, such as *Geobacter*[23, 27, 28].

In this study, we bypass the cultivation barrier by isolating enzymes directly from a wastewater treatment plant, offering a native "ecological" snapshot of catabolic enzymes from *Methanothrix* species. With this proof of concept, we established a framework that will encourage future native studies of environmental proteins.

Results

Enrichment of a *Methanothrix* population from active sludge.

Anaerobic sludge from a wastewater treatment plant is a mesocosm that shelters a microbial ecosystem excelling at degrading organic matter, with *Methanothrix* species comprising a large fraction of the archaeal population[26]. We sampled active sludge from a local wastewater treatment plant twice within four months, referred to as samples A and B (Fig. 1). Metagenomic analysis was performed on each sample to evaluate the microbial population composition and to obtain the sequences of genes of interest. Qualitative information about the composition of the microbial population was derived from these analyses, as read abundance can only be interpreted as an indicator of diversity and relative abundance. As a first observation, the taxonomic composition remained relatively stable between the two sampling time points. Secondly, *Methanothrix* constitutes a third of the archaeal population. A large fraction of reads remained unassigned, pointing towards a high species diversity within both samples (Fig. 1). The reads from both samples were merged and assembled, followed by genome binning and taxonomic assignment. Two metagenome-assembled genomes (MAGs) were assigned to *Methanothrix*, and a total of nine high-quality MAGs were retrieved (Extended Data Table 1).

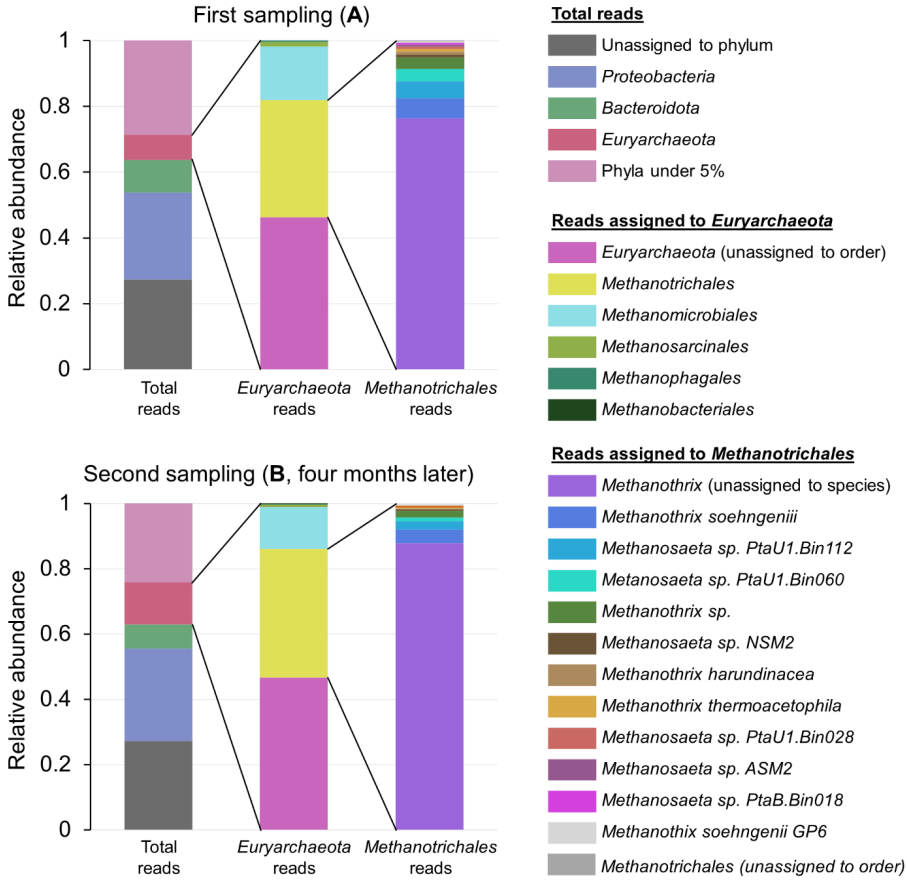


Figure 1. Activated sludge metagenome from two sampling points (sample A and B). From left to right: relative reads assigned at the phylum level with groups below 5 % collapsed, relative reads assigned on the order level to *Euryarchaeota,* and relative reads assigned to species within *Methanothrichales*. An analysis of the species diversity revealed a majority of bacteria (89 %/84 %, for A/B samples) with only small fractions of *Archaea* (9 %/15 %), *Eukaryota* (1 %/0.6 %) and viruses (0.5 %/0.5 %) in both samples. *Methanothrix* constitutes 35 %/39 % of *Euryarchaeota*, which themselves make up 86 %/84 % of the archaeal population. *Methanothrix* and *Methanosaeta* belong to the same genus.

Activated sludge contains a high degree of microbial heterogeneity and recalcitrant inorganic/organic particles (*e.g.*, flocculates) that would hamper the purification of proteins of interest. To circumvent the strain on the chromatography equipment and streamline the purification process, a rough microbial separation was established (Fig. 2a). Following the dilution of the sludge, the mechanical enrichment protocol consisting of serial centrifugation steps was applied to fractionate the microbial population in terms of their physical properties (*e.g.*, planktonic, filamentous, or aggregates). At that stage, long *Methanothrix* filaments were separated from bulkier particles and single cells. Fluorescence *in situ* hybridization (FISH) using a specific probe confirmed a high enrichment of filamentous *Methanothrix* cells in the enriched fraction (Fig. 2b). Most fluorescent signals specific to *Methanothrix* were lost after sonication, confirming successful cell lysis via sonication.

Isolation and identification of the main catabolic enzyme of *Methanothrix* from the physical enrichment.

The particle-free brown soluble fraction of the cell lysate was passed on an anionic exchange chromatography column to capture the exploitable pool of protein (Fig. 3, and Extended Data Fig.1a-c). Catabolic enzymes were expected to be predominant in the different fractions composing the chromatographic separation, among them the methane-generating enzyme complex named Methyl-coenzyme M reductase (MCR, constituted of an $\alpha_2\beta_2\gamma_2$ arrangement). Because of its natural high abundance and its distinctiveness, MCR represented an adequate protein model for this study. Indeed, MCR-containing fractions were detected by the yellow color from its cofactor F_{430} and its profile on sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE, denaturing conditions) constituted of a three-band pattern (α, β and γ subunits). Pooled MCR fractions were passed on a hydrophobic interaction chromatography column (Fig. 3). While the resulting SDS-PAGE profile suggests a high-grade purified MCR, the protein fraction might still harbor a mixture since MCR-encoding metagenomics reads

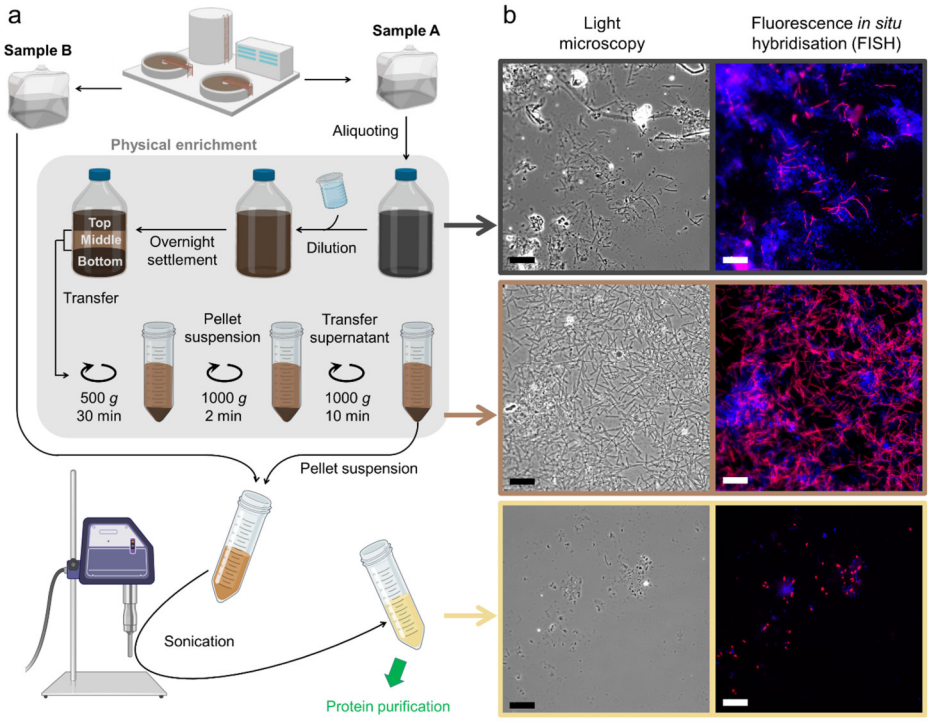


Figure 2. Cell lysate preparation for protein purification. (**a**) Scheme of the sample preparation for protein purification through optional physical enrichment (for sample A). (**b**) Microscopic images of the untreated sludge (top), resuspended microbial pellet treated by physical enrichment before (middle) and after sonication (bottom). Scale bars represent 20 µm. Biological samples applied for microscopic observation have different cellular concentrations between the different treatments. Microbial cells are colored in blue (DAPI staining), and *Methanothrix* cells in pink (fluorescent probes).

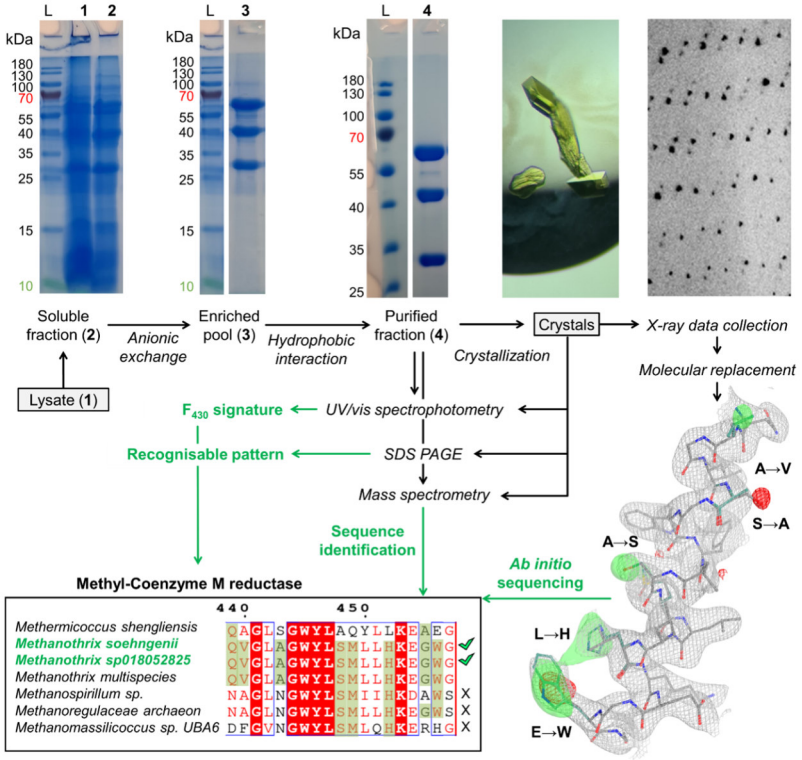


Figure 3. Scheme of the purification and strategies to identify *Methanothrix* MCR from activated sludge. From left to right, the top panel shows the SDS-PAGE of samples from the purification steps (L. Ladder, 1. total extract, 2. soluble fraction, 3. selected fraction during anionic exchange chromatography, 4. selected fraction during hydrophobic interaction chromatography), obtained MCR crystals and their diffraction pattern. Complete gel images are shown in Extended Data Fig. 1. Bottom right panel displays an example of sequencing in the electron density. A stretch of MCR modelled with the sequence of *Methermicoccus shengliensis* MCR (PDB 7NKG) is shown as a grey cartoon and sticks. The correct model of *Methanothrix* MCR is superposed and shown in teal. The $2F_o–F_c$ map is shown at 1 σ as a grey mesh, and the $F_o–F_c$ at 4 and −3 σ are shown in green and red, respectively. Positions of sequence deviation indicated by the electron density are marked. The alignment of the stretch in all MCRs found in the metagenome is shown on the bottom left panel. Green boxes indicate agreement with the electron density, and sequences that were excluded due to disagreement with the electron density are marked with an X, while fitting sequences are marked with a green hook.

were affiliated to *Methanothrix*, *Methanospirillum*, *Methanoregulaceae* and *Methanosassiliicoccus* from the original biological sample (Supplementary Table S1). To isolate a unique MCR population from possible variants, we utilized the selective process of crystallization.

Harvested crystals from this preparation exhibited the same typical UV/visible spectrum as a previously characterized MCR homologue (Extended Data Fig. 2a-b). The signature is compatible with the F_{430}-carrying enzyme in its $MCR_{ox1\text{-}silent}$ (Ni^{2+}) state. X-ray diffraction data was collected (Extended Data Table 2), and two independent strategies were tested to identify which MCR was trapped in the crystal. In the first strategy, the leftover protein in solution or washed crystals were passed on SDS-PAGE, and the corresponding bands were excised for mass spectrometry (MS) analyses. The resulting spectra searched against a reference database constructed with the obtained metagenome confirmed the presence of two MCRs annotated to *Methanothrix soehngenii* and *Methanothrix sp01805282* (from here on, abbreviated as *M. soehngenii*/*M. sp.*, respectively) in an approximative equal amount based on the peptide counts (Supplementary Table S2). In the alternative second strategy, the crystal structure was solved by molecular replacement using the homologue from *Methermicoccus shengliensis* (PDB 7NKG) as a template. The electron density map quality from the initial model refined to 1.98-Å allowed an accurate sequencing of the three subunits based on the available MCR sequences from the metagenome (Fig. 3). A detailed inspection of the electron density map confirmed a mix of different sidechains at certain positions (Extended Data Fig. 3) indicating a mixture of *M. soehngenii* and *M. sp.* MCRs. According to the data analyses, both methods (MS and sequencing in the electron density map) concluded an unexpected mix of *M. soehngenii* and *M. sp.* MCRs.

An overlay of two MCRs from *Methanothrix* arranged as a shamrock in the crystal. The two MCR sequences within the protein crystal are highly similar (94.99 %, 95.85 % and 95.29 % amino acid identity for the α-, β- and γ-subunit respectively, with 100 % coverage). Only a few of these substitutions are located at the crystalline packing contacts, explaining why the two complexes coexist in the same crystal and overlay perfectly (Fig. 4a, Supplementary Fig. S1). While a similar phenomenon of substitutional/occupational disorder is known in supramolecular chemistry[29], it has been rarely reported in protein crystallography and, on a smaller basis, referred to as microheterogeneity. Generating the final protein model that contains both MCR sequences required a particular procedure. First, we modelled the two MCR sequences separately, followed by refinement and validation, ignoring artefacts in the electron density caused by the alternative MCR sequence. Then, both structures were set to an occupancy of 50 % and overlaid as alternative conformers, resulting in a final mixed model in which further structural analyses were performed.

The final structural model contains three complete MCRs (*i.e.*, the asymmetric unit content), each MCR being an overlay of *M. soehngenii* and *M. sp.* The three MCRs are arranged as a shamrock, a triangular unit that was initially interpreted as a packing artefact. However, the trimeric organization was systematically observed in three other crystalline forms (Extended Data Fig. 4a-b and Supplementary Fig. S2). The cavities providing access to the active site are fully accessible to the solvent, allowing substrate diffusion (Fig. 4b). A trimeric MCR has never been structurally reported so far but has previously been proposed for the MCR of ANME-2d based on a blue native PAGE experiment[30]. While the above-mentioned arguments suggest that the shamrock shape MCR would be the physiological state in *Methanothrix*, deeper investigations through packing analyses (see materials and methods), analytical size exclusion chromatography (see materials and methods), high-resolution native PAGE (Extended Data Fig. 5a-b), and electron microscopy (Extended Data Fig. 5c-g), could not confirm

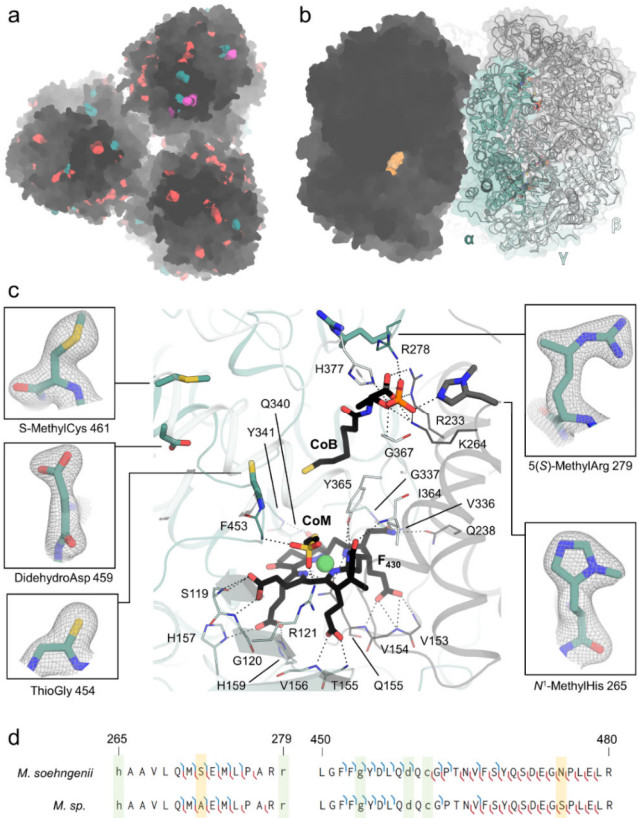


Figure 4. MCR structure of *Methanothrix*. (**a**) MCR trimer (black surface) top view with substitutions between *M. soehngenii/M. sp.* in red, crystal packing contacts in teal and substitution colocalizing with residues at the interface of the crystal packing in magenta. (**b**) Side view of the trimeric MCR, with one MCR shown as a black surface, its active site entrance highlighted in orange, and the second in a cartoon with transparent surface color coded with one αβγ unit in grey, and the second with each unit independently colored. (**c**) Detail of the active site with a transparent cartoon and residues interacting with ligands via hydrogen bonds (black dashes) shown in the sticks. Color coding of the chains is the same as (b) with oxygen, nitrogen, phosphorous, sulfur and nickel colored red, blue, orange, yellow, and green, respectively. PTMs are shown in surrounding panels with the $2F_o-F_c$ electron density map (grey mesh) contoured at 1 σ. (**d**) PTM-containing peptides detected by MS. Green boxes highlight PTM positions, and yellow boxes highlight sequence deviations between *M. soehngenii* and *M. sp.* See Extended Data Table 3 for coverage statistics.

with certitude the trimeric state. Since these in-solution methods are unable to mimic the high-intracellular MCR concentrations (*e.g.*, the trimeric organization might fall apart under protein concentration below 1 $mg.mL^{-1}$ and need high concentrations to stay associated), more investigations will have to be carried out with *in situ* methods or experimental studies across close lineages to confirm the existence of the trimeric MCR state.

A conserved organization with an active site equipped by five post-translational modifications.

Despite a few structural deviations (Extended Data Fig. 6a-c), comprising minor deletions and insertions as well as the presence of phosphate at the dimeric interface (Extended Data Fig. 6d-e), both MCR structures from *Methanothrix* show a highly conserved architecture compared to structural homologs (Supplementary Table S3). On the same line, the electrostatic surface analyses revealed similar characteristics as previously shown for MCR[31], with the entry of the catalytic cavity being mainly positively charged and a negatively charged outer region (Supplementary Fig. S3). Together, these specificities, located far from the active site and its entrance, are unlikely to affect the enzyme activity.

The active site contains a conventional F_{430}, coenzyme M and coenzyme B bound in the same fashion as previously described[31] (Fig. 4, Supplementary Fig. S4). It is identical between both *Methanothrix* structures and highly conserved with homologs since substitutions in the second coordination shell have no impact on the coordination of the cofactor and substrates (Supplementary Fig. S5). A putative gas channel (*e.g.*, to release methane), observed the Ethyl-CoM reductase homologue is not detectable[32] (Supplementary Fig. S6).

Methanogenic MCRs are known to harbor a plethora of post-translational modifications (PTMs) that are not correlated to phylogeny[33, 34] and whose functions are still unknown.

Examination of the electron density maps together with MS, showed the presence of a N^1-methylhistidine 265, 5(*S*)-methylarginine 279, thioglycine 454, didehydroaspartate 459 and S-methylcysteine 561 in both sequences (Fig. 4c-d, Extended Data Table 3).

MCR isolation directly from active sludge collected months later.

Following the successful isolation and structural elucidation of *Methanothrix* MCR from a physically enriched sample, the pipeline was tested on raw activated sludge collected four months later from the same site (sample B, Fig. 1). Due to the higher amount of mineral and organic particles that can damage the chromatography matrices, the raw cell suspension was subjected to centrifugation followed by ultracentrifugation. The purification protocol was optimized by including size exclusion chromatography and following the absorbance at 425 nm. The final enriched fraction containing MCR in the majority corresponds to about 1 % of the total cellular protein content after sonication (Extended Data Fig. 1d-g). High-resolution native PAGE exhibited a comparable profile as sample A (Extended Data Fig. 5b).

Since the metagenome confirmed the presence of the same *Methanothrix* species, we were confident that the results could be reproduced using the identical crystallization protocol. Exploitable crystals were harvested for X-ray diffraction. Dissolved crystals were passed on SDS-PAGE (Extended Data Fig. 1g) and concomitantly analyzed via MS to confirm the MCR identity. Compared to the MS analyses from sample A, the peptide analyses from the dissolved crystals of sample B present a predominance of *M. sp.* over *M. soehngenii* (Supplementary Table S2). Accordingly, the crystallographic data collected for these crystals and refined to 2.5-Å matched perfectly with the sequence from *M. sp.* In the absence of artefacts in the electron density map that would witness an MCR mixture, only *M. sp*. MCR was modelled for the sample B. Despite a slight deviation caused by an internal readjustment of the trimeric assembly (Extended Data Table 2, Extended Data Fig. 4, and Supplementary Fig. S2 and S7), the structures

are almost identical between samples A and B (Supplementary Data Table S3) and contain the same PTMs (Extended Data Fig. 7, except the methyl-Arg containing peptide was not covered in MS). Compared to sample A, harboring the fully reduced form of CoM and CoB, sample B contains a mixture of the reduced and oxidized states (*i.e.*, the reaction product CoM-S-S-CoB). This allows a depiction of the contacts when the product is bound to the protein (Extended Data Fig. 7) and illustrates the reproducibility of the approach.

Discussion

The isolation and study of catabolic enzymes from uncultured microorganisms have mainly been achieved by recombinant protein expression with limitations regarding the requirement of native machinery for the correct folding and insertion of complex metallocofactors and post-translational modifications. In contrast, characterizing proteins from native microbial communities will provide unbiased key information on the physiological function and reaction mechanisms critical to understanding cellular processes and metabolic roadmaps[35].

Here, we show that isolating proteins from a native microbial community is technically feasible despite the challenges of the biofilm matrix processing, biomass complexity, and the diversity of enzyme variants with close homology (*e.g.*, > 95 % identity). By performing a mechanical enrichment and using only simple chromatography equipment, we captured the core enzyme of acetoclastic methanogenesis from understudied *Methanothrix*, an important player in the environment and for biotechnology[15, 36], converting the fermentation product acetate to methane and CO_2. The presented work echoes the structural elucidation of the MCR from ANME-1 isolated from Black Sea mats[37], involved in methane oxidation in anoxic sediments. Instead of a physical enrichment, Shima and colleagues relied on the natural enrichment of microbial mats that constitute a defined stable population, well adapted to a particular

ecological niche. Our study pushed the frontiers of the concept further by isolating an enzyme from a raw activated sludge without an enrichment step. Here, the intrinsic microbial diversity of wastewater digesters covers a whole range of biological matter degraders, with its extreme complexity captured in the metagenomics snapshot (Fig. 1). More importantly, we also showed the reproducibility of the method since both samples were collected at four months of difference. Together, the results highlight the possibility of working with an even more heterogeneous and complex sample similar to environmental conditions. Therefore, targeting biological samples with reduced microbial diversity, like microbial mats, aggregates, flocks, or environmental sites favoring specialists through extreme abiotic factors or specific chemical compositions, are promising subjects. Soil samples harboring a wide diversity of microbes will be more challenging, except if mechanical enrichments are available by exploiting the cell morphology and the physical composition to decrease microbial heterogeneity.

The difficulty also lies with subspecies sharing protein sequences of very high identity that cannot be separated over the purification and even not during crystallization, as attested by our structural data and mass spectrometry identification. Based on the high stability of MCR (224 kcal.mol^{-1} ΔG^{diss}), it is unlikely that the variants from both *Methanothrix* species have interchanged subunits over the isolation process. However, since intersubunit swapping has been observed in recombinant systems[11], it cannot be excluded from future environmental studies, where fragile complexes outside their cellular environment might recompose with other proteins from close species.

Microheterogeneity in terms of sequences is a rare event in protein crystallization. For instance, two crambin isoforms were trapped within the same crystal that overlap perfectly (PDB 1EJG)[38]. More cases of microheterogeneity in terms of sequences or PTMs might have been overlooked because of insufficient resolution (*e.g.*, Crambin being at subatomic resolution), and MS is a potent approach to verifying the precise content of a protein sample, as it has been revealed in profiling the PTMs of each MCR

variant. Notably, this PTM profile is the same as the ones reported in *Methanosarcina barkeri* and *Methanothermobacter marburgensis* (the latter harboring an additional methylglutamine) and differs from the closest homologue *M. shengliensis*, illustrating once more the absence of correlation between sequence similarity and PTMs and reflecting the need of experimental data to attest their presence.

Despite the differences in PTM profile and a few structural details, MCRs from *Methanothrix* and homologous structures share the same overall organization and catalytic site, highlighting the evolutionary pressure on the sophisticated CH_4-generating reaction. It also indicates that the solved structures, captured close to "environmental" conditions, are consistent with structural homologs from cultivated methanogens. A molecular feature between the structures from samples A and B is still important when considering the physiological context: compared to sample A, sample B presents a partial occupancy of the reaction product CoM-S-S-CoB. This oxidized form was previously obtained by controlling the cellular redox state of hydrogenotrophic methanogens[39-41] before harvesting. In this case, cells have been harvested similarly, and the partially trapped oxidized state might reflect a difference in cell physiology between the two sampling points (*e.g.*, starvation versus feast, symbiotic interaction versus non-cooperation).

Proposed to be the rate-limiting step in methanogenesis[42] and a key enzyme of microbial carbon cycling during the biological matter degradation, MCR is also a pertinent model to establish the presented pipeline. Indeed, MCR is highly abundant, has a globular shape with a robust scaffold preventing decay and aggregation over time, a recognizable pattern on SDS-PAGE and a spectrophotometric signature of F_{430} cofactor that facilitated traceability during purification and crystallization. Nevertheless, we argue that the method and its concept can be extended to other enzymes from environmental samples as long as they are highly expressed and there are means for tracking them during the purification (*e.g.*, unique intrinsic properties and/or enzymatic assays). For

example, environmental enzymes used to acquire elementary components and conserve energy, such as nitrogenases in soils, bacteriorhodopsin in salt lakes, CAZymes in the sea during microbial blooms, CO_2-fixation machinery or ATP synthase apparatus, are important biological processes that could be applied to mimicry chemistry and could be investigated through the approach.

The key of the presented method lies in crystallization, representing the quintessence of protein separation. Crystallization benefited from constant technological progress, such as the extension of the crystallization space, miniaturization and high throughput robotics, seeding techniques, or crystallophores[43]. On the other hand, crystallography capabilities expanded through modern synchrotron beamline specificities, X-ray detectors, real-time experiments, or microcrystal electron diffraction techniques if crystals are limited in size and diffraction quality. Protein crystals can be utilized for analyses beyond structural studies, *e.g.,* measuring spectroscopic data to gain information about the cofactors[44]. Dissolved crystals can be further used for activity assays[45], spectroscopic analysis, or cryo-electron microscopy.

To conclude, the approach offers the opportunity to study the native catabolic enzymes from metabolically active environmental microbes that are more prone to maintaining cellular functions in contrast to lab-grown cultures optimized for fast growth. Venturing forth to apply the presented approach to other members of the microbial dark matter would open a new field at the interface of environmental microbiology and biochemistry to discover and understand enzymes involved in so far uncharted metabolic pathways.

Material and Methods

Sampling

Digested sludge from a 1,800 m^3 methanogenic digestion tower was obtained from a sample port as part of a recirculation unit at the local wastewater treatment plant in Osterholz-Scharmbeck, Germany. 10 L and 15 L activated sludge were collected on the 5th of December 2021 and the 10th of April 2022, respectively. The first batch of sludge was immediately used for the physical enrichment protocol, while the second batch was used for direct MCR purification.

Physical enrichment of *Methanothrix* from activated sludge

10 L of activated sludge were diluted 1:3 with 1/20 seawater (2 $g.l^{-1}$ Wigands sea salt) and well mixed. The whole procedure presented below has been performed aerobically. The diluted sludge was distributed into 5 L bottles and left at 21 °C overnight to allow the settling of particles. Overpressure was avoided by the addition of a needle for pressure relief. The floating particles were discarded, and the cell suspension was decanted. The remaining settled particles were also discarded. Cells were centrifuged at 500 × *g* for 30 min, and the supernatant was discarded. The pellet was resuspended in a few milliliters of the supernatant and centrifuged again in 45-50 mL portions at 1,000 × *g* for 2 min. The supernatant was transferred into 50 mL falcon tubes, and the pellet containing particles with nearly no grey surface was discarded. The supernatant was centrifuged again at 1,000 × *g* for 10 min. The resulting greyish pellet without black particles was stored at −80 °C until further use.

MCR purification from physical enrichment

The purification was performed during a microbiology practical by master students in four separate groups. The described protocol corresponds to a single group. One gram of wet cells was thawed between fingers, 4 mL of buffer A (50 mM Tris/HCl pH 8.0 and 2 mM dithiothreitol (DTT)) was added for pellet resuspension and gently vortexed.

The cell suspension was sonicated with an MS73 probe (SONOPULS Bandelin, Germany) at 75 % power, six cycles of 20 seconds with a 1-minute break in between on ice. The lysate was split into 4 x 1.25 mL and centrifuged for 30 minutes at 13,000 × *g* at 4 °C. The supernatant, containing an average of 18 mg of protein, was filtered through a 0.2 µm filter (Sartorius, Germany) and loaded at 1 $mL.min^{-1}$ onto 2 x 1 mL Q Sepharose HP column (GE Healthcare, Germany) equilibrated with buffer A. The elution has been performed by collecting 5 mL fractions of increasing NaCl concentration with the following steps: 200, 240, 280, 320, 360 and 1000 mM NaCl. The fractions of interest were pooled (280 mM fraction and adjacent fractions, varying slightly per group). One volume of buffer B (50 mM Tris/HCl pH 8.0, 2 M $(NH_4)_2SO_4$, 2 mM DTT) was added to the pool, followed by filtration through a 0.45 µm filter (Sartorius, Germany). The sample was loaded on a 1 mL Phenyl Sepharose HP column (GE Healthcare, Germany) equilibrated with buffer B. The sample was eluted stepwise with 1, 0.8, 0.6, 0.4, 0.2 and 0 M $(NH_4)_2SO_4$ and 4 mL per step were collected. The fractions of interest were pooled (0.8, 0.6, 0.4 M for three groups, in one case 0.4, 0.2 and 0 M). The pools of all four groups were combined and the buffer was exchanged for a storage buffer (50 mM Tris/HCl pH 8.0, 10 % v/v glycerol, 2 mM DTT) by ultrafiltration (concentrated with a 30-kDa centricon (Sartorius, Germany) with a final 1:32 dilution). The sample was then concentrated to a final volume of ~400 µL. The protein concentration of 30 $mg.mL^{-1}$ was estimated via Bradford assay. For each cell lysis and purification step, 15 µL was taken for analysis on SDS-PAGE. High-resolution clear native PAGE was performed as described in Lemaire *et al.*[46].

MCR purification directly from activated sludge

5 L of activated sludge collected from the wastewater treatment plant was freshly processed to purify MCR directly without physical enrichment. The biomass was harvested by centrifugation at 500 × g for 60 min at 4 °C, yielding a total of 1,680 g pellets, aliquoted in ~300 g cell pellets. Cell pellets were frozen in liquid N_2 and stored

at −80 °C until further use. 285 g of activated sludge pellet was thawed under warm water and diluted with 230 mL of buffer A. All further steps were performed on ice when possible. The cell lysate was homogenized by sonication with 10 cycles comprising one minute at 75 % intensity followed by a 1.5 min break (probe KE76, SONOPULS Bandelin). Cell debris, as well as other contaminations from the sludge, were removed via centrifugation (15,000 × *g*, 45 min at 4 °C) followed by ultracentrifugation of the supernatant (100,000 × *g*, 90 min, 4 °C). The supernatant, containing 96 mg of protein, was passed twice through a 0.2 µm filter (Sartorius, Germany) and loaded on a 5 mL Q Sepharose high-performance column (GE Healthcare, Germany) equilibrated with buffer A. A gradient of 0 to 0.5 M NaCl was applied for 60 min with a flow rate of 2 mL.min^{-1}. Fractions of 1.5 mL were collected. MCR eluted between 0.27 and 0.35 M NaCl. The MCR fractions were pooled (~ 25 mL) and diluted with 100 mL buffer B. The sample was passed through a 0.2 µm filter (Sartorius, Germany) and split into two pools. Each pool was successively loaded on a 1 mL Phenyl Sepharose high-performance column (GE Healthcare, Germany), previously equilibrated with buffer B. The loaded column was washed with 1 M $(NH_4)_2SO_4$ followed by elution of MCR using a gradient from 1 to 0 M $(NH_4)_2SO_4$ for 40 min at a flow rate of 0.6 mL.min^{-1}. Fractions of 0.7 mL were collected. MCR eluted between 0.62 and 0.37 M $(NH_4)_2SO_4$, and the respective fractions were pooled. The pooled MCR fractions of both runs were merged and concentrated in a 10-kDa-cutoff filter (Merck Millipore, Germany). The buffer was exchanged for storage buffer through ultrafiltration and concentrated to a volume of 300 µL. The sample was injected into a Superose 6 30/100 GL (GE Healthcare, Germany), previously equilibrated in a storage buffer. MCR eluted at a flow rate of 0.4 mL.min^{-1} in a sharp Gaussian peak at an elution volume of 20 mL. Considering the column specifics, the elution volume of 20 mL is aberrant for a protein complex estimated between 250-kDa (single MCR species) and 750-kDa (trimeric species) and is probably due to the interaction of the protein with the

chromatography matrix. The fractions containing MCR were concentrated in a 100-kDa-cutoff filter, and the final quantity of protein, 0.6 mg in 20 µL final volume, was estimated by the Bradford assay (Bio-Rad, Germany).

UV/Visible spectrophotometry

A UV/Vis spectrum of the sample purified from the physical enrichment was recorded with a Cary 60 (Agilent Technology, USA) at RT. The sample was measured at 1 and 4 $\mu g.\mu L^{-1}$. As controls, data previously published for MCR purified from *Methermicoccus shengliensis* (as described in Kurth & Müller *et al.*[31]) and *Methanothermococcus thermolithotrophicus* (as described in Hahn *et al.*[32]) are overlaid.

The crystallized MCR purified directly from sludge, as well as some crystals from *M. thermolithotrophicus* MCR (obtained as described in[47]) were measured. Here, UV-visible absorption spectra were measured on the crystals at 100 K on the *ic*OS setup using a DH-2000-BAL deuterium-halogen lamp (Ocean Optics) as the reference light source[48]. The lamp was connected to the incoming objective of the *ic*OS setup via a 200-µm-diameter fiber. A QE65 Pro spectrometer (Ocean Optics) was connected to the outgoing objective of the *ic*OS setup through a 400-µm-diameter fiber for data collection. A UV/Vis absorption spectrum was recorded on the mounted crystal. Absorbance values from 300-500 nm were normalized to the maximum value.

Transmission electron microscopy

MCR in solution and crystallized MCR were analyzed for its oligomerization state. MCR crystals were transferred to 0.5 µL crystallization solution (0.2 M NH_4Cl, 100 mM MES pH 6.0 and 20 % w/v PEG 6,000) and diluted 1:10 in 50 mM Tris pH 8.0 and 2 mM DTT. The in-solution sample was diluted to 1 $mg.mL^{-1}$ in 50 mM Tris pH 8.0, 0.02 % glutaraldehyde and 2 mM DTT and incubated for 30 min at RT. Finally, the sample was diluted to 0.03 $g.L^{-1}$. Both samples were adsorbed onto a copper grid

(FCF200-Cu) for 3 min. Grids were imaged at 200 kV with a JEOL JEM-2100 Plus transmission electron microscope.

Metagenome sequencing

Metagenome sequencing was performed on a Sequel IIe PacBio at the Max Planck-Genome center of Cologne. The Hifi reads were demultiplexed, ultra-low adapter trimmed, and deduplicated with SMRTLink v11.0.0. Samples A and B were merged and assembled with hifiasm-meta v0.13-r308 with default parameters[49]. PacBio metagenomics pipelines were used downstream for genome binning and taxonomic assignment. The 'Hifi-MAG-Pipeline' was used to identify high-quality MAGs from the metagenome assembly. In short, reads were aligned to previously assembled contigs with minimap2 v2.17[50, 51]. Binning was performed with MetaBAT2 v2.15[52] and DAS Tool v1.1.3[53]. Bin quality was assessed with CheckM v1.1.3[54, 55]., and taxonomy was assigned to the metagenomes using the search engine diamond v2.0.15, GTDB release207[56], and GTDB-Tk v2.0.0[57, 58]. Species diversity was analyzed with the PacBio metagenomics pipelines 'Taxonomic-Functional-Profiling-Protein' and 'MEGAN-RMA-Summary' using MEGAN v6.24.1[59].

Fluorescence *in situ* hybridization microscopy

The samples for fluorescence *in situ* hybridization (FISH) were fixed with 2 % (v/v) formaldehyde for 1 h at 21 °C. 1 µL of sample was pipetted into a reaction well on a glass slide (Paul Marienfeld GmbH & Co. KG, Germany). After air drying, the samples were washed with deionized water and air dried again. The cells were embedded in 0.1 % (w/v) low electroendosmosis (LE) agarose (Biozyme Biotech Trading GmbH, Austria) by adding 20 µL to a well and air drying overnight. The samples were hybridized with the tetralabelled FISH probe MX825[60] labelled with ATTO 594 (5' TCG CAC CGT GGC CGA CAC CTA GC 3') and newly designed helper oligonucleotides (5' GGC TTC CCT ACG GCA CCG ACA ACG G 3', 5' GAG TAT

CGT TTA CGG CTA GGA CTA CCC GGG 3', 5' GGA TGC TTC ACA GTA CGA ACT GAC GAC GGC 3', 5' GGA TGC TTC ACA GTA CGAACT GAC GAC GGC 3') in a hybridization buffer containing 50 % (v/v) formamide, 900 mM NaCl, 20 mM Tris/HCl (pH 8.0) and 0.01 % SDS (w/v) for 2.5 h at 46 °C. Afterwards, the samples were washed at 48 °C with washing buffer (28 mM NaCl, 20 mM Tris/HCl pH 8.0, 5 mM ethylenediaminetetraacetic acid (EDTA) and 0.01 % (w/v) SDS) for 15 min and in autoclaved deionized water for 1 min. The dried samples were covered with embedding medium citifluor-vectashield (4:1 v/v) with 1 $\mu g.mL^{-1}$ 4′,6′-diamidino-2-phenylindole (DAPI). Coverslips were added, and the samples were sealed with clear nail polish. The cells were visualized using an epifluorescence microscope (Zeiss Axioimager.D2, Carl Zeiss Microscopy, Germany).

Crystallization

MCR purified from the physical enrichment was immediately crystallized at a final concentration of 25 $mg.mL^{-1}$. The best MCR crystal was obtained by the sitting drop method at 20 °C in polystyrene MRC 2-well plates (SWISSCI, UK) containing 90 µL of crystallization solution in the reservoir. 0.7 µL protein was mixed with 0.7 µL reservoir solution. Yellow crystals appeared after a few days in the following crystallization conditions: 0.1 M Tris pH 8.5, 25 % w/v polyethylene glycol 3,350. Prior to freezing to liquid nitrogen, the crystals were soaked in the crystallization solution supplemented with 15 % v/v glycerol as a cryoprotectant for a few seconds.

The two other structures at low resolution presented in Extended Data Fig. 4 have been crystallized from the same preparation, in the following conditions: Alternative form 1 (orthorhombic form) was obtained from a crystallization solution containing 25 % w/v polyethylene glycol 3,350. The crystals were soaked in the crystallization solution supplemented with 20 % v/v glycerol for a few seconds prior freezing. Alternative form 2 (monoclinic form) was obtained with a slightly modified purification protocol in which the Q sepharose step elution was performed at 320 mM NaCl, and the Phenyl

Sepharose with a gradient step of 0.5 and 0 mM $(NH_4)_2SO_4$. Protein was crystallized at 23 mg.mL^{-1} in the same way as others. Crystals were obtained in 0.1 M Malic acid/MES/Tris (MMT) buffer pH 5, 25 % w/v polyethylene glycol 1,500. Crystals were soaked in 15 % glycerol for a few seconds prior freezing.

MCR purified directly from sludge was immediately crystallized at a concentration of 30 mg.mL^{-1}. The best MCR crystal was obtained by the sitting drop method at 20 °C in a polystyrene (SWISSCI) clover plate containing 100 µL of crystallization solution in the reservoir. 1.4 µL MCR at a concentration of 30 mg.mL^{-1} was mixed with 0.7 µL reservoir solution. Yellow plate-shaped crystals appeared after a few weeks in the following crystallization conditions: 0.1 M Tris pH 8.5, 25 % w/v polyethylene glycol 3,350. Prior freezing to liquid nitrogen, the crystals were soaked in the crystallization solution supplemented with 15 % v/v ethylene glycol as a cryoprotectant for a few seconds.

Mass spectrometry acquisition

Both samples were analyzed via mass spectrometry to determine the identity of the protein composing the final pools. For the MCR fraction purified from the physical enrichment, 11 µL was extracted from the crystallization plate and passed on SDS PAGE. For the sample purified directly from the sludge, MCR crystals were recovered after data collection, dissolved, and passed on SDS PAGE. The bands corresponding to the MCR α-, β- and γ-subunits were cut out and processed separately for in-gel digestion with a protocol adapted from Kurth & Müller *et al*. 2021[31].

Protein extraction from SDS-PAGE gels was started by excising the bands of interest and cutting them into 3 mm x 3 mm pieces. For destaining, the gel pieces were washed in different solvents/buffers and incubated for 10 minutes at room temperature, followed by removing the liquid. Solvents and buffers were added in the following order: 20 µL of acetonitrile, 20 µL of 50 mM ammonium bicarbonate, 20 µL of 30 % isopropanol with 50 mM ammonium bicarbonate buffer and 30 mM thioglycolic acid, 20 µL of

acetonitrile. These washing steps were repeated twice until the Coomassie blue stain was removed. For reduction and alkylation, 20 µL of 10 mM DTT in 50 mM ammonium bicarbonate buffer were added, the gel pieces were incubated for 30 minutes at 56 °C, and afterwards, the liquid was removed. Followed by another series of addition of solvent/buffers, incubation at room temperature for 10 minutes, and removal of the liquid in the order of 20 µL acetonitrile, 20 µ of iodoacetamide in 50 mM ammonium bicarbonate buffer, 20 µL acetonitrile, 20 µL ammonium bicarbonate buffer, 20 µL acetonitrile, 20 µL ammonium bicarbonate buffer. After this final washing round, the sample was subjected to trypsin digestion by adding 20 µL of 5 ng.µL^{-1} trypsin in 50 mM ammonium bicarbonate buffer and incubating for 30 minutes at room temperature. Then, 20 µL of ammonium bicarbonate buffer was added to the samples (without removing the trypsin liquid) and incubated overnight at 37 °C. Samples were sonicated for 20 seconds, 20 µL of 0.1 % v/v formic acid diluted in uHPLC water was added and incubated for 20 minutes at room temperature. The liquid was transferred to a nanoLC vial, and 20 µL of acetonitrile was added to the gel pieces and incubated for another 30 minutes at room temperature. Finally, the liquid from the gel pieces was combined with the previous liquid in the nanoLC vial and frozen at −80 °C until further use.

The samples were loaded in loading solvent A (98 % uHPLC water, 2 % acetonitrile, 0.05 % trifluoroacetic acid) onto a 300 m i.d. x 5 mm trap cartridge column packed with Acclaim PepMap100 C18, 5 um 100 Å (Thermo Fisher, 174500) using an Ultimate 3000 RSLCnano Liquid Chromatograph (Thermo Fisher Scientific, USA). The trap was connected to a 70 um x 50 cm analytical EASY-Spray column packed with PepMap RSLC C18, 2 um material, 100 Å (ES803, Thermo Fisher Scientific, USA), which was heated to 35 °C via the integrated heating module. The analytical column was connected via an Easy Spray source to a Q Exactive HF-X Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, USA). Peptides were separated on the analytical column at a flow rate of 300 nl.min^{-1} using a 140 min gradient. The gradient

went from 95 % buffer A (uHPLC water, 0.1 % formic acid) to 50 % buffer B (20 % uHPLC water, 0.1 % formic acid, 80 % acetonitrile) in 102 minutes, then to 99 % buffer B in 38 minutes. Eluting peptides were ionized via electrospray ionization (ESI) and analyzed in Q Exactive HF. Full scans were acquired in the Orbitrap at 70,000 resolution. The 15 most abundant precursor ions were selected in a data-dependent manner, isolated with the quadruple with a 1.2 m/z isolation window size, fragmented in the HCD cell with an NCE of 30, and measured in the Orbitrap at 17,500 resolution. Singly charged ions were excluded from MS/MS analysis. Dynamic exclusion was set to 25 seconds.

Protein identification

A custom database was generated by predicting protein sequences from MAGs sequenced from the sludge. Annotated protein sequences from the MAGs, as well as the unbinned contigs (annotated via PROKKA with the --metagenome and --kingdom Archaea settings[61]), were included with the addition of the cRAP protein sequence database (http://www.thegpm.org/crap) of common laboratory contaminants. The MS/MS spectra were searched against this database with the Sequest HT node in Proteome Discoverer vs. 2.5 (Thermo Fisher Scientific). The following parameters were used: trypsin (full), maximum two missed cleavages, 10 ppm precursor mass tolerance, 0.5 Da fragment mass tolerance, and maximum of 3 equal dynamic modifications per peptide: oxidation on N, W (+ 15.995 Da), Deaminated on N, Q, R (+0.984 Da), Thioglycine on G (+15.977 Da), Didehydro on D (-2.016 Da), Methyl on C, H, Q, R (+ 14.016 Da), acetylation on the protein N terminus (+42.011 Da). False discovery rates (FDRs) for peptide spectral matches were calculated and filtered using the Percolator Node in Proteome Discoverer. Percolator was run with a maximum delta Cn 0.05, a strict target FDR of 0.01, and a relaxed target FDR of 0.05 and validation based on q-value. The Protein FDR Validator Node in Proteome Discoverer was used to calculate

q-values for inferred proteins based on the results from a search against a target-decoy database (generally automatically by Proteome Discoverer by reversing the input sequences). Proteins with a q-value of <0.01 were categorized as high-confidence identifications, and proteins with a q-value of <0.01 were categorized as medium-confidence identifications. Post-translational modifications predicted for specific MS2 spectra were visually inspected both within Proteome Discoverer and the online peptide spectral annotator tool to validate the presence of beta and gamma ions that would confirm the presence of the PTM in multiple fragments of a single peptide (http://www.interactivepeptidespectralannotator.com/PeptideAnnotator.html).

Data processing, model building, refinement and validation

Data was collected at 100 K at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) synchrotron on BM07-FIP2 beamline for MCR obtained from the physical enrichment (sample A) and at PROXIMA-1 from the Source optimisée de lumière d'énergie intermédiaire du LURE (SOLEIL, Saint Aubin, France) synchrotron for MCR obtained directly from the sludge (sample B). Both datasets were processed and scaled with *autoPROC*[62]. The initial structure obtained from sample A was solved by using *PHENIX*[63] with the crystal structure from *Methermicoccus shengliensis* (PDB 7NKG) as a template. The correct sequences from the metagenome were identified via MS or by manual inspection of the electron density profile. An AlphaFold 2 model[64] was generated for the three different subunits and overlaid to the initial model from the molecular replacement. The structure of the MCR from sample B was solved by molecular replacement with the MCR model from sample A. Both models were further manually built with *COOT*[65] and refined with *PHENIX.* The electron density analyses and the refinement were assisted by the six non-crystallography symmetries from the three MCR complexes composing the asymmetric unit. Translational-libration-screw (TLS) was applied over the refinement, and hydrogens were added in the riding position in the last refinement cycles except for the MCR from sample B. Models were validated

through the MolProbity server[66]. Figures were generated with PyMOL (2.2.0 Schrödinger, LLC, USA).

Since the structure of the MCR isolated from sample A contains both *M. soehngenii* and *M. sp.*, two modelling strategies were tested to best reflect the crystal content. In the first strategy, the backbone and identical residues of *M. soehngenii* and *M. sp.* were modelled, and then the differing residues were inserted in the same position with an occupancy of 0.5. Alternatively, we modelled both sequences separately, and after refinements, each model was set to either alternative conformation A/B with an occupancy of 0.5. Finally, both models were merged and deposited into a single model. The latter method, recommended by the Protein Data Bank, was applied and deposited. For the case of the MCR structure from sample B, the more prominent sequence of *M. sp.* was modeled. Two additional datasets were refined to 4 Å in order to verify the formation of an MCR trimer in other crystalline forms/space groups. Both datasets were processed with *autoPROC*, and the two structures were solved by molecular replacement using *PHASER* with *M. sp.* MCR model from sample A as a template. A single refinement was run with *PHENIX* with default parameters.

Alignments were performed with MAFFT 7[67] and ESPript 3.0[68]. BioRender.com was used for scheme generation. The stability of MCR complexes was calculated via the PDBePISA server[69] (accessed 6th of June 2024). The server suggests that a trimeric assembly would be unstable under standard conditions, and the weak contacts at the MCR-MCR interfaces led us to perform further investigations to validate the assembly in solution.

Acknowledgements, author contributions and data availability

We would like to thank the Max Planck Institute for Marine Microbiology and the Max Planck Society for their continuous support. We are also grateful for the beam time allocation at Soleil and the support of the Proxima I beamline staff. Christina Probian and Ramona Appel from the Microbial Metabolism laboratory provided fantastic technical assistance to the study. We would like to thank the staff from the wastewater treatment plant in Osterholz-Scharmbeck for providing the sludge samples. Furthermore, we congratulate the IMPRS Marmic class 2026 (Ella Case, Mahum Farhan, Margaret Menke, Louison Nicolas-Asselineau, Leonard Rößler, Kinga Sántha, Sebastián Silva, Nicole von Possel, Yueqing Wang) who purified *Methanothrix* MCR from the physical enrichment during the microbial metabolism practical. Finally, we are thankful to Susanne Erdmann's (Institut für Microbiologie, Universität Innsbruck, Austria) help with electron microscopy collection and interpretation.

Sludge sampling and physical enrichment were performed by J.H., MCR, purification and crystallization from the physical enrichment was performed by M-C.M., M.J., N.M., T.W., MCR purification and crystallization directly from the sludge was performed by M.J. Structure refinement was performed by M.J., M-C.M., T.W. FISH microscopy was performed by A.L. X-ray data were collected by S.E., T.W., M.J. M-C.M. and N.M. *In cristallo* spectrophotometry has been collected and interpreted by S.E. Mass-spectrometry analyses, as well as peptide and PTM analysis, were performed by G.D., M.L. Metagenome sequencing and diversity analysis were carried out by S.D. and B.H. Transmission electron microscopy was performed by M.J., M-C. M., N.M., T.W. The manuscript was prepared by M-C.M, M.J and T.W with contributions of all co-authors.

Metagennomic data is currently being deposited on the Sequence Read Archive (SRA) under the BioProject number PRJNA1189308. Protein structures are currently being

deposited on the PDB. Mass-spec data is currently being deposited on the EMBL Proteomics Identification Database (PRIDE). Data will be made available upon publication or upon request.

Extended data

Extended Data Table 1. Genome assembly retrieved from the activated sludge.

Bin Name	Average Depth	Bin Complete-ness	Bin Conta-mination	GTDB_Taxonomy
p5537AB_bin.full.49	**43**	**100**	**0.65**	**d__Archaea;p__Halobacteriota;c__Methanosarcinia;o__Methanotrichales;f__Methanotrichaceae;g__Methanothrix;s__Methanothrix soehngenii**
p5537AB_bin.full.12	**31**	**84.97**	**3.27**	**d__Archaea;p__Halobacteriota;c__Methanosarcinia;o__Methanotrichales;f__Methanotrichaceae;g__Methanothrix;s__Methanothrix sp018052825**
p5537AB_bin.circ.1	24	96.77	6.18	d__Bacteria;p__Desulfobacterota;c__Syntrophia;o__Syntrophales;f__Smithellaceae;g__UBA8904;s__
p5537AB_bin.full.124	9	96.7	4.4	d__Bacteria;p__Cloacimonadota;c__Cloacimonadia;o__Cloacimonadales;f__Cloacimonadaceae;g__;s__
p5537AB_bin.full.128	11	86.06	4.3	d__Bacteria;p__Bacteroidota;c__Bacteroidia;o__Bacteroidales;f__TTA-H9;g__TTA-H9;s__TTA-H9 sp002071415
p5537AB_bin.full.147	8	98.39	0	d_Archaea;p_Thermoplasmatota;c_Thermoplasmata;o_Methanomassiliicoccales;f__Methanomassiliicoccaceae;g_UBA6;s_UBA6 sp002067635
p5537AB_bin.full.52	16	94.54	1.68	d__Bacteria;p__Desulfobacterota_G;c__Syntrophorhabdia;o__Syntrophorhabdales;f__Syntrophorhabdaceae;g__MWEV01;s__
p5537AB_bin.full.7	7	96.43	0	d__Bacteria;p__Caldisericota;c__Caldisericia;o__JAAYUI01;f__JAAYUI01;g__JAAYUI01;s__JAAYUI01 sp017992475
p5537AB_bin.full.87	11	88.46	0.85	d_Bacteria;p_Acidobacteriota;c_Thermoanaerobaculia;o_Thermoanaerobaculales;f_Thermoanaerobaculaceae;g_JAAYVZ01;s__JAAYVZ01 sp012516935

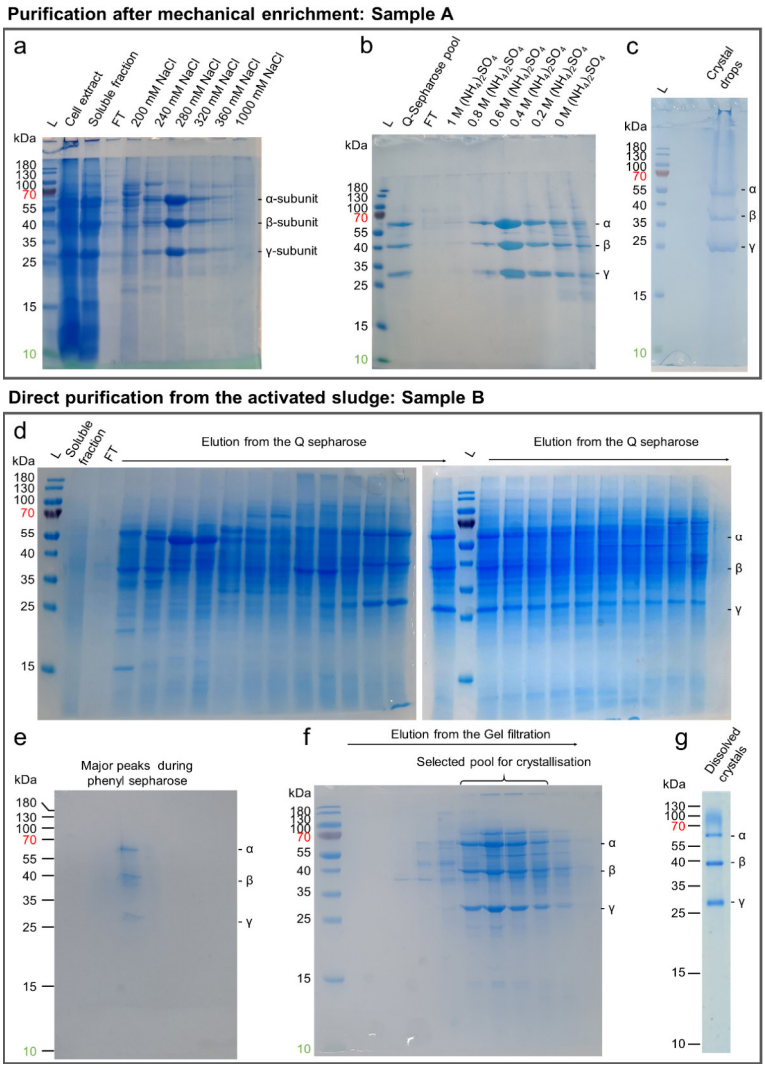
Extended Data Table 2. X-ray analysis statistics for *Methanothrix* MCR

	***Methanothrix* MCR mix from physical enrichment**	**MCR *M. soehngenii* from physical enrichment**	**MCR *M. sp.* from physical enrichment**	***Methanothrix* MCR directly from sludge**
Data collection				
Wavelength (Å)	0.97951	/	/	0.97856
Space group	*P* 1 2_1 1	/	/	*P* 1 2_1 1
Resolution (Å)	131.92 – 1.98 (2.21 – 1.98)	/	/	19.89 – 2.50 (2.72 – 2.50)
Cell dimensions				
a, b, c (Å)	116.39, 230.73, 143.63	/	/	118.18, 231.50 145.98
α, β, γ (°)	90, 113.30, 90	/	/	90, 112.60, 90
R_{merge} (%)[a]	15.4 (123.2)	/	/	14.7 (115.4)
R_{pim} (%)[a]	6.3 (49.2)	/	/	5.4 (42.9)
$CC_{1/2}$ [a]	0.996 (0.713)	/	/	0.997 (0.662)
I/σ_I[a]	9.3 (1.5)	/	/	10.0 (1.9)
Spherical completeness[a]	71.1 (12.8)	/	/	69.5 (15.3)
Ellipsoidal completeness[a]	92.2 (58.7)	/	/	94.3 (70.1)
Redundancy[a]	6.9 (7.1)	/	/	8.1 (7.9)
Nr. unique reflections[a]	342,102 (17,105)	/	/	173,000 (8,651
Refinement				
Resolution (Å)	59.73 – 1.98	59.73 – 1.98	59.73 – 1.98	19.89 – 2.50
Number of reflections	340,205	340,189	340,200	172,957
R_{work}/R_{free}[b] (%)	18.56/20.50	18.09/20.00	17.93/20.22	18.82/19.99
Number of atoms				
Protein	112,849	56,446	56,403	56,384
Ligands/ions	1,360	676	680	711
Solvent	2,535	2,636	2,632	397
Mean B-value ($Å^2$)	39.65	39.49	39.74	66.86
Molprobity clash score	2.45	2.03	2.28	6.20
Ramachandran plot				
Favoured regions (%)	96.72	96.74	96.72	96.75
Outlier regions (%)	0.08	0.08	0.08	0.35
rmsd[c] bond lengths (Å)	0.006	0.006	0.006	0.013
rmsd[c] bond angles (°)	0.850	0.827	0.874	1.541
PDB ID code	-	/	/	-

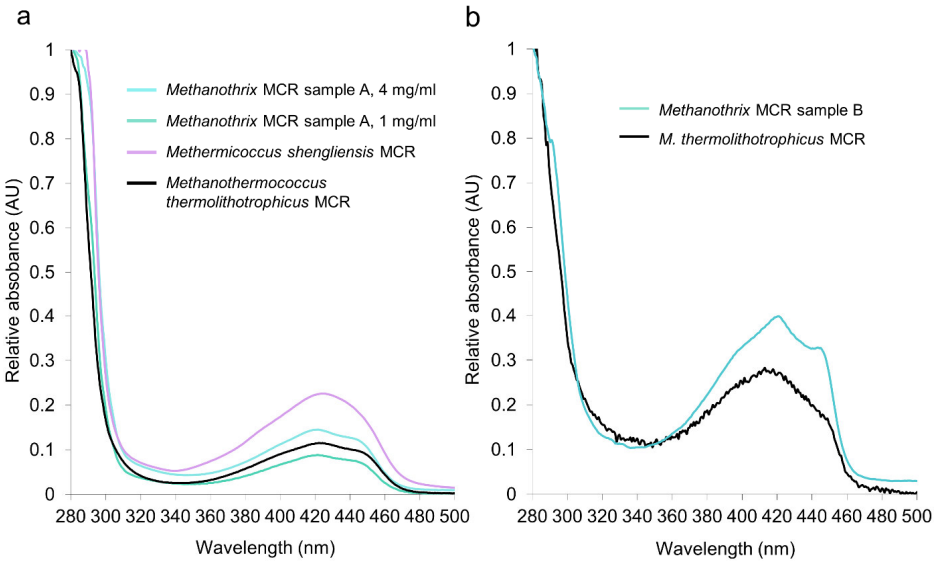
[a] Values relative to the highest resolution shell are within parentheses. [b] Rfree was calculated as the Rwork for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

Extended Data Table 3. Peptide coverage of post-translational modification (PTM) positions in combination with sequence-specific residues

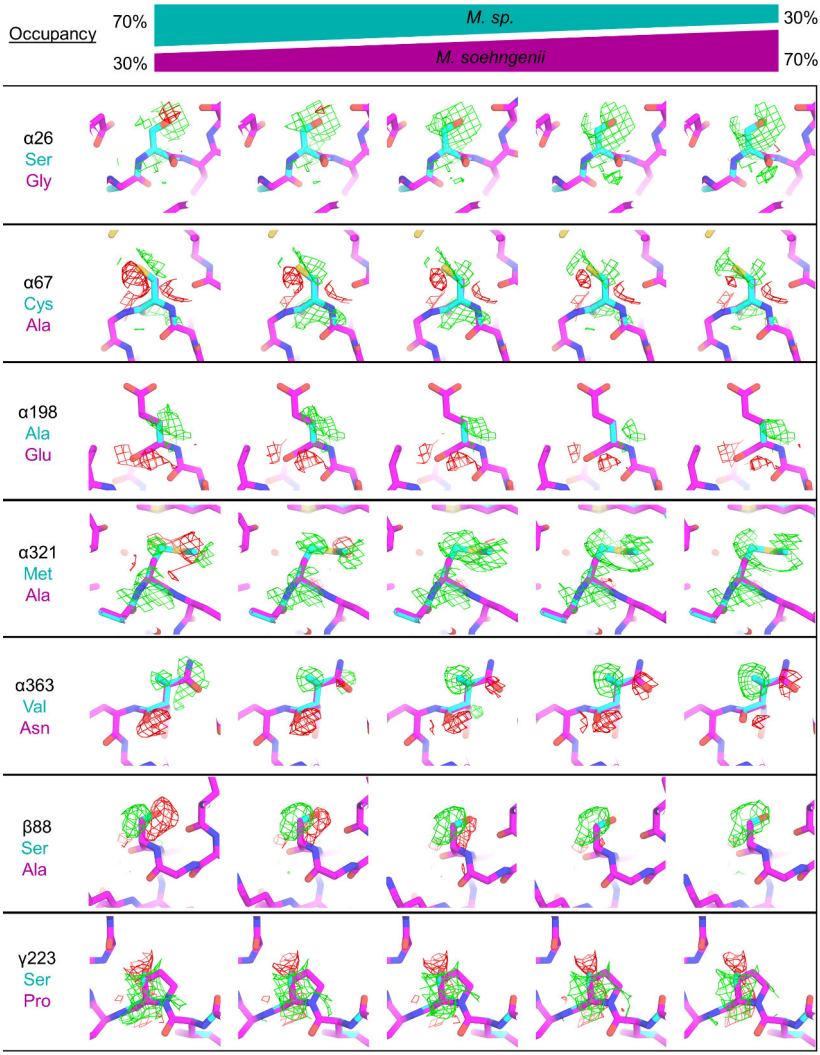
	Methyl-histidine 265	Methyl-Arginine 279	Thio-glycine 454	Didehydro-aspartate 459	Methyl-cysteine 461	454, 459, 461 on one peptide
Purified from the physical enrichment (sample A)						
Number of peptides specific to *M. sp.*	36	2	57	48	35	2
Number of peptides specific to *M. soehngenii*	102	1	54	50	44	3
Directly purified from sludge (sample B)						
Number of peptides specific to *M. sp.*	2	Not covered	102	100	55	8
Number of peptides specific to *M. soehngenii*	39	Not covered	43	36	19	5



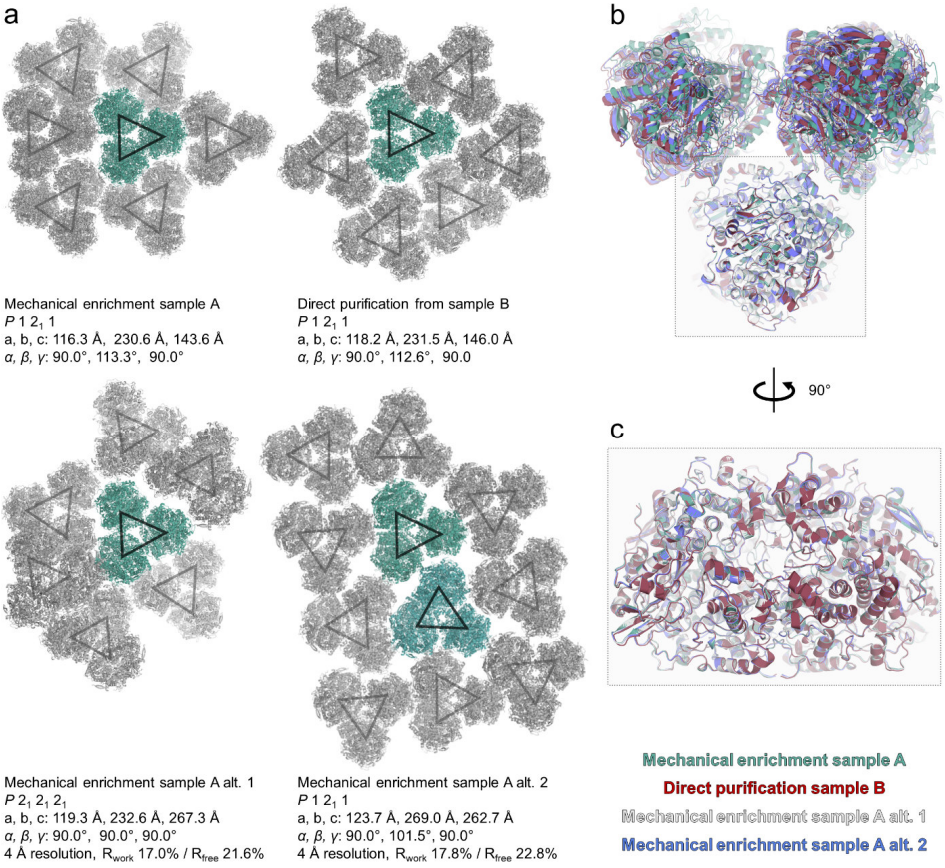
Extended Data Figure 1. SDS-PAGE profiles of the MCR native purification. (**a-b**) Q-sepharose (**a**) and Phenyl-sepharose (**b**) fractions from the physical enrichment. (**c**) MCR in crystallization solution purified from the physical enrichment. (**d-g**) Q-sepharose (d), Phenyl-sepharose (**e**), and gel filtration (**f**) fractions from processed sludge. (**g**) Dissolved MCR crystals of MCR purified from the sludge. L and FT stand for Ladder and Flow through, respectively.



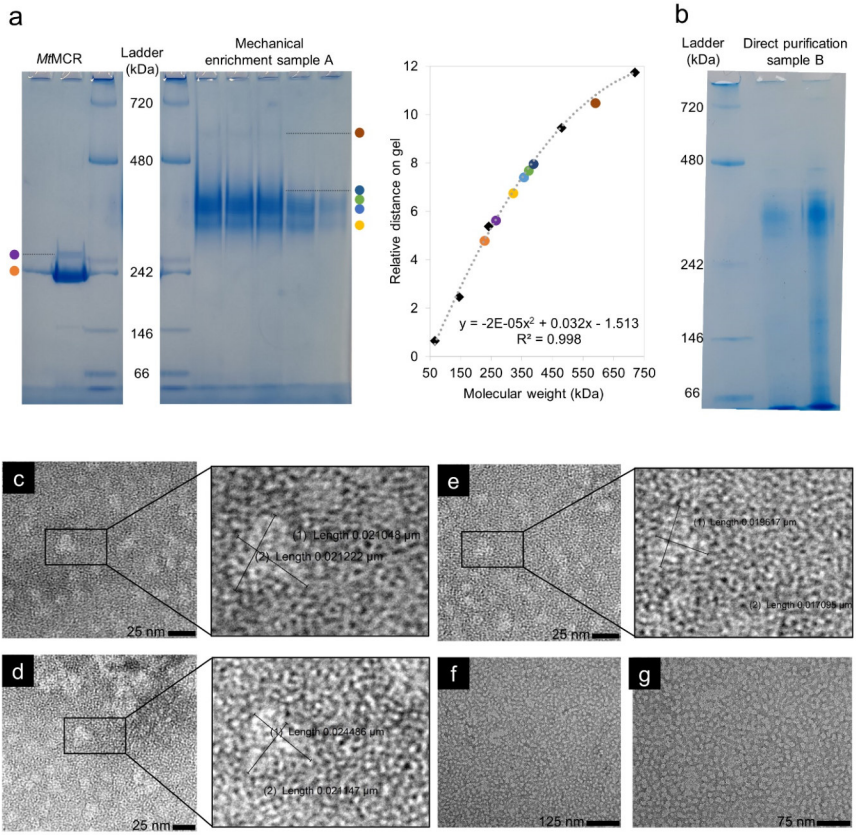
Extended Data Figure 2. UV/Vis-spectra of *Methanothrix* MCR. (**a**) The absorbance of soluble *Methanothrix* MCR and homologs normalized at 280 nm. Data for *Methermicoccus shengliensis* was plotted by Kurth *et al*. 2021[31] and for *Methanothermococcus thermolithotrophicus* by Hahn *et al.* 2021[32]. (**b**) The relative UV/Vis absorption spectrum was recorded on mounted crystals by the *ic*OS lab. Absorbance was normalized at 280 nm.



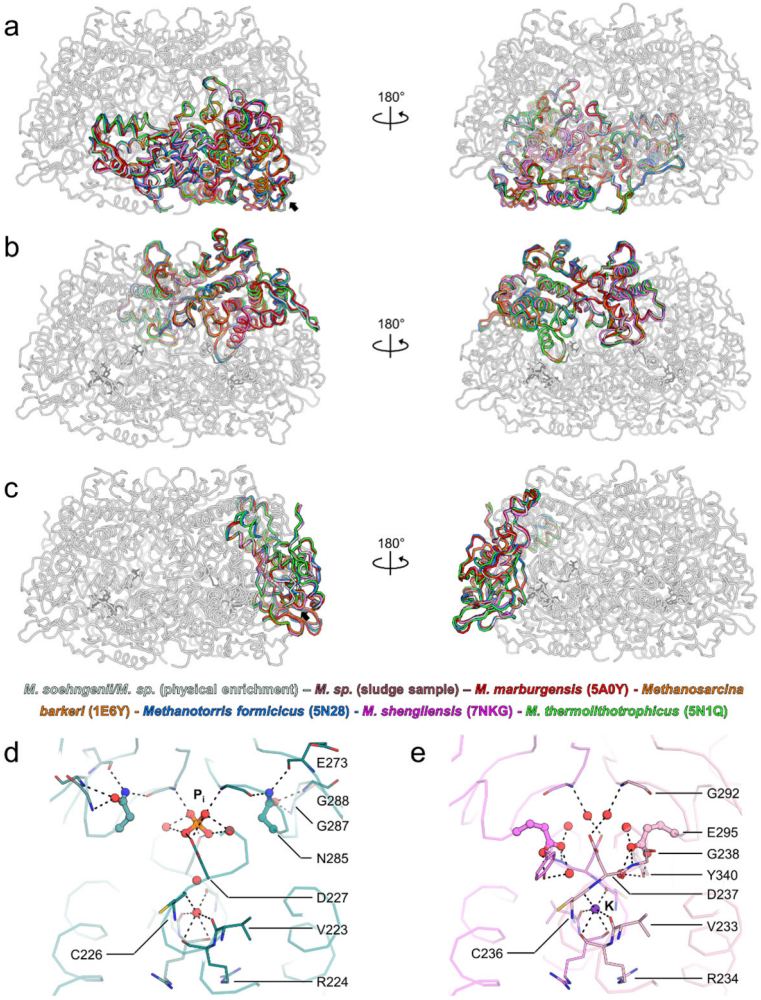
Extended Data Figure 3. Modelling of the different *Methanothrix* MCR sequences at different occupancies. *M. soehngenii* MCR is shown in teal and *Methanothrix sp.* in pink sticks with nitrogen, oxygen, and sulfur atoms colored in blue, red and yellow, respectively. The non-crystallographic symmetry (NCS) difference map is contoured at 2.5 and −2.5 σ in green and red, respectively. Based on this observation, a sequence ratio of 50 %/50 % was used for the final model.



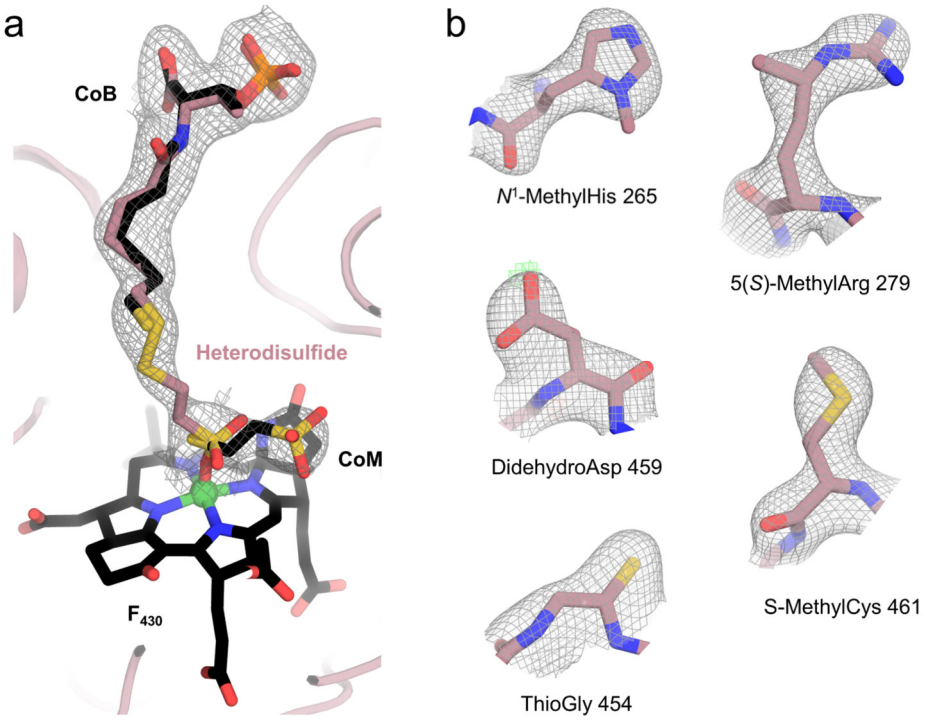
Extended Data Figure 4. Trimeric arrangements of MCR are redundant in other crystalline forms. (**a**) MCR crystal packing from different crystalline forms. All obtained crystalline forms systematically display the triangular arrangement. The asymmetric unit content is shown in teal, and the symmetry mates are grey cartoons. Symmetry mates were generated within 8-Å and are shown from the vertical plane. Triangles highlight MCR trimers. (**b**) MCR superposition of trimeric MCR from the crystalline forms shown in (**a**). Trimers are superposed on only one MCR, which is shown in a framed box. (**c**) Front view of the MCR superposition on a single unit from the crystalline forms shown in (**a**). Proteins are represented as cartoons. While independent MCR superposes fairly well, a shift is observable in the trimeric arrangement. This observation suggests that the MCR-MCR contacts are extensible and can recompose to some degree (Supplementary Fig. S2 and S7).



Extended Data Figure 5. Oligomerization of MCR on native PAGE and transmission electron microscopy. (**a**) Native PAGE of *M. thermolithotrophicus* MCR and *Methanothrix* MCR from the physical enrichment (sample A). Molecular weights were estimated by plotting the run distance of the gel bands. A major smeared double band at ~320 and 360-390 kDa, with a minor band at ~590 kDa at high protein quantities, were observed for *Methanothrix* (expected molecular weight of one and three MCRs being ~269.7 kDa and ~808.9 kDa, respectively). The observed pattern differed from the MCR model from the hydrogenotrophic methanogen *M. thermolithotrophicus* migrating at 228.6 kDa without a smear exhibiting an expected (2(αβγ)) composition. (**b**) Native PAGE of *Methanothrix* MCR purified from the sludge. (c-g) Transmission electron microscopy images of *Methanothrix* MCR. (**c-e**) Dissolved MCR crystals with zoom on spotted trimers, including measurements. (**f-g**) Soluble MCR. Mostly, single MCR and only very few trimeric entities on dissolved crystals are observable



Extended Data Figure 6. Superposition of *Methanothrix* MCR with structural homologs. (**a-c**) Superposition of the α (**a**), β (**b**) and γ (**c**) subunits of structural homologs on MCR *M. soehngenii*/*M. sp.* Arrows highlight a shortened loop at the α N-terminal and an extension close to the γ C-terminal wrapping around α. (**d-e**) Dimeric interface comparison between *Methanothrix* (d) and *M. shengliensis* (**e**) MCRs. Chains are in ribbons, coordinating residues in lines and interaction in grey dashes. The waters are red balls. Phosphate incorporation might be due to substituting the Asp with an Asn at position 285.



Extended Data Figure 7. Ligands and PTMs modelled in *Methanothrix* MCR directly purified from activated sludge. (**a**) Coenzyme B (CoB), coenzyme M (CoM) and F_{430} are shown as black sticks, and the heterodisulfide (CoM-S-S-CoB) is shown as pale red sticks. The protein backbone is shown as a pale red cartoon. Oxygen, nitrogen, sulfur, phosphorous, and nickel are colored red, blue, yellow, orange, and green, respectively. The $2F_o-F_c$ surrounding the CoM, CoB, and heterodisulfide is contoured at 1 σ and shown as grey mesh. (**b**) PTMs in the *Methanothrix* MCR directly purified from sludge are shown as sticks. $2F_o-F_c$ map for each PTM is contoured at 1 σ. Oxygen, nitrogen, and sulfur are colored red, blue, and yellow.

Supplementary materials

Supplementary table S1. MCR BLAST hits of the metagenome reads. Continued on next page.

Query	Subject Annotation	lenght	e-value	% ident.
Methanothrix thermoacetophila MCRα ABK14360.1	[Methanothrix MULTISPECIES] coenzyme-B sulfoethylthiotransferase subunit alpha	559	0.0	89.1
	[Methanothrix sp018052825] Methyl-coenzyme M reductase subunit alpha	559	0.0	88.4
	[Methanothrix soehngenii] Methyl-coenzyme M reductase subunit alpha	559	0.0	87.3
	[Methanospirillum sp.] coenzyme-B sulfoethylthiotransferase subunit alpha	562	0.0	64.9
	[Methanoregulaceae archaeon] coenzyme-B sulfoethylthiotransferase subunit alpha	561	0.0	67.0
	[Methanomassiliicoccus_sp._UBA6] Methyl-coenzyme M reductase I subunit alpha	558	0.0	62.0

Query	Subject Annotation	lenght	e-value	% ident.
Methanothrix thermoacetophila MCRβ ABK14363.1	[Methanothrix sp018052825] Methyl-coenzyme M reductase subunit beta	434	0.0	86.2
	[Methanothrix sp.] coenzyme-B sulfoethylthiotransferase subunit beta	434	0.0	85.9
	[Methanothrix soehngenii] Methyl-coenzyme M reductase subunit beta	434	0.0	85.5
	[Methanospirillum hungatei] coenzyme-B sulfoethylthiotransferase subunit beta	432	0.0	62.7
	[Methanoregulaceae archaeon] coenzyme-B sulfoethylthiotransferase subunit beta	432	0.0	64.4
	[Methanomassiliicoccus_sp._UBA6] Methyl-coenzyme M reductase I subunit beta	431	0.0	58.9
	[Methanospirillum sp.] coenzyme-B sulfoethylthiotransferase subunit beta	431	0.0	55.0

Query	Subject Annotation	lenght	e-value	% ident.
Methanothrix thermoacetophila MCRγ ABK14361.1	[Methanothrix sp018052825] Methyl-coenzyme M reductase subunit gamma	255	0.0	86.7
	[Methanothrix sp.] coenzyme-B sulfoethylthiotransferase subunit gamma	255	0.0	86.7
	[Methanothrix soehngenii] Methyl-coenzyme M reductase subunit gamma	255	0.0	85.9
	[Methanoregulaceae archaeon] coenzyme-B sulfoethylthiotransferase subunit gamma	258	0.0	57.8
	[Methanospirillum sp.] coenzyme-B sulfoethylthiotransferase subunit gamma	229	0.0	57.2
	[Methanomassiliicoccus sp. UBA6] Methyl-coenzyme M reductase II subunit gamma	227	0.0	58.2

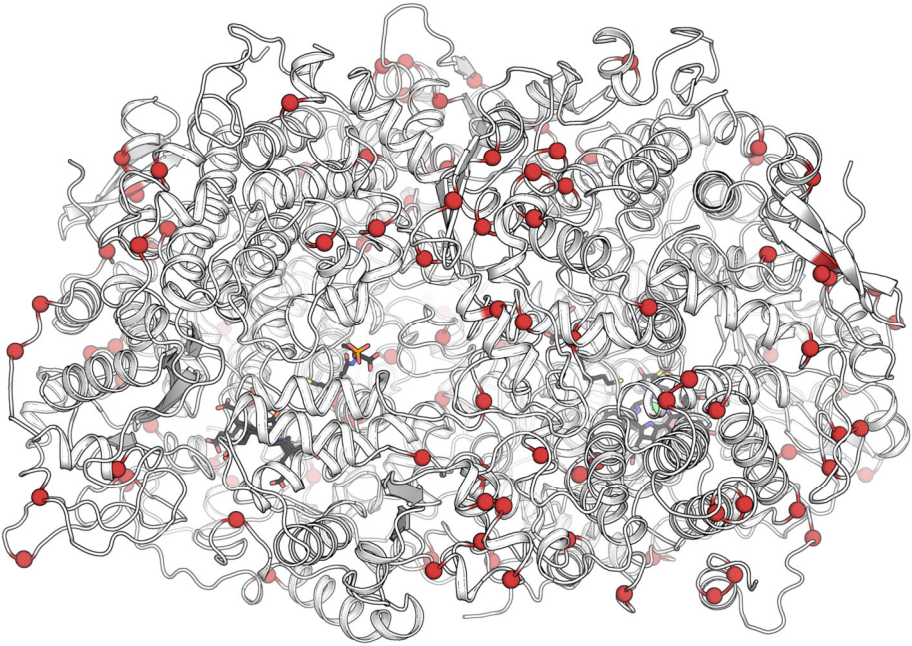
Supplementary table S2. Mass spectrometry peptide spectral matches.

MCR purified from the physical enrichment protocol	Peptide Spectral Matches (PSMs)					
Accession	**α**	**β**	**γ**	**α (duplicate)**	**β (duplicate)**	**γ (duplicate)**
Sample ID	**Alpha_11**	**Beta_13**	**Gamma_15**	**Alpha_21**	**Beta_23**	**Gamma_25**
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_alpha	**1966**	570	175	**2067**	568	233
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	**1904**	566	214	**1892**	512	247
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_beta	338	**1229**	61	321	**1331**	92
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta	313	**1241**	79	297	**1305**	107
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_gamma	229	127	**1094**	230	124	**1114**
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_gamma	236	179	**955**	191	157	**882**

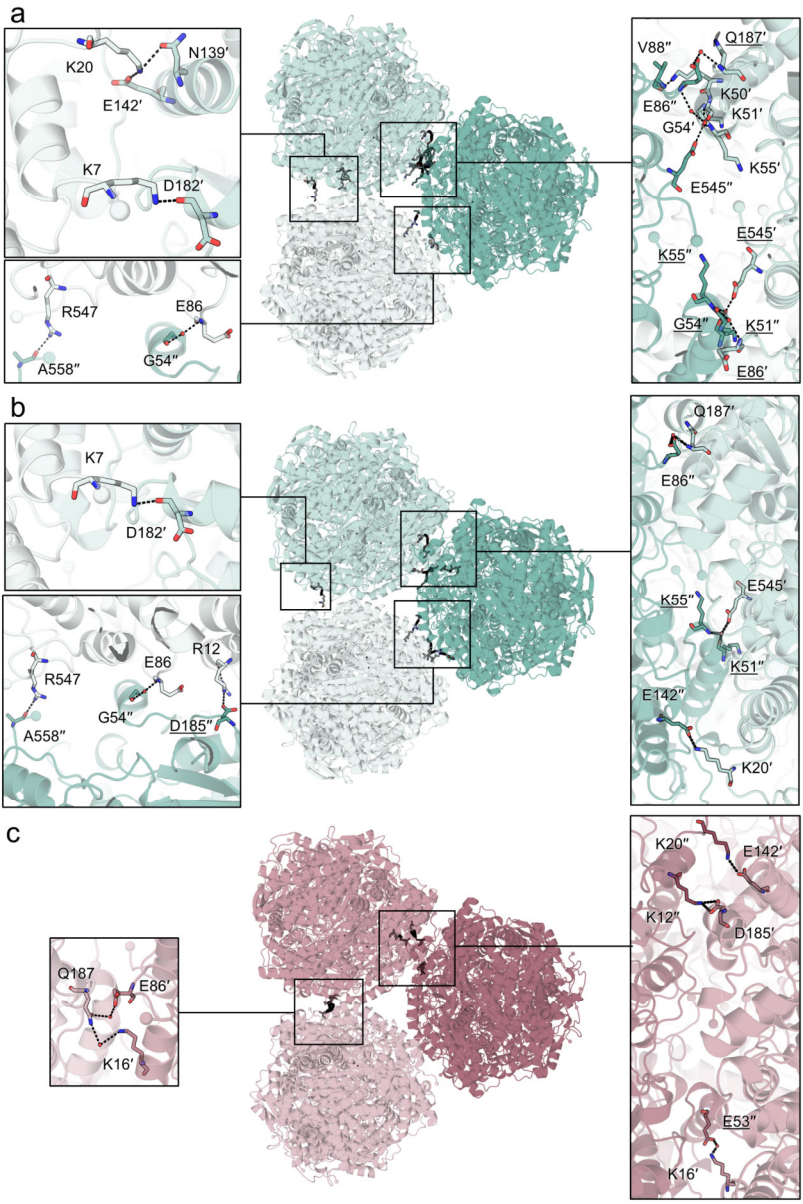
MCR directly purified from the sludge	Peptide Spectral Matches (PSMs)		
Accession	**α**	**β**	**γ**
Sample ID: 2ndExtraction	**Alpha_03**	**Beta_05**	**Gamma_07**
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_alpha	**604**	5	1
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	**1185**	19	3
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_beta	108	**790**	2
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta	217	**1459**	3
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_gamma	29	21	**249**
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_gamma	70	75	**700**

Supplementary Table S3. Sequence and structural identity of MCR homologs. Sequences of homologs were obtained from each PDB entry, which is indicated in brackets.

α chain	Seq ID	Coverage	RMSD (Å)	RMSD (Å, atoms)
Methanothrix sp. (sludge sample)	94.99%	100%	0.208	534
Methanothrix sp. (physical enrichment sample)	94.99%	100%	0	522
Methermicoccus shengliensis (7NKG)	67.14%	99%	0.371	461
Methanosarcina acetivorans (8GF5)	65.72%	98%	0.589	461
Methanosarcina barkeri (1E6Y)	65.19%	98%	0.399	461
Methanotorris formicicus (5N28)	63.02%	98%	0.51	481
Methanothermobacter marburgensis (5A0Y)	62.50%	98%	0.457	460
Methanothermococcus thermolithotrophicus (5N1Q)	60.75%	98%	0.489	472
β chain	Seq ID	Coverage	RMSD (Å)	RMSD (Å, atoms)
Methanothrix sp. (sludge sample)	95.85%	100%	0.183	416
Methanothrix sp. (physical enrichment sample)	95.85%	100%	0	422
Methermicoccus shengliensis (7NKG)	67.51%	100%	0.394	360
Methanosarcina acetivorans (8GF5)	64.98%	100%	0.618	402
Methanosarcina barkeri (1E6Y)	64.65%	99%	0.592	406
Methanotorris formicicus (5N28)	58.00%	99%	0.561	406
Methanothermobacter marburgensis (5A0Y)	57.04%	99%	0.588	408
Methanothermococcus thermolithotrophicus (5N1Q)	56.28%	99%	0.632	410
γ chain	Seq ID	Coverage	RMSD (Å)	RMSD (Å, atoms)
Methanothrix sp. (sludge sample)	95.29%	100%	0.176	237
Methanothrix sp. (physical enrichment sample)	95.29%	100%	0	246
Methermicoccus shengliensis (7NKG)	63.92%	100%	0.332	203
Methanosarcina acetivorans (8GF5)	61.96%	98%	0.517	222
Methanosarcina barkeri (1E6Y)	60.24%	99%	0.461	222
Methanothermobacter marburgensis (5A0Y)	50.99%	98%	0.657	214
Methanotorris formicicus (5N28)	48.84%	99%	0.557	204
Methanothermococcus thermolithotrophicus (5N1Q)	46.51%	99%	0.886	229

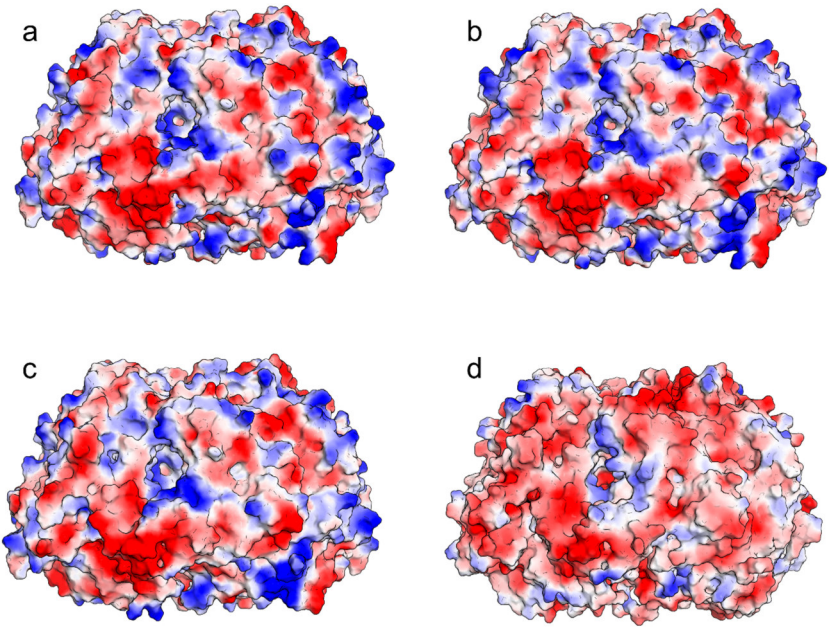


Supplementary Figure S1. Sequence differences between *M. soehngenii* and *M. sp. Methanothrix* MCR are displayed as a cartoon with ligands as black sticks. Oxygen, nitrogen, sulfur, phosphorous and nickels are colored red, blue, yellow, orange and green. Residues that are identical between both sequences are colored in grey, and residues that differ have their Cα highlighted by a red ball.

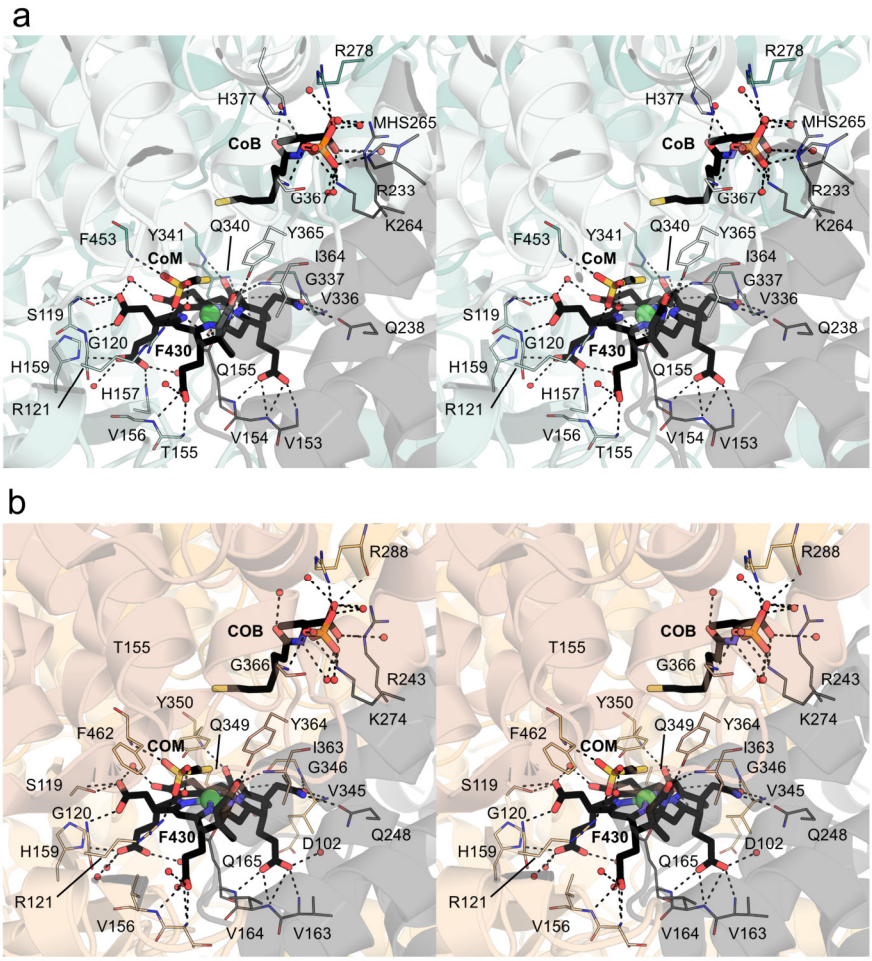


Supplementary Figure S2. Description on the following page

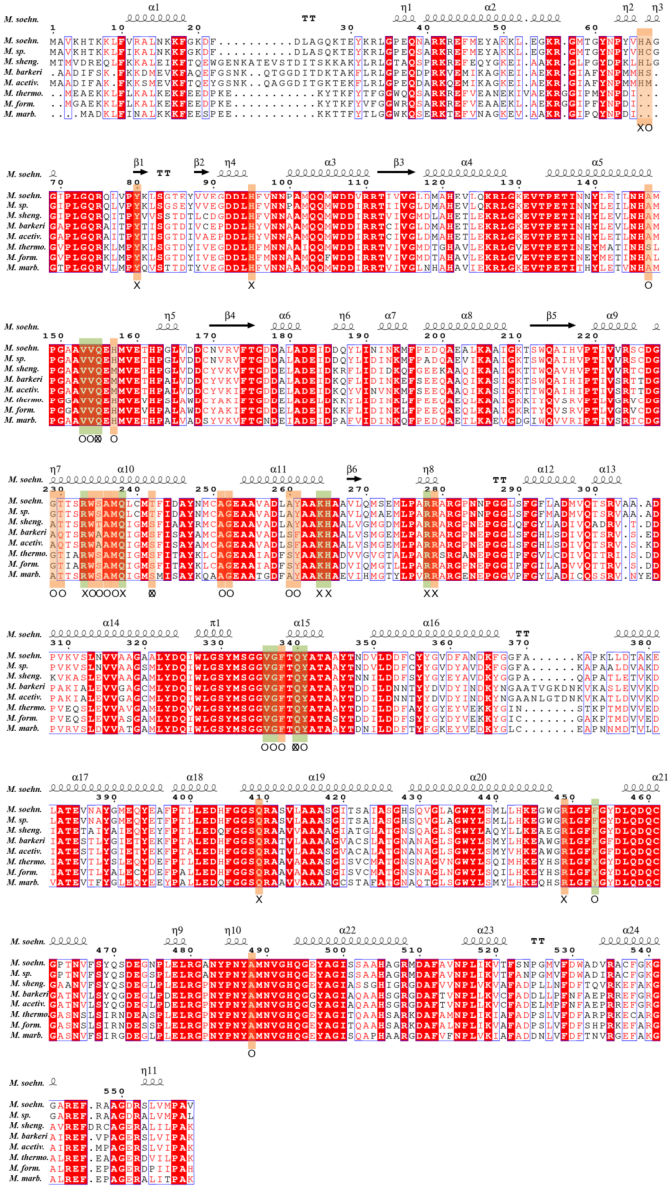
Supplementary Figure S2. Trimeric interface analyses between the different *Methanothrix* MCR. (**a**) Overview of the *M. sp.* MCR trimer purified from mechanical enrichment (sample A) with each unit shown in cartoon colored in a different shade of cyan. (**b**) Overview of the *M. soehngenii* MCR trimer purified from mechanical enrichment (sample A) with each unit shown in cartoon colored in a different shade of cyan. (**c**) Overview of the *M. sp.* MCR trimer directly purified from sludge (sample B) with each unit shown in cartoon colored in a different shade of red. For all, close-ups of the interfaces are shown in framed panels. Residues forming polar contacts (including up to one water) between MCRs are shown as sticks with dashed lines. Oxygen and nitrogen are colored red and blue, respectively. Waters are shown as red spheres. Prime marks indicate residues belonging to different MCRs, and underlined residues belong to varying chains within the same MCR dimer. All residues belong to an α-chain.



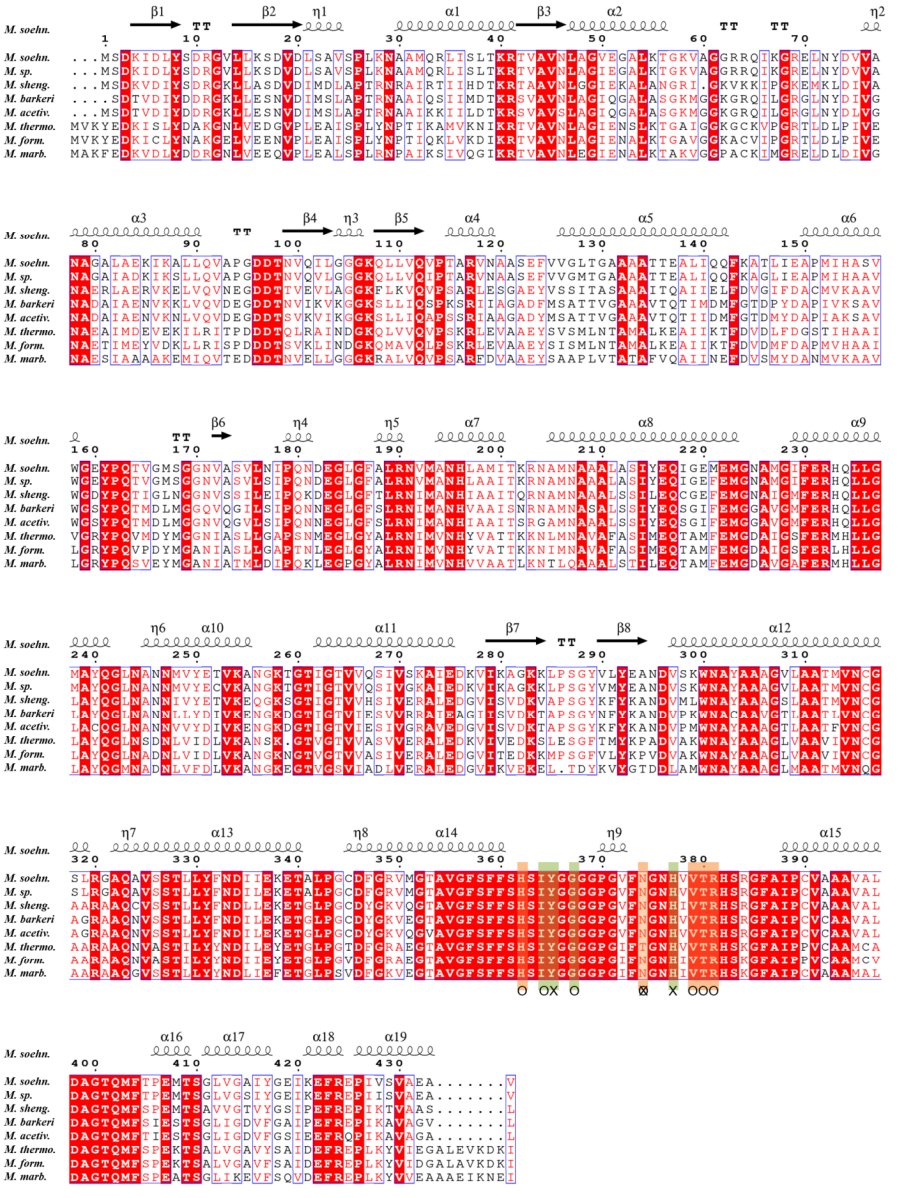
Supplementary Figure S3. Surface charge profile. (**a**) *Methanothrix* MCR purified from the mechanical enrichment modelled with the *M. sp.* sequence or (**b**) the *M. soehngenii* sequence. (**c**) MCR purified directly from the sludge, compared to (**d**) *M. shengliensis* MCR (7NKG). The electrostatic potential is colored from red (negatively charged) to blue (positively charged).



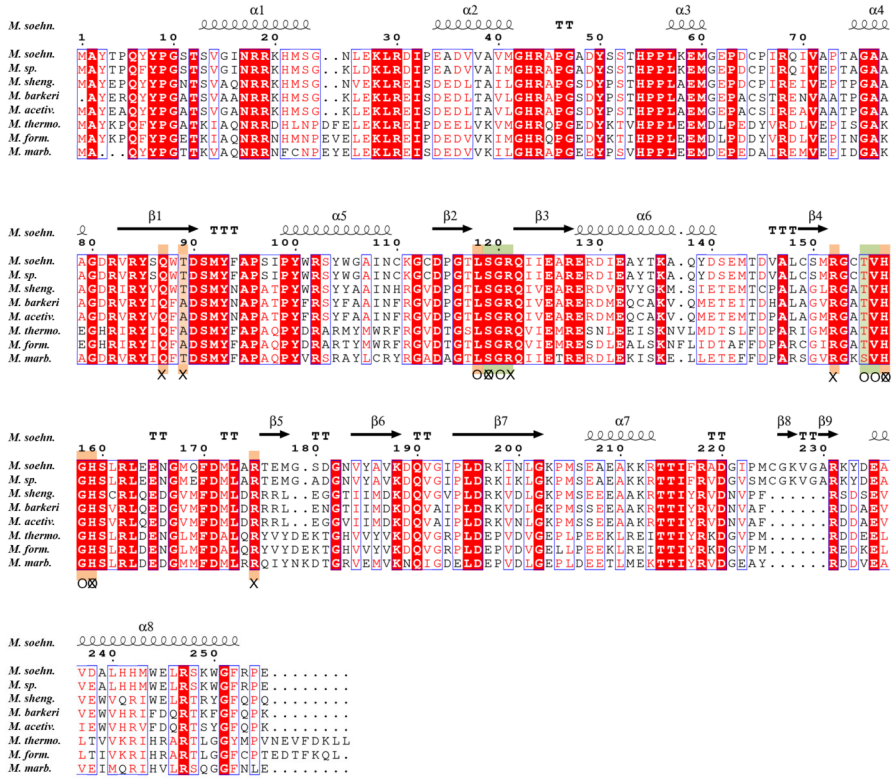
Supplementary Figure S4. MCR ligand binding site. (**a**) Stereo images of the ligand binding site in *Methanothrix* MCR purified from the physical enrichment and (**b**) *M. shengliensis* (7NKG). The protein is depicted as a transparent cartoon with different chains in different shades. The ligands F_{430} (F430), Coenzyme M (COM) and Coenzyme B (COB) are depicted as black sticks. Oxygen, nitrogen, sulfur, phosphorous, and nickel are colored red, blue, yellow, orange and green, respectively. Polar contacts are shown as black dashes.



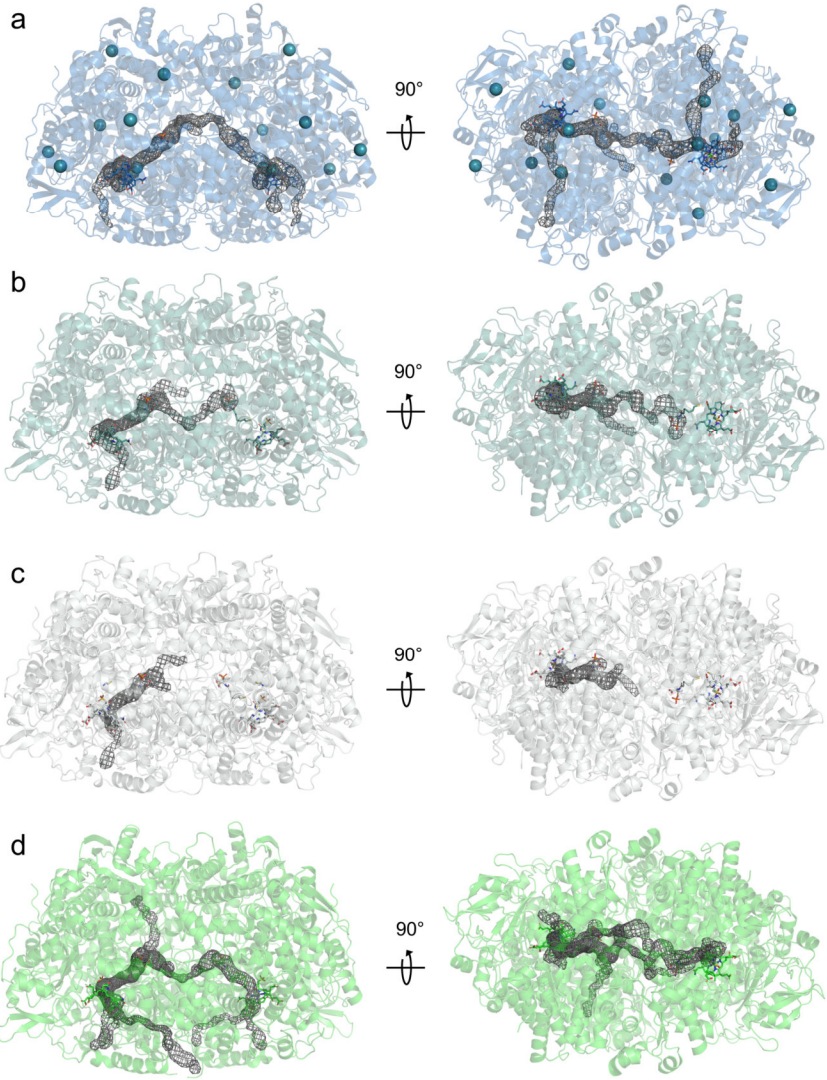
Supplementary Figure S5. Continued on next page.



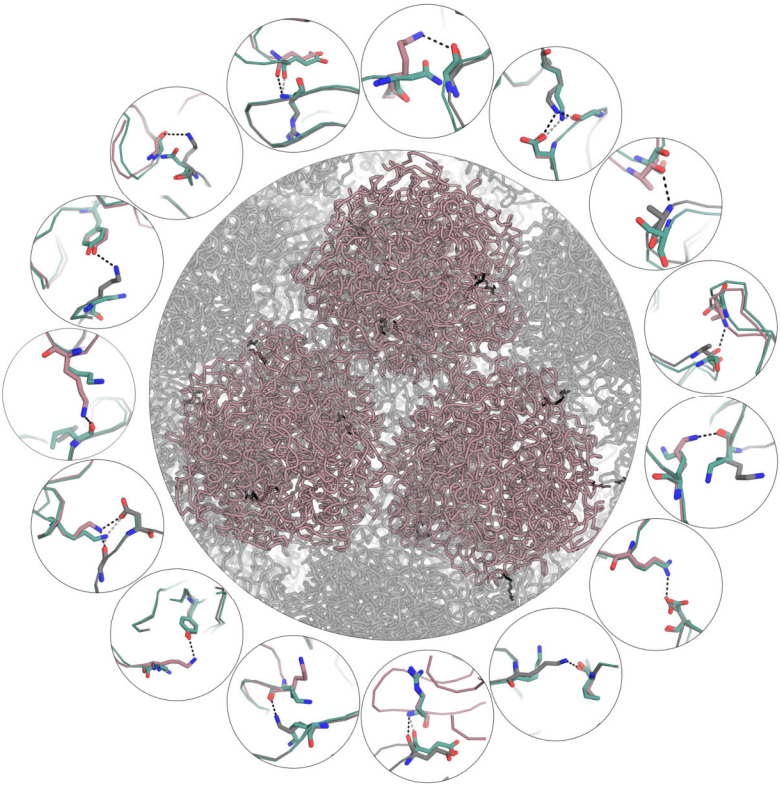
Supplementary Figure S5. Continued on next page.



Supplementary Figure S5. Sequence alignment of the α, β and γ subunits of MCR from *Methanothrix*. Residues coordinating F_{430}, CoM and/or CoB for the primary and secondary coordination shell (based on *Methanothrix* MCR) are highlighted in green and orange, respectively. Coordination via the main chain or side chain is labelled with an O and/or X, respectively. *M. soehng.* stands for *Methanothrix soehngenii*, *M. sp.* stands for *Methanothrix sp01805282*, *M. sheng.* stands for *Methermicoccus shengliensis*, *M. barkeri* stands for *Methanosarcina barkeri*, *M. acetiv.* stands for *Methanosarcina acetivorans*, *M. thermo.* stands for *Methanothermococcus thermolithotrophicus*, *M. form.* stands for *Methanotorris formicicus*, *M. marb.* stands for *Methanothermobacter marburgensis*.



Supplementary Figure S6. Channel prediction in different MCRs. Gas channels are predicted with Caver using a probe radius of 1 Å. *Candidatus* Ethanoperedens thermophilum (**a**, blue), *M. soehngenii* (**b**, teal), *M. sp.* (**c**, white), and *M. thermolithotrophicus* (5N1Q) (**d**). Models are in cartoons, with predicted channels as a black mesh. Xe atoms from *Ca.* E. thermophilum (7B2C) are shown as spheres.



Supplementary Figure S7. Packing contacts in *Methanothrix* MCR purified from sludge compared to the sequence of *M. soehngenii*. Individual *M. soehngenii* MCR (from sample A) was superposed onto each *M. sp.* MCR (from sample B), constituting the asymmetric unit of the crystal to investigate if sequence differences would cause exclusion from the packing. The central panel shows the MCR trimer of the asymmetric unit in light red and residues in the packing as grey ribbons. Residues forming packing contacts are shown as sticks, and contacts as dashed lines. The outer panels display zoom-ins on each contact and additionally the superposed residues of *M. soehngenii*. Polar contacts formed by *M. sp.* are shown in black and by *M. soehngenii* in grey. Oxygen and nitrogen are colored red and blue, respectively. This analysis shows that packing contacts revealed no clashes when replicated with the sequence of *M. soehngenii*. It suggests that the absence of MCR *M. soehngenii* in the crystal from sample B is due to its low quantities in the purified pool rather than its exclusion upon crystal formation.

References

1. Lloyd, K. G., Steen, A. D., Ladau, J., Yin, J., Crosby, L. Phylogenetically novel uncultured microbial cells dominate earth microbiomes. *mSystems* **3**, doi:10.1128/mSystems.00055-18 (2018).
2. Amann, R., Fuchs, B. M. Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. *Nat. Rev. Microbiol.* **6**, 339-348, doi:10.1038/nrmicro1888 (2008).
3. Singh, J., Behal, A., Singla, N., Joshi, A., Birbian, N., Singh, S., Bali, V., Batra, N. Metagenomics: Concept, methodology, ecological inference and recent advances. *Biotechnol J.* **4**, 480-94, doi:10.1002/biot.200800201 (2009).
4. Sorek, R., Cossart, P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat. Rev. Genet.* **11**, 9-16, doi:10.1038/nrg2695 (2010).
5. Schneider, T., Riedel, K. Environmental proteomics: Analysis of structure and function of microbial communities. *Proteomics* **10**, 785-798, doi:10.1002/pmic.200900450 (2010).
6. Aguiar-Pulido, V., Huang, W. R., Suarez-Ulloa, V., Cickovski, T., Mathee, K., Narasimhan, G. Metagenomics, metatranscriptomics, and metabolomics approaches for microbiome analysis. *Evol. Bioinf.* **12**, 5-16, doi:10.4137/Ebo.S36436 (2016).
7. Rosano, G. L., Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **5**, doi:10.3389/fmicb.2014.00172 (2014).
8. Albers, S. V., Jonuscheit, M., Dinkelaker, S., Urich, T., Kletzin, A., Tampé, R., Driessen, A. J. M., Schleper, C. Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Appl. Environ. Microbiol.* **72**, 102-111, doi:10.1128/Aem.72.1.102-111.2006 (2006).
9. Fink, C., Beblawy, S., Enkerlin, A. M., Mühling, L., Angenent, L. T., Molitor, B. A shuttle-vector system allows heterologous gene expression in the thermophilic methanogen *Methanothermobacter thermoautotrophicus* ΔH. *mBio* **12**, doi:10.1128/mBio.02766-21 (2021).
10. Xu, Q., Du, Q., Gao, J., Chen, L., Dong, X. Z., Li, J. A robust genetic toolbox for fine-tuning gene expression in the CO_2-Fixing methanogenic archaeon *Methanococcus maripaludis*. *Metab. Eng.* **79**, 130-145, doi:10.1016/j.ymben.2023.07.007 (2023).
11. Shao, N. N., Fan, Y., Chou, C. W., Yavari, S., Williams, R. V., Amster, I. J., Brown, S. M., Drake, I. J., Duin, E. C., Whitman, W. B., et al. Expression of divergent methyl/alkyl coenzyme M reductases from uncultured archaea. *Commun. Biol.* **5**, doi:10.1038/s42003-022-04057-6 (2022).
12. Ritthausen, H. Krystallinische Eiweisskörper aus verschiedenen Oelsamen. *J. Prakt. Chem.* **23**, 481-486, doi:10.1002/prac.18810230141 (1881).
13. Lyu, Z., Rotaru, A. E., Pimentel, M., Zhang, C. J., Rittmann, S. K. M. R. Editorial: The methane moment - Cross-boundary significance of methanogens: Preface. *Front. Microbiol.* **13**, doi:10.3389/fmicb.2022.1055494 (2022).
14. Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W., Hedderich, R. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **6**, 579-91, doi:10.1038/nrmicro1931 (2008).
15. Smith, K. S., Ingram-Smith, C. *Methanosaeta*, the forgotten methanogen? *Trends Microbiol.* **15**, 150-5, doi:10.1016/j.tim.2007.02.002 (2007).
16. Huser, B. A., Wuhrmann, K., Zehnder, A. J. B. *Methanothrix soehngenii* gen. nov. sp. mov., a new acetotrophic non-hydrogen-oxidizing methane bacterium. *Arch. Microbiol.* **132**, 1-9, doi:10.1007/Bf00690808 (1982).
17. Jetten, M. S. M., Stams, A. J. M., Zehnder, A. J. B. Methanogenesis from acetate: A comparison of the acetate metabolism in *Methanothrix soehngenii* and *Methanosarcina* spp. *FEMS Microbiol. Rev.* **8**, 181-197, doi:10.1111/j.1574-6968.1992.tb04987.x (1992).
18. Touzel, J. P., Prensier, G., Roustan, J. L., Thomas, I., Dubourguier, H. C., Albagnac, G. Description of a new strain of *Methanothrix soehngenii* and rejection of *Methanothrix concilii* as a synonym of *Methanothrix soehngenii*. *Int. J. Sys. Bacteriol.* **38**, 30-36, doi:10.1099/00207713-38-1-30 (1988).
19. Ferry, J. G. Methane from acetate. *J. Bacteriol.* **174**, 5489-5495, doi:10.1128/Jb.174.17.5489-5495.1992 (1992).

20. Jeris, J. S., McCarty, P. L. The biochemistry of methane fermentation using C^{14} tracers. *J. Water Pollution Cont. Fed.* **37**, 178-192, doi:- (1965).
21. Mah, R. A., Hungate, R. E., Ohwaki, K. Acetate, a key intermediate in methanogenesis. In *Microbial Energy Conversion* (eds Schlegel, H. G.; Barnea, J.) 97-106 (Pergamon, 1977).
22. Fey, A., Conrad, R. Effect of temperature on carbon and electron flow and on the archaeal community in methanogenic rice field soil. *Appl. Environ. Microb.* **66**, 4790, doi:10.1128/Aem.66.11.4790-4797.2000 (2000).
23. Holmes, D. E., Shrestha, P. M., Walker, D. J. F., Dang, Y., Nevin, K. P., Woodard, T. L., Lovley, D. R. Metatranscriptomic evidence for direct interspecies electron transfer between *Geobacter* and *Methanothrix* species in methanogenic rice paddy soils. *Appl. Environ. Microbiol.* **83**, doi:10.1128/AEM.00223-17 (2017).
24. Angle, J. C., Morin, T. H., Solden, L. M., Narrowe, A. B., Smith, G. J., Borton, M. A., Rey-Sanchez, C., Daly, R. A., Mirfenderesgi, G., Hoyt, D. W., et al. Methanogenesis in oxygenated soils is a substantial fraction of wetland methane emissions. *Nat. Commun.* **8**, doi:10.1038/s41467-017-01753-4 (2017).
25. Tveit, A. T., Urich, T., Frenzel, P., Svenning, M. M. Metabolic and trophic interactions modulate methane production by Arctic peat microbiota in response to warming. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E2507-16, doi:10.1073/pnas.1420797112 (2015).
26. Vítezová, M., Kohoutová, A., Vítez, T., Hanisáková, N., Kushkevych, I. Methanogenic microorganisms in industrial wastewater anaerobic treatment. *Processes* **8**, doi:10.3390/pr8121546 (2020).
27. Rotaru, A. E., Shrestha, P. M., Liu, F. H., Shrestha, M., Shrestha, D., Embree, M., Zengler, K., Wardman, C., Nevin, K. P., Lovley, D. R. A new model for electron flow during anaerobic digestion: direct interspecies electron transfer to *Methanosaeta* for the reduction of carbon dioxide to methane. *Energ. Environ. Sci.* **7**, 408-415, doi:10.1039/c3ee42189a (2014).
28. Wang, L. Y., Nevin, K. P., Woodard, T. L., Mu, B. Z., Lovley, D. R. Expanding the diet for DIET: electron donors supporting direct interspecies electron transfer (DIET) in defined co-cultures. *Front. Microbiol.* **7**, 236, doi:10.3389/fmicb.2016.00236 (2016).
29. Müller, P. Disorder. In *Crystal Structure Refinement: A Crystallographer's Guide to SHELXL* Ch. 5, 56-96 (Oxford University Press, 2006).
30. Berger, S., Cabrera-Orefice, A., Jetten, M. S. M., Brandt, U., Welte, C. U. Investigation of central energy metabolism-related protein complexes of ANME-2d methanotrophic archaea by complexome profiling. *Biochim. Biophys. Acta Bioenerg.* **1862**, 148308, doi:10.1016/j.bbabio.2020.148308 (2021).
31. Kurth, J. M., Müller, M. C., Welte, C. U., Wagner, T. Structural insights into the methane-generating enzyme from a methoxydotrophic methanogen reveal a restrained gallery of post-translational modifications. *Microorganisms* **9**, doi:10.3390/microorganisms9040837 (2021).
32. Hahn, C. J., Lemaire, O. N., Kahnt, J., Engilberge, S., Wegener, G., Wagner, T. Crystal structure of a key enzyme for anaerobic ethane activation. *Science* **373**, 118-121, doi:10.1126/science.abg1765 (2021).
33. Lemaire, O. N., Wagner, T. A Structural view of alkyl-coenzyme M reductases, the first step of alkane anaerobic oxidation catalyzed by archaea. *Biochemistry* **61**, 805-821, doi:10.1021/acs.biochem.2c00135 (2022).
34. Kahnt, J., Buchenau, B., Mahlert, F., Kruger, M., Shima, S., Thauer, R. K. Post-translational modifications in the active site region of methyl-coenzyme M reductase from methanogenic and methanotrophic archaea. *FEBS J.* **274**, 4913-21, doi:10.1111/j.1742-4658.2007.06016.x (2007).
35. Lemaire, O. N., Wegener, G., Wagner, T. Ethane-oxidising archaea couple CO_2 generation to F_{420} reduction. *Nat Commun* **15**, 9065, doi:10.1038/s41467-024-53338-7 (2024).
36. Jetten, M. S. M. *Acetate metabolism in Methanothrix soehngenii* Dissertation. Wageningen University and Research, (1991).
37. Shima, S., Krueger, M., Weinert, T., Demmer, U., Kahnt, J., Thauer, R. K., Ermler, U. Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature* **481**, 98-101, doi:10.1038/nature10663 (2012).

38. Jelsch, C., Teeter, M. M., Lamzin, V., Pichon-Pesme, V., Blessing, R. H., Lecomte, C. Accurate protein crystallography at ultra-high resolution: Valence electron distribution in crambin. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 3171-3176, doi:10.1073/pnas.97.7.3171 (2000).
39. Grabarse, W., Mahlert, F., Duin, E. C., Goubeaud, M., Shima, S., Thauer, R. K., Lamzin, V., Ermler, U. On the mechanism of biological methane formation: Structural evidence for conformational changes in methyl-coenzyme M reductase upon substrate binding. *J. Mol. Biol.* **309**, 315-330, doi:10.1006/jmbi.2001.4647 (2001).
40. Ragsdale, S. W. Biochemistry of methyl-coenzyme M reductase: The nickel metalloenzyme that catalyzes the final step in synthesis and the first step in anaerobic oxidation of the greenhouse gas methane. In *The Metal-Driven Biogeochemistry of Gaseous Compounds in the Environment* Vol. 14 (eds Kroneck, P.; Torres, M.) 125-145 (Springer, 2014).
41. Thauer, R. K. Methyl (Alkyl)-coenzyme M reductases: Nickel F-430-containing enzymes involved in anaerobic methane formation and in anaerobic oxidation of methane or of short chain alkanes. *Biochemistry* **58**, 5198-5220, doi:10.1021/acs.biochem.9b00164 (2019).
42. Ermler, U. On the mechanism of methyl-coenzyme M reductase. *Dalton Trans.*, 3451-3458, doi:10.1039/b506697b (2005).
43. Roux, A., Talon, R., Alsalman, Z., Engilberge, S., D'Aléo, A., Di Pietro, S., Robin, A., Bartocci, A., Pilet, G., Dumont, E., et al. Influence of divalent cations in the protein crystallization process assisted by lanthanide-based additives. *Inorg. Chem.* **60**, 15208-15214, doi:10.1021/acs.inorgchem.1c01635 (2021).
44. Ronda, L., Bruno, S., Bettati, S., Storici, P., Mozzarelli, A. From protein structure to function via single crystal optical spectroscopy. *Front. Mol. Biosci.* **2**, doi:10.3389/fmolb.2015.00012 (2015).
45. Vögeli, B., Engilberge, S., Girard, E., Riobé, F., Maury, O., Erb, T. J., Shima, S., Wagner, T. Archaeal acetoacetyl-CoA thiolase/HMG-CoA synthase complex channels the intermediate via a fused CoA-binding site. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3380-3385, doi:10.1073/pnas.1718649115 (2018).
46. Lemaire, O. N., Müller, M. C., Kahnt, J., Wagner, T. Structural rearrangements of a dodecameric ketol-acid reductoisomerase isolated from a marine thermophilic methanogen. *Biomolecules* **11**, doi:10.3390/biom11111679 (2021).
47. Wagner, T., Wegner, C. E., Kahnt, J., Ermler, U., Shima, S. Phylogenetic and structural comparisons of the three types of methyl coenzyme M reductase from *Methanococcales* and *Methanobacteriales*. *J. Bacteriol.* **199**, doi:10.1128/JB.00197-17 (2017).
48. von Stetten, D., Giraud, T., Carpentier, P., Sever, F., Terrien, M., Dobias, F., Juers, D. H., Flot, D., Mueller-Dieckmann, C., Leonard, G. A., et al. *In crystallo* optical spectroscopy (*ic*OS) as a complementary tool on the macromolecular crystallography beamlines of the ESRF. *Acta Crystallogr. D* **71**, 15-26, doi:10.1107/S139900471401517x (2015).
49. Cheng, H., Jarvis, E. D., Fedrigo, O., Koepfli, K. P., Urban, L., Gemmell, N. J., Li, H. Haplotype-resolved assembly of diploid genomes without parental data. *Nat. Biotechnol.* **40**, 1332-1335, doi:10.1038/s41587-022-01261-x (2022).
50. Li, H. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* **37**, 4572-4574, doi:10.1093/bioinformatics/btab705 (2021).
51. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100, doi:10.1093/bioinformatics/bty191 (2018).
52. Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., Wang, Z. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359, doi:10.7717/peerj.7359 (2019).
53. Sieber, C. M. K., Probst, A. J., Sharrar, A., Thomas, B. C., Hess, M., Tringe, S. G., Banfield, J. F. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat. Microbiol.* **3**, 836-843, doi:10.1038/s41564-018-0171-1 (2018).
54. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043-55, doi:10.1101/gr.186072.114 (2015).

55. Buchfink, B., Reuter, K., Drost, H. G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366-368, doi:10.1038/s41592-021-01101-x (2021).
56. Parks, D. H., Chuvochina, M., Rinke, C., Mussig, A. J., Chaumeil, P. A., Hugenholtz, P. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* **50**, D785-D794, doi:10.1093/nar/gkab776 (2022).
57. Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., Parks, D. H. GTDB-Tk: A toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925-1927, doi:10.1093/bioinformatics/btz848 (2020).
58. Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., Parks, D. H. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**, 5315-5316, doi:10.1093/bioinformatics/btac672 (2022).
59. Huson, D. H., Albrecht, B., Bagci, C., Bessarab, I., Gorska, A., Jolic, D., Williams, R. B. H. MEGAN-LR: new algorithms allow accurate binning and easy interactive exploration of metagenomic long reads and contigs. *Biol. Direct* **13**, 6, doi:10.1186/s13062-018-0208-7 (2018).
60. Raskin, L., Stromley, J. M., Rittmann, B. E., Stahl, D. A. Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. *Appl. Environ. Microbiol.* **60**, 1232-40, doi:10.1128/aem.60.4.1232-1240.1994 (1994).
61. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069, doi:10.1093/bioinformatics/btu153 (2014).
62. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., Bricogne, G. Data processing and analysis with the *autoPROC* toolbox. *Acta Crystallogr. D* **67**, 293-302, doi:10.1107/S0907444911007773 (2011).
63. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. *Acta Crystallogr. D* **75**, 861-877, doi:10.1107/S2059798319011471 (2019).
64. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
65. Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
66. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12-21, doi:10.1107/S0907444909042073 (2010).
67. Katoh, K., Rozewicki, J., Yamada, K. D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160-1166, doi:10.1093/bib/bbx108 (2019).
68. Robert, X., Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-W324, doi:10.1093/nar/gku316 (2014).
69. Krissinel, E., Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774-97, doi:10.1016/j.jmb.2007.05.022 (2007).

CHAPTER VII. Crystallomics on activated sludge from wastewater treatment

The following results are preliminary, and studies have not been completed. However, insights from the experiments are valuable and therefore presented briefly.

Main

The previous work (Ch. VI) demonstrated that enzyme purification from a sample with a high microbial diversity is feasible. Therefore, in the next step a larger scale "shotgun" purification was performed. Biomass from the second sampling point, previously described in Ch. VI, was utilized (sample B). The same physical enrichment protocol was performed, consisting of a short centrifugation to exclude inorganic and large particles, followed by centrifugation to pellet the medium sized *Methanothrix* filaments and finally a long centrifugation step yielding the remaining smaller *Methanothrix* filaments as well as the small cell fraction. The last pellet was utilized for the purification described in this chapter. The primary focus was to explore whether other *Methanothrix* enzymes besides MCR could be captured. Additionally, initial insights into the effectiveness of the method to capture enzymes from other members of the microbial community would also be valuable. Furthermore, a direct comparison of the approach performed on a pure culture (see Ch. II) was expected to yield insights into limitations and possibilities resulting from a more diverse sample.

7 g of cell pellet (yielding 48 mg soluble protein after ultracentrifugation) were used. Subsequent anionic exchange chromatography, hydrophobic interaction chromatography and optionally size exclusion chromatography were used to separate the proteins (Fig. 1). All steps were performed under anaerobic conditions to prevent damage to oxygen-sensitive proteins. Ultimately, 15 fractions were obtained (Table S1) and compared to the results of Ch. II. As expected, final fractions exhibited a higher

degree of contamination on SDS-PAGE compared to the final fractions from the pure culture, reflecting the more diverse microbial community. Multiple fractions were colored brown or yellow (Table S1), indicating the presence of prosthetic groups (*e.g.*, Fe-S clusters, heme-groups) or cofactors (*e.g.*, flavins, F_{430}). Prominent SDS-PAGE bands of the final fractions were analyzed via mass spectrometry (Fig. 2, Table 1, Table S2). A high amount of peptide hits was obtained in accordance with a higher protein diversity, of which about two-thirds of the top results were annotated to *Methanothrix*, reflecting the enrichment in the cell pellet.

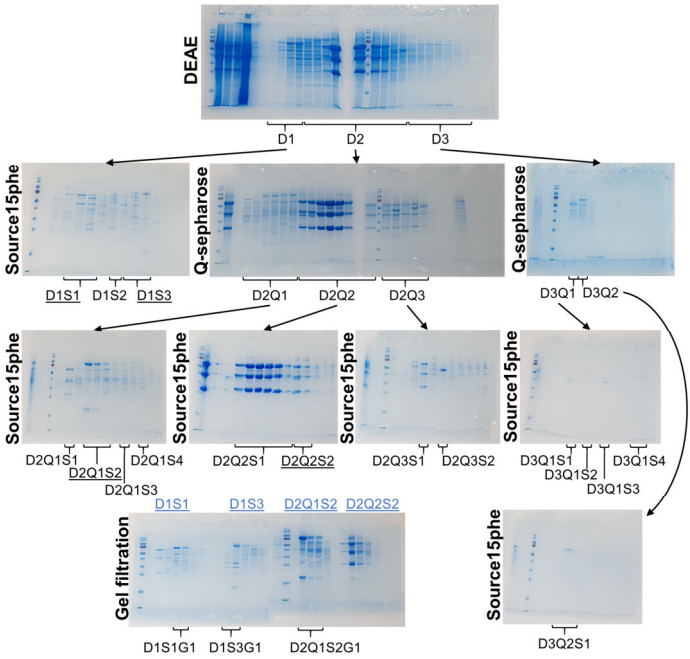


Figure 1. Crystallomics on mechanically enriched biomass from wastewater treatment sludge. SDS-PAGE of protein purification steps. Curly brackets indicate pooled fractions, and the pool identifier and arrows indicate the purification path of each pool. Underlined identifiers mark pools selected for final gel filtration chromatography. For the gel filtration (bottom left), injected fractions are not marked by arrows but indicated in blue.

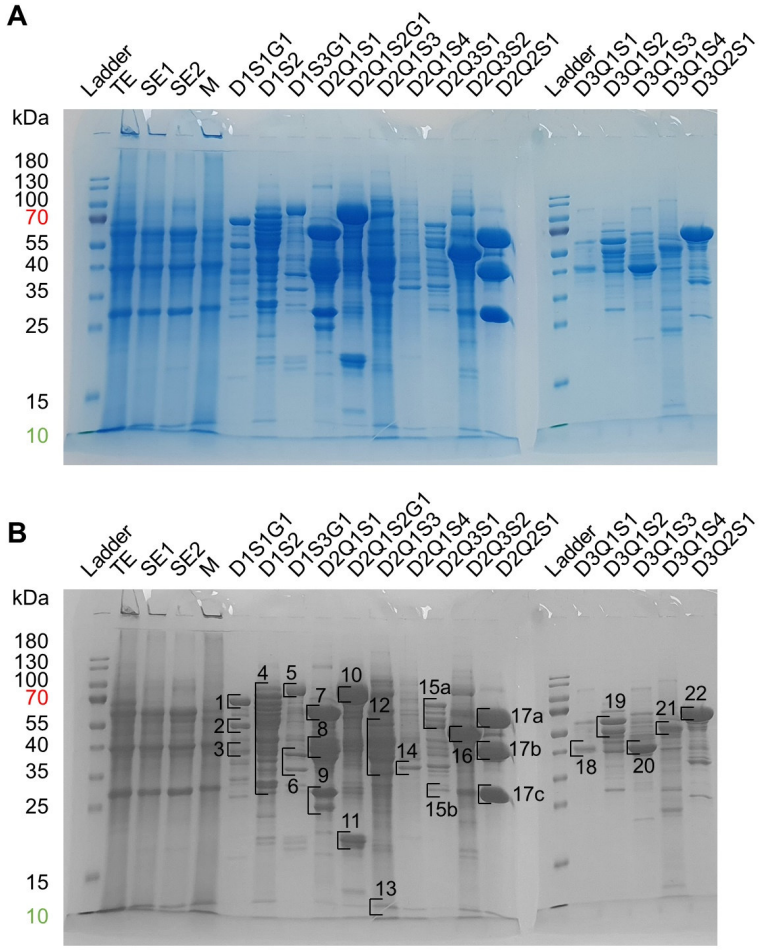


Figure 2. Final protein fractions. (A) SDS-PAGE of the final purification fractions. TE=total extract, SE1=soluble extract after centrifugation, SE2=soluble extract after ultracentrifugation, M=resuspended pellet after ultracentrifugation containing membrane fragments. TE, SE1, SE2, M were loaded at 10 µg and all other fractions at 5 µg. Fraction identifiers are in accordance with Fig. 1. (**B**) Gel bands selected for MS analysis. Cut bands are indicated with brackets and band number. Gel bands 15a-b as well as 17a-c were processed as one.

Table 1. Best mass spectrometry hits for each excised gel band. Hits with the highest number of peptide spectral matches (PSM) and unique peptides. Unbinned hits have been manually annotated via BLAST. The top ten hits are listed in Table S2.

Gel band	Best hit
1	Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase
2	Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_gamma
3	Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_delta_1
4	Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha
5	Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha
6	unbinned_Glutaconate_CoA-transferase_subunit_A BLAST: TPA: CoA-transferase [Syntrophales bacterium] HPX82384.1 Cov:100% ID:99.71%
7	unbinned_hypothetical_protein (coenzyme-B sulfoethylthiotransferase subunit alpha [Methanoregulaceae archaeon])
8	unbinned_Acyl-CoA_dehydrogenase (Select seq gb|MBP8980195.1| acyl-CoA dehydrogenase family protein [Syntrophobacterales bacterium]) BLAST: TPA: acyl-CoA dehydrogenase family protein [Syntrophales bacterium] HNY7220 Cov:100% ID:100%
9	Methanothrix_soehngenii_Triosephosphate_isomerase
10	Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha
11	Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_epsilon
12	unbinned_Glutaconate_CoA-transferase_subunit_A BLAST: TPA: CoA-transferase [Syntrophales bacterium] HPX82384.1 Cov:100% ID:99.71%
13	Methanothrix_soehngenii_hypothetical_protein BLAST: conserved hypothetical protein [Methanothrix soehngenii GP6] AEB69607.1 Cov:1(ID:100%
14	unbinned_Malate_dehydrogenase BLAST: malate dehydrogenase [Bacillota bacterium] MDI9449352.1 Cov:99% ID:99.36%
15	Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta_2
16	Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta
17	Methanothrix_sp018052825 adn soehngenii_Methyl-coenzyme_M_reductase_subunit_alpha subunit beta + subunit gamma
18	Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta
19	Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta_2
20	Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta
21	Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta_2
22	no clear hit

Among the hits were the Acetyl-CoA-synthetase, multiple subunits of the Acetyl-CoA-decarbonylase/synthase complex and MCR which confirmed that the technique has a higher chance to capture highly expressed enzymes of the catabolism when applied to microbial communities thriving under natural conditions. In contrast, the previous application of crystallomics on the pure culture of *M. infernus* had retrieved multiple anabolic enzymes, which are expected to be of higher importance for cells rapidly dividing under laboratory conditions.

Ten final fractions contained sufficient protein for crystallization (>60 µg total protein) and crystals were obtained for eight of them (Fig. 3). Some samples yielded crystals with different colors and morphologies (*e.g.*, D2Q1S2G1, D2Q1S3), showcasing that crystallization functioned as a final purification step, separating proteins from the mixture. The first crystals that appeared were fished and shipped for X-ray diffraction collection, and exploitable data was obtained for D2Q1S1 (at 2.9 Å resolution), D2Q1S3 (at 4 Å resolution) and D1S2 (at ~ 3 Å resolution, Table S3). The processing of the latter two has been on hold due to time constraints.

D2Q1S1 was identified as MCR based on the SDS-PAGE pattern and color. Since the chromatographic profile differed from the previously obtained *Methanothrix* MCR (Ch. VI) the fraction was further investigated. The structure was solved by molecular replacement with *Methanothrix* MCR and the asymmetric unit contained a trimer of MCRs as previously indicated by blue native PAGE for MCR of ANME-2d[1] and observed for the *Methanothrix* MCR structure. Noticeably, inspection of the electron density showed some larger deviations of the backbone density (Fig. 3C). A comparison of all MCR sequences obtained in the metagenome revealed a sequence stretch including a matching insertion in only two sequences belonging not to *Methanothrix* but to *Methanospirillum sp.* and *Methanoregulaceae sp.* To determine the correct sequence, AlphaFold 2 models of both sequences were generated and compared to the density. Due to time restrictions, a detailed analysis of the electron density fit has not

yet been carried out, however a brief evaluation indicated a better fit for the sequence from *Methanospirillum*. The previous MS analysis of prominent gel bands in the fraction indicated the presence of MCR-α from *Methanoregulaceae archaeon* and multiple *Methanothrix* members but not *Methanospirillum sp.* while the β- and γ-subunits were not among the ten most abundant hits. This could be due to the high degree of contamination obscuring signals or improper digestion. Alternatively, it might reflect the crystallization bias as MCR from *Methanospirillum* might be less abundant in the sample but crystallize better than other variants. To obtain a clearer result, fresh and already processed crystals were retrieved, dissolved, and sent for MS analysis (Fig. 3B), identification is currently in progress. Additionally, identification and structure processing for the fractions D2Q1S3 and D1S2 will follow in the near future.

Besides MCR, various proteins of interest were detected by MS including the acetate-activating enzymes of *Methanothrix*. Among the top hits was the acetyl-CoA synthetase of *M. soehngenii* in fraction D1S1G1. The enzyme has been described to form an α_2 assembly[2] and the closest structural homolog on the PDB is the recombinant AMP-forming Acetyl-CoA synthetase from *Chloroflexota bacterium* (8RPK), sharing a sequence identity of 50 %. Furthermore, all five subunits (αβγδε) of the acetyl-CoA decarbonylase/synthase complex (ACDS) from *Methanothrix sp.* could be detected. The large complex was likely destabilized during purification resulting in the capture of subcomplexes in different fractions. The γ- and δ-chain and the α- and ε-chain were detected together in the fractions D1S1G1 and D2Q1S2G1 respectively, suggesting they might have retained their interaction. The closest structural homologs were found to be: the F_{420}-reducing carbon monoxide dehydrogenase component ($\alpha_2\varepsilon_2\zeta_2$) from *Ca.* Ethanoperedens thermophilum (PDB: 8RIU) sharing 51.18 % and 49.43 % sequence identity for the α- and ε-chain respectively; the CODH/ACS ($\alpha_2\varepsilon_2\beta$) from *Methanosarcina thermophila* (9C0S) sharing 63.82 % sequence identity for the β-subunit; the corrinoid and [4Fe4S] cluster-containing protein (CoFeSP) from

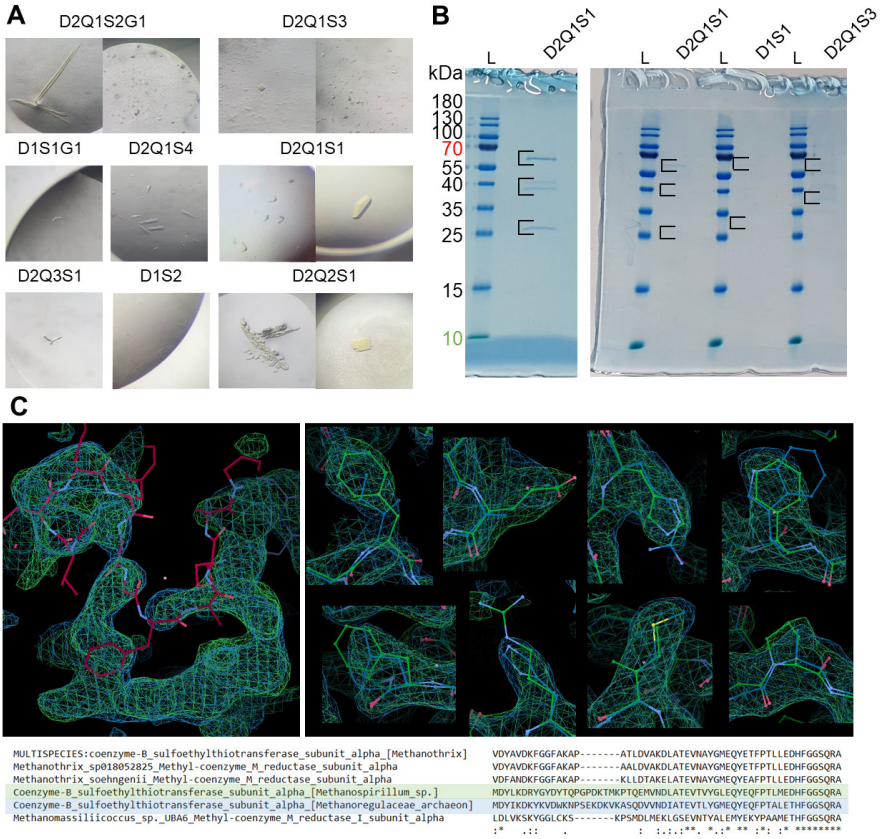


Figure 3. Crystallized fractions and identification of MCR. (**A**) Crystals obtained from the different final purified fractions. (**B**) SDS-PAGE of dissolved crystals. The first D2Q1S1 sample consists of fresh crystals from the plate while the second sample consist of crystals retrieved from the beamline after the X-ray crystallography experiment. Brackets indicate extracted bands for MS analysis. (**C**) Illustration of the identification of the obtained MCR. A backbone mismatch was found in the density in accordance with the sequence alignment and a more detailed inspection of single sidechains was conducted to distinguish between the two remaining sequences. The $2F_o$-F_c electron density (contoured at ~ 1σ) and model (sticks) for processing with the AlphaFold 2 models containing the *Methanospirillum* or *Methanoregulaceae* sequence are colored in green or blue, respectively while the *Methanothrix* model is shown in red. Atoms are colored in red for O, blue for N, and yellow for S.

Carboxydothermus hydrogenoformans (4C1N) sharing 34.85 % sequence identity for the γ-chain; and the corrinoid iron-sulfur protein (CFeSP) from *Moorella thermoacetica* (4DJD) sharing 37.06 % sequence identity for the δ-chain. Structures of both the acetyl-CoA synthetase and the ACDS complex or subcomponents could give valuable insights into the metabolism of *Methanothrix* and will be prioritized in future purification and crystallization efforts. This could be facilitated by the reduction of contamination with non-*Methanothrix* enzymes by utilizing the physical enrichment pellet with the highest *Methanothrix* content, excluding the small cell fraction. In contrast, focusing studies on fractions excluding *Methanothrix* (*e.g.*, through filtration) could give insights into other microbial members of the wastewater treatment sludge, participating in organic matter degradation.

In addition to the soluble enzymes, the membrane fraction was also investigated. The pellet obtained after centrifugation of the cell lysate was discarded as it consisted of bigger particles and large cell envelope fragments. However, the soluble extract was subjected to ultracentrifugation to harvest smaller membrane fragments forming micelles. The obtained pellet was washed to reduce contamination with proteins from the soluble extract, followed by treatment with different detergents, to extract proteins from the membrane fragments (Fig. 4A). Good results were obtained with n-dodecyl β-maltoside (DDM) at 100 x critical micelle concentration (CMC) and octylphenol ethoxylate (Triton-X) at 10 and 100 x CMC and 3-[(3-cholamido-propyl)dimethylammonio]-1-propanesulfonate (CHAPS) at 10 x CMC. The tested DDM and Triton-X fractions were pooled and passed on a Q-sepharose column to evaluate whether the extracted proteins could be separated by chromatography. A reduced number of peaks with low absorbance at 280 nm were observed in the chromatogram. Due to the low quantities of remaining sample, all fractions were precipitated by trichloroacetic acid (TCA) and loaded on SDS-PAGE for visualization (Fig. 4B). Bands of interest were excised and analyzed via mass spectrometry (Table

S4). The membrane origin of the sample is reflected well by the most prominent hits, which are dominated by ATP synthase subunits and S-layer protein domain-containing proteins. However, proteins of the soluble extract were also captured like the Ketol-acid reductoisomerase and MCR-α. The presence of these enzymes can be explained by either contamination with the soluble extract (especially as MCR was observed to be highly abundant) or the formation of large molecular weight complexes that are affected by ultracentrifugation rather than unknown interactions with the membrane. Compared to the soluble extract, a higher fraction of proteins appears to be annotated as hypothetical, potentially reflecting that membrane proteins tend to be less well studied as their characterization is more challenging.

These initial results corroborate the conclusions of Ch. VI, supporting the feasibility of enzyme extraction from heterogeneous samples through crystallomics. The utilized sample originated from the second sampling and has undergone a slightly modified physical enrichment. The obtained cell pellet contained *Methanothrix* filaments and the small cell fraction and therefore exhibits a different composition compared to both samples utilized in Ch. VI. Multiple enzymes could be extracted and crystallization was successful despite a higher degree of contamination in the final protein fractions. The capture of MCR from *Methanospirillum/Methanoregulaceae* demonstrates that enzymes can be extracted from different organisms within the sample. Moreover, it also illustrates the importance of protein abundance. MCR is expected to be among the most expressed enzymes in methanogens, a feature aiding in its isolation. In comparison to the adapted targeted purification of MCR (Ch. VI) where crystals could be obtained from 1 g of cells, a higher sample amount will be required to obtain sufficient quantities of less abundant enzymes. Furthermore, it became apparent that MS analysis might be more effective when carried out on crystallized samples than purified fractions. The diverse protein composition within the final fractions results in numerous MS hits, complicating the accurate identification of the correct sequence for model building. In

contrast, washed and dissolved crystals should yield pure protein which facilitates accurate identification. Alternatively, adapted more stringent data processing parameters might improve analyses, an approach worthwhile to explore in further studies.

Furthermore, these results support the concept that extracting membrane proteins from heterogeneous native biomass is possible, and experiments utilizing the optimal detergent and more biomass should be pursued. Targets of high interest could for example be membrane complexes from *Methanothrix* or partners involved in energy conservation and DIET. Structural investigation could be performed through cryoEM which is more suitable than X-ray crystallography for membrane proteins. Additionally, the cell envelope fraction could also be utilized in order to study proteins of the S-layer or *Methanothrix* sheath proteins. In summary, these results demonstrate promising potential and encourage further exploration of the approach.

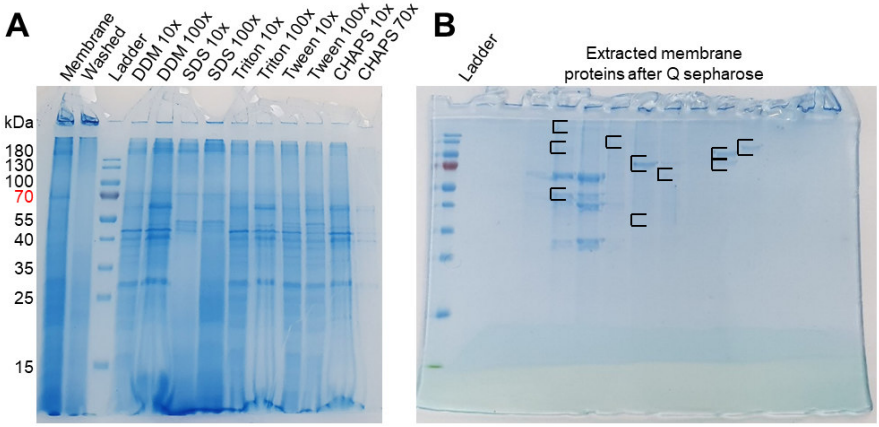


Figure 4. Extraction of proteins from the membrane fraction. (**A**) SDS-PAGE of membrane fractions tested with different detergents at x-fold critical micelle concentration (10x, 70x, 100x). (**B**) SDS-PAGE of concentrated elution fractions from anion exchange chromatography. Brackets indicate samples sent for MS analysis.

Supplementary material

Sampling and physical enrichment

The sample described in Ch. VI from the second sampling point was utilized. The same purification protocol as described in chapter VI was applied to the biomass. After the last centrifugation an additional centrifugation step was added. The supernatant was centrifuged for 15 min at 40,000 × *g* yielding a pellet including *Methanothrix* and smaller microbial cells (observed by light microscopy, data not shown).

Cell lysis

Cell lysis was carried out in an anaerobic tent under a N_2/CO_2 atmosphere (90:10 %). 7.18 g enriched biomass was filled up to 50 mL with IEC A buffer (50 mM Tricine pH 8, 2 mM dithiothreitol (DTT)) and split into two. Each sample was lysed by sonication (72 %, 6 x 20 sec with 1 min on an icepack between cycles, probe KE73 (Bandelin, Germany)) followed by centrifugation for 30 min, at 13,000 × *g* and 4 °C. Then the supernatant was transferred to ultracentrifugation vials and centrifuged for 90 min at 150,237.1 × *g* and 4 °C. The pellet containing smaller membrane fragments and very large protein complexes (M) was resuspended in a small volume of IEC A and stored at −80 °C for later treatment.

Protein purification

The soluble extract after ultracentrifugation was transferred into an anaerobic tent with a N_2/H_2 atmosphere (97:3 %) under yellow light. The sample was filtered through a 0.2 µm filter (Sartorius, Germany) and loaded on a 3 x 5 mL HiTrap™ DEAE Sepharose FF column (Cytiva, Sweden). Proteins were eluted over a gradient of 0-100 % IEC B (1M NaCl, 50 mM Tricine pH 8, 2 mM DTT) for 100 min at 2 mL.min^{-1}. Fractions were combined based on SDS-PAGE and absorbance at 254, 280, and 420 nm. Three pools were generated: D1 75-140 mM, D2 150-310 mM and D3 325-440 mM NaCl.

Pools D2 and D3 were diluted 1:3 with IEC A and filtered. D2 was the passed on a 2 x 5 mL HiTrap Q HP™ column (Cytiva, Sweden) and eluted with a gradient of 0-50 % IEC B over 100 min at 1 mL.min^{-1}. Fractions were pooled into three pools: D2Q1 160-260 mM, D2Q2 270-330 mM and D2Q3 340-420 mM NaCl. Pool D3 was also passed on Q HP and eluted with a gradient of 20-70 % IEC B over 100 min followed by 70-90 % for 40 min. Four pools were obtained: D3Q1 400-430 mM, D3Q2 450-480 mM, D3Q3 500-530 mM, D3Q4 570-590 mM NaCl. Pool D1 was not passed on QHP due to a smaller protein quantity and instead diluted 1:3 in HIC B (25 mM Tris pH 7.6, 2 M $(NH_4)_2SO_4$, 2mM DTT), filtered, directly loaded on a Source™15 Phe 4.6/100 PE column (Cytiva, Sweden) and eluted at 75-0 % HIC B. Three pools were formed: D1S1 1060-920 mM, D1S2 760-640 mM and D1S3 620-360 mM $(NH_4)_2SO_4$. All pools from the Q-sepharose runs of D2 and D3 were also passed separately on the Source™15 Phe after 1:3 dilution with HIC B. All elutions were carried out at 75-0 % HIC B for 60 min at 1 mL.min^{-1}. Following pools were formed: D2Q1S1 1040-960 mM, D2Q1S2 860-660 mM, D2Q1S3 620-460 mM, D2Q1S4 390-320 mM, D2Q2S1 1000-660 mM, D2Q2S2 640-500 mM, D2Q3S1 980-920 mM, D2Q3S2 740-640 mM, D3Q1S1 1500-990 mM, D3Q1S2 940-820 mM, D3Q1S3 690-620, D3Q1S4 15-0, D3Q2S1 1020-820 mM $(NH_4)_2SO_4$. D3Q3 and D3Q4 did not yield exploitable fractions due to insufficient protein content. Finally pools D1S1, D1S3, D2Q1S2 and D2Q2S2 were passed on a Superdex™ 200 Increase 10/300 GL (Cytiva, Sweden) gel filtration column at 0.4 mL.min^{-1}. Fractions were combined into three final pools: D1S1G1 12.5-15.7 mL, D1S3G1 10.0-12.0 mL, D2Q1S2G1 10.4-13.8 mL. All samples were washed 1:1000 in GF buffer (25 mM Tris pH 7.6, 10 % v/v glycerol, 2 mM DTT), and concentrated to ~80 µL (apart from fraction D2Q2S1 that was identified as MCR and concentrated to 500 µL) using a concentrator cutoff (100 kDa or 30 kDa, Sartorius, Germany) based on the size of the fraction on SDS-PAGE. Protein concentrations were estimated via Bradford assay (see Table S1). Fractions D3Q1S1, D3Q1S2, D3Q1S3, D3Q1S4 and

D3Q2S1 were deemed to contain too little protein for crystallization. Finally, all fractions were loaded on SDS-PAGE, TE, SE1, SE2, M were loaded at 10 µg and all other fractions at 5 µg.

Crystallization

Fractions were crystallized on 96-Well MRC 2-Drop Crystallization Plates in polystyrene (SWISSCI, UK) with an OryxNano (Douglas instruments, UK) with the crystallization kits PEP, PACT+, Wizard 1-2 or Wizard 3-4 (Jena Bioscience, Germany) with drops of 0.5 µL protein and 0.5 µL crystallization solution. Crystals were stored at 20 °C under anaerobic conditions (N_2/H_2, gas ratio of 97:3). Since MCR was obtained in higher quantity, three plates were prepared instead of one. Exploitable crystals were obtained for D2Q1S1 (0.1 M 1:2:2 – DL-malic acid:MES:Tris base (MMT buffer) pH 5, 25 % w/v PEG 1500, soaked in 20 % glycerol for cryoprotection), D1S2 (45 % w/v Pentaerythritol propoxylate (17/8 PO/OH), 100 mM MES pH 6.5, 300 mM KCl) and D2Q1S3 (2 M $(NH_4)_2SO_4$, 100 mM MES pH 6.5, 200 mM NaCl, soaked in 30 % glycerol). Fished crystals were immediately flash-frozen in liquid nitrogen.

Data collection and processing

The X-ray diffraction was collected at the beamline PXII of SOLEIL Synchrotron (Saclay, France) (Table S3). Data for D2Q1S1 was processed with autoPROC[3] and molecular replacement using the previously obtained structure for *Methanothrix* MCR from the physical enrichment (with the *Methanothrix soehngenii* sequence) was performed with Molrep (CCP4 [4]). AlphaFold 2[5] models of the separate α-, β-, and γ-chain were generated for the *Methanoregulaceae* and *Methanospirillum* sequences detected in the genome (see Ch. VI)). The full asymmetric unit content was then generated by superposing each chain to the initial model followed by one round of rigid body refinement with Refmac (CCP4)[4] using the default settings.

Extraction of proteins from the membrane pool

The membrane pool was thawed and split into two ultracentrifugation vials. The volume was filled up to 6 mL with solubilization buffer (90 mL buffer A + 10 mL buffer B (A: 50 mM TrisHCl pH 8, 2 mM DTT, B: 50 mM TrisHCl pH 7.6, 1 M NaCl, 2 mM DTT), followed by ultracentrifugation at 150,237.1 × *g* for 90 min at 4 °C. The pellet was resuspended in solubilization buffer and split into 10 Eppendorf tubes with about 90 µL volume each. Five different detergents at 10x and 100x critical micelle concentration final (DDM 0.09 % w/v and 0.87 % w/v, SDS 2.36 % w/v and 23.65 % w/v, Triton-X 0.19 % w/v and 1.94 % w/v, Tween20 0.07 % w/v and 0.72 % w/v and Chaps 4.0 % w/v and 39.97 % w/v) were added and filled up to 1 mL with solubilization buffer. Samples were incubated over night at 4 °C in a slowly rotating holder. On the next day, samples were filled up to 5 mL with solubilization buffer and ultracentrifugated at 150,237.1 × *g* for 90 min at 4 °C. Supernatant was concentrated to 100 µL using a 10 kDa cutoff. All samples were passed on SDS-PAGE (boiled and unboiled) to evaluate which detergent gave the best results. All DDM and Triton-X fractions were pooled, diluted 1:3 in buffer A with 0.05 % DDM and passed on a 1 mL Q HP column. Proteins were eluted at 0-80 % buffer B with 0.05 % DDM for 80 min at 0.5 $mL.min^{-1}$. To concentrate the obtained protein fractions 1:4 100 % w/v trichloroacetic acid was added to each fraction, incubated for 1 h at 4 °C. Then the fractions were spinned down at 16,000 × *g* for 5 min. The pellets were resuspended in 30 µL SDS-PAGE loading dye, incubated for 15 min at RT and loaded on SDS-PAGE without boiling.

MS analysis

Major gel bands of interest were cut out of the gel and processed for MS analysis as described in chapter VI by Grace D'Angelo. Selected unbinned hits were manually annotated via BLAST.

Table S1. Protein content of the final purification fractions estimated by Bradford assay

Fraction	Estimated final protein amount (µg)	Description
D1S1G1	84.8	clear
D1S2	320.0	yellowish
D1S3G1	128.0	yellowish
D2Q1S1	414.4	yellowish
D2Q1S2G1	528.8	brown
D2Q1S3	492.0	yellowish
D2Q1S4	219.2	clear
D2Q2S1 (MCR)	8,700	yellow
D2Q3S1	153.6	slightly yellowish
D2Q3S2	323.2	slightly brown
D3Q1S1	0	-
D3Q1S2	16.8	clear
D3Q1S3	26.4	clear
D3Q1S4	17.6	clear
D3Q2S1	54.4	clear
TE	104,500	brown
SE (after centrifugation)	75,500	brown
SE 2 (after ultracentrifugation)	48,000	brown

Table S2 (next page): Top ten peptide hits for protein fractions purified from the soluble extract. Entries with a molecular weight mismatch are greyed out, best hits per gel band are shown in green. #PSM and #unique peptides are colored from red to yellow to green for 0-200-400 and 0-20-40 respectively. Selected unbinned entries have been manually annotated with BLAST. Hits that mismatch with the molecular weight indicated by SDS-PAGE are shown in grey.

Accession Gel 1	# PSMs	# Unique Peptides	MW [kDa]
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	4475	85	75.1
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	341	21	78.9
unbinned_Acetyl-coenzyme_A_synthetase	239	3	74.5
unbinned_Acetylene_hydratase	124	30	79.6
Methanothrix_sp018052825_Acetyl-coenzyme_A_synthetase	87	5	74.5
Methanothrix_sp018052825_Acetyl-coenzyme_A_synthetase	84	18	74.2
Methanothrix_soehngenii_hypothetical_protein	67	16	64.3
unbinned_Nicotinate_dehydrogenase_subunit_B	55	24	78.2
unbinned_hypothetical_protein	49	28	79.4
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	46	6	73.8
Accession Gel 2	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_gamma_1	1566	18	52.1
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_gamma_1	808	1	52
Methanothrix_soehngenii_hypothetical_protein	720	33	56.2
unbinned_Acetyl-CoA_acetyltransferase	141	20	48.9
Methanothrix_sp018052825_hypothetical_protein	122	5	53.8
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	67	21	75.1
unbinned_Leucine-,_isoleucine-,_valine-,_threonine-,_and_alanine-binding_protein_precursor	52	12	49.2
unbinned_hypothetical_protein	26	9	30.3
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta_2	20	6	51.7
Methanothrix_sp018052825_V-type_ATP_synthase_beta_chain	7	2	50.9
Accession Gel 3	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_delta_1	648	16	48.5
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_delta_1	290	1	48.4
unbinned_Leucine-,_isoleucine-,_valine-,_threonine-,_and_alanine-binding_protein_precursor	206	22	40.2
unbinned_Leucine-,_isoleucine-,_valine-,_threonine-,_and_alanine-binding_protein_precursor	111	18	49.2
unbinned_Leucine-,_isoleucine-,_valine-,_threonine-,_and_alanine-binding_protein_precursor	103	19	42.5
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	59	16	75.1
unbinned_Phosphoserine_aminotransferase	41	7	14.2
unbinned_Glyceraldehyde-3-phosphate_dehydrogenase	25	13	36.1
Synthrophorhabdus_sp._Soluble_hydrogenase_42_kDa_subunit	20	9	41.4
unbinned_1,5-anhydro-D-fructose_reductase	18	6	44.3
Accession Gel 4	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Acetyl-coenzyme_A_synthetase	476	31	74.2
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha	451	54	88.5
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	448	36	78.9
unbinned_hypothetical_protein	403	28	62.4
unbinned_hypothetical_protein	382	30	46.1
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_gamma_1	350	18	52.1
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	333	50	75.1
unbinned_hypothetical_protein	259	15	28.9
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_delta_1	223	14	48.5
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_gamma_1	203	2	52
Accession Gel 5	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha	4106	109	88.6
unbinned_Catalase-peroxidase	374	31	80.6
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	323	31	75.1
Methanothrix_sp018052825_Acetyl-coenzyme_A_synthetase	310	29	74.2
unbinned_Catalase-peroxidase	168	3	81.2
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	97	1	73.8
Synthrophorhabdus_sp_Acetyl-coenzyme_A_synthetase	85	27	73.9
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	69	1	74
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	59	11	78.9
Methanothrix_sp018052825_Acetyl-coenzyme_A_synthetase	44	21	75.1
Accession Gel 6	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
unbinned_Glutaconate_CoA-transferase_subunit_A BLAST: TPA: CoA-transferase [Syntrophales bacterium] HPX82384.1 Cov:100% ID:99.71%	1301	35	39
unbinned_Glutaconate_CoA-transferase_subunit_B	1111	24	33.2

Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha	434	56	88.6
unbinned_Glutaconate_CoA-transferase_subunit_B	315	2	33.2
unbinned_Malate_dehydrogenase	254	24	32.9
unbinned_Soluble_hydrogenase_42_kDa_subunit	174	20	27.7
unbinned_Lactate_utilization_protein_A	146	23	39.3
unbinned_Acyl-CoA_dehydrogenase	140	28	41.4
unbinned_putative_oxidoreductase_YcjS	115	21	38.1
unbinned_Soluble_hydrogenase_42_kDa_subunit	100	8	14.2
Accession Gel 7	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
unbinned_hypothetical_protein (coenzyme-B sulfoethylthiotransferase subunit alpha [Methanoregulaceae archaeon])	158	8	62.4
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_alpha	82	10	60.9
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	71	10	60.8
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	64	24	75.1
Synthrophorhabdus_sp._Aspartate--tRNA_ligase	29	19	66.7
unbinned_Oligopeptide-binding_protein_AppA_precursor	27	17	71.8
Synthrophorhabdus_sp._Acetyl-coenzyme_A_synthetase	23	16	73.9
unbinned_hypothetical_protein	15	7	94.3
Methanothrix_sp018052825_Acetyl-coenzyme_A_synthetase	15	5	74.2
Caldisericia_bacterium_Methylmalonyl-CoA_mutase	14	10	57.2
Accession Gel 8	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
unbinned_Acyl-CoA_dehydrogenase (Select seq gb|MBP8980195.1| acyl-CoA dehydrogenase family protein [Syntrophobacterales bacterium]) BLAST: TPA: acyl-CoA dehydrogenase family protein [Syntrophales bacterium] HNY72207.1 Cov:100% ID:100%	424	25	44
unbinned_Leucine-,_isoleucine-,_valine-,_threonine-,_and_alanine-binding_protein_precursor	248	27	45.2
unbinned_3-oxoadipate_CoA-transferase_subunit_A	157	21	39.8
unbinned_Leucine-,_isoleucine-,_valine-,_threonine-,_and_alanine-binding_protein_precursor	110	26	41.2
unbinned_hypothetical_protein	97	18	41.8
unbinned_Soluble_hydrogenase_42_kDa_subunit	83	21	40.2
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_delta_1	67	10	48.5
unbinned_Soluble_hydrogenase_42_kDa_subunit	65	16	40.1
Methanothrix_sp018052825_Glutamate-1-semialdehyde_2,1-aminomutase	63	20	41.3
unbinned_Inositol_2-dehydrogenase	59	20	49.4
Accession Gel 9	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_soehngenii_Triosephosphate_isomerase	657	23	23
Methanothrix_soehngenii_hypothetical_protein	398	12	25.9
Methanothrix_sp018052825_2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate_synthase	265	19	28.1
unbinned_Glutaconate_CoA-transferase_subunit_B	261	20	28.6
unbinned_Glutaconate_CoA-transferase_subunit_B	235	12	27.9
unbinned_hypothetical_protein	175	17	21.3
unbinned_Hydroxypyruvate_isomerase	157	25	32.6
Methanothrix_soehngenii_hypothetical_protein	90	20	27.5
unbinned_hypothetical_protein	89	7	27.2
unbinned_Glutaconate_CoA-transferase_subunit_B	69	3	29
Accession Gel 10	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha	1258	55	88.6
unbinned_Pyruvate,_phosphate_dikinase	86	40	97
unbinned_N,N'-diacetylchitobiose_phosphorylase	41	25	93
unbinned_Xanthine_dehydrogenase_molybdenum-binding_subunit	24	16	80.2
unbinned_Gluconolactonase_precursor	23	16	83.3
Bacteroidetes_bacterium_Alpha-xylosidase	22	16	92
Methanothrix_sp018052825_Acetyl-coenzyme_A_synthetase	19	10	74.2
Methanothrix_soehngenii_Acetyl-coenzyme_A_synthetase	18	10	75.1
Synthrophorhabdus_sp_Acetyl-coenzyme_A_synthetase	18	13	73.9
unbinned_hypothetical_protein	14	11	88.8
Accession Gel 11	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_epsilon	718	15	20.9
unbinned_Glutaconate_CoA-transferase_subunit_B	63	11	33.2
unbinned_Glutaconate_CoA-transferase_subunit_A	60	14	39

unbinned_Arginine_biosynthesis_bifunctional_protein_ArgJ	59	11	40.1
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha	57	17	88.6
unbinned_Acyl-CoA_dehydrogenase	53	18	41.3
unbinned_Acyl-CoA_dehydrogenase	36	14	44
Methanothrix_sp018052825_hypothetical_protein	34	11	19.6
unbinned_hypothetical_protein	22	9	62.4
Methanothrix_soehngenii_Type_A_flavoprotein_FprA	20	10	44.3
Accession Gel 12	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
unbinned_Glutaconate_CoA-transferase_subunit_A BLAST: TPA: CoA-transferase [Syntrophales bacterium] HPX82384.1 Cov:100% ID:99.71%	323	22	39
unbinned_hypothetical_protein BLAST: TPA: acetyl-CoA hydrolase/transferase C-terminal domain-containing protein [Syntrophales bacterium] HOH73308.1 Cov:100% ID:99.61%	316	29	57.2
unbinned_Glutamate_dehydrogenase	272	32	46
unbinned_hypothetical_protein	173	24	46.1
unbinned_Acyl-CoA_dehydrogenase	168	28	41.4
unbinned_Purine-binding_protein_precursor	159	19	38.5
unbinned_D-ribose-binding_periplasmic_protein_precursor	153	25	36.1
unbinned_Acyl-CoA_dehydrogenase	129	29	41.3
unbinned_Succinyl-CoA_ligase_[ADP-forming]_subunit_beta	128	28	40.2
Methanothrix_soehngenii_Homoisocitrate_dehydrogenase	126	30	45.5
Accession Gel 13	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha	353	33	88.6
unbinned_Glutaconate_CoA-transferase_subunit_A	137	9	39
Methanothrix_sp018052825_hypothetical_protein	121	8	25.2
Methanothrix_soehngenii_hypothetical_protein BLAST: conserved hypothetical protein [Methanothrix soehngenii GP6] AEB69607.1 Cov:100% ID:100%	64	11	14.7
unbinned_Nitrogen_regulatory_protein_P-II	55	1	13.5
Methanothrix_sp018052825_V-type_ATP_synthase_subunit_D	46	8	25.8
Methanothrix_soehngenii_putative_peroxiredoxin	44	10	27.1
Methanothrix_sp018052825_Type-2_serine--tRNA_ligase	36	7	57
unbinned_hypothetical_protein	35	6	10.4
Methanothrix_sp018052825_hypothetical_protein	34	3	12.2
Accession Gel 14	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
unbinned_Malate_dehydrogenase BLAST: malate dehydrogenase [Bacillota bacterium] MDI9449352.1 Cov:99% ID:99.36%	2149	39	32.9
unbinned_Lactate-binding_periplasmic_protein_precursor	672	24	39.8
unbinned_Putative_2-aminoethylphosphonate-binding_periplasmic_protein_precursor	120	14	36.6
unbinned_Carbon_monoxide_dehydrogenase_medium_chain	93	15	31.2
unbinned_Glutaconate_CoA-transferase_subunit_B	82	13	33.2
unbinned_hypothetical_protein	80	11	35
unbinned_hypothetical_protein	70	5	37.4
unbinned_Autoinducer_2-binding_protein_LsrB_precursor	52	9	21.3
unbinned_Branched-chain-amino-acid_aminotransferase_2	50	12	39.9
Methanothrix_soehngenii_GDP-mannose_4,6-dehydratase	50	12	36.7
Accession Gel 15	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta_2	449	27	51.7
unbinned_putative_ABC_transporter-binding_protein_precursor	192	19	49.7
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	158	9	51.7
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	118	2	52.3
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_alpha	67	4	60.9
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	31	1	60.8
Methanothrix_soehngenii_Proline--tRNA_ligase	27	10	66.2
unbinned_hypothetical_protein	16	4	15.1
unbinned_2-amino-4,5-dihydroxy-6-one-heptanoic_acid-7-phosphate_synthase	14	8	27.7
unbinned_Glutaconate_CoA-transferase_subunit_B	13	8	28.6
Accession Gel 16	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	4060	65	51.7
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta_2	1466	20	51.7

Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	1051	29	52.3
Methanothrix_soehngenii_Hydroxyacylglutathione_hydrolase	214	30	50.5
unbinned_putative_glycine_dehydrogenase_(decarboxylating)_subunit_2	176	27	52.9
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	128	21	60.8
Methanothrix_soehngenii_Pyruvate_carboxylase_subunit_A	111	19	54.1
unbinned_hypothetical_protein	100	27	48.8
Methanothrix_soehngenii_hypothetical_protein	93	26	80.7
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_alpha	73	34	88.6
Accession Gel 17	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	**1486**	**24**	60.8
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_alpha	**1408**	**23**	60.9
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta	**1269**	**26**	45.4
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_gamma	**1089**	**15**	28.5
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_beta	**1041**	**26**	45.4
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_gamma	**1032**	**18**	28.5
unbinned_hypothetical_protein	107	6	62.4
unbinned_hypothetical_protein	78	7	46.1
unbinned_hypothetical_protein	57	5	63
unbinned_Glutaconate_CoA-transferase_subunit_B	53	17	28.6
Accession Gel 18	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta	11	3	45.4
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_beta	8	2	45.4
Methanothrix_soehngenii_hypothetical_protein	8	1	38.1
Methanothrix_sp018052825_Aspartate-semialdehyde_dehydrogenase	6	4	37.1
unbinned_Multidrug_resistance_protein_MdtB	3	1	94.5
Methanothrix_soehngenii_Aspartate-semialdehyde_dehydrogenase	3	2	37.2
Methanothrix_sp018052825_Glyceraldehyde-3-phosphate_dehydrogenase	2	2	41.7
Methanomassiliicoccus_sp._UBA6_Inositol_2-dehydrogenase/D-chiro-inositol_3-dehydrogenase	2	1	36
unbinned_hypothetical_protein	2	1	27.6
Holophagae_bacterium_UvrABC_system_protein_A	1	1	103.3
Accession Gel 19	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta_2	**1177**	**42**	51.7
Methanothrix_sp018052825_Aspartate-semialdehyde_dehydrogenase	917	24	37.1
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	517	18	51.7
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	377	17	52.3
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	176	20	60.8
unbinned_Threonine--tRNA_ligase	84	28	68.3
unbinned_Oligopeptide-binding_protein_AppA_precursor	59	3	46.2
unbinned_Malate_dehydrogenase	53	12	32.9
coenzyme_A_transferase	52	20	49.7
unbinned_hypothetical_protein	31	18	72.9
Accession Gel 20	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	**356**	**28**	51.7
unbinned_Acyl-CoA_dehydrogenase	**206**	**33**	42.2
Methanothrix_sp018052825_Aspartate-semialdehyde_dehydrogenase	194	14	37.1
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta_2	180	3	51.7
unbinned_D-malate_dehydrogenase_[decarboxylating]	118	18	41.1
unbinned_Serine_hydroxymethyltransferase	94	18	52.7
unbinned_Acyl-CoA_dehydrogenase	93	26	41.3
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta	74	22	45.4
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	47	6	52.3
unbinned_3-hydroxy-3-methylglutaryl-coenzyme_A_reductase	37	14	42.1
Accession Gel 21	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta_2	**1023**	**34**	51.7
Methanothrix_sp018052825_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	389	8	51.7
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_alpha	305	12	60.9
Methanothrix_soehngenii_Acetyl-CoA_decarbonylase/synthase_complex_subunit_beta	299	1	52.3
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	243	11	60.8
unbinned_hypothetical_protein	46	3	12.8
Methanothrix_soehngenii_hypothetical_protein	19	12	56.7

unbinned_NAD(P)H-quinone_oxidoreductase_subunit_I,_chloroplastic	11	5	76.3
Methanothrix_soehngenii_V-type_ATP_synthase_alpha_chain	9	5	63.6
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_beta	7	4	45.4
Accession Gel 22	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta	4	2	45.4
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_alpha	3	1	60.9
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	3	1	60.8
Methanothrix_soehngenii_Methyl-coenzyme_M_reductase_subunit_beta	3	1	45.4
unbinned_hypothetical_protein	3	1	24.6
Smithella_sp._putative_cation-transporting_ATPase_F	2	1	97.6
unbinned_hypothetical_protein	2	1	8.5
unbinned_DNA_polymerase_I,_thermostable	1	1	90.5
unbinned_hypothetical_protein	1	0	85.8
Synthrophorhabdus_sp._DNA_translocase_FtsK	1	0	75.4

Table S3. X-ray crystallography data collection statistics

	D2Q1S1
Data collection	
Wavelength (Å)	0.97625
Space group	C2
Resolution (Å)	128.8 − 2.4 (2.7 − 2.4)
Cell dimensions	
a, b, c (Å)	237.160 155.451 206.266
α, β, γ (°)	90.0 103.8 90.0
R_{merge}(%)[a]	9.0 (110.5)
R_{pim} (%)[a]	5.2 (62.2)
$CC_{1/2}$ [a]	0.995 (0.561)
I/σ[a]	8.8 (1.5)
Spherical completeness[a]	76.9 (17.1)
Ellipsoidal completeness[a]	94.1 (64.7)
Redundancy[a]	3.9 (3.9)
Nr. unique reflections[a]	205,286 (10,264)

[a] Values relative to the highest resolution shell are within parentheses

Table S4 (next page): Top peptide hits for protein fractions purified from the membrane fractions. Best hits per gel band are shown in green. #PSM and #unique peptides are colored from red to yellow to green for 0-20-40 and 0-5-10 respectively**.** Selected unbinned entries have been manually annotated with BLAST.

Accession Gel 1	# PSMs	# Unique Peptides	MW [kDa]
Methanothrix_sp018052825_V-type_ATP_synthase_beta_chain	52	16	50.9
Methanothrix_sp018052825_V-type_ATP_synthase_alpha_chain	42	15	63.3
unbinned_hypothetical_protein BLAST: TPA: hypothetical protein [Candidatus Kapaibacterium sp.] HPI18950.1 Cov:100% ID:92.37%	20	13	202.3
unbinned_4-hydroxybenzoyl-CoA_reductase_subunit_alpha	5	3	88
Methanothrix_soehngenii_V-type_ATP_synthase_alpha_chain	5	1	63.6
unbinned_hypothetical_protein	4	1	180.8
unbinned_hypothetical_protein	4	2	178.4
unbinned_hypothetical_protein	4	4	178.2
unbinned_hypothetical_protein	4	1	62.4
unbinned_DNA-directed_RNA_polymerase_subunit_beta	3	1	158.7
Accession Gel 2	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Bacteroidetes_bacterium_Ferrienterobactin_receptor_precursor	54	19	121.3
unbinned_hypothetical_protein	17	9	116.7
unbinned_4-hydroxybenzoyl-CoA_reductase_subunit_alpha	14	7	88
unbinned_hypothetical_protein	13	7	125.2
Methanothrix_sp018052825_V-type_ATP_synthase_alpha_chain	12	6	63.3
unbinned_Vitamin_B12_transporter_BtuB	11	8	116
Methanothrix_sp018052825_V-type_ATP_synthase_beta_chain	10	6	50.9
Bacteroidetes_bacterium_hypothetical_protein	8	2	56.1
unbinned_hypothetical_protein	7	5	94.3
Bacteroidetes_bacterium_hypothetical_protein	6	2	54.3
Accession Gel 3	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_V-type_ATP_synthase_beta_chain	561	29	50.9
unbinned_V-type_ATP_synthase_beta_chain	39	10	51.2
unbinned_hypothetical_protein	21	10	54.9
Bacteroidetes_bacterium_hypothetical_protein	14	8	58.1
unbinned_V-type_ATP_synthase_beta_chain	14	5	50.8
unbinned_hypothetical_protein	10	2	51.2
Methanothrix_soehngenii_CoB--CoM_heterodisulfide_reductase_iron-sulfur_subunit_D	6	3	50.8
Methanothrix_soehngenii_hypothetical_protein	6	4	46.4
unbinned_ATP_synthase_subunit_beta	5	2	50.4
Methanomassiliicoccus_sp_UBA6_hypothetical_protein	5	2	39.1
Accession Gel 4	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
unbinned_hypothetical_protein (Methanoregulaceae_archaeon_DUF3821_domain-containing protein) BLAST: TPA: MEMAR_RS02690 family S-layer glycoprotein [Methanoregulaceae archaeon] Cov:100% ID:100%	62	14	89.5
Bacteroidetes_bacterium_hypothetical_protein	24	8	68.5
unbinned_Vitamin_B12_transporter_BtuB	12	6	117.3
Bacteroidetes_bacterium_hypothetical_protein	16	6	73.7
unbinned_hypothetical_protein	14	6	73.7
Methanothrix_sp018052825_V-type_ATP_synthase_alpha_chain	6	5	63.3
Methanothrix_sp018052825_hypothetical_protein	6	4	165.7
Bacteroidetes_bacterium_Ferrienterobactin_receptor_precursor	7	4	121.3
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta	7	4	45.4
Bacteroidetes_bacterium_hypothetical_protein	14	4	18.7

Accession Gel 5	# PSMs	# Unique Peptides	MW [kDa]
unbinned_hypothetical_protein Methanosaeta _sp. PtaU1.Bin112 _S-layer_family_duplication_domain_containing) BLAST: TPA: S-layer protein domain-containing protein [Methanothrix sp.] Cov:100% ID:96.7%	143	22	78.1
unbinned_hypothetical_protein	29	10	89.5
unbinned_Oligopeptide-binding_protein_OppA_precursor	12	9	67
unbinned_hypothetical_protein	6	1	12.4
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta	5	3	45.4
unbinned_Protease_1_precursor	1	1	108.4
unbinned_putative_subtilase-type_serine_protease_precursor	1	1	73.2
unbinned_putative_subtilase-type_serine_protease_precursor	1	1	72.8
unbinned_hypothetical_protein	1	1	71.9
unbinned_Oligopeptide-binding_protein_AppA_precursor	1	1	69.2
Accession Gel 6	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_soehngenii_Ketol-acid_reductoisomerase	39	9	36.6
unbinned_hypothetical_protein	16	5	78.1
Methanothrix_sp018052825_Ketol-acid_reductoisomerase	15	5	36.8
unbinned_Ketol-acid_reductoisomerase	11	5	36.5
unbinned_Ketol-acid_reductoisomerase	11	4	36.5
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_beta	10	3	45.4
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	6	1	60.8
unbinned_F420-dependent_glucose-6-phosphate_dehydrogenase	6	2	35.4
Methanothrix_sp018052825_Tetrahydromethanopterin_S-methyltransferase_subunit_H	6	3	33.7
unbinned_F420-dependent_glucose-6-phosphate_dehydrogenase	5	1	37.2
Accession Gel 7	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_sp018052825_Methyl-coenzyme_M_reductase_subunit_alpha	21	2	60.8
unbinned_sn-glycerol-3-phosphate-binding_periplasmic_protein_UgpB_precursor	16	6	51.9
Methanothrix_soehngenii_V-type_ATP_synthase_alpha_chain	12	4	63.6
Methanothrix_sp018052825_V-type_ATP_synthase_alpha_chain	11	4	63.3
unbinned_putative_ABC_transporter-binding_protein_precursor	6	2	64.5
Methanomassiliicoccus_sp_UBA6_V-type_ATP_synthase_alpha_chain	4	0	63.7
unbinned_hypothetical_protein	4	2	52.3
unbinned_Periplasmic_dipeptide_transport_protein_precursor	3	1	68.7
unbinned_hypothetical_protein	2	2	89.5
unbinned_hypothetical_protein	2	1	66.2
Accession Gel 8	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
Methanothrix_soehngenii_hypothetical_protein BLAST: S-layer protein domain-containing protein [Methanothrix sp.] Cov:100% ID:93.87%	989	34	77.6
unbinned_hypothetical_protein	134	19	95
unbinned_hypothetical_protein	75	2	94.5
Methanothrix_soehngenii_hypothetical_protein	70	14	119.8
unbinned_hypothetical_protein	33	2	48.1
unbinned_hypothetical_protein	24	5	78.1
Methanothrix_sp018052825_hypothetical_protein	15	5	118.4
unbinned_Vitamin_B12_transporter_BtuB	15	8	117.1
unbinned_putative_D,D-dipeptide-binding_periplasmic_protein_DdpA_precursor	5	2	75.4
Methanothrix_soehngenii_hypothetical_protein	4	1	28.3

Accession Gel 9	# PSMs	# Unique Peptides	MW [kDa]
Methanothrix_soehngenii_hypothetical_protein BLAST: S-layer protein domain-containing protein [Methanothrix sp.] WP_348305550.1 Cov:100% ID:93.87%	96	13	77.6
unbinned_putative_subtilase-type_serine_protease_precursor	7	4	73.2
unbinned_S-methyl-5'-thioadenosine_phosphorylase	7	1	28.4
unbinned_putative_subtilase-type_serine_protease_precursor	5	3	72.8
Methanothrix_soehngenii_hypothetical_protein	3	2	119.8
unbinned_hypothetical_protein	2	1	19.8
unbinned_hypothetical_protein	1	1	162.6
unbinned_hypothetical_protein	1	1	95
unbinned_hypothetical_protein	1	1	78.1
unbinned_Stage_V_sporulation_protein_D	1	1	66.1
Accession Gel 10	**# PSMs**	**# Unique Peptides**	**MW [kDa]**
BLAST: S-layer protein domain-containing protein [Methanothrix sp.] WP_348305550.1 Cov:100 ID:93.87%	55	14	77.6
BLAST: TPA: ABC transporter substrate-binding protein [Anaerolineaceae bacterium] HQO96929.1 Cov:100% ID:99.88%	50	8	95
unbinned_hypothetical_protein	23	1	94.5
Methanothrix_soehngenii_hypothetical_protein	3	2	119.8
unbinned_hypothetical_protein	3	0	48.1
unbinned_Heterocyst_differentiation_ATP-binding_protein_HepA	1	1	57.6
unbinned_hypothetical_protein	1	1	30.7
unbinned_hypothetical_protein	1	1	28.4
unbinned_Dihydroorotate_dehydrogenase_B_(NAD(+)),_electron_transfer_subunit	1	1	28
unbinned_hypothetical_protein	1	1	27.5

References

1. Berger, S., Cabrera-Orefice, A., Jetten, M. S. M., Brandt, U., Welte, C. U. Investigation of central energy metabolism-related protein complexes of ANME-2d methanotrophic archaea by complexome profiling. *Biochim. Biophys. Acta Bioenerg.* **1862**, 148308, doi:10.1016/j.bbabio.2020.148308 (2021).
2. Jetten, M. S., Stams, A. J., Zehnder, A. J. Isolation and characterization of acetyl-coenzyme A synthetase from *Methanothrix soehngenii*. *J. Bacteriol.* **171**, 5430-5, doi:10.1128/jb.171.10.5430-5435.1989 (1989).
3. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., Bricogne, G. Data processing and analysis with the *autoPROC* toolbox. *Acta Crystallogr. D Struct. Biol.* **67**, 293-302, doi:10.1107/S0907444911007773 (2011).
4. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., et al. Overview of the *CCP4* suite and current developments. *Acta Crystallogr. D Struc. Biol.* **67**, 235-242, doi:10.1107/S0907444910045749 (2011).
5. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zídek, A., Potapenko, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).

CHAPTER VIII. Discussion

The aim of this thesis was to evaluate the crystallomics approach and explore the possibility of expanding it to environmental samples, with the goal of studying enzymes of the microbial dark matter. The presented chapters detail incremental steps of applying the approach of native protein purification to samples with increasing microbial heterogeneity, yielding novel scientific insights and highlighting strengths and weaknesses of the approach.

Scientific insights gained from the native approach

Utilizing the native protein purification method in the presented chapters has been instrumental in gaining scientific insights that might not have been uncovered with other approaches.

Under control – studying different GS regulation systems

Chapter III describes the different regulation systems in glutamine synthetases from two methanogens. The purification protocol was obtained through the application of the crystallomics approach by Engilberge *et al.*[1] creating the basis for native purification, crystallization, and further analysis of two archaeal GS, while other studies have required recombinant expression[2-4].

Apart from the GS from *M. thermolithotrophicus* an additional enzyme could be isolated throughout the experiments. The ketol-acid reductoisomerase (KARI) was previously observed to co-purify with GS. Both enzymes were separated into pure fractions through optimization of the purification protocol. KARI was then further investigated in parallel and described by by Lemaire *et al.*[5]. The approach, therefore, allowed us to not only isolate the target but also an additional enzyme. This highlights the value of the crystallomics approach as parallel purification remains feasible even in targeted subsequent purification steps, yielding additional insights.

Furthermore, the enzymatic studies and more thorough analysis of the regulation mechanism in GS was facilitated by the native approach. Initially, the enzyme was found to be inactive, which contradicted previous reports of GS activities in archaea[4, 6]. If the typical recombinant approach had been used, it would have been reasonable to suspect common problems such as incorrect folding or insolubility that interfere with enzyme activity. However, because the enzyme was natively purified and freshly prepared, these concerns were ruled out, prompting a more in-depth analysis of the regulation system revealing the complete inactivation of the enzyme in the absence of 2OG.

During the preparation of the manuscript a recent publication emerged describing the structure of recombinantly expressed GS from *M. mazei* (*Mm*GS). *Mm*GS was found to be present mostly in a dimeric state while the dodecameric complex is stabilized by the binding of the P_{II} regulatory protein GlnK, representing a system of GS regulation through modulation of the oligomeric state[2]. However, GlnK binding to the dodecameric GS could not be reproduced in a following study, which instead reported 2OG to be the trigger for complex assembly[3]. Moreover, activation of *Mm*GS by 2OG binding in an allosteric site, causing conformational changes that translate to the active site, as well as feedback inhibition by glutamine, were observed. This is in accordance with the proposed presence of these regulation mechanisms in *Methanosarcinales* group 2 (including *M. mazei*) based on sequence conservation (see Ch. III). Contrastingly, no alternative assemblies apart from the dodecameric state were observed in *M. thermolithotrophicus* or *M. shengliensis*. While this could be due to the carryover of 2OG from the cell extract stabilizing the complex, it is unlikely, as the protein fractions were diluted heavily during chromatography. Furthermore, crystallization of the apo states was successful, displaying fully assembled dodecameric complexes in the absence of GlnK or 2OG. Further studies on archaeal GS will be required to determine if regulation through modulation of the oligomeric state is a unique feature in *M. mazei*

or presents an additional regulation system for archaeal GS and native enzyme structures could provide essential insights to solve this question.

Besides the presented work, other components of the methanogenic N-assimilation have been elucidated using the native approach in the MicroMet group. Previous work could show the release of GlnK regulatory proteins from the native membrane fraction of *M. thermolithotrophicus* upon Mg^{2+}, ATP and 2OG addition, supporting their conserved mechanism and role as ammonium transporter regulators in methanogens[7]. Moreover, the nitrogenase regulator $NifI_{1,2}$ (also a PII family member) was purified from both native and recombinant sources and structural analysis revealed the formation of a dimer of hetero-trimers consisting of 2:1 $NifI_1$:$NifI_2$ [8]. Further investigations resolved the mechanism of Nitrogenase inactivation through the native purification and structural elucidation of the $NifDKI_{1,2}$ complex from *M. infernus*[9]. Finally, the capture of GlnK from *M. infernus* (Ch. II) further supports the potential of the native approach for the study of methanogenic N-assimilation regulation systems.

Snapshots of archaeal C-assimilation enzymes

Applying the untargeted shotgun approach to the hyperthermophile *M. infernus,* yielded two enzymes of the C-assimilation pathway. Enolase catalyzes the conversion of phosphoenolpyruvate and H_2O to 2-phospho-D-glycerate, a step of gluconeogenesis for the generation of carbohydrates that are used as glycogen storage or for biosynthetic purposes. To date, only one archaeal enolase structure has been deposited on the PDB and remains to be described (PDB: 2PA6). Furthermore, the enzyme 2-oxoglutarate:ferredoxin oxidoreductase, which catalyzes the reversible reaction of CO_2 reduction by ferredoxin and succinyl-CoA to form 2OG, was captured. A reproducible purification protocol was established and allowed more detailed follow up studies of the enzyme. A structure of the apo state (lacking the TPP cofactor) was obtained, which in combination with further structural studies, could give insights into the structural

remodeling during the cofactor binding process. Activity assays indicated that TPP was only partially lost in the purified enzyme fraction and reconstitution with TPP should stabilize the complex and facilitate subsequent crystallization attempts for obtaining the holo state. Previous structural studies have only been conducted for the recombinant OGOR from *Magnetococcus marinus.* A more detailed comparison between the two structures could also reveal adaptations to different physiological temperatures as the optimal growth was observed at 25 °C for *M. marinus*[10] in contrast to 85 °C for *M. infernus*[11]. Moreover, *Mi*OGOR possesses a δ-subunit, usually absent in 2OG utilizing homologs and therefore poses as an intriguing anchor for phylogenetic analysis of the enzyme's evolution within archaea.

MCRs from species utilizing three different methane metabolisms

The native approach was successfully applied to capture MCR, not just from cultivated organisms, but also from organisms that are challenging or impossible to grow in pure cultures, by utilizing samples with increasing microbial complexity. Overcoming the cultivation limitations allowed the study of this enzyme across various new organisms with varying environmental contexts. In total, six new MCR structures were obtained (Ch. IV-VI), increasing the data pool of MCR structures from different species by 75 %. The described variants originate from organisms employing diverse methane-related pathways: methoxydotrophic and acetoclastic methanogenesis as well as methanotrophy. All obtained MCRs shared a conserved overall structure as well as active site, highlighting the strong structural restraints resulting from the intricate reaction mechanism. The substrates CoM and CoB were observed at the canonical positions. Additionally, the heterodisulfide could be observed in *Methanothrix* MCR purified directly from the sludge, potentially due to the direct purification from the complex environment. All structures contained the classical F_{430} cofactor.

Furthermore, various PTMs could be observed in the high-resolution structures, as the MCRs were purified from their native source ensuring correct installation. While being conserved, the physiological roles of the modifications have not yet been solved. Previous studies have deciphered that PTM installation occurs independently from each other, with *S*-methylcysteine (SMC) being installed by a SAM-dependent methyltransferase designated McmA[12], thioglycine (GL3) by TfuA and YcaO[13], 5-C-(*S*)-methylarginine (AGM) by the methanogenesis marker protein Mmp10[14] and 2-C-(S)-methylglutamine by a cobalamin-dependent radical SAM glutamine C-methyltransferase[15]. The installation of all other PTMs remains unsolved. Studies of deletion mutants lacking one or multiple of these PTMs (SMC, GL3, AGM) showed that neither is essential for growth[12-14] and no effects on the overall enzyme structure or active site could be detected. Combinatorial mutants displayed several significant growth phenotypes, including growth defects in mutants lacking GL3 and AGM when grown on low free energy-yielding substrates, and SMC lacking mutants at temperatures lower and higher than the optimal[12, 13]. However, the triple mutant was found to grow better than the wild type under nearly all tested conditions. It has been proposed that PTMs form hard-to-predict physiologically relevant interactions potentially related to catalysis[12]. The growth-rate comparison of closely related species, as an indicator for *in vivo* activity, revealed slightly faster growth (1.6-fold rate) for the species harboring an MCR with a didehydroaspartate (DYA) instead of a canonical aspartate. This suggests that the modification may confer an ecological advantage over other methanogens[16].

The observed PTM patterns and numbers varied drastically between the obtained MCR variants (Table 1.). However, clear distribution patterns following phylogeny could not be identified, modifications differ even between closely related species. For instance, *M. shengliensis* MCR contains only a minimal amount of PTMs compared to *M. acetivorans* and *M. barkeri* MCR. In contrast, the closely related ANME-2d MCRs possess the highest known number. Furthermore, the DYA is not shared among

Table 1. Overview of MCR PTMs. Growth temperature according to DSMZ[17]. Blue entries were determined by MS in the absence of an experimental structure[18], the grey entry refers to ECR.

Organism	Substrate(s)	T (°C)	PTMs								PDB codes
			GL3	SMC	MHS	AGM	MGN	DYA	TRX	Other	
M. marburgensis	$H_2 + CO_2$ [19]	65	X	X	X	X	X	X			1MRO, 1HBU…
M. formicicus	$H_2 + CO_2$, formate [20]	70	X		X	X	X		X		5N28, 5N2A
M. wolfei	$H_2 + CO_2$, formate [19]	60	X	X	X	X	X				5A8K, 5A8W
M. thermo-lithotrophicus	$H_2 + CO_2$, formate [21]	65	X		X	X	X				5N1Q
M. kandleri	$H_2 + CO_2$ [22]	98	X		X	X	X				1E6V
M. barkeri	$H_2 + CO_2$, $H_2 + CH_3OH$, MMA, DMA, TMA, acetate [23, 24]	37	X	X	X	X		X			1E6Y
M. acetivorans	Acetate, CH_3OH, MMA, DMA, TMA [25, 26]	37	X	X	X	X		X			8GF5, 8GF6
ANME-1	(CH_4)	-	X		X					0AF, MHO	3SQG
Ca. Ethanoperedens thermophilum	Ethane [27]	50[27]	X		X	X	X			HIC, I2M, SMC[X]	7B1S, 7B2C
M. jannaschii	$H_2 + CO_2$ [28]	80	X		X	X	X				-
M. voltae	$H_2 + CO_2$, formate [29]	35	X	X	X		X				-
M. thermophilus	$H_2 + CO_2$, formate [30]	55	X	X		X	X				-
M. shengliensis	CH_3OH, MMA, TMA, MAC [31]	65	X		X	X					7NKG
ANME-2d O	CH_4 [32]	30[32]	X	X	X	X		X	X	MEH	This work
ANME-2d V	CH_4 *	21*	X	X	X	X		X	X	MEH	This work
ANME-2c	CH_4 *	20*	X	X	X	X	X	X	X		This work
M. soehngenii	Acetate [33]	37	X	X	X	X		X			This work
Methanothrix sp.	(Acetate)*	(37)	X	X	X	X		X			This work

Abbreviations: (MMA, DMA, TMA) mono-, di- tri-methylamine, (MAC) methoxylated aromatic compounds, (TRX) 6-hydroxy-tryptophan, (MHS) 1-*N*-methylhistidine, (SMC) *S*-methylcysteine, (AGM) 5-C-(*S*)-methylarginine, (MGN) 2-C-(*S*)-methylglutamine, (GL3) thioglycine, (DYA) didehydroaspartate, (0AF) 7-hydroxy-L-tryptophane, (MHO) S-oxymethionine, (HIC) 4-methyl-histidine, (I2M) 3-methylisoleucine, (MEH) 3(S)-methylhistidine. Additional *M. marburgensis* PDB entries: 3POT, 7B2H, 3M2R, 3M1V, 3M2U, 3M2V, 3M30, 3M32, 5G0R, 5A0Y, 5A8R, 7SUC, 7SXM, 1HBM, 1HBN, 1HBO. * See presented chapter. ( ) Inferred. [X] Presence at a non-canonical position.

M. marburgensis and *M. wolfeii* belonging to the same genus. Alternatively, temperature could correlate with certain modifications as combinatory effects on the thermal stability have been observed, and SMC was proposed to play a role in the adaptation to mesophilic conditions[12]. Closer examination of growth temperatures and PTM patterns revealed the presence of SMC in organisms growing at up to 65 °C and absence of the modification at higher temperatures, this exceeds the mesophilic range and raises the question if the proposed adaptive role of SMC might be unique to *M. acetivorans*. Furthermore, 2-C-(*S*)-methylglutamine (MGN) is absent in multiple species growing at temperatures below 50 °C with the exceptions of *M. voltae* and ANME-2c. Moreover, DYA is only found at temperatures below 37 °C with the exception of *M. marburgensis*. Due to the still limited number of experimental MCR structures from different species, it remains unclear if the observed trends and exceptions reflect true correlations between temperature and PTMs or are just coincidental.

Finally, the type of methanogenesis could be a factor, as cellular energy stress might drive the development of adaptations. As previously mentioned, the deletion of several PTMs resulted in changed growth rates and yields for various high- and low-energy-yielding substrates[12]. Noticeably, among the structurally characterized MCRs is the low content of PTMs in MCR from the methoxydotrophic methanogen *M. shengliensis* and contrastingly the high amount of PTMs in ANME-2d MCR and ECR, which each also contain unique PTMs as does MCR from ANME-1. Considering the energetics, it has been proposed that the net energy gain for methanogenesis on methoxylated substrates is twice as high as on methylated substrates[34]. Furthermore, *M. shengliensis* was isolated from an oil field[31] and shown to be able to utilize more than 30 types of methoxylated aromatic compounds[25]. This high efficiency and substrate availability might result in low cellular stress, reducing evolutionary pressure on MCR. In contrast, sulfate-driven AOM through reverse methanogenesis operates at the limits of energetic feasibility

($\Delta G°' = -21$ kJ.mol^{-1})[35]. Nitrate dependent AOM is energetically more feasible ($\Delta G°' = -521$ kJ.mol^{-1})[35]. However, slow growth with doubling times of weeks, and AOM rates comparable to AOM with sulfate, have been reported for *Ca.* M. nitroreducens[36]. Possible explanations could be that energy yield might not scale to doubling times, the presence of unknown growth limiting factors or limitation by the MCR reaction[37]. Latter has been shown to be endergonic under standard conditions ($\Delta G^{o'} = 30$ kJ.mol^{-1})[35]. Additionally, microbial activity at mud volcanoes has been reported to be limited by hydrocarbon and sulfate transport and to be affected by kinetic limitations like heat and disruptive mud expulsions[38]. Increased energetic stress might explain the PTM profiles in the studied alkanotrophs. Additionally, both ECR and ANME-1 MCR contain a modified F_{430} cofactor which might be another adaptation for the fine-tuning of MCR activity. The presence of different PTM patterns is likely the result of longer evolutionary adaptations opposed to a short term reaction to an environmental change. For instance, MCR purified from *M. shengliensis* growing on methoxylated aromatic compounds versus methanol did not exhibit a different PTM pattern (data not shown, *in preparation*). Further studies will be necessary to explore MCRs from more organisms to confirm or contradict a correlation between PTM patterns and methane metabolism.

Intriguingly, besides a high total number of PTMs in ANME-2d, an entirely novel modification was detected: the 3(*S*)-methylhistidine. This PTM has not yet been described in any protein structure, according to the authors knowledge. Furthermore, the modification stands out as it is placed on the γ-subunit, a novel feature, as all previously reported PTMs are positioned on the α-chain of MCR. This finding expands our wealth of knowledge about MCR PTMs and prompts reevaluation of previous studies, especially MS analyses of PTMs have focused solely on the α-subunit[18]. Given the discovery of this new feature, future studies should make sure to investigate all subunits for PTMs in order to not overlook any modifications. Recent recombinant attempts to study ANME-2 MCR in *Methanococcus maripaludis* struggled with subunit

exchange, resulting in the formation of hybrid-complexes of ANME-2 and host MCR[39]. Furthermore, it was observed that the employed heterologous expression system resulted in the installation of PTMs identical to those of the expression host MCR[39]. Therefore, a recombinant MCR would not accurately reflect the native PTM pattern, but would instead harbor the modifications determined by installation machinery encoded by the expression host genome. This highlights the importance for native protein purification for studying unbiased natural features of enzymes. Ultimately, the identification of a novel PTM on a different subunit in ANME-2d MCR highlights the greatest advantage of the crystallomics approach: the ability to uncover novel protein features from challenging-to-isolate microorganisms that are not easily reachable with conventional techniques.

Considerations for the application of crystallomics

While setting up the crystallomics pipeline, it became apparent that the procedure can be approached using different strategies. In the "untargeted shotgun purification" native proteins are purified in parallel from one sample in an exploratory way. Thereby, an overview of the main proteins in the cellular extract, as well as an initial purification and crystallization strategy, is obtained (Ch. II and VII). In order to perform additional and more detailed analyses, the purification of a specific target can then be repeated in a targeted second purification round (Ch. III). Alternatively, in an "adapted targeted purification", a potential target can be identified at the start and then followed by characteristic features through the purification as decisions about the purification path are made on the spot (Ch. IV, V, VI). These various strategies are adaptable to different research goals and share the native approach that facilitates the study of proteins from difficult-to-access microorganisms.

Table 2. Strengths, weaknesses and strategies of crystallomics

Aspect	Strength	Weakness	Strategies to overcome weakness
Sample quantity and quality	- Samples can be heterogeneous, pure cultures are not required - Proteins from less abundant microbial populations can be captured - Samples as low as 1 g can be utilized	- Higher protein diversity increases contamination during purification and background signals in MS - Low abundancy of the organism of interest complicates enzyme enrichment - Obtaining sufficient material from enrichments or environmental samples is a bottleneck	- More extensive purification protocols and stringent selection for the pool preparation could improve selective protein enrichment - Biological samples with a naturally high content of the organism of interest (biofilms, mats, aggregates, samples from selective environments…) can be utilized or physical enrichment methods can be performed
Output quantity, identity, diversity	- Extraction of multiple proteins from a single sample - High success rate of crystallization - Besides soluble proteins, membrane fractions and metabolites can also be studied	- Untargeted purification can yield proteins that are of less interest	- Target enzymes can be followed by natural features (absorbance, gel pattern, activity, abundance changes)
Crystallization	- Acts as final purification step to isolate pure protein populations out of a mixture	- Some proteins might not generate crystals or might not diffract - Bias towards well-crystallizing proteins	- Complementary cryo-EM analysis can extend results - Small crystals can be exploited with microED
Target protein quantity	- Capture of highly expressed proteins like catabolic enzymes is facilitated	- Capture of less expressed proteins is impeded (*e.g.*, transcriptional regulators, anabolic enzymes in energy and nutrient limited environments, specialized biosynthetic enzymes, …)	
Reproducibility	- Purification pipelines can be repeated successfully	- Protocols must be adapted to different biological samples	

Aspect	Strength	Weakness	Strategies to overcome weakness
Protein identification	- High resolution structural data can be used for protein identification through SAD phasing, fold-analysis and *ab initio* sequencing - MS on crystals yields clear identification	- Low electron density resolution doesn't allow for *ab initio* sequencing - Interpretation of MS results for fractions with a high protein diversity is difficult	
Complementary techniques	- Computational models accelerate structure solving and refinement - In solution and *in cristallo* UV/Vis spectrophotometry - SDS- and hrCN-PAGE - CryoEM and microED (see above)	- Protein sequences are required - Sequencing and MS analysis increase costs and processing time	- Coordinated collaborations can streamline the process and provide optimal analyses for maximum output
Organization and workload		- High number of samples increases the risk of generating mistakes - Prioritization during fraction pooling might require the omission of fractions with low-abundance or high protein diversity, leading to some loss of biological data	- Samples can be organized through a system for the assignment of identifiers - Dividing work between scientists would optimize output quality and quantity
Discovery potential	- Novel enzymes, functions, catalysts, cofactors and features like PTMs, bound substrates or intermediates, regulation mechanisms and Partners can be identified - Results can be unpredictable	- Results can be unpredictable	

It should be considered that, while the more targeted strategies are rather straightforward, the untargeted approach generates a rapidly increasing workload. Once the cells are lysed, purification needs to be carried out as quickly as possible to retain native enzyme properties (*e.g.*, activity, structural integrity, intrinsic (metallo)cofactors, …) and increase the chances of crystallization. During the purification the number of samples increases rapidly as each step splits the protein pools further. It is essential to maintain a consistent organizational system in order to keep track of fractions. In the performed experiments a general identifier was given for each purification attempt (*e.g.*, Minf_A) followed by the fraction identifier named after the chromatography step and pool number (*e.g.*, D1Q2S1 refers to the first pool of DEAE, the second pool of Q-sepharose and the first pool of Souce15phe in the order of purification).

Moreover, if single protein populations are split during fraction pooling this will result in redundancy in the final protein fractions, and it is advised to regularly compare SDS-PAGE profiles between samples and steps. Due to the high number of treated protein fractions, pooling will be prioritized for fractions containing distinct proteins of high quantity, while those containing a higher protein diversity at lower quantity are discarded. However, when working with valuable biomass that is challenging to obtain (*e.g.*, a slow growing enrichment culture, environmental samples), it would be advisable (if possible) to divide labor between multiple scientists and use additional chromatography equipment. If multiple chromatography and SDS-PAGE steps can be run in parallel, more fractions could be processed and therefore maximize the output. Additionally, implementation of chromatography automation and high throughput crystallization would be beneficial.

The untargeted simultaneous purification approach yielded 22 final fractions from 1.9 g of soluble protein obtained from the *M. infernus* pure culture (Ch. II) while 15 fractions were obtained from 48 mg of soluble protein of the enriched *Methanothrix* pellet (Ch. VII). The difference in starting material was reflected in the final protein

amounts, which were on average ~ 900 µg and ~ 200 µg for *M. infernus* and *Methanothrix* respectively (excluding the main MCR fraction). Nevertheless, crystallization succeeded in both attempts with crystals forming in 77 % and 80 % of the tested fractions. This success rate was unexpectedly high as crystallization is generally regarded as largely empirical, requiring patience, perseverance, and experience from scientists[40]. It is possible that the advances in chromatography equipment and resin material, as well as improved crystallization screens, contributed these remarkable results. However, it should be considered that there is likely a bias towards well-crystallizing proteins. For example, GlnK was previously observed to crystallize well and fast[7], explaining why the small regulatory protein could be captured from *M. infernus* (Ch. II). Proteins from a mixed sample will typically self-assemble into a pure crystal, a feature that is exploited by the crystallomics approach as a final protein purification step. Different protein populations can crystallize independently, potentially yielding multiple different crystals from one mixed sample. However, proteins that crystallize more readily may appear sooner, across a broader range of conditions and at lower concentrations. In contrast, proteins with lower crystallization efficiency might appear later, in fewer conditions and require higher concentrations. The formation of crystals by one protein within a mixed sample will alter the composition of the crystallization solution (*e.g.,* salt and protein concentration, pH), potentially shifting the equilibrium to reach conditions favorable for the formation of crystals by other proteins. In any case, it is extremely difficult to predict which proteins from a mixture might crystallize.

Another noteworthy aspect during purification is the influence of the abundance of single proteins within the soluble extract. More abundant proteins appear as thicker gel bands and larger chromatogram peaks, which can be followed more easily throughout the purification. On one hand this can be beneficial when targeting catabolic enzymes, especially those that catalyze a rate limiting metabolic step like MCR[41], as they are

expected to be highly expressed due to their significant physiological role. For instance, while the untargeted shotgun purification on an isolated microbe grown under laboratory conditions (Ch. II) could capture multiple anabolic enzymes, the same approach yielded a higher fraction of catabolic enzymes when applied to a more complex community grown in more natural conditions (Ch. VII). The physiological expression levels must reflect cellular needs, and the study of more natural systems could give insights into processes with a higher relevance under environmental conditions as opposed to optimized laboratory conditions. On the other hand, less expressed enzymes might not be captured with the technique.

The effects of increasing sample heterogeneity

A main aim of this work was to explore the possibility of expanding the native purification approach to environmental samples. Therefore, samples with increasing degrees of microbial diversity were utilized.

The technique was established with a pure culture as a starting point, producing satisfying results. In the next step the microbial heterogeneity was increased by working with enrichment cultures. The utilized ANME-2c and ANME-2d enrichments included a mixture of organisms with ANME-2 accounting for 15-50 % read-based abundance. MCR was chosen as a target and purification was carried out successfully. The content of ANME within the samples was comparable to the microbial mats utilized by Shima *et al.*[42], which contained ~47 % ANME-1 within the pink mat layer (based on CARD-FISH)[43]. The difficulty of working with a mixture of microbes was counterbalanced by the natural extremely high abundance of MCR, therefore, no major negative effect on protein purification was observed. To explore to which degree sample complexity could be pushed, the approach was tested on an even more heterogeneous sample. The wastewater treatment sludge contained a much higher microbial diversity, and the target genus *Methanothrix* only accounted for a small fraction (3-5 %) of metagenome reads.

To facilitate the initial purification, an attempt was made to develop a physical enrichment that utilized size separation of cells by serial centrifugation to separate larger particles and smaller cells from medium sized *Methanothrix* filaments. The resulting pellet contained a larger fraction of *Methanothrix* cells and MCR purification could be carried out. Subsequently, it was tested if purification was also possible directly from the sample without the physical enrichment step. Here, a larger amount of biomass was required, and a higher strain on the purification equipment was observed. Final fractions were less pure. However, crystallization was successful nonetheless. An unexpected finding was the presence of two MCRs with different sequences (from different *Methanothrix* strains) within the same crystal. While co-crystallization is a common observation, in this case we did not find two different adjacent proteins within one asymmetric unit, but an overlay/average of the two sequences resulting from a mix of version A and B within the crystal. Similar observations have been made in a few other studies including the proteins crambin from *Crambe abyssinica*[44] and Lili-Mip from *Diploptera punctata*[45]. Both proteins are rather small and sequence differences occur for 4 % and 16 % of all residues (in carmbin and Lili-Mip respectively), while 5 % of all residues differ between both *Methanothrix* MCRs. If the sequence differences cause only minimal structural deviations, identical crystal contacts can still be formed. This phenomenon presents both advantageous and disadvantageous implications. An averaged structure might result in the loss of detail between the two included variants and could complicate correct identification and structure building. However, if a sufficient resolution is obtained, both structures can be resolved, yielding more structural insights per sample.

The preliminary untargeted approach on the sludge sample yielded multiple promising fractions, but compared to the experiment on the pure culture it was observed that the obtained fractions were far less pure, likely due to the larger diversity of proteins in the soluble fraction. More chromatography steps or a more stringent strategy during fraction

pooling might help to increase purity. Another effect of the increased heterogeneity was observed during MS analysis of the gel bands as a high protein diversity complicated the identification of isolated enzymes. Utilizing the purer fractions obtained from washed and dissolved crystals should facilitate identification. Additionally, because MS analysis relies on computationally predicting peptide masses from the reference genome to compare with experimentally obtained data, higher sequence diversity results in a higher computational effort. An incomplete reference genome database could also affect the accurate detection of peptides as well as the identification of PTMs[46]. Interpretation of MS results from complex samples should, therefore, be approached with caution.

Notably, cell lysis was found to be an additional step that affects protein purification from a mixed sample (Ch. VI). Analysis of the total extract, after sonication, with FISH revealed that *Methanothrix* cells were successfully lysed, however some intact DAPI-stained cells could still be observed. It is possible that microbial cells with more sturdy cell envelopes (*e.g.*, gram-positive bacteria) withstood the lysis method. This would exclude some bacterial proteins from the following purification and increase the relative abundance of *Methanothrix* proteins in the soluble extract, acting as a selective filtering method. The approach of differential lysis could facilitate the study of archaeal proteins. Since the archaeal and bacterial cell envelopes differ[47], various selective lysis methods could be employed. For instance, methanogens are less sensitive to lysozyme[48] treatment which is regularly employed for the degradation of the bacterial peptidoglycan cell wall[49]. Therefore, pre-treatment of biomass with lysozyme (and detergents) could be applied for the lysis of bacterial cells. The bacterial soluble proteins could then be discarded followed by lysis of archaeal cells through sonication and osmotic shock resulting in soluble extract with a high content of archaeal proteins. Alternatively, treatment with proteaseK or achromopeptidase was shown to lyse methanogenic cells with an S-layer or sheath[48] while the bacterial cell envelope should be unaffected. Selective cell lysis could reduce contamination with bacterial proteins

(or archaeal proteins if the target organism is a bacterium), and feasibility should be explored in further studies.

The obtained results show that the technique can be applied to samples with an increased microbial diversity. Enzyme purification and crystallization can be achieved despite a higher sample complexity, which encourages further studies on heterogeneous environmental samples.

Complementary techniques

For a successful application of crystallomics, techniques beyond protein purification and crystallization have been found to be required or beneficial.

Obtaining the protein sequence is critical to identify targets and generate structural models. When working with a cultured organism the genome is available, but for mixed microbial samples, such as microbial enrichments or environmental samples, metagenomic analyses are indispensable. They allow to get an overview about the microbial diversity as well as the sequences of the majority of potential enzymes in the sample.

To solve the crystal structure of an unidentified protein, SAD phasing can be employed (see Ch. II). In some cases of *de novo* solved structures, the resolution of the data might allow backbone modeling to identify the enzyme by its fold, which can then be combined with selecting the correct sequence from a pool of potential sequences by examining single residues in the density. However, this requires a high resolution and is only possible when the pool of potential sequences is limited (*e.g.*, four P_{II} sequences in a pure culture of *M. infernus*). This is not the case in a diverse sample with plenty of species potentially harboring a plethora of similar proteins. Therefore, MS greatly contributes to identifying enriched proteins and their source organisms.

If the protein can be identified before data processing, molecular replacement can be performed with published homologous models. In the absence of a homologous structure, computed structural models can be utilized. Computational protein modeling has had a great impact on the field of structural biology and bears the potential for a great symbiosis with the approach of crystallomics. Computational models accelerate the process of solving structures through molecular replacement and subsequent refinement efforts. In turn, crystallomics yields new experimental models that enhance the databases on which algorithms are trained. While the most recent Nobel prize was awarded to key contributors to computational protein structure prediction[50], it is important to acknowledge that this advancement builds upon decades of experimentally determined structures. Latter provide essential insights into reaction mechanisms, conformational states and novel native features like cofactors and PTMs that cannot be unveiled through modeling.

While crystallomics relies on crystallization as the last purification step, it cannot be excluded that obtained crystals will be too small for conventional X-ray crystallography. In this case, microcrystals could be analyzed via microED. Furthermore, if obtained crystals do not diffract, cryoEM might be an alternative to model the structure as dissolved crystals could yield pure fractions for further analysis. Moreover, advances in EM technology (software and hardware) allow the processing of increasingly mixed samples and, similar to crystallomics, shotgun EM has been proposed as a technique for the study of enzymes from a heterogeneous mix. In a proof of concept, proteins from human cell lysate were separated by chromatography followed by MS and EM analysis[51]. This approach could be an alternative to crystallomics and potentially also applied to purified fractions that are not yielding crystals at all. Furthermore, some proteins were successfully extracted from the membrane fraction (Ch. VII) and cryo-EM may be more suitable than X-ray crystallography for their structure determination.

As most of these complementary techniques require extensive equipment and expert knowledge, the presented approach would greatly benefit from collaborations or even the establishment of a research consortium. Designating experts for different steps of the pipeline - sampling, sequencing, processing of genomic data, protein extraction and purification, crystallization, mass-spectrometry analysis, X-ray diffraction data collection and processing, (cryo-EM), structure refinement, interpretation and enzymatic analyses - would streamline the processing of results and allow the best interpretation.

Reproducibility and strategies for specific targets

Given the exploratory nature of the crystallomics approach and the challenge of natively enriching multiple targets, is crucial to determine reproducibility for in-depth enzyme characterization.

In the presented work, GS and OGOR could be repeatedly purified from pure cultures. Moreover, MCR from *Methanothrix* was purified twice from very heterogeneous wastewater treatment sludge. In the latter case, reproducibility could still be achieved when working with samples collected at different time points, even though the microbial population had undergone some changes. These results show that once a purification pipeline is established the target enzyme can be successfully repurified using the same protocol. It should be noted though that reproducibility is not guaranteed when switching the biological system, as illustrated by the different purification protocols required to purify GS from *M. thermolithotrophicus* and *M. shengliensis*.

Adapted targeted purification of specific targets was facilitated by specific aiding features. MCR is known to be highly abundant in the cell, has a very recognizable SDS-PAGE profile and the F_{430} cofactor that absorbs at a special wavelength (430 nm). Similarly, OGOR has a distinct brown color caused by Fe-S clusters that can be followed by absorbance (410-415 nm) and a specific pattern on SDS-PAGE due to the

four distinct subunits. In contrast GS does have none of these features and reproducible purification relied entirely on the precise protocol and a distinct elution volume during size exclusion chromatography due to the large dodecameric assembly.

However, utilizing unique features might aid in future studies to either allow easier tracking of the target enzyme or to identify a potential target during the first purification attempt. Exploitable features are, as previously mentioned, a unique SDS-PAGE profile or unique absorbance features caused by cofactors or intrinsic metals. Furthermore, enzymatic activity might be an exploitable feature. Enzyme assays could be performed on the final purification fractions to identify the presence of the target. A variety of labeled substrates for in-solution activity assays are available. Fluorogenic and chromogenic substrates allow the detection of various enzyme classes through fluorescence or color changes. Targets include proteases, glycosidases, lipases, esterases, oxidoreductases and oxidases[52]. Alternatively, in-gel zymography allows to detect enzyme activity directly on SDS-PAGE or native PAGE by tracking substrate conversion via various visualization methods like clearing zones, color or fluorescence changes and different stains[53]. Targets for this include hydrolases acting on various substrates, such as proteases[54], esterases[55], lipases[56], cellulases[57], glycosidases[58] and nucleases[59]. This approach has been demonstrated for the purification of the carbon-monoxide dehydrogenase/acetyl-CoA synthase complex from *Clostridium autoethanogenum* where viologen dependent oxidase activity (causing red signals on hrCN-PAGE) helped to follow the enzyme through purification[60]. Similarly, the dissimilatory sulfite reductase from *Ca.* Desulfofervidus auxilii was followed through purification on hrCN-PAGE with heme staining[61], detecting heme-assicoated peroxidase activity[62]. However, especially in-solution enzymatic studies are less likely to succeed during early purification steps as the high protein diversity causes unpredictable background activities that will likely distort results. Therefore, later stages of the purification are more suited for obtaining clearer results. Performing

activity assays at extreme conditions (*e.g.,* high temperatures) could contribute to the detection of extremophilic enzymes which are in high demand for biotechnological applications as their adapted properties make them ideal biocatalysts (see below).

Furthermore, Western blot analyses could be utilized for the detection of specific target proteins. The approach has traditionally depended on the synthesis of specific antibodies through the recombinant expression of the target protein, injection into an animal to induce an immune response[63]. This method is time and cost intensive and cannot be applied for novel proteins. However, recently the development of AI-based tools for the generation of antibodies has started. For instance, the PALM-H3 and A2binder computational models are able to generate antibody sequences targeting a specific antigen followed by the prediction of binding specificity and affinity for optimization[64]. The proposed antibodies can be synthesized and applied for Western blot protein analysis bypassing the requirement for natural sources. As AI applications are developing rapidly, more conventional techniques may become accessible for application on novel enzymes.

Another strategy to identify a potentially interesting target could be the utilization of the expression changes. In the case of lab-grown cultures and enrichments, it is possible to change the growth conditions in order to observe differences in the soluble extract proteins. For example switching an organism from NH_3 to N_2 will cause a shift in the expression pattern [65]. Differentially expressed enzymes related to N-assimilation could be identified as new bands on SDS-PAGE and then be followed during purification. This can not be applied to environmental samples, even though it would be interesting to compare the soluble extract from different sampling time points (*e.g.*, day vs night, summer vs winter). However, larger shifts in the environmental conditions will likely alter the microbial community in a way that utterly changes the soluble protein profile obscuring any single protein abundance changes.

Transferability to other microbial samples

While there are various bacterial model organisms, archaea remain less well studied. Crystallomics could be applied to environmental samples containing uncultured archaea to give insights about their diverse metabolisms as well as their interaction with other microbes. However, when considering potential samples, the abundance of the microbe(s) of interest should be considered. Purification of *Methanothrix* MCR was successful even though the group only accounted for a fraction of the sample (~3-5 % metagenomics reads). However, it is likely that protein purification from less abundant microbes would be more challenging. Therefore, biomass containing a larger fraction of the target organism could be utilized. A promising way to study archaea of the MDM could be to target microbial hubs with larger archaeal populations like biofilms, microbial mats or aggregates.

This was demonstrated by Shima *et al.*[42] for the purification of MCR from ANME-1 and ANME-2 enriched microbial mats found at the sediment-water interface in the Black Sea[43]. Other environments of interest could, for example, be the sediments of cold acid mine drainage streams that are dominated by uncultivated *Thermoplasmatales* (58 % 16S rRNA amplicon reads)[66]. Similarly, in acidic ferric iron mats from Yellowstone geothermal systems, *Ca.* Geoarchaeota accounts for up to half of the total metagenomic reads. Their metabolic capabilities have been proposed to include novel strategies for energy conservation and respiration with adaptations to low pH, high temperature, low oxygen, and high reduced iron conditions[67]. Additionally, *Ca.* Marsarchaeota are also abundant in acidic microaerobic Fe(III)-oxide mats and are potentially mediating iron reduction[68]. Furthermore, ANME do not just form consortia with SRB in the deep sea[69] but have also been found to form biofilms in sediment pockets at arctic cold seep sites[70] and formation of consortia has also been reported for related alkanotrophs[27, 71]. Moreover, high amounts of uncultured *Euryarchaeota* and *Ca*. Thaumarchaeota have been detected in interior zones of deep-sea hydrothermal sulfide chimneys[72].

While some MDM members have been obtained in enrichment cultures[27, 71, 73-79] many species still evade enrichment attempts[80, 81]. An alternative to samples with a naturally high target abundance or lab-grown enrichment cultures could be the employment of physical enrichment methods like the demonstrated centrifugation-based size separation of *Methanothrix* filaments. Previous studies aiming to cultivate MDM members include methods that could also be employed for crystallomics.

For example, the extremely small uncultivated *Ca.* Micrarchaeota with cells <0.006 μm^3 was highly enriched through filtration from acid mine drainage pools[82] and the near single species biofilm droplets of *Ca.* Altiarchaeum hamiconexum has been harvested by an *in situ* biofilm trapping system, utilizing polystyrene nets as artificial attachment material, in streams originating from deep subsurface layers of cold anoxic freshwater springs[83]. The latter organism is highly interesting due to its unique features like a double membrane, a proposed novel archaeal acetyl-CoA pathway and highly structured cell surface hooks (hami)[84, 85]. In another sophisticated study, uncultivated aerobic organoheterotrophic bacteria could be grown in diffusion growth chambers, placed on top of their natural habitat intertidal marine sediments, that allowed the exchange of metabolites with the surrounding microbial community[86]. Furthermore, flow cytometry-based cell sorting might also have the potential for studying uncultivated microbes. Studies have shown that methanogenic cells can be quantified and sorted by exploiting the natural fluorescence signal of the F_{420} cofactor[87] and protocols to sort uncultivated environmental microbes based on cellular properties like size, fluorescence and granularity have been developed for single cell sequencing[88]. Even though the current maximal throughput (*e.g.*, sorting speed up to >70.000 events per second[89]) would not yield exploitable cell amounts for protein extraction, instruments and protocols are constantly improving so the approach might become feasible in some years.

It is noteworthy that, in addition to a sufficient abundance of the microbe of interest, the total quantity of biomass must also be considered. In the presented work, successful

purifications were carried out with a few grams of cells the lowest being 3.21 g of organic/inorganic ANME-2c enrichment particles (Ch. V) and 1 g pellet B from the activated sludge utilized for a successful pre-trial (Ch. VI, section "Data processing, model building, refinement and validation"). Obtaining sufficient biomass from enrichments or environmental samples will be challenging and poses as a bottleneck to the technique. Nonetheless, the presented results demonstrate that even limited sample quantities can be effectively utilized, underscoring the technique's potential to study uncultivated microbes.

Studying new biocatalysts from extreme environments

Microorganisms from extreme environments are excellent candidates for the study of new enzymes and functions, as adaptations to extreme temperatures and pH are desirable for biotechnological adaptations[90]. Extremophilic enzymes are optimal biocatalysts as their adaptations can withstand the harsh industrial conditions often required by various production processes[91]. Furthermore, they increase efficiency, produce fewer toxic by-products and are more environmentally friendly and sustainable compared to purely chemical processes. The demand for new biocatalysts is increasing and the global enzyme market size was estimated at 60.48 billion USD in 2023[92]. Prominent examples are *e.g.*, the thermophilic *Taq* and *Pfu* DNA polymerases of *Thermus aquaticus* isolated from Yellowstone hydrothermal vents and *Pyrococcus furiosus* isolated from geothermally heated marine sediments[93-95]. These enzymes constituted the basis for enormous advances in molecular biology and genetic engineering. On the other hand, psychrophilic amylases, proteases, lipases, pectinases, mannanases, and cellulases are used in laundry detergent. Lowering the required temperature severely reduces energy costs and carbon-footprints[96].

Bioprospecting (discovery and commercialization of novel products from biological sources) of extremozymes is usually carried out with culture-dependent approaches[91].

Function based screening is typically based on the generation of gene libraries followed by recombinant expression. Enzymatic activity is detected on agar plates or in crude cell lysates by the addition of chromophore/fluorophore-labeled substrates[97]. However, the MDM has been recognized to hold major potential for the discovery of new biocatalysts. While genomic tools have made the MDM more accessible than ever, they have only contributed to the discovery of relatively few new extremozymes.

Sequence-based screening is limited by the unreliability of annotations for novel and understudied taxa as alterations of the amino-acid composition, resulting from the adaptation to extreme conditions, complicate homology-based searches[91]. In a recent study, environmental activity-based protein profiling was successfully applied to identify active thermostable serine hydrolases from two hot spring communities. Through specific probes, active enzymes could be labeled and identified through MS[98].

Since the crystallomics approach aims to study enzymes of uncultivated microbes, application to communities from extreme environments could aid in the discovery of novel biocatalysts. Successful protein extraction from *M. infernus* illustrates the feasibility of applying this approach to hyperthermophilic archaea. The purification of thermophilic enzymes is facilitated by their intrinsic robustness, retaining thermostability even at low temperatures[99], which is ideal for the crystallomics approach. Alternatively, adjustments to the purification pipeline could be made to generate suitable conditions for other types of extremozymes. The study of psychrophilic enzymes could be performed in a temperature-controlled environment or with a chromatography system that includes a cooling unit. Acido- or alkaliphilic enzymes could be purified in buffers with an appropriate pH as modern chromatography columns can accommodate a wider pH range. If successful, this approach could contribute to the direct structural investigation of novel biocatalyst targets.

Conclusion and outlook

The approach of crystallomics aims to study enzymes from uncultivated microbes through native purification ultimately employing crystallization as the final protein separation step. While higher sample diversity and limited biomass quantities are hurdles for the protein purification and crystallization, this work demonstrated the potential of the technique by successful native protein purification from various samples with differing degrees of heterogeneity. Based on the presented studies some more imminent targets could be the MCR of medium-/long-chain alkane oxidizing archaea. Exploiting already established enrichments[71, 75, 77] could give insights into the adaptations of MCR towards larger substrates. Detailed biological knowledge might allow the reversal of the reaction for the production of short-chain alkanes as biofuel[27]. This could be followed up by investigating the divergent MCR variants identified outside the *Euryarchaeota*. Examining enrichments[100] or metagenomically identified hotspots[101] could lead to entirely novel insights about this globally relevant enzyme. On the other hand, further investigation of the proteome of *Methanothrix,* including the specialized acetate activation step and membrane complexes involved in energy conservation, might contribute to better understanding of their metabolism at the molecular level. Furthermore, the study of enzymes involved in DIET could provide foundational knowledge about how electrons derived from syntrophic interactions are accepted and funneled into the cellular metabolism. These insights might deepen our comprehension of microbial interactions in natural environments and might facilitate more accurate modeling of *Methanothrix* contributions to methane emissions in climate-relevant ecosystems[102, 103].

The results presented here demonstrate the potential of extracting and characterizing enzymes directly from small amounts of heterogeneous environmental samples. Therefore, this thesis encourages scientists to apply more exploratory methods to study novel enzymes, especially from the MDM. Despite some noticed challenges, direct

protein purification from environmental samples can help to discover enzymes with unique structural and functional features not accessible by other techniques. These enzymes have great potential to expand our understanding of biochemical diversity, inspire innovations in biotechnology, and offer insights into metabolic adaptations to extreme environments. Combination of the crystallomics approach with state of the art crystal-based techniques like time-resolved crystallography has the potential to decipher novel enzymatic reactions at the atomic level. Bypassing the limitations of cultivation, this new biochemical knowledge obtained through the crystallomics approach could aid in better understanding of the unexplored microbial realm at the functional, metabolic, and atomic scale. Ultimately, the findings of this work support an adventurous approach to studying enzymes that embraces the unknown and uncharacterized, where fantastic discoveries are waiting.

References

1. Engilberge, S., Wagner, T., Santoni, G., Breyton, C., Shima, S., Franzetti, B., Riobé, F., Maury, O., Girard, E. Protein crystal structure determination with the crystallophore, a nucleating and phasing agent. *J. Appl. Crystallogr.* **52**, 722-731, doi:10.1107/S1600576719006381 (2019).
2. Schumacher, M. A., Salinas, R., Travis, B. A., Singh, R. R., Lent, N. *M. mazei* glutamine synthetase and glutamine synthetase-GlnK1 structures reveal enzyme regulation by oligomer modulation. *Nat. Commun.* **14**, 7375, doi:10.1038/s41467-023-43243-w (2023).
3. Herdering, E., Reif-Trauttmansdorff, T., Kumar, A., Habenicht, T., Hochberg, G., Bohn, S., Schuller, J., Schmitz, R. A. 2-oxoglutarate triggers assembly of active dodecameric *Methanosarcina mazei* glutamine synthetase. *eLife*, doi:10.7554/elife.97484.1 (2024).
4. Pedro-Roig, L., Camacho, M., Bonete, M. J. Regulation of ammonium assimilation in *Haloferax mediterranei*: Interaction between glutamine synthetase and two GlnK proteins. *Biochim. Biophys. Acta* **1834**, 16-23, doi:10.1016/j.bbapap.2012.10.006 (2013).
5. Lemaire, O. N., Müller, M. C., Kahnt, J., Wagner, T. Structural rearrangements of a dodecameric ketol-acid reductoisomerase isolated from a marine thermophilic methanogen. *Biomolecules* **11**, doi:10.3390/biom11111679 (2021).
6. Ehlers, C., Weidenbach, K., Veit, K., Forchhammer, K., Schmitz, R. A. Unique mechanistic features of post-translational regulation of glutamine synthetase activity in *Methanosarcina mazei* strain Gö1 in response to nitrogen availability. *Mol. Microbiol.* **55**, 1841-54, doi:10.1111/j.1365-2958.2005.04511.x (2005).
7. Müller, M.-C., Wagner, T. The oxoglutarate binding site and regulatory mechanism are conserved in ammonium transporter inhibitors GlnKs from *Methanococcales*. *Int. J. Mol. Sci.* **22**, doi:10.3390/ijms22168631 (2021).
8. Maslać, N., Müller, M.-C., Bolte, P., Wagner, T. *Molecular basis of nitrogenase regulation by a P_{II} family protein in a hyperthermophilic methanogen. In: Structural and biochemical characterization of N_2-assimilation system and its regulation in thermophilic methanogenic archaea.* Dissertation by Maslać, N., University Bremen, doi:10.26092/elib/2553, (2023).
9. Maslać, N. *Structural and biochemical characterisation of N_2-assimilation system and its regulation in thermophilic methanogenic archaea* Dissertation. University Bremen, doi:10.26092/elib/2553, (2023).
10. Bazylinski, D. A., Williams, T. J., Lefèvre, C. T., Berg, R. J., Zhang, C. L., Bowser, S. S., Dean, A. J., Beveridge, T. J. *Magnetococcus marinus* gen. nov., sp. nov., a marine, magnetotactic bacterium that represents a novel lineage (*Magnetococcaceae* fam. nov., *Magnetococcales* ord. nov.) at the base of the *Alphaproteobacteria*. *Int. J. Syst. Evol. Microbiol.* **63**, 801-808, doi:10.1099/ijs.0.038927-0 (2013).
11. Jeanthon, C., L'Haridon, S., Reysenbach, A. L., Vernet, M., Messner, P., Sleytr, U. B., Prieur, D. *Methanococcus infernus* sp. nov., a novel hyperthermophilic lithotrophic methanogen isolated from a deep-sea hydrothermal vent. *Int. J. Syst. Bacteriol.* **48 Pt 3**, 913-9, doi:10.1099/00207713-48-3-913 (1998).
12. Nayak, D. D., Liu, A. D., Agrawal, N., Rodriguez-Carerro, R., Dong, S. H., Mitchell, D. A., Nair, S. K., Metcalf, W. W. Functional interactions between posttranslationally modified amino acids of methyl-coenzyme M reductase in *Methanosarcina acetivorans*. *PLoS Biol.* **18**, doi:10.1371/journal.pbio.3000507 (2020).
13. Nayak, D. D., Mahanta, N., Mitchell, D. A., Metcalf, W. W. Post-translational thioamidation of methyl-coenzyme M reductase, a key enzyme in methanogenic and methanotrophic Archaea. *eLife* **6**, doi:10.7554/eLife.29218 (2017).
14. Lyu, Z., Shao, N. N., Chou, C. W., Shi, H., Patel, R., Duin, E. C., Whitman, W. B. Posttranslational methylation of arginine in methyl coenzyme M reductase has a profound impact on both methanogenesis and growth of *Methanococcus maripaludis*. *J. Bacteriol.* **202**, doi:10.1128/JB.00654-19 (2020).
15. Gagsteiger, J., Jahn, S., Heidinger, L., Gericke, L., Andexer, J. N., Friedrich, T., Loenarz, C., Layer, G. A cobalamin-dependent radical SAM enzyme catalyzes the unique C-methylation of glutamine in methyl-coenzyme M reductase. *Angew. Chem. Int. Edit.* **61**, doi:10.1002/anie.202204198 (2022).

16. Wagner, T., Kahnt, J., Ermler, U., Shima, S. Didehydroaspartate modification in methyl-coenzyme M reductase catalyzing methane formation. *Angew. Chem. Int. Ed Engl.* **55**, 10630-3, doi:10.1002/anie.201603882 (2016).
17. *DSMZ. German collection of microorganisms and cell cultures. Front page.*, <https://www.dsmz.de> (accessed 22.11.2024).
18. Kahnt, J., Buchenau, B., Mahlert, F., Krüger, M., Shima, S., Thauer, R. K. Post-translational modifications in the active site region of methyl-coenzyme M reductase from methanogenic and methanotrophic archaea. *FEBS J.* **274**, 4913-21, doi:10.1111/j.1742-4658.2007.06016.x (2007).
19. Wasserfallen, A., Nölling, J., Pfister, P., Reeve, J., Conway de Macario, E. Phylogenetic analysis of 18 thermophilic *Methanobacterium* isolates supports the proposals to create a new genus, *Methanothermobacter* gen. nov., and to reclassify several isolates in three species, *Methanothermobacter thermautotrophicus* comb. nov., *Methanothermobacter wolfeii* comb. nov., and *Methanothermobacter marburgensis* sp. nov. *Int. J. Syst. Evol. Microbiol.* **50 Pt 1**, 43-53, doi:10.1099/00207713-50-1-43 (2000).
20. Takai, K., Nealson, K. H., Horikoshi, K. *Methanotorris formicicus* sp. nov., a novel extremely thermophilic, methane-producing archaeon isolated from a black smoker chimney in the Central Indian Ridge. *Int. J. Syst. Evol. Microbiol.* **54**, 1095-1100, doi:10.1099/ijs.0.02887-0 (2004).
21. Huber, H., Thomm, M., König, H., Thies, G., Stetter, K. O. *Methanococcus thermolithotrophicus*, a novel thermophilic lithotrophic methanogen. *Archives of Microbiology* **132**, 47-50, doi:10.1007/Bf00690816 (1982).
22. Kurr, M., Huber, R., König, H., Jannasch, H. W., Fricke, H., Trincone, A., Kristjansson, J. K., Stetter, K. O. *Methanopyrus kandleri*, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110°C. *Arch. Microbiol.* **156**, 239-247, doi:10.1007/Bf00262992 (1991).
23. Hippe, H., Caspari, D., Fiebig, K., Gottschalk, G. Utilization of trimethylamine and other *N*-methyl compounds for growth and methane formation by *Methanosarcina barkeri*. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 494-498, doi:10.1073/pnas.76.1.494 (1979).
24. Kulkarni, G., Mand, T. D., Metcalf, W. W. Energy conservation via hydrogen cycling in the methanogenic archaeon *Methanosarcina barkeri*. *mBio* **9**, doi:10.1128/mBio.01256-18 (2018).
25. Mayumi, D., Mochimaru, H., Tamaki, H., Yamamoto, K., Yoshioka, H., Suzuki, Y., Kamagata, Y., Sakata, S. Methane production from coal by a single methanogen. *Science* **354**, 222-225, doi:10.1126/science.aaf8821 (2016).
26. Sowers, K. R., Baron, S. F., Ferry, J. G. *Methanosarcina acetivorans* sp. nov., an acetotrophic methane-producing bacterium isolated from marine sediments. *Appl. Environ. Microbiol.* **47**, 971-978, doi:10.1128/Aem.47.5.971-978.1984 (1984).
27. Hahn, C. J., Laso-Pérez, R., Vulcano, F., Vaziourakis, K. M., Stokke, R., Steen, I. H., Teske, A., Boetius, A., Liebeke, M., Amann, R., et al. "*Candidatus* Ethanoperedens," a thermophilic genus of *Archaea* mediating the anaerobic oxidation of ethane. *mBio* **11**, doi:10.1128/mBio.00600-20 (2020).
28. Jones, W. J., Leigh, J. A., Mayer, F., Woese, C. R., Wolfe, R. S. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. *Arch. Microbiol.* **136**, 254-261, doi:10.1007/Bf00425213 (1983).
29. Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R., Wolfe, R. S. Methanogens: reevaluation of a unique biological group. *Microbiol. Rev.* **43**, 260-96, doi:10.1128/mr.43.2.260-296.1979 (1979).
30. Rivard, C. J., Smith, P. H. Isolation and characterization of a thermophilic marine methanogenic bacterium, *Methanogenium thermophilicum* sp. nov. *Int. J. Syst. Bacteriol.* **32**, 430-436, doi:10.1099/00207713-32-4-430 (1982).
31. Cheng, L., Qiu, T. L., Yin, X. B., Wu, X. L., Hu, G. Q., Deng, Y., Zhang, H. *Methermicoccus shengliensis* gen. nov., sp. nov., a thermophilic, methylotrophic methanogen isolated from oil-production water, and proposal of *Methermicoccaceae* fam. nov. *Int. J. Syst. Evol. Microbiol.* **57**, 2964-2969, doi:10.1099/ijs.0.65049-0 (2007).

32. Wissink, M., Glodowska, M., van der Kolk, M. R., Jetten, M. S. M., Welte, C. U. Probing denitrifying anaerobic methane oxidation via antimicrobial intervention: implications for innovative wastewater management. *Environ. Sci. Technol.* **58**, 6250-6257, doi:10.1021/acs.est.3c07197 (2024).
33. Jetten, M. S. M., Stams, A. J. M., Zehnder, A. J. B. Methanogenesis from acetate: A comparison of the acetate metabolism in *Methanothrix soehngenii* and *Methanosarcina* spp. *FEMS Microbiol. Rev.* **8**, 181-197, doi:10.1111/j.1574-6968.1992.tb04987.x (1992).
34. Kurth, J. M., Nobu, M. K., Tamaki, H., de Jonge, N., Berger, S., Jetten, M. S. M., Yamamoto, K., Mayumi, D., Sakata, S., Bai, L., et al. Methanogenic archaea use a bacteria-like methyltransferase system to demethoxylate aromatic compounds. *ISME J.* **15**, 3549-3565, doi:10.1038/s41396-021-01025-6 (2021).
35. Scheller, S., Ermler, U., Shima, S. Catabolic pathways and enzymes involved in anaerobic methane oxidation. In *Anaerobic Utilization of Hydrocarbons, Oils, and Lipids* (ed Boll, M.) 31-59 (Springer International Publishing, 2020).
36. Raghoebarsing, A. A., Pol, A., van de Pas-Schoonen, K. T., Smolders, A. J. P., Ettwig, K. F., Rijpstra, W. I. C., Schouten, S., Damsté, J. S. S., Op den Camp, H. J. M., Jetten, M. S. M., et al. A microbial consortium couples anaerobic methane oxidation to denitrification. *Nature* **440**, 918-921, doi:10.1038/nature04617 (2006).
37. McGlynn, S. E. Energy metabolism during anaerobic methane oxidation in ANME archaea. *Microbes Environ.* **32**, 5-13, doi:10.1264/jsme2.ME16166 (2017).
38. Felden, J., Lichtschlag, A., Wenzhöfer, F., de Beer, D., Feseker, T., Ristova, P. P., de Lange, G., Boetius, A. Limitations of microbial hydrocarbon degradation at the Amon mud volcano (Nile deep-sea fan). *Biogeosciences* **10**, 3269-3283, doi:10.5194/bg-10-3269-2013 (2013).
39. Shao, N. N., Fan, Y., Chou, C. W., Yavari, S., Williams, R. V., Amster, I. J., Brown, S. M., Drake, I. J., Duin, E. C., Whitman, W. B., et al. Expression of divergent methyl/alkyl coenzyme M reductases from uncultured archaea. *Commun. Biol.* **5**, doi:10.1038/s42003-022-04057-6 (2022).
40. McPherson, A., Gavira, J. A. Introduction to protein crystallization. *Acta Cryst. F* **70**, 2-20, doi:10.1107/S2053230X13033141 (2014).
41. Wongnate, T., Sliwa, D., Ginovska, B., Smith, D., Wolf, M. W., Lehnert, N., Raugei, S., Ragsdale, S. W. The radical mechanism of biological methane synthesis by methyl-coenzyme M reductase. *Science* **352**, 953-8, doi:10.1126/science.aaf0616 (2016).
42. Shima, S., Krueger, M., Weinert, T., Demmer, U., Kahnt, J., Thauer, R. K., Ermler, U. Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature* **481**, 98-101, doi:10.1038/nature10663 (2012).
43. Krüger, M., Blumenberg, M., Kasten, S., Wieland, A., Känel, L., Klock, J. H., Michaelis, W., Seifert, R. A novel, multi-layered methanotrophic microbial mat system growing on the sediment of the Black Sea. *Environ. Microbiol.* **10**, 1934-47, doi:10.1111/j.1462-2920.2008.01607.x (2008).
44. Jelsch, C., Teeter, M. M., Lamzin, V., Pichon-Pesme, V., Blessing, R. H., Lecomte, C. Accurate protein crystallography at ultra-high resolution: Valence electron distribution in crambin. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 3171-3176, doi:10.1073/pnas.97.7.3171 (2000).
45. Banerjee, S., Coussens, N. P., Gallat, F. X., Sathyanarayanan, N., Srikanth, J., Yagi, K. J., Gray, J. S. S., Tobe, S. S., Stay, B., Chavas, L. M. G., et al. Structure of a heterogeneous, glycosylated, lipid-bound, *in vivo*-grown protein crystal at atomic resolution from the viviparous cockroach *Diploptera punctata*. *IUCrJ* **3**, 282-293, doi:10.1107/S2052252516008903 (2016).
46. Kim, M. S., Zhong, J., Pandey, A. Common errors in mass spectrometry-based analysis of post-translational modifications. *Proteomics* **16**, 700-714, doi:10.1002/pmic.201500355 (2016).
47. Albers, S. V., Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414-426, doi:10.1038/nrmicro2576 (2011).
48. Kubota, K., Imachi, H., Kawakami, S., Nakamura, K., Harada, H., Ohashi, A. Evaluation of enzymatic cell treatments for application of CARD-FISH to methanogens. *J. Microbiol. Methods* **72**, 54-9, doi:10.1016/j.mimet.2007.10.006 (2008).

49. Danaeifar, M. New horizons in developing cell lysis methods: A review. *Biotechnol. Bioeng.* **119**, 3007-3021, doi:10.1002/bit.28198 (2022).
50. Nobel Prize Outreach, A. B., Press release. 2024.
51. Verbeke, E. J., Mallam, A. L., Drew, K., Marcotte, E. M., Taylor, D. W. Classification of single particles from human cell extract reveals distinct structures. *Cell. Rep.* **24**, 259-268 e3, doi:10.1016/j.celrep.2018.06.022 (2018).
52. Rehdorf, J., Pelzer, A., Eck, J. Enzyme identification and screening: Activity-based methods. In *Introduction to Enzyme Technology* (eds Jaeger, K.-E.; Liese, A.; Syldatk, C.) 125-146 (Springer International Publishing, 2024).
53. Vandooren, J., Geurts, N., Martens, E., Van den Steen, P. E., Opdenakker, G. Zymography methods for visualizing hydrolytic enzymes. *Nat. Methods* **10**, 211-20, doi:10.1038/nmeth.2371 (2013).
54. Valenzuela, B., Solís-Cornejo, F., Araya, R., Zamorano, P. Isolation and characterization of *Thermus thermophilus* strain ET-1: An extremely thermophilic bacterium with extracellular thermostable proteolytic activity isolated from El Tatio geothermal field, Antofagasta, Chile. *Int. J. Mol. Sci.* **24**, doi:10.3390/ijms241914512 (2023).
55. Ortega-de la Rosa, N. D., Romero-Borbón, E., Rodríguez, J. A., Camacho-Ruiz, A., Córdova, J. Cloning, expression, chracterization and immobilization of a recombinant carboxylesterase from the halophilic archaeon, *Halobacterium salinarum* NCR-1. *Biomolecules* **14**, doi:10.3390/biom14050534 (2024).
56. Dab, A., Hasnaoui, I., Mechri, S., Allala, F., Bouacem, K., Noiriel, A., Bouanane-Darenfed, A., Saalaoui, E., Asehraou, A., Wang, F., et al. Biochemical characterization of an alkaline and detergent-stable Lipase from *Fusarium annulatum* Bugnicourt strain CBS associated with olive tree dieback. *PLoS One* **18**, e0286091, doi:10.1371/journal.pone.0286091 (2023).
57. Paul, M., Das Mohapatra, P. K., Thatoi, H. Purified cellulase-mediated simultaneous sugar utilization by *Bacillus albus* isolated from Similipal, Odisha, India. *J. Basic Microbiol.* **63**, 759-780, doi:10.1002/jobm.202300006 (2023).
58. Bharwad, K., Shekh, S., Singh, N. K., Patel, A., Joshi, C. Heterologous expression and biochemical characterization of novel multifunctional thermostable α-amylase from hot-spring metagenome. *Int. J. Biol. Macromol.* **242**, 124810, doi:10.1016/j.ijbiomac.2023.124810 (2023).
59. Ding, J., Xu, N., Wang, J., He, Y., Wang, X., Liu, M., Liu, X. Plancitoxin-1 mediates extracellular trap evasion by the parasitic helminth *Trichinella spiralis*. *BMC Biol.* **22**, 158, doi:10.1186/s12915-024-01958-2 (2024).
60. Lemaire, O. N., Wagner, T. Gas channel rerouting in a primordial enzyme: Structural insights of the carbon-monoxide dehydrogenase/acetyl-CoA synthase complex from the acetogen *Clostridium autoethanogenum*. *Biochim. Biophys. Acta. Bioenerg.* **1862**, 148330, doi:10.1016/j.bbabio.2020.148330 (2021).
61. Jespersen, M. *Structural and biochemical characterization of enzymes used by anaerobic energy extremophiles for sulfur and carbon transformation* Dissertation. University Bremen, doi:10.26092/elib/2515, (2023).
62. Thomas, P. E., Ryan, D., Levin, W. An improved staining procedure for the detection of the peroxidase activity of cytochrome *P*-450 on sodium dodecyl sulfate polyacrylamide gels. *Anal. Biochem.* **75**, 168-176, doi:10.1016/0003-2697(76)90067-1 (1976).
63. Nelson, D., Cox, M. *Lehninger biochemie*. 3 edn, (Springer, 2005).
64. He, H. H., He, B., Guan, L., Zhao, Y., Jiang, F., Chen, G. X., Zhu, Q. G., Chen, C. Y. C., Li, T., Yao, J. H. *De novo* generation of SARS-CoV-2 antibody CDRH3 with a pre-trained generative large language model. *Nat. Commun.* **15**, doi:10.1038/s41467-024-50903-y (2024).
65. Maslać, N., Sidhu, C., Teeling, H., Wagner, T. Comparative transcriptomics sheds light on remodeling of gene expression during diazotrophy in the thermophilic methanogen *Methanothermococcus thermolithotrophicus*. *mBio* **13**, e02443-22, doi:10.1128/mbio.02443-22 (2022).
66. Korzhenkov, A. A., Toshchakov, S. V., Bargiela, R., Gibbard, H., Ferrer, M., Teplyuk, A. V., Jones, D. L., Kublanov, I. V., Golyshin, P. N., Golyshina, O. V. Archaea dominate the microbial community

in an ecosystem with low-to-moderate temperature and extreme acidity. *Microbiome* **7**, 11, doi:10.1186/s40168-019-0623-8 (2019).
67. Kozubal, M. A., Romine, M., Jennings, R., Jay, Z. J., Tringe, S. G., Rusch, D. B., Beam, J. P., McCue, L. A., Inskeep, W. P. Geoarchaeota: A new candidate phylum in the Archaea from high-temperature acidic iron mats in Yellowstone National Park. *ISME J.* **7**, 622-34, doi:10.1038/ismej.2012.132 (2013).
68. Jay, Z. J., Beam, J. P., Dlakić, M., Rusch, D. B., Kozubal, M. A., Inskeep, W. P. Marsarchaeota are an aerobic archaeal lineage abundant in geothermal iron oxide microbial mats. *Nat. Microbiol.* **3**, 732-740, doi:10.1038/s41564-018-0163-1 (2018).
69. Boetius, A., Ravenschlag, K., Schubert, C. J., Rickert, D., Widdel, F., Gieseke, A., Amann, R., Jørgensen, B. B., Witte, U., Pfannkuche, O. A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* **407**, 623-6, doi:10.1038/35036572 (2000).
70. Gründger, F., Carrier, V., Svenning, M. M., Panieri, G., Vonnahme, T. R., Klasek, S., Niemann, H. Methane-fuelled biofilms predominantly composed of methanotrophic ANME-1 in Arctic gas hydrate-related sediments. *Sci. Rep.* **9**, 9725, doi:10.1038/s41598-019-46209-5 (2019).
71. Zehnle, H., Laso-Pérez, R., Lipp, J., Riedel, D., Merino, D. B., Teske, A., Wegener, G. *Candidatus* Alkanophaga archaea from Guaymas Basin hydrothermal vent sediment oxidize petroleum alkanes. *Nat. Microbiol.* **8**, 1199-1212, doi:10.1038/s41564-023-01400-3 (2023).
72. Schrenk, M. O., Kelley, D. S., Delaney, J. R., Baross, J. A. Incidence and diversity of microorganisms within the walls of an active deep-sea sulfide chimney. *Appl. Environ. Microbiol.* **69**, 3580-92, doi:10.1128/AEM.69.6.3580-3592.2003 (2003).
73. Chen, S. C., Musat, N., Lechtenfeld, O. J., Paschke, H., Schmidt, M., Said, N., Popp, D., Calabrese, F., Stryhanyuk, H., Jaekel, U., et al. Anaerobic oxidation of ethane by archaea from a marine hydrocarbon seep. *Nature* **568**, 108-111, doi:10.1038/s41586-019-1063-0 (2019).
74. Holler, T., Widdel, F., Knittel, K., Amann, R., Kellermann, M. Y., Hinrichs, K. U., Teske, A., Boetius, A., Wegener, G. Thermophilic anaerobic oxidation of methane by marine microbial consortia. *ISME J.* **5**, 1946-56, doi:10.1038/ismej.2011.77 (2011).
75. Laso-Pérez, R., Wegener, G., Knittel, K., Widdel, F., Harding, K. J., Krukenberg, V., Meier, D. V., Richter, M., Tegetmeyer, H. E., Riedel, D., et al. Thermophilic archaea activate butane via alkyl-coenzyme M formation. *Nature* **539**, 396-401, doi:10.1038/nature20152 (2016).
76. Nauhaus, K., Treude, T., Boetius, A., Kruger, M. Environmental regulation of the anaerobic oxidation of methane: A comparison of ANME-I and ANME-II communities. *Environ. Microbiol.* **7**, 98-106, doi:10.1111/j.1462-2920.2004.00669.x (2005).
77. Zhou, Z., Zhang, C. J., Liu, P. F., Fu, L., Laso-Pérez, R., Yang, L., Bai, L. P., Li, J., Yang, M., Lin, J. Z., et al. Non-syntrophic methanogenic hydrocarbon degradation by an archaeal species. *Nature* **601**, 257-262, doi:10.1038/s41586-021-04235-2 (2022).
78. Luo, F., Devine, C. E., Edwards, E. A. Cultivating microbial dark matter in benzene-degrading methanogenic consortia. *Environ. Microbiol.* **18**, 2923-36, doi:10.1111/1462-2920.13121 (2016).
79. Hu, H., Natarajan, V. P., Wang, F. Towards enriching and isolation of uncultivated archaea from marine sediments using a refined combination of conventional microbial cultivation methods. *Mar. Life Sci. Technol.* **3**, 231-242, doi:10.1007/s42995-021-00092-0 (2021).
80. Ferris, M. J., Ruff-Roberts, A. L., Kopczynski, E. D., Bateson, M. M., Ward, D. M. Enrichment culture and microscopy conceal diverse thermophilic *Synechococcus* populations in a single hot spring microbial mat habitat. *Appl. Environ. Microbiol.* **62**, 1045-50, doi:10.1128/aem.62.3.1045-1050.1996 (1996).
81. Lok, C. Mining the microbial dark matter. *Nature* **522**, 270-3, doi:10.1038/522270a (2015).
82. Baker, B. J., Tyson, G. W., Webb, R. I., Flanagan, J., Hugenholtz, P., Allen, E. E., Banfield, J. F. Lineages of acidophilic archaea revealed by community genomic analysis. *Science* **314**, 1933-5, doi:10.1126/science.1132690 (2006).
83. Henneberger, R., Moissl, C., Amann, T., Rudolph, C., Huber, R. New insights into the lifestyle of the cold-loving SM1 euryarchaeon: natural growth as a monospecies biofilm in the subsurface. *Appl. Environ. Microbiol.* **72**, 192-199, doi:10.1128/AEM.72.1.192-199.2006 (2006).

84. Probst, A. J., Weinmaier, T., Raymann, K., Perras, A., Emerson, J. B., Rattei, T., Wanner, G., Klingl, A., Berg, I. A., Yoshinaga, M., et al. Biology of a widespread uncultivated archaeon that contributes to carbon fixation in the subsurface. *Nat. Commun.* **5**, 5497, doi:10.1038/ncomms6497 (2014).
85. Rudolph, C., Wanner, G., Huber, R. Natural communities of novel archaea and bacteria growing in cold sulfurous springs with a string-of-pearls-like morphology. *Appl. Environ. Microbiol.* **67**, 2336-44, doi:10.1128/AEM.67.5.2336-2344.2001 (2001).
86. Kaeberlein, T., Lewis, K., Epstein, S. S. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **296**, 1127-9, doi:10.1126/science.1070633 (2002).
87. Lambrecht, J., Cichocki, N., Hübschmann, T., Koch, C., Harms, H., Müller, S. Flow cytometric quantification, sorting and sequencing of methanogenic archaea based on F_{420} autofluorescence. *Microb. Cell. Fact.* **16**, 180, doi:10.1186/s12934-017-0793-7 (2017).
88. Rinke, C., Lee, J., Nath, N., Goudeau, D., Thompson, B., Poulton, N., Dmitrieff, E., Malmstrom, R., Stepanauskas, R., Woyke, T. Obtaining genomes from uncultivated environmental microorganisms using FACS-based single-cell genomics. *Nat. Protoc.* **9**, 1038-48, doi:10.1038/nprot.2014.067 (2014).
89. *Thermo Fisher. Bigfoot spectral cell sorter specifications*, <https://www.thermofisher.com/de/de/home/life-science/cell-analysis/flow-cytometry/flow-cytometers/bigfoot-spectral-cell-sorter/specifications.html> (accessed 17.11.2024).
90. Prieto-Barajas, C. M., Valencia-Cantero, E., Santoyo, G. Microbial mat ecosystems: Structure types, functional diversity, and biotechnological application. *Electron. J. Biotechnol.* **31**, 48-56, doi:10.1016/j.ejbt.2017.11.001 (2018).
91. Mesbah, N. M. Industrial biotechnology based on enzymes from extreme environments. *Front. Bioeng. Biotechnol.* **10**, 870083, doi:10.3389/fbioe.2022.870083 (2022).
92. Enzymes Market Size. Share & Trends Analysis Report By Product (Carbohydrase, Proteases, Polymerases & Nucleases), By Type (Industrial, Specialty), By Source (Plants, Animals), By Region, And Segment Forecasts, 2023–2030. (2023).
93. Chien, A., Edgar, D. B., Trela, J. M. Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J. Bacteriol.* **127**, 1550-7, doi:10.1128/jb.127.3.1550-1557.1976 (1976).
94. Lundberg, K. S., Shoemaker, D. D., Adams, M. W., Short, J. M., Sorge, J. A., Mathur, E. J. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Gene* **108**, 1-6, doi:10.1016/0378-1119(91)90480-y (1991).
95. Fiala, G., Stetter, K. O. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. *Arch. Microbiol.* **145**, 56-61, doi:10.1007/BF00413027 (1986).
96. Kumar, A., Mukhia, S., Kumar, R. Industrial applications of cold-adapted enzymes: Challenges, innovations and future perspective. *3 Biotech* **11**, 426, doi:10.1007/s13205-021-02929-y (2021).
97. Goddard, J.-P., Reymond, J.-L. Enzyme assays for high-throughput screening. *Curr. Opin. Biotechnol.* **15**, 314-322, doi:10.1016/j.copbio.2004.06.008 (2004).
98. Ninck, S., Klaus, T., Kochetkova, T. V., Esser, S. P., Sewald, L., Kaschani, F., Bräsen, C., Probst, A. J., Kublanov, I. V., Siebers, B., et al. Environmental activity-based protein profiling for function-driven enzyme discovery from natural communities. *Environ. Microbiome* **19**, 36, doi:10.1186/s40793-024-00577-2 (2024).
99. Vieille, C., Zeikus, G. J. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1-43, doi:10.1128/MMBR.65.1.1-43.2001 (2001).
100. Buessecker, S., Chadwick, G. L., Quan, M. E., Hedlund, B. P., Dodsworth, J. A., Dekas, A. E. Mcr-dependent methanogenesis in *Archaeoglobaceae* enriched from a terrestrial hot spring. *ISME J.* **17**, 1649-1659, doi:10.1038/s41396-023-01472-3 (2023).
101. Evans, P. N., Parks, D. H., Chadwick, G. L., Robbins, S. J., Orphan, V. J., Golding, S. D., Tyson, G. W. Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science* **350**, 434-438, doi:10.1126/science.aac7745 (2015).
102. Angle, J. C., Morin, T. H., Solden, L. M., Narrowe, A. B., Smith, G. J., Borton, M. A., Rey-Sanchez, C., Daly, R. A., Mirfenderesgi, G., Hoyt, D. W., et al. Methanogenesis in oxygenated soils is a

substantial fraction of wetland methane emissions. *Nat. Commun.* **8**, doi:10.1038/s41467-017-01753-4 (2017).

103. Holmes, D. E., Shrestha, P. M., Walker, D. J. F., Dang, Y., Nevin, K. P., Woodard, T. L., Lovley, D. R. Metatranscriptomic evidence for direct interspecies electron transfer between *Geobacter* and *Methanothrix* species in methanogenic rice paddy soils. *Appl. Environ. Microbiol.* **83**, doi:10.1128/AEM.00223-17 (2017).

Acknowledgements

I would like to thank Dr. Tristan Wagner for welcoming me in the MicroMet group and teaching me about the wonderful world of enzymes, crystals and X-ray crystallography. I admire your contagious love for science. Thank you for always being a supportive and dedicated supervisor. I would also like to thank you for creating a safe and professional work environment. Finally, thank you for allowing me to work on this adventurous project.

I would like to thank the Micromet group, current and previous members, for the welcoming and friendly work environment. Thank you for always listening to my questions and giving valuable advice.

I would like to thank all committee members for taking the time to evaluate this work.

I would like to thank the MPG for funding my research.

Thank you to all collaborators that I had the joy working with. In particular I would like to thank Dr. Sylvain Engilberge for assistance during plenty of X-ray diffraction experiments and teaching me about S-SAD. Furthermore, I would like to thank Dr. Jens Harder for always having an open ear for questions and giving valuable advice.

Thank you to Stian Torset, Isabella Wilkie, Dr. Olivier Lemaire, Lauren Queiss, Disha Jawadekar, Pauline Bolte, Natalie-S. Müller and Florian Saade for proofreading.

I would like to thank the Bavarian State Library for help with getting hold of the original depictions of crystals by Hünefeld 1840.

I would like to thank my marmic cohort and honorary rats for being the best friends I ever had. I truly enjoyed every unhinged lunch conversation and every scientific discussion. Doing experiments together has always been a joy. I am glad that starting marmic led me to meet yall. Thank you for dealing with my silliness and always supporting me in hard times.

Vielen Dank an meine Familie dafür, dass ihr mein Studium von Beginn bis Ende unterstützt habt. Danke, dass ihr Interesse an meiner Arbeit habt und über Archaeen and Proteine lernen möchtet auch wenn sie wie Bakterien und Konfetti-girlanden aussehen. Danke, dass Ihr stolz auf mich seid, ich versuche mein Bestes. Ich hab euch lieb.

Und zuletzt möchte ich mich bei Florian bedanken. Du bist mein Leuchtfeuer. Danke, dass du zu jeder Zeit für mich da bist und auf mein mentales und physisches Wohl achtest. Ohne dich hätte ich es nicht geschafft. Danke für deine Geduld an langen Arbeitstagen und während des Schreibens an dieser Arbeit. Danke, dass du immer an mich geglaubt hast, ich bin so froh dich zu haben ♥.

Appendix

Universität Bremen | Fachbereich 02 | Postfach 33 04 40, 28334 Bremen

Universität Bremen
Fachbereich 2
Prüfungsamt Chemie
z. Hd. Frau Frauke Ernst
Leobener Straße

28359 Bremen
Deutschland

Universität Bremen

Prüfungsamt
Chemie

Frauke Ernst
Geschäftsstelle
Fachbereich 02
Leobener Str. / NW2
D-28359 Bremen

Verwaltungspavillon 06
Tel. 0421 218-62802
Fax 0421 218-9862802
frauke.ernst@uni-bremen.de
www.fb2.uni-bremen.de

Versicherung an Eides Statt

Name, Vorname	Müller, Marie-Caroline
Matrikel-Nr.	4243937
Straße	
Ort, PLZ	Bremen

Ich, Marie-Caroline Müller

versichere an Eides Statt durch meine Unterschrift, dass ich die vorstehende Arbeit selbständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich dem Sinne nach aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe.
Ich versichere an Eides Statt, dass ich die vorgenannten Angaben nach bestem Wissen und Gewissen gemacht habe und dass die Angaben der Wahrheit entsprechen und ich nichts verschwiegen habe.
Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung.

Ort, Datum, Unterschrift: _____

Universität Bremen

Universität Bremen | Fachbereich 02 | Postfach 33 04 40, 28334 Bremen

Erklärungen zur elektronischen Version und zur Überprüfung einer Dissertation

Hiermit betätige ich gemäß §7, Abs. 7, Punkt 4, dass die zu Prüfungszwecken beigelegte elektronische Version meiner Dissertation identisch ist mit der abgegebenen gedruckten Version.

Ich bin mit der Überprüfung meiner Dissertation gemäß §6 Abs. 2, Punkt 5 mit qualifizierter Software im Rahmen der Untersuchung von Plagiatsvorwürfen einverstanden.

Ort, Datum, Unterschrift: _____

Declaration on the contribution of the candidate for the presented chapters

CHAPTER II. Crystallomics on *Methanocaldococcus infernus*
The chapter represents preliminary results.
Conceptual design: 90 %
Data acquisition and experimental work: 80 %
Analysis and interpretation of results: 90 %
Preparation of figures and tables: 100 %
Writing of the chapter: 95 %

CHAPTER III. Differences in regulation mechanisms of glutamine synthetases from methanogenic archaea unveiled by structural investigations.
Published in Communications biology
Conceptual design: 50 %
Data acquisition and experimental work: 80 %
Analysis and interpretation of results: 70 %
Preparation of figures and tables: 80 %
Writing of the manuscript: 50 %
Preparation of the first draft by M-C..M.

CHAPTER IV. Structural insights into the methane-generating enzyme from a methoxydotrophic methanogen reveal a restrained gallery of post-translational modifications.
Published in Microorganisms
Conceptual design: 0 %
Data acquisition and experimental work: 30 %
Analysis and interpretation of results: 30 %
Preparation of figures and tables: 40 %
Writing of the manuscript: 30 %

CHAPTER V. Atomic resolution structures of methane-fixing enzymes from marine and freshwater ANME-2 depict a conserved organisation with numerous post-translational modifications.
Manuscript in preparation, not seen by all coauthors
Conceptual design: 30 %
Data acquisition and experimental work: 50 %
Analysis and interpretation of results: 90 %

Preparation of figures and tables: 70 %
Writing of the manuscript: 60 %
Preparation of the first draft by M-C. M.

CHAPTER VI. Native isolation of the CH_4-generating enzyme from a mesocosm – steps towards microbial environmental structural biology.
Manuscript in preparation.
Conceptual design: 30 %
Data acquisition and experimental work: 40 %
Analysis and interpretation of results: 40 %
Preparation of figures and tables: 40 %
Writing of the manuscript: 30 %
Preparation of the first draft by M-C.M.

CHAPTER VII. Crystallomics on activated sludge from wastewater treatment.
The chapter represents preliminary results.
Conceptual design: 90 %
Data acquisition and experimental work: 80 %
Analysis and interpretation of results: 90 %
Preparation of figures and tables: 100 %
Writing of the chapter: 90 %

While working on this project, I found myself amazed by the natural beauty of protein structures. This inspired me to attempt to capture these amazing complexes in some artistic ways that could be appreciated by non-biologists. The following pictures are a result of that effort and were generated with Blender 3.2.

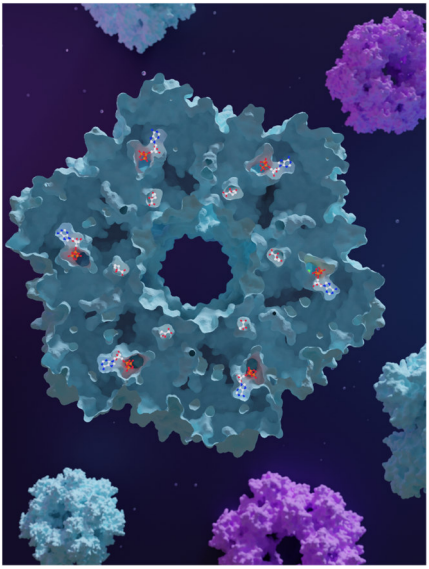
1. Surface of the glutamine synthetase including internal active- and allosteric site pockets filled with ligands. A modified version was submitted with the Ch. III manuscript and featured on the Communications' Biology homepage in January 2024.

2. Active site of ANME-2d MCR containing F_{430}, CoM, CoB and seven PTMs.

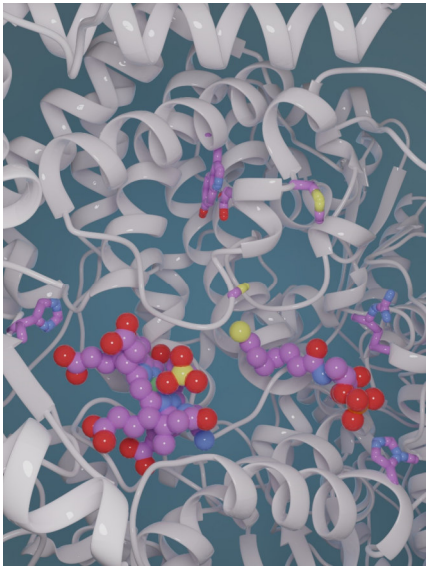
3. F_{430}, Coenzyme M and Coenzyme B.

4. Backbone of GlnK from *M. infernus.*

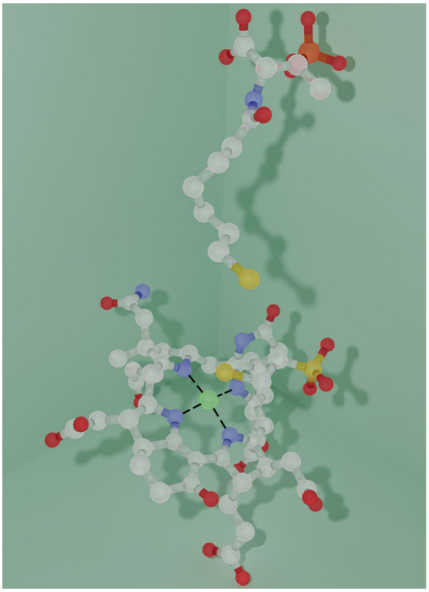
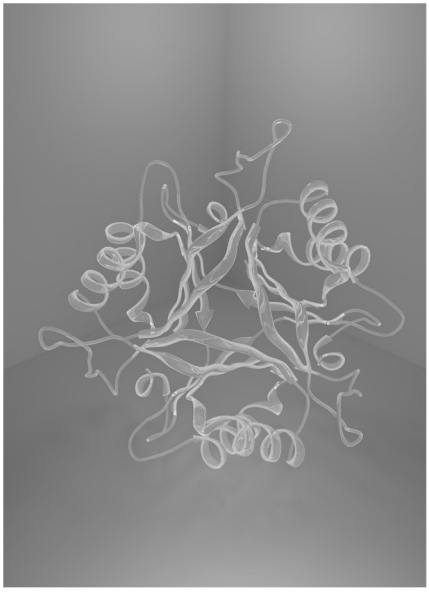
1.



2.



3.



4.