# Linking Metabolic Capacity and Molecular Biology of *Methylocystis* sp. Strain SC2 by a Newly Developed Proteomics Workflow

## Dissertation

*„Kumulativ“*

Zur Erlangung des Grades eines
Doktor der Naturwissenschaften
(Dr. rer.nat.)

des Fachbereichs Biologie

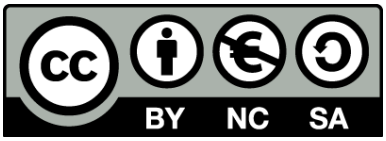
der Philipps-Universität Marburg

Vorgelegt von

**Anna Hakobyan**

Aus Gyumri, Armenien

Marburg | 2020



Die vorliegende Dissertation wurde von November 2015 bis Dezember 2019 am Max-Planck-Institut für terrestrische Mikrobiologie in Marburg unter Leitung von Herrn PD Dr. Werner Liesack angefertigt.

Vom Fachbereich Biologie

der Philipps-Universität Marburg

(Hochschulkennziffer 1180) als Dissertation

angenommen am **13. Dezember, 2019**

Erstgutachter(in): Herr PD Dr. Werner Liesack

Zweitgutachter(in): Herr Prof. Dr. Lennart Randau

Weitere Mitglieder der Prüfungskommission:

Herr Prof. Dr. Andreas Brune

Herr Prof. Dr. Uwe G. Maier

Tag der Disputation: **17. April, 2020**

**Die in dieser Dissertation beschriebenen Ergebnisse sind in folgenden Publikationen veröffentlicht bzw. zur Veröffentlichung vorgesehen:**

**Hakobyan A.[#],** Zhu J.[#], Glatter T., Liesack, W. (***under review***) Hydrogen utilization by *Methylocyctis* sp. strain SC2 expands the known metabolic versatility of type IIa methanotrophs. ([#]equal contribution)

**Hakobyan A.**, Schneider M.B., Liesack W., Glatter T. (**2019**) Efficient tandem LysC/trypsin digestion in detergent conditions. *Proteomics*, 19(20)**,** e1900136.

Bordel S., Rodríguez Y., **Hakobyan A.**, Rodríguez E., Lebrero R., Muñoz R. (**2019**) Genome scale metabolic modeling reveals the metabolic potential of three Type II methanotrophs of the genus *Methylocystis*. *Metabol Eng*, 54, 191

**Hakobyan A.**, Liesack W.*, Glatter T.* (**2018**) Crude-MS Strategy for in-depth proteome analysis of the methane-oxidizing *Methylocystis* sp. strain SC2. *J Proteome Res*, 17 (9), 3086 (*corresponding authors)

## Erklärung

Ich versichere, dass meine Dissertation mit dem Titel „Linking metabolic capacity and molecular biology of *Methylocystis* sp. strain SC2 by a newly developed proteomics workflow" selbstständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat keinen sonstigen Prüfungszwecken gedient.

Marburg, den 23. April, 2020

Anna Hakobyan

*Dedicated to my son, my husband, and my parents*

*"I was taught that the way of progress was neither swift nor easy."*

- Marie Curie

## Summary

Microbial methane oxidation is one of the fundamental processes in global methane cycle. Methane-oxidizing bacteria, or methanotrophs, are the major biological sink for the methane produced from anthropogenic and natural sources. Our model organism, *Methylocystis* sp. strain SC2, is one of the best-studied representatives of alphaproteobacterial (type IIa) methanotrophs. Proteobacterial methanotrophs possess a unique cell architecture characterized by intracytoplasmic membranes (ICMs). The cellular amount of the ICMs is increasing with methanotrophic activity. The presence of ICMs makes molecular biology approaches, but in particular global proteomics, highly challenging. In this study, we therefore aimed to develop an efficient proteomics workflow for strain SC2 and to apply this state-of-the-art tool for investigation of the strain SC2 response to environmental factors. To successfully develop the proteomics workflow, we particularly focused on an efficient solubilization and digestion of the integral membrane proteins of strain SC2 for further downstream analysis. We introduced the so-called crude-MS proteomics workflow, upon assessing and optimizing all the major steps in the proteomics workflow, including cell lysis, protein solubilization, and protein digestion. Our new SC2 proteomics workflow greatly increased not only the protein quantification accuracy (mean coefficient of variation 3.2 %) but also the proteome coverage to 62%, with up to 10-fold increase in the detection intensity of membrane-associated proteins.

Previous studies have shown that the LysC/trypsin tandem digestion resulted in higher coverage of fully cleaved tryptic peptides than a trypsin-only digestion. Therefore, the development of our optimized proteomics workflow involved the application of the LysC/trypsin tandem digestion in detergent environment to increase the SC2 proteome coverage. Prior to publication of our crude-MS approach, all systematic assessments of LysC/trypsin proteolysis were conducted in chaotropic environments, like urea. As a spin-off, we therefore initiated a follow-up study to compare the efficiency of the LysC/trypsin tandem digestion in detergent environments (e.g., SDC, SLS) relative to chaotropic environments. The study revealed that the LysC/trypsin tandem digestion could be efficiently carried out not only in chaotropic environments but also in MS-compatible detergent environments. In fact, the

LysC/trypsin tandem digestion in both environments resulted in a higher coverage of fully cleaved peptides than the trypsin-only digestion.

After successful development of the crude-MS proteomics workflow, we used this high-throughput method to assess the molecular response of strain SC2 to the availability of hydrogen as a potentially alternative energy source. Starting point of this research was the knowledge that strain SC2 and other *Methylocystis* spp. possess the genetic potential to produce various hydrogenases. In fact, the addition of 2% hydrogen to the headspace atmosphere led, under limiting concentrations of methane and oxygen, to the complete hydrogen consumption by strain SC2. Concurrently, the SC2 biomass yield was significantly increased, while the methane consumption rate was significantly decreased. Global proteome analyses revealed that the addition of hydrogen induced an increase in the production of Group 1d and Group 2b [NiFe]-hydrogenases, and hydrogenase accessory proteins. Notably, the upregulation of the Group 1d, 2b [NiFe]-hydrogenases was concomitantly linked to a reconstruction of the energy metabolism in strain SC2.

In another project, genome-scale metabolic modeling and growth experiments were applied to show that strain SC2 has the capacity to utilize acetate through the glyoxylate assimilation pathway. In addition, the study revealed that in type II methanotrophs, energy demand for methane oxidation is covered by complex I of the electron transport chain.

In summary, our research demonstrates how to experimentally link the metabolic potential of *Methylocystis* sp. strain SC2 with the underlying proteome complexity. Thus, the newly developed highly reproducible SC2 proteomics workflow represents a high-throughput method that makes it possible to achieve in future research an understanding of the molecular adaptation mechanisms of strain SC2 to environmental change.

# Zusammenfassung

Die mikrobielle Methanoxidation ist eine der grundlegenden Prozesse im globalen Methankreislauf. Methanoxidierende Bakterien oder Methanotrophe sind die wichtigste biologische Senke von Methan, welches sowohl natürlichen als auch anthropogenen Quellen entstammt. Unser Modellorganismus, *Methylocystis* sp. Stamm SC2, ist einer der am besten untersuchten Vertreter der den *Alphaproteobacteria* zugehörigen Typ IIa Methanotrophen. Methanotrophe Bakterien bilden in Abhängigkeit von ihrer Aktivität intracytoplasmatische Membranen aus. Diese für Mikroorganismen einzigartige Zellstruktur stellt in Hinblick auf Proteomanalysen dieser Bakterien eine erhebliche Herausforderung dar. In der vorliegenden Arbeit war es daher mein vorrangiges Ziel, die Methodik für eine effiziente SC2-Proteomanalyse zu entwickeln. Zellaufschluss und Solubilisierung der integralen Membranproteine waren daher ein zentrales Element meiner Untersuchungen, um diese nach anschließendem Verdau dem massenspektroskopischen Nachweis zugänglich zu machen. Meine neu erarbeitete Methodik erlaubte nicht nur die Signalintensität membranassoziierter Proteine um den Faktor 10 zu erhöhen, sondern führte auch im Vergleich zu vorhergehenden Verfahren zu einer signifikanten Verbesserung der Reproduzierbarkeit in der Proteinquantifizierung (mittlerer Variationskoeffizient 3,2%). Frühere Berichte wiesen darauf hin, dass der Tandem-LysC/Trypsin-Verdau zu einer höheren Abdeckung mit vollständig gespaltenen tryptischen Peptiden führt als ein Verdau nur durch Trypsin. Daher wurde im Rahmen der Entwicklung einer für Stamm SC2 effizienten Methodik solch ein Tandem-LysC/Trypsin-Verdau angewandt, um damit die Abdeckung des SC2-Proteoms zu erhöhen. Der Tandem-LysC/Trypsin-Verdau wird traditionell in einer chaotropen Umgebung (z.B. Harnstoff) durchgeführt. Ich habe daher in einer Folgestudie die Effizienz des Tandem-LysC/Trypsin-Verdaus in der Umgebung von Detergenzien (z.B. SDC, SLS) untersucht, da dies die methodische Durchführung erheblich vereinfachen würde. Es zeigte sich, dass der Tandem-LysC/Trypsin Verdau nicht nur in chaotroper Umgebung sondern auch in MS-kompatibler Detergenzien-Umgebung effizient durchführbar ist. Im Vergleich zu einem Verdau ausschließlich mit Trypsin kommt es auch mit dieser Variante zu einer höheren Abdeckung mit vollständig gespaltenen Peptiden.

Die neu entwickelte Methodik zur effizienten Analyse des SC2-Proteoms wurde dann in Untersuchungen zur Frage eingebunden, ob Wasserstoff von Stamm SC2 als alternative Energiequelle genutzt werden kann. Ausgangspunkt dieser Fragestellung war die Kenntnis, dass Stamm SC2, aber auch andere *Methylocystis* spp., das genetische Potential zur Produktion eines breiten Spektrums an Hydrogenasen hat. Die Zugabe von 2% Wasserstoff zur Atmosphäre führte unter limitierenden Bedingungen an Methan und Sauerstoff nicht nur zu einer vollständigen Verwertung des Wasserstoffs sondern auch zu einem signifikant höheren Biomasseertrag bei gleichzeitig verringerter Verbrauchsrate an Methan. Die Analyse des SC2-Proteoms zeigte, dass die Verwertung des Wasserstoffs mit einer signifikant erhöhten Produktion von [NiFe]-Hydrogenasen der Gruppen 1d und 2b sowie von akzessorischen Hydrogenase-Proteinen einherging. Die Verwertung von Wasserstoff als alternative Energiequelle durch *Methylocystis* spp. könnte für die Erzielung hoher Biomasseerträge in biotechnologischen Fragestellungen von Nutzen sein.

In einer weiteren Studie wurde unter Anwendung Genom-basierter metabolischer Modelle gezeigt, dass Stamm SC2 das Potenzial hat Acetat mittels des Glyoxylat-Stoffwechselwegs als Substrat zu nutzen. Ferner belegten die Forschungsergebnisse, dass die Elektronen notwendig für die Oxidation von Methan zu Methanol durch den Komplex I der Atmungskette gestellt werden.

In der Summe ermöglichten die hier beschriebenen Arbeiten Experimente zum metabolischen Potenzial und der Molekularbiologie von Stamm SC2 in sinnvoller Weise miteinander zu verknüpfen. Mit der effizienten Proteomanalyse von Stamm SC2 steht jetzt ein schnelles und reproduzierbares Verfahren zur Verfügung um in Folgestudien auf molekularbiologischer Ebene vertiefte Kenntnisse zu den Anpassungsmechanismen von Stamm SC2 auf Umweltveränderungen zu erzielen.

# **Chapter 1:** Introduction

# 1 Introduction

In order to provide conceptual background information for the research chapters presented in my doctoral thesis, I give in **Chapter 1** (Introduction) an overview of the processes involved in methane cycling followed by various aspects of methanotrophy. The different types of methanotrophs, their habitats, the biochemistry of their methane oxidation and carbon assimilation pathways, and the molecular techniques available for research on methanotrophs are discussed. In the end of the **Chapter 1,** I introduce our model organism *Methylocystis* sp. strain SC2 and present the objectives of my study.

## 1.1. Global methane budget

Methane ($CH_4$) is the most abundant organic compound in the Earth's atmosphere. It is considered as a greenhouse gas, along with water vapor ($H_20$), carbon dioxide ($C0_2$), nitrous oxide ($N_2O$), ozone ($0_3$), and the chlorofluorocarbon compounds (Wahlen, 1993). Methane is known to be the second most significant greenhouse gas in the atmosphere with a global warming potential ($GWP_{100}$) 28-36 times greater than $CO_2$ (Myhre et al., 2014; Lau et al., 2015; Hakobyan et al., 2018). With its current atmospheric concentration of 1.85 ppmv, $CH_4$ contributes approx. 15% to the total greenhouse effect (Cubasch et al., 2014; Saunois et al., 2016). Therefore, the methane cycle is one of the most important biogeochemical processes of the earth's system.

Atmospheric $CH_4$ can be of biogenic, thermogenic, pyrogenic, or mixed origin (**Figure 1**). Biogenic sources of $CH_4$ include methane-producing microorganisms (methanogens) that live in anaerobic habitats, such as wetlands, rice fields, lakes with a low concentration of dissolved oxygen, permafrost, the digestive system of ruminants and termites, and organic waste (composts, sewage, landfills). Thermogenic methane, formed over millions of years, refers to fossil fuels. It releases to the atmosphere through natural processes, in particular from the development of coal, oil and natural gas deposits. Pyrogenic $CH_4$ is the product of incomplete combustion of biomass, fossil fuels and biofuels, as well as soil organic matter during forest fires. Mixed methane

usually has a geological origin or is released from gas hydrates (Kirschke et al., 2013; Turner et al., 2019)

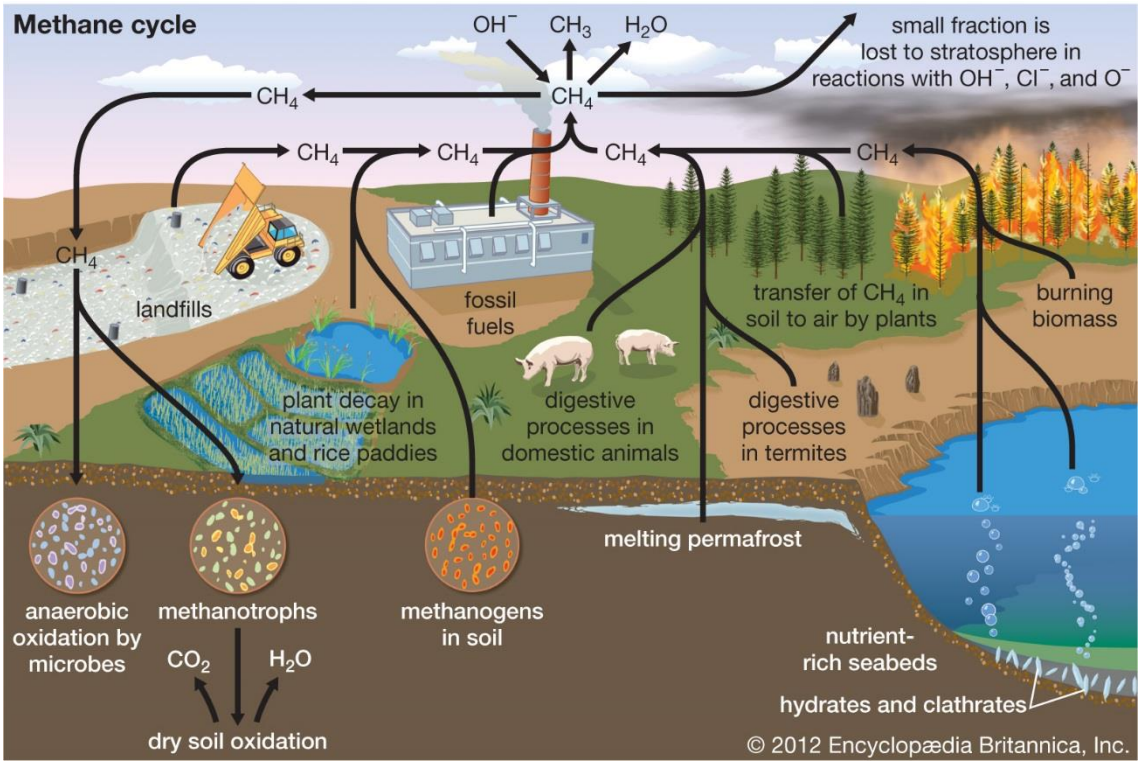



**Figure 1** | Atmospheric methane budget. The balance of methane sources and sinks create the global methane budget. Biogenic methane is produced by the activity of methanogens in wetlands, landfill sites, agricultural soils, and termites. In contrast, thermogenic methane is generated during long-lasting geological natural processes such as coal, oil, and natural gas production. Pyrogenic methane is formed through forest fires, incomplete combustion of fossil fuels, biomass and biofuels. Finally, deep sea hydrates are considered as a mixed source of methane. The largest methane sinks are the $OH^-$ radicals in troposphere and methane-oxidizing bacteria in oxic soils. *Adopted from Encyclopædia Britannica, Inc.*

Starting from the industrial era, over the past 15 years, the atmospheric concentration of methane has significantly increased (Cubasch et al., 2014). Following a relative stable period between 1990 and 2006, the methane concentration in the atmosphere has been rapidly increasing since 2007 (Dlugokencky et al., 2011), thereby leading to an imbalance in the global methane budget (Rigby et al., 2008; Kirschke et al., 2013; Saunois et al., 2016). Due to the acceleration of the methane emission to the atmosphere over the last years, the level of mitigation is currently most closely consistent with RCP (Representative concentration pathways) 8.5 W $m^{-2}$ (Fujino et al., 2006; Riahi et al., 2007; van Vuuren et al., 2007) (**Figure 2**). According to estimates of IPCC (The Intergovernmental Panel on Climate Change) under the RCP 8.5 scenario, the methane emissions will increase throughout the 21$^{st}$ century and will reach 888 Tg $CH_4$ per year

by the end of the century (Cubasch et al., 2014; Intergovernmental Panel Climate Change, 2014).

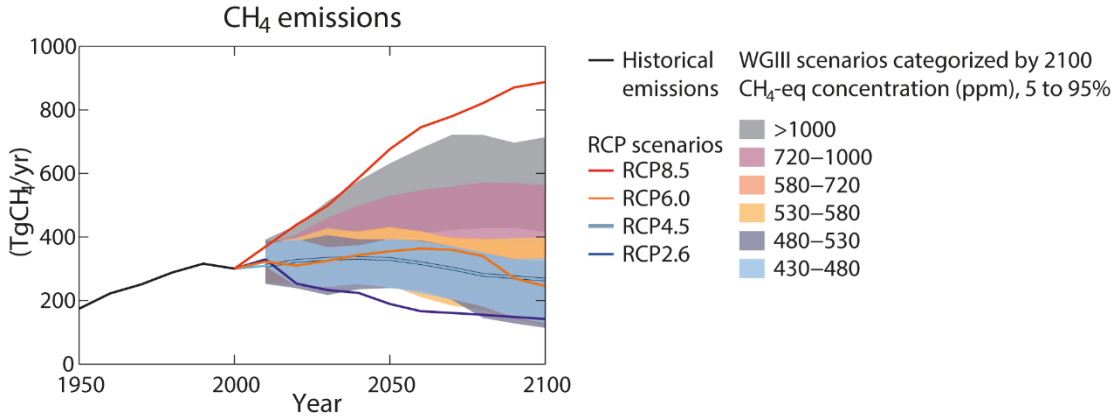



**Figure 2** | Projection of atmospheric methane concentrations and annual methane emissions for the four Representative Concentration Pathway (RCP) scenarios (based on (Meinshausen et al., 2011)). *Adopted from IPCC "Climate Change 2014 Synthesis Report".*

Terrestrial sources (natural and anthropogenic) of methane counts for two-thirds of the global methane emissions; in particular, being dominated by emissions from wetlands that are calculated to be 127-202 Tg $CH_4$ g per year (Saunois et al., 2016). More than half of global terrestrial methane emissions are contributed by anthropogenic activities, including agriculture and waste, biomass burning, and fossil fuel production (**Figure 1**).

The balance between methane sources and sinks defines the global methane budget (Saunois et al., 2016). The geochemical reactions that lead to the removal of methane from the atmosphere are much better quantified than the biotic and abiotic processes that drive the methane emissions into the atmosphere. The most important sink in the methane cycle is the methane reaction with the hydroxyl ($OH^-$) radicals, which are produced photochemically in the atmosphere. Here, $CH_4$ is oxidized to CO and ultimately to $CO_2$ and $H_20$ (Cicerone and Oremland, 1988; Wahlen, 1993). Methane is also oxidized by methanotrophic microorganisms, and its oxidation occurs both under aerobic and anaerobic conditions (**Figure 1**). Methanotrophs act as a natural biofilter, the activity of which significantly reduces the emission of methane to the atmosphere in various environmental settings. They are widespread both in aerobic interfaces of methanogenic environments and in upland soils. Inhabiting oxic/anoxic soil interfaces, methanotrophs reoxidize 20% to 90% of the methane produced in natural wetlands or

submerged rice paddies by methanogenic archaea before its emission to the atmosphere. In upland soils, methanotrophs act as a sink for 28-38 Tg per year of atmospheric methane (Conrad, 1996; Knief and Dunfield, 2005; Allen, 2016; Saunois et al., 2016).

## 1.2. Methane-oxidizing microorganisms or methanotrophs

Methanotrophs are a subgroup of methylotrophs and are usually utilizing methane as their sole source of carbon and energy. Methanotrophic bacteria are found in three phyla; *Proteobacteria*, *Verrucomicrobia* and candidate phylum NC10. *Proteobacteria* are the most widespread and most well-studied methanotrophs. Verrucomicrobial methanotrophs develop only in extremely acidic (pH 1.8–5.0) geothermal (up to 82 ° C) habitats (Dunfield et al., 2007; Op den Camp et al., 2009; Sharp et al., 2014). The methanotrophic bacteria of the newly described candidate phylum NC10 oxidize methane aerobically, but combine this oxidation step with denitrification under anoxic conditions (Ettwig et al., 2010).

In addition, it has been shown that ammonia oxidizers are capable of converting methane to methanol by the activity of ammonia monooxygenase. This enzyme is a homolog of pMMO. However, these bacteria cannot grow on methane because methanol cannot be further oxidized to gain energy (Hyman and Wood, 1983; Jones and Morita, 1983).

The term “methanotrophic bacteria” traditionally defines a group of aerobic bacteria that are structurally and functionally specialized in the use of methane as their sole source of carbon and energy (Hanson and Hanson, 1996). Nowadays, the term “aerobic methanotrophic bacteria” or “aerobic methanotrophs” is preferred in order to make a distinction between microorganisms that aerobically (bacteria) or anaerobically (archaea) oxidize methane.

### *1.2.1. Aerobic methanotrophs*

The first isolate of methanotrophic bacteria, “*Bacillus methanicus*“(later renamed *Methylomonas methanica*), was isolated by Söngen at the beginning of the 20$^{th}$ century (Söngen, 1907; Dedysh and Knief, 2018). Over the next several decades, the known diversity of methanotrophs remained limited and consisted of only a few cultures. In

1970, after extensive work on the isolation and description of more than 100 isolates, the first classification of methanotrophic bacteria was proposed (Whittenbury et al., 1970). All methanotrophs known at that time belonged to the later-described *Proteobacteria* phylum and were divided into two subgroups, termed type I and type II. The phylum *Proteobacteria* was established by Carl Woese in 1987 (Woese, 1987).

Several characteristic features were used for the differentiation of the proteobacterial methanotrophs into type I and type II methanotrophs. As a rule, type I methanotrophs use the ribulose monophosphate pathway (RuMP) for the assimilation of formaldehyde and have intracytoplasmic membranes (ICMs) organized in the form of stacks of flattened vesicles that fill most of the cell and are oriented perpendicular to the cell membrane (**Figure 3**). In contrast, type II methanotrophs use the serine pathway for formaldehyde assimilation, and their ICMs are stacks of flattened vesicles oriented parallel to the outer membrane and located around the cell perimeter (**Figure 3**). In addition, type I and II methanotrophs were proposed to differ in the predominant fatty acids, with 16 (type I) and 18 (type II) carbon atoms (Hanson and Hanson, 1996; Trotsenko and Murrell, 2008). Later, the comparative analysis of 16S rRNA genes confirmed the phylogenetic distinction of type I and type II methanotrophs. Type I methanotrophs belong to the class *Gammaproteobacteria*, while type II methanotrophs are members of the class *Alphaproteobacteria*.

**Figure 3** | Electron micrograph of a cross-section of the Type I methanotroph *Methylomonas methanica* (left) showing characteristic bundles of intracytoplasmic membranes (ICMs) inside the cell and a typical Type II methanotroph *Methylosinus trichosporium* (right) showing characteristic ICMs arranged around the periphery of the cell.

Until recently, the entire diversity of methanotrophic *Gammaproteobacteria* was combined in the *Methylococcaceae* family. Initially, five genera of methanotrophic bacteria have been described within the *Alphaproteobacteria*: *Methylosinus,*

*Methylocystis, Methylocella, Methyloferula* and *Methylocapsa* (Bowman et al., 1993; Dedysh et al., 2000; Dedysh et al., 2002; Vorobev et al., 2011). More recently, the description of new proteobacterial methanotrophs on genus and species levels has called the traditional cellular and biochemical criteria for the differentiation between type I and type II methanotrophs into question. While the main pathway for carbon assimilation is still a valid feature, a number of other features are no longer exclusive to either type I or type II methanotrophs (Knief, 2015). Due to these contradictory features, the classification into type I and type II methanotrophs does not anymore accurately reflect the existing diversity of methanotrophic bacteria. Therefore, the terms "type I" and "type II" are nowadays considered only as synonyms for methanotrophs of the classes *Gamma-* and *Alphaproteobacteria*. Consequently, gammaproteobacterial methanotrophs have been divided into three new groups. Members of the *Methylococcaceae* family were categorized as type Ia and type Ib, while those of the *Methylothermaceae* family were assigned as type Ic (**Figure 4**) (Knief, 2015). The seperation of *Methylococcaceae* family into type Ia and Ib originated from the previous assignment of *Methylococcus capsulatus* as type X (Hanson and Hanson, 1996; Bowman, 2006). In addition to the RuMP pathway, *M. capsulatus* is able to express ribulose bisphosphate carboxylase, a key enzyme in the Calvin – Benson cycle. Subsequently, *Methylococcus*-like methanotrophs were renamed from type X to type Ib (Knief, 2015). Among all type I methanotrophs, *Methylococcaceae* is the biggest family that includes more than 17 genera of well-described methanotrphs (Knief, 2015). Members of the *Methylothermaceae* inhabit various extreme habitats (Hirayama et al., 2014). The family *Crenotrichaceae* comprises unusual filamentous methanotrophic bacteria such as "*Candidatus* Crenothrix polyspora" (Stoecker et al., 2006) and "*Candidatus* Clonothrix fusca" (Vigliotta et al., 2007).

The alphaproteobacterial methanotrophs have also been divided into two subgroups: type IIa (*Methylocystaceae*) and type IIb (*Beijerinckiaceae*) (Deng et al., 2013; Dumont et al., 2014). The moderately acidophilic methanotrophs of the family *Beijerinkiaceae* differ from type IIa methanotrophs in a number of cellular characteristics. For example, cells of *Methylocella* and *Methyloferula* do not possess ICMs, while in *Methylocapsa*, the ICMs are always located on one side of the cell. Representatives of *Methylocella* and *Methylocapsa* also lack the previously considered

indicator $C_{18}$ fatty acid for type II methanotrophs. In addition, representatives of the genera *Methylocella* and *Methyloferula* lack pMMO, which is present in all other aerobic methanotrophs (Dedysh et al., 2000; Vorobev et al., 2011).

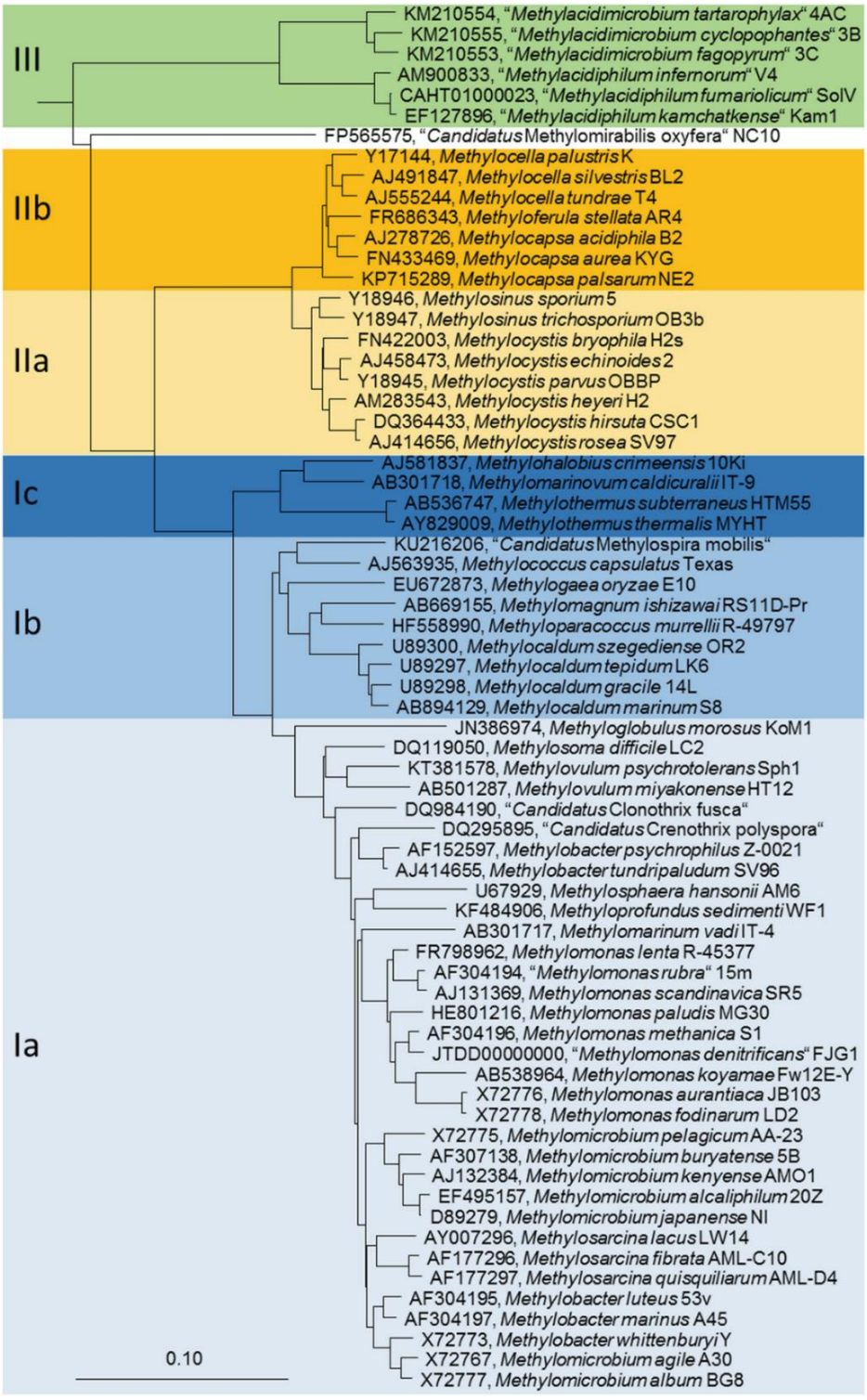



**Figure 4** | Phylogeny of described aerobic methanotrophic bacteria based on 16S rRNA gene sequences. A neighbor-joining tree was calculated with Jukes Cantor correction based on 1556 nucleotide positions using the ARB software package. The tree was rooted with sequences of methanogenic *Archaea*

(AB301476, M60880, AB065296, AM114193, AB196288). The scale bar displays 0.10 changes per nucleotide position. *Adopted from (Dedysh and Knief, 2018).*

The variety of known methanotrophs has expanded significantly thanks to the description of methanotrophs affiliated with the *Verrucomicrobia* phylum. Their identification was reported in three independent research papers published in 2007-2008 (Dunfield et al., 2007; Pol et al., 2007; Islam et al., 2008). In fact, the reports by Dunfield et al. (2007) and Pol et al. (2007) were published side-by-side in the same issue of Nature. Representative isolates of the verrucomicrobial methanotrophs were obtained from remote geothermal habitats in New Zealand, Southern Italy, and Kamchatka. They all showed high similarity between their 16S rRNA gene sequences (> 98%) and were described as a new genus, "*Methyloacidiphilum*" (Op den Camp et al., 2009). More recently, three other strains isolated from volcanic soil were proposed to represent a new genus, "*Methylacidimicrobium*" within the *Methylacidiphilaceae* family (van Teeseling et al., 2014).

### *1.2.2. Biochemistry of aerobic methanotrophs*

A common characteristic of all aerobic methanotrophs is the ability to oxidize methane with oxygen to carbon dioxide and water with the formation of methanol, formaldehyde, and formate as intermediates. The methane monooxygenase (MMO) enzyme, which is produced only in methanotrophic bacteria, is responsible for the oxidation of $CH_4$ to $CH_3OH$. This enzyme exists in two forms: particulate (pMMO) and soluble (sMMO). pMMO is a copper-containing enzyme that is associated with the ICMs of methanotrophic bacteria, while sMMO is a cytoplasmic non-heme iron enzyme complex (Ross and Rosenzweig, 2017). Only a few methanotrophs possess both enzymes. In methanotrophs that produce both forms of MMO, copper plays a key role in regulating the mutual expression of the genes encoding pMMO and sMMO, as well as in regulating the activity of these enzymes (Murrell et al., 2000b; Semrau et al., 2010).

Except *Methylocella* and *Methyloferula*, all known methanotrophs possess pMMO which is expressed at high (> 0.2 μM) content of copper ions in the medium. pMMO has a fairly narrow substrate specificity, oxidizing alkanes and alkenes with a length of not more than 5 carbon atoms. It consists of PmoB, PmoA, and PmoC subunits organized as a $\alpha_3\beta_3\gamma_3$ trimer (**Figure 5A**) (Lieberman and Rosenzweig, 2005). Based on

crystallographically modeled pMMO protein, it was proposed that the active site is localized in the di-copper center of the PmoB subunit (Culpepper and Rosenzweig, 2012). Genes encoding pMMO are organized in *pmoCAB* operon and are usually located on chromosomal DNA (Murrell et al., 2000a). In the genome of most methanotrophs, two identical copies of these genes are present, and a third copy of the *pmoC* gene is also found. The *pmoA* gene sequence is highly conserved and therefore widely used as a functional marker for the identification of methanotrophic bacteria in environmental samples (McDonald and Murrell, 1997).

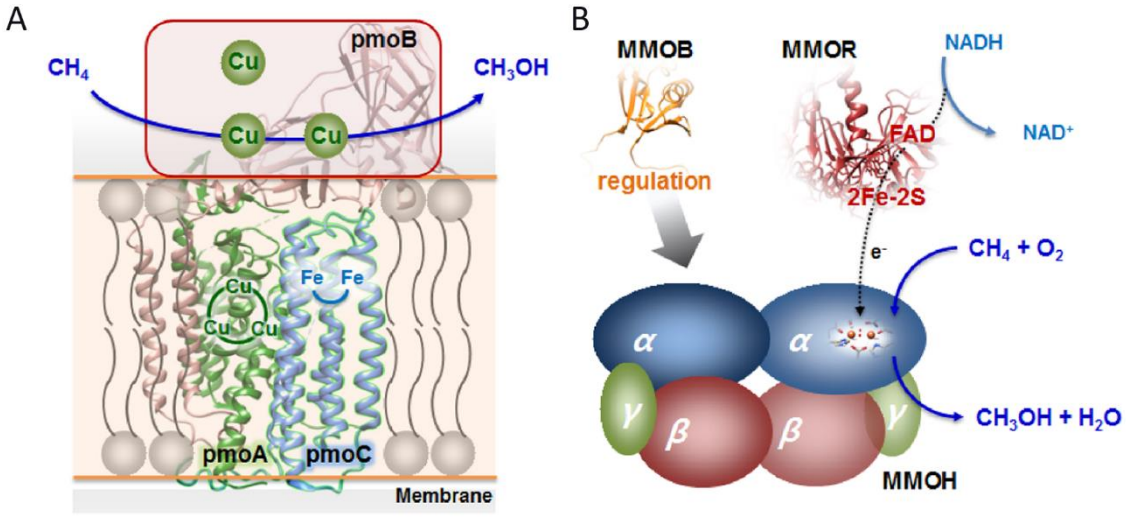



**Figure 5** | **(A)** Crystallographic structure of integral membrane-bound particulate methane monooxygenase (pMMO), which catalyzes the methane-to-methanol conversion in the majority of methanotrophs. The enzyme is shown to be associated with the ICMs of methanotrophs. It consists of the PmoA, PmoB, and PmoC subunits. The catalytic activity of pMMO is localized in the di-copper site of the PmoB subunit. **(B)** Structural overview of *Methylococcus capsulatus* (Bath) soluble methane monooxygenase (sMMO) consisting of three subunits. *Adopted from (Hwang et al., 2014).*

sMMO is present in various representatives of type II methanotrophs as well as in *Methylococcus capsulatus* and some other strains of type I methanotrophs. It catalyzes the pyridine nucleotide-dependent oxidation of methane by molecular oxygen to methanol and, contrary to pMMO, has a broad substrate range. In addition to methane, sMMO can catalyze the oxidation of alkanes, alkenes, aromatic hydrocarbons and other multicarbon substrates. sMMO is a three-component enzyme complex consisting of hydroxylase, reductase, and regulatory protein (**Figure 5B**). Hydroxylase is responsible for the oxidation of methane and consists of 3 subunits organized in a $\alpha_2\beta_2\gamma_2$ dimer and is encoded by the *mmoX, mmoY*, and *mmoZ* genes. Reductase and regulatory proteins

are encoded by *mmoB* and *mmoC* genes. The sMMO operon also contains genes encoding a protein of unknown function (*mmoD*), which may play a role in the copper-dependent switch between pMMO and sMMO expression. Genes encoding sMMO are arranged in *mmoRGXYBZDC* order (Murrell et al., 2000a; Semrau et al., 2010; Sirajuddin and Rosenzweig, 2015).

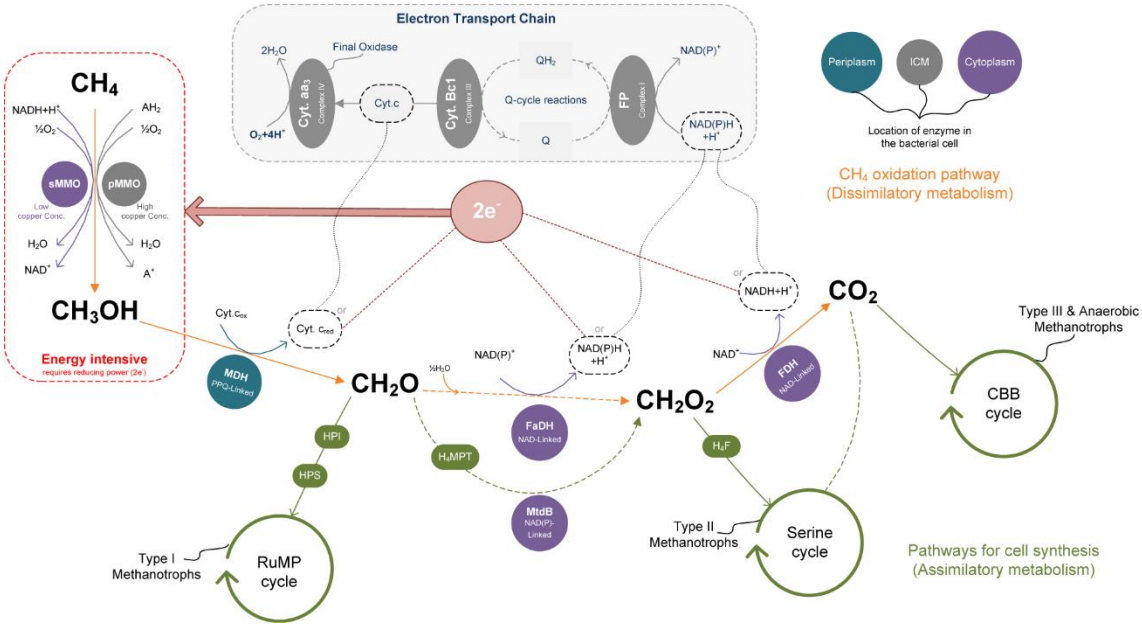



**Figure 6** | Different methane oxidation pathways in methanotrophs. Typical type I methanotrophs use RuMP pathway for C1 assimilation, while type II methanotrophs use serine pathway. In *Verrucomicrobia*, Calvin-Benson cycle (CBB) is used. Abbreviations: pMMO, particulate methane monooxygenase; sMMO, soluble methane monooxygenase; RuMP, ribulose monophosphate; MDH, methanol dehydrogenase; $H_4$MPTP, methylene tetrahydromethanopterin pathway; FADH, formaldehyde dehydrogenase; FDH, formate dehydrogenase. *Adopted from (AlSayed et al., 2018)*

Methanol, formed from methane, is subsequently oxidized by the methanol dehydrogenase (MDH) enzyme located in the periplasmic space of the cell (**Figure 6**). The prosthetic group of this enzyme is pyrroloquinolinequinone (PQQ) (Trotsenko and Murrell, 2008). MDH has a tetrameric $\alpha_2\beta_2$ structure. Each $\alpha$-subunit carries a $Ca^{2+}$ ion and one PQQ molecule (Anthony and Williams, 2003). Genes encoding the large subunits of MDH (*mxaF*) are highly conserved in both methanotrophs and methylotrophs (Bastien et al., 1989; McDonald and Murrell, 1997).

After methanol oxidation, the oxidation of formaldehyde is the next essential step of the conversion of methane to carbon dioxide (**Figure 6**). Most of the reduction equivalents for energy conservation and growth are formed by the oxidation of

formaldehyde to formate and finally to carbon dioxide. Alternatively, formaldehyde may be assimilated by one of the two pathways of primary C1 assimilation – RuMP pathway or serine pathway - and may also be used as a precursor for the synthesis of key intermediates of the central cell metabolism (**Figure 6**).

The RuMP pathway produces mainly sugar phosphate intermediates, while the serine pathway includes amino acids, CoA derivatives and TCA intermediates. RuMP pathway is operated in the metabolism of monocarbon compounds only by *Gammaproteobacteria* methanotrophs and some other methylotrophs that are not capable of growth on methane. In contrast, in type II methanotrophs belonging to *Alphaproteobacteria*, the assimilation of formaldehyde proceeds along the serine pathway (Hanson and Hanson, 1996). In the first reaction of the serine pathway catalyzed by serine hydroxymethyl transferase, formaldehyde combines with glycine to form serine. Then, serine is transaminated with the participation of hydroxylate as an acceptor of the amino group leading to the formation of hydroxypyruvate and glycine. Hydroxypyruvate is reduced by hydroxypyruvate reductase to glycerate. Next, glycerate kinase catalyzes the addition of ATP phosphate to form 2-phosphoglycerate. Further reactions of the conversion of 2-phosphoglycerate to phosphoenolpyruvate, the fixation of carbon dioxide catalyzed by phosphoenolpyruvate carboxylase, and the reduction of oxaloacetate to malate are characteristic of many heterotrophic bacteria (Hanson and Hanson, 1996). In the second stage of the metabolic pathway, derived acetyl-CoA is converted to glyoxylate (Kalyuzhnaya et al., 2015).

*Verrucomicrobia* methanotrophs also contain pMMOs, but they lack the formaldehyde assimilation pathways known for aerobic methanotrophs (**Figure 6**). Verrucomicrobial carbon required for growth and biomass synthesis is obtained by $CO_2$ fixation through the Calvin – Benson cycle (Op den Camp et al., 2009).

### *1.2.3. "High-affinity" and "low-affinity" methanotrophs*

Aerobic methanotrophs are widespread in nature and usually colonize environments, in which both methane and oxygen are available. Such habitats include wetlands, rice paddies, fuel fossils, coal mines, and solid waste landfills, where intensive methanogenesis occurs. They can be also found in thermal springs, soils, bogs, plant

rhizosphere and may even be endosymbionts of marine invertebrates (Hanson and Hanson, 1996; Dedysh et al., 2003; Knief, 2015).

Rice cultivation is a major source of atmospheric methane and its contribution to the global methane budget may increase due to the growing demand for rice. Studies have shown that rice fields are colonized by a highly diverse methanotrophic community, including members of the genera *Methylomonas, Methylobacter, Methylomicrobium, Methylococcus, Methylocaldum, Methylocystis* and *Methylosinus* (Cai et al., 2016). Freshwaters and sediments harbor a variety of different methanotrophs as well. In particular type I methanotrophs, including *Methylomonas, Methylobacter, Methylosarcina, Methylococcus* and *Methylosoma*, are predominant in these environments (Costello et al., 2002). These methanotrophs, thriving in methane-rich environments, are called conventional or "low-affinity" methanotrophs.

In upland soils, methanotrophic bacteria act as a biological sink for atmospheric methane. Their activity accounts for a small but a significant part of the global methane budget. In contrast to methanotrophs in the oxic-anoxic interfaces of methanogenic environments, the methanotrophs that are active in dry, well aerated upland soils are able to consume methane with high apparent affinity (Kolb et al., 2005). This suggests that methanotrophs in upland soils are oligotrophic and possess a mechanism of high-affinity methane uptake to survive on the trace levels of atmospheric methane.

Cultivation-independent research has shown that conventional type II methanotrophs are present in most upland soils. These studies provided increasing evidence that, in these soils, the most abundant methanotrophs belong to *Methylocystis* spp. and to so-called upland soil cluster (USCα) methanotrophs. The first USCα representative was only recently isolated into pure culture (Tveit et al., 2019). In a number of type II methanotrophs, two different *pmoA* genes were found: the "standard" *pmoA* or *pmoA1* and the "new" *pmoA* or *pmoA2* (Dunfield et al., 2002; Yimga et al., 2003). The latter is absent in representatives of the genera *Methylobacter, Methylomicrobium, Methylomonas, Methylococcus* and *Methylocaldum* (type I methanotrophs). Comparative sequence analysis showed that all *pmoA2* sequences form a coherent cluster clearly distinct from *pmoA1* sequences of type I and type II methanotrophs, and from *amoA* sequences of the *Nitrosomonas-Nitrosospira* group (Yimga et al., 2003).

*Methylocystis* spp., which were shown to harbor both *pmoA1* and *pmoA2* genes, possess a complete *pmoCAB2* gene cluster, in addition to two copies of the standard *pmoCAB1* operon (Ricke et al., 2004). Enzymes encoded by the *pmoCAB1* (pMMO1) and *pmoCAB2* (pMMO2) genes have different affinities for $CH_4$. The expression of pMMO1 takes place at elevated methane concentrations, while pMMO2 is even capable of oxidizing atmospheric $CH_4$ (Baani and Liesack, 2008). Therefore, those bacteria responsible for atmospheric methane oxidation are usually called ''high-affinity methane oxidizers'' or ''atmospheric methane oxidizers'' (Dunfield, 2007). The pMMO2 provides type II methanotrophs with a selective advantage, enabling them to survive under methane conditions at which most type I methanotrophs do not thrive.

### *1.2.4. Methanotrophs in hypoxic and anaerobic environments*

Recently, metabolically active type I methanotrophs have been increasingly detected in hypoxic and even anaerobic habitats, where they are assumed to play a leading role in the oxidation of methane (Blees et al., 2014; Milucka et al., 2015; Danilova et al., 2016; Oswald et al., 2016). Methanotrophic *Gammaproteobacteria* (type I) are universally recognized as strict aerobes capable of maintaining activity even in microaerobic (hypoxic) conditions. This ability is especially important for nitrogen-fixing type I methanotrophs, which are more sensitive to oxygen under conditions of nitrogen fixation than type II methanotrophs.

Under hypoxic conditions, conventional type I methanotrophs of the *Methylococcaceae* family (*Methylobacter* spp.) are often found to co-occur with methylotrophic bacteria (utilizing methanol and methylamines as substrates for growth, but not methane) of the *Methylophilaceae* family (*Methylotenera* spp.) (Chistoserdova, 2015; Savvichev et al., 2017). The availability of oxygen is one of the main factors determining the partnership of these microbes in the process of methane oxidation (Oshkin et al., 2015). Furthermore, the ability to oxidize methane, coupled with the reduction of nitrate, was recently discovered in representatives of the newly described gammaproteobacteria *Methylomonas denitrificans*. Denitrification ability has also been demonstrated in the genus *Methylotenera*. Nitrogen metabolism, and especially its denitrification ability, is thought to provide a selective advantage for both *Methylobacter* and *Methylotenera* genera in environments with low oxygen content.

Metabolically active type I methanotrophs were also detected in the anaerobic zone of stratified freshwater lakes. Their population size correlated with the methane-oxidizing activity (Milucka et al., 2015; Oswald et al., 2016). The research hypothesis was that these aerobic gammaproteobacterial methanotrophs may have a trophic relationship with oxygenic phototrophs, e.g. cyanobacteria, under anaerobic conditions. The cyanobacteria develop with low light source below the oxycline zone, thereby forming oxygen which is immediately consumed by aerobic methanotrophs to oxidize methane and is therefore not detected during the biochemical measurements. This process of "photodependent methane oxidation" has been experimentally confirmed for the stratified freshwater lakes Cadano, Rotsee (Switzerland) and La Cruz (Spain). Overall, the above examples cast doubt on the "strictly" aerobic nature of methanotrophic gammaproteobacteria (Chistoserdova, 2015).

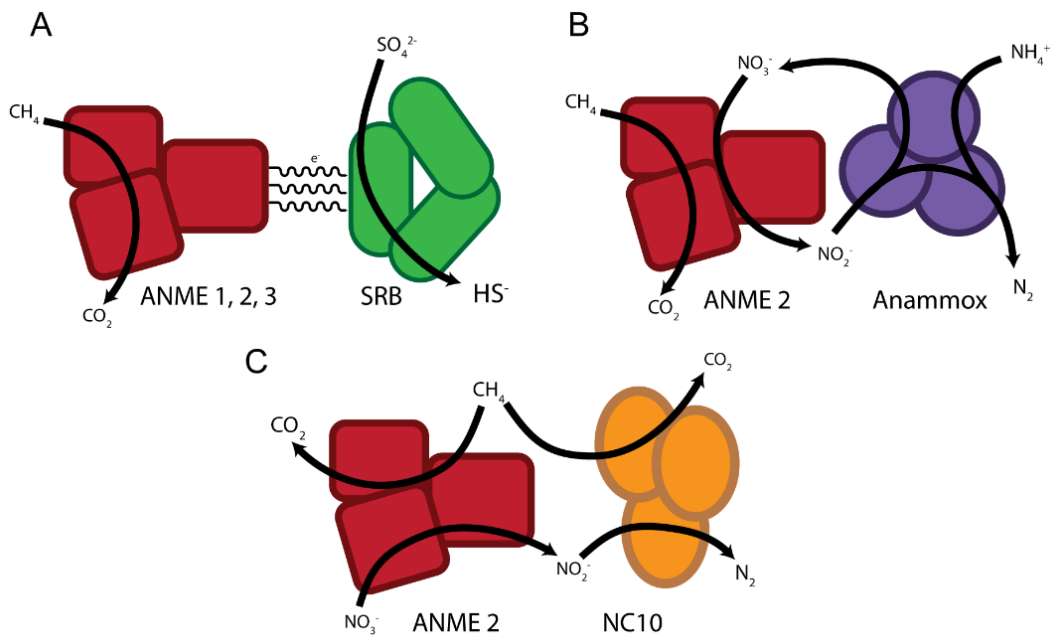



**Figure 7** | Three mechanisms of anaerobic oxidation of methane. **(A)** The first mechanism is mediated by a consortium composed of anaerobic methanotrophic (ANME) archaea and sulfate-reducing bacteria (SRB). The archaeal partner may belong to the ANME clades 1, 2a, 2b, 2c, and 3. The oxidation of methane occurs in the ANME where electrons are passed directly to the SRB, which performs sulfate reduction. **(B)** The second mechanism links methane oxidation with nitrate reduction, mediated by consortia of ANME archaea and anammox bacteria. **(C)** The third mechanism also links methane oxidation with nitrate reduction but is mediated by ANME archaea and NC10 bacteria. Unlike the first two mechanisms, both the ANME archaea and NC10 bacteria compete for methane. *Adopted from Wikipedia.*

The process of anaerobic oxidation of methane has long been known (Reeburgh, 1976; Hoehler et al., 1994; Boetius et al., 2000; Valentine and Reeburgh, 2000). According to some estimates, about 90% of methane formed in the oceans is removed due to anaerobic methane oxidation (Knittel and Boetius, 2009). Anaerobic oxidation of methane is carried out by archaea of three phylogenetic clusters belonging to the phylum *Euryarchaeot*a: ANME-1, ANME-2 and ANME-3 (ANaerobic MEthan-oxidising archaea, ANME) (**Figure 7**). Archaea of the ANME-1 and ANME-2 clusters are very widespread, while representatives of the ANME-3 cluster are found mainly in deep-sea mud volcanoes. Through syntrophic partnership with the bacterial component, the ANME archaea carry out sulfate- (**Figure 7A**) and nitrate-dependent (**Figure 7B**) anaerobic oxidation of methane. It is performed by the reverse methanogenesis mechanism, which is catalyzed by methyl coenzyme M reductase (Mcr) (Michaelis et al., 2002; Knittel and Boetius, 2009; Haroon et al., 2013; Timmers et al., 2016; Timmers et al., 2017). There is increasing evidence that metal oxides - manganese (IV), iron (III) and chromium (VI) – are also used as electron acceptors in anaerobic oxidation of methane, but the mechanisms and agents of these processes have not yet been reliably determined (Ettwig et al., 2016; Fu et al., 2016; Timmers et al., 2017).

It is also necessary to take into account the so-called “trace methane oxidation (TMO)” when assessing the extent of methane oxidation under anaerobic conditions (Timmers et al., 2017). The occurrence of this process was first demonstrated by Zender and Brock in 1979; using $^{14}CH_4$ they observed the ability of pure cultures of methanogens to form $^{14}CO_2$ (Zehnder and Brock, 1979). In mixed communities, the rates of TMO were higher than in pure cultures, and the increased partial pressure of methane stimulated this process. However, unlike anaerobic oxidation of methane by ANME archaea, the rates of TMO never exceed the rates of methanogenesis. TMO occurs both in the presence and absence of external electron acceptors, but only in the background of general methanogenesis. Therefore, it is most likely to be the result of a reverse flow by individual enzymes of the methanogenic pathway (Pimenov et al., 2014; Timmers et al., 2017).

An intermediate position between aerobic and anaerobic methanotrophy is occupied by representatives of the candidate phylum NC10, namely “Ca. *Methylomirabilis oxyfera*”. These unique bacteria carry out nitrite-dependent anaerobic

oxidation of methane by intracellular formation of oxygen due to the dismutation of two NO molecules (**Figure 7C**) (Ettwig et al., 2010). NO-dismutase is a homologue of NO-reductase in denitrification process. Given that "Ca. *Methylomirabilis oxyfera*" bacteria are obligate anaerobes, they are not able to use external oxygen. Though, intracellularly formed oxygen can be partially used for the oxidation of methane along the route known for conventional aerobic methanotrophs with the participation of pMMO. The rest of the oxygen produced is believed to be used for normal cellular respiration (Shen et al., 2015).

### *1.2.5. Facultative methanotrophy and alternative energy sources for methanotrophs*

The majority of aerobic methanotrophs can use methane as their only source of carbon and energy. Thus, they are "obligate methanotrophs". Some of the obligate methanotrophs can also utilize the intermediates of methane oxidation, such as methanol, formate, and formaldehyde and, in addition, methylamines. Interestingly, after several years of research, it was finally shown that methanotrophs are also able to utilize multicarbon compounds as carbon and energy source; therefore being facultative methanotrophs. The facultative lifestyle was observed among several members of genera *Methylocystis, Methylocella* and *Methylocapsa* (Dedysh et al., 2005; Dunfield et al., 2010; Belova et al., 2011; Dedysh and Dunfield, 2011; Im et al., 2011). Among these, the research on *Methylocella* was the pioneering study, which provided the first valid evidence for facultative methanotrophy. *Methylocella* is an unusual methanotroph with respect to both its cell architecture and preference for carbon sources. They lack ICMs and possess only sMMO and, in addition, preferably use multicarbon sources instead of methane (acetate, pyruvate, succinate, malate, and ethanol) (Dedysh et al., 2005).

Later, facultative methanotrophy was also shown to occur in *Methylocystis* spp. and *Methylocapsa* spp. (Dunfield et al., 2010; Belova et al., 2011; Im et al., 2011). In comparison to *Methylocella*, the latter organisms possess both a conventional cell architecture (ICMs) and pMMO. Though, in the absence of methane, they can use acetate and/or ethanol, and maintain pMMO expression even without methane. In summary, methanotrophs are able to survive in the oligotrophic environments where methane is limited or close to atmospheric concentrations. However, it remains to be

elucidated how frequently and which kind of alternative energy sources methanotrophic bacteria prefer to use in order to survive periods of methane starvation.

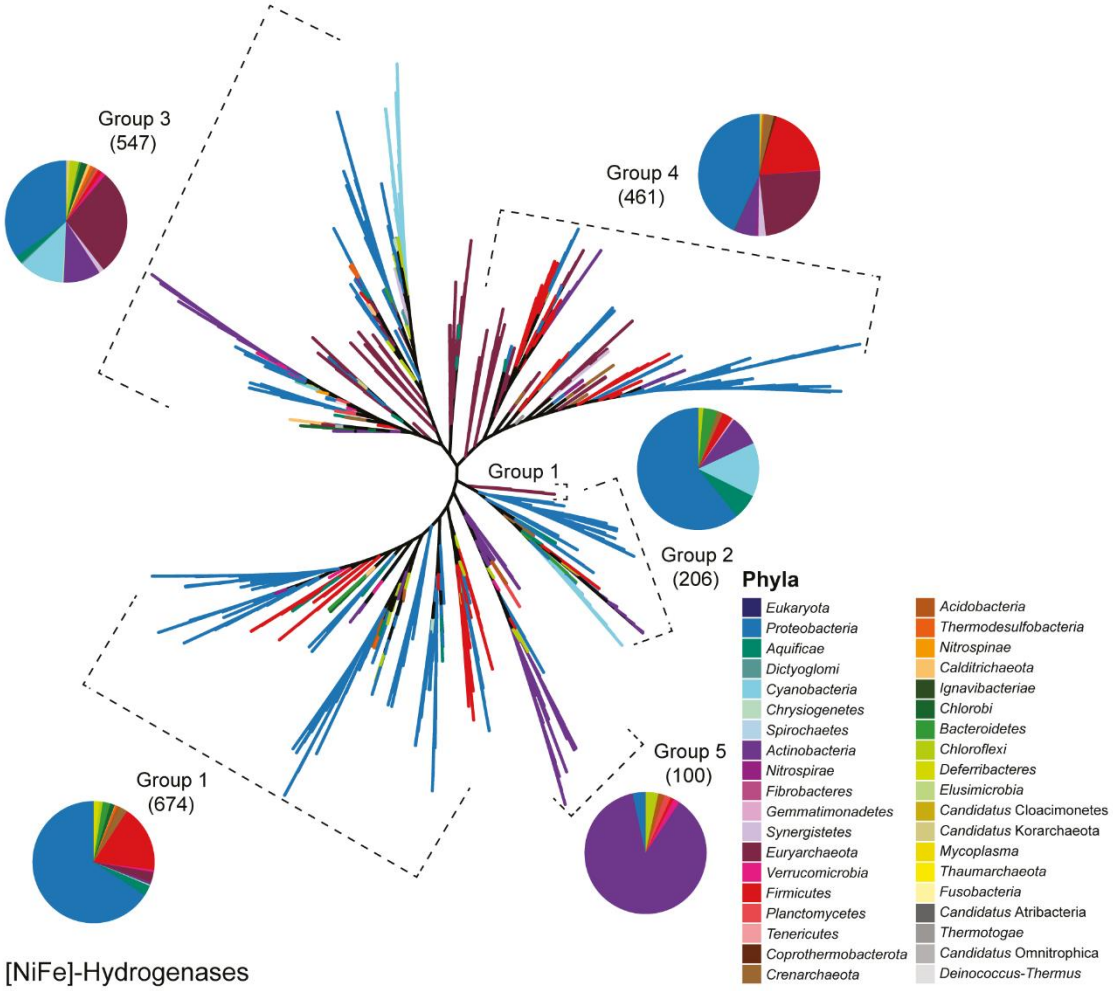



**Figure 8** | Consensus tree of 1,988 [NiFe]-hydrogenase sequences, with an alignment length of 1,850 amino acids. Branch colors represent taxonomic classification at the phylum level, while pie charts show the relative abundance of each phylum within hydrogenase subgroups (except for *Eukaryota* being a domain). Dashed brackets show which tree branches belong to a specific hydrogenases group. Numbers between parentheses depict the amount of sequences within that group. *Adopted from (Piche-Choquette and Constant, 2019).*

Recently, several physiological and ecological studies have demonstrated that hydrogen is a ubiquitous and easily accessible energy source for a wide range of microbial taxa in soil ecosystems (Piche-Choquette and Constant, 2019). Hydrogen is continuously produced during abiogenic and biogenic processes on our planet, particularly during biochemical processes like fermentation, nitrogen fixation, and photoproduction (Morita, 1999). Using $H_2$ as an effortless energy source, many microorganisms can prevail the fluctuations in the availability of carbon and energy sources that occur due

to changes in environmental conditions (Hoehler and Jorgensen, 2013). In upland soils and in oxic-anoxic layers of the soil, where hydrogen is present in high concentrations, $H_2$ oxidation is carried out by so-called *Knallgas* bacteria (Thauer, 2011). In addition, several groups of soil microorganisms were shown to oxidize even atmospheric concentrations of $H_2$ (ca. 0.53 ppmv), thus, being "high-affinity" hydrogen oxidizers (Schafer et al., 2013). *Knallgas* bacteria metabolize hydrogen using the hydrogenase enzyme and the electron-transport chain (ETC), resulting in the generation of ATP (Stephenson and Stickland, 1931).

Hydrogenases are metalloenzymes and can be classified as [Fe]-hydrogenases, [FeFe]-hydrogenases, and [NiFe]-hydrogenases. The most widespread hydrogenases in aerobic soil bacteria belong to [NiFe]-hydrogenases (Vignais and Billoud, 2007) (**Figure 8**). The group of [NiFe]-hydrogenases is quite heterogeneous and is divided into 5 subgroups, having either membrane-bound or cytosolic localization (Piche-Choquette and Constant, 2019). The majority of [NiFe]-hydrogenases are $O_2$ tolerant and have low catalytic activity.

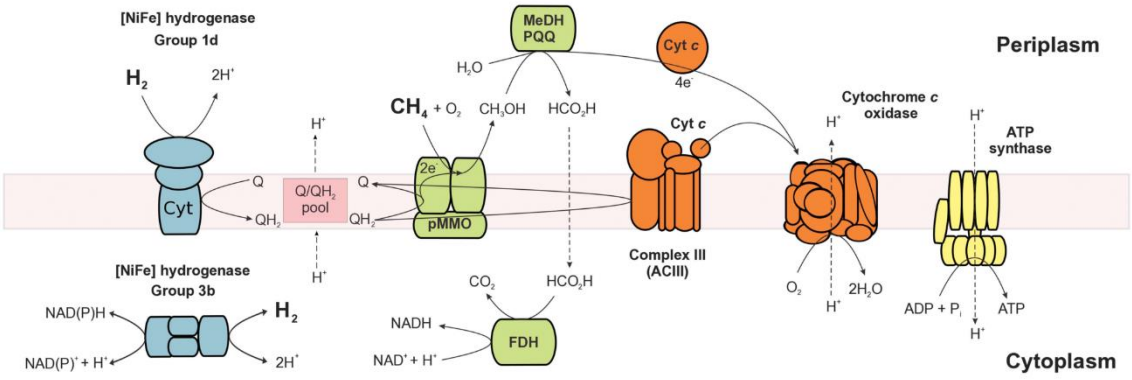



**Figure 9** | Proposed model of methane ($CH_4$) and hydrogen ($H_2$) oxidation in *Methylacidiphilum* sp. RTK17.1. During mixotrophic growth, the oxidation of both $H_2$ and $CH_4$ yields reducing equivalents in the form of reduced quinones (quinol-$QH_2$). A large proton-motive force is generated and sufficient ATP is produced for growth via an $H^+$-translocating $F_1F_O$-ATP synthase. Some of the quinol generated through $H_2$ oxidation provides the electrons necessary for pMMO catalysis. Following $CH_4$ oxidation by pMMO, ensuing reactions catalyzed by a XoxF-type methanol dehydrogenase (MeDH) and formate dehydrogenase (FDH) contribute additional reductant (cyt c and NADH) into the respiratory chain for ATP production and growth. NADH reduced through the activities of the formate dehydrogenase and $H_2$-dependent group 3b [NiFe]-hydrogenase is used to support $CO_2$ fixation through the Calvin–Benson–Bassham cycle. *Adopted from (Carere et al., 2017).*

A recent genomic survey of hydrogenase distribution among microbial taxa demonstrated that genes encoding hydrogenases are widely distributed in

methanotrophic bacteria (Carere et al., 2017). Several hypotheses on the contribution of hydrogen oxidation in methanotrophs were published. The oxidation of hydrogen in methanotrophic bacteria has been predicted to contribute reducing energy for methane oxidation (Hanczar et al., 2002), to recycle endogenous hydrogen produced during nitrogen fixation (Chen and Yoch, 1987) and to drive the non-productive oxidation of chlorinated solvents (Shah et al., 1995). Most recently, the newly isolated thermoacidophilic verrucomicrobial methanotroph *Methylacidiphilum* sp. RTK17.1 was demonstrated to oxidize hydrogen for mixotrophic growth and persistence (**Figure 9**). In strain RTK17.1 hydrogen oxidation enhanced mixotrophic growth yields during methanotrophic growth, under both oxygen-replete and oxygen-limiting conditions. In addition, hydrogen acted as an alternative energy source in the absence of methane oxidation. Therefore, it was proposed that $H_2$ can be an important electron donor for the methanotrophs to meet energy and reductant demands in response to fluctuations in $CH_4$ and oxidant availability (Carere et al., 2017). In addition, hydrogen oxidation increases the cell biomass yield and thus provides new opportunities for methanotrophs to be used in biotechnological applications.

## 1.3. Molecular biology methods to study methane-oxidizing bacteria

Nowadays, molecular studies of methanotrophs are primarily based on modern "omics" approaches. Accordingly, genomics and transcriptomics and, to lesser extent, proteomics and metabolomics are the main methods used in methanotrophs research.

Genomics is a molecular biology technique, which focuses on the sequencing of genomes, i.e. the total DNA of a cell or organism (Baldrian and Lopez-Mondejar, 2014). With the development of next-generation sequencing techniques, bacterial and archaeal genomes can now be sequenced within few hours. During last few years many genomes of type I and type II methanotrophs were made available in public databases, including the alphaproteobacterial methanotrophs *Methylosinus trichosporium* OB3b (Stein et al., 2010), *Methylocystis parvus* OBBP (del Cerro et al., 2012), *Methylocystis* sp. strain Rockwell (Stein et al., 2011), *Methylocystis* sp. strain SC2 (Dam et al., 2012a), and the facultative *Methylocella silvestris* BL2 (Chen et al., 2010). Among gammaproteobacterial methanotrophs, *Methylococcus capsulatus* Bath (Ward et al., 2004), *Methylomicrobium album* BG8 (Kits et al., 2013), *Methylomicrobium alcaliphilum*

20Z (Vuilleumier et al., 2012), and many others were genome-sequenced. Moreover, the genome sequences of the acidophilic *Verrucomicrobia* members *Methylacidiphilum infernorum* V4 (Hou et al., 2008) and *Methylacidiphilum fumariolicum* strain SolV (Khadem et al., 2012) were made available.

Upon the availability of high-quality genome sequences of methanotrophs, comparative genomics started to be commonly used for the study of genetic differences between closely related strains of a single species, species of a particular genus, or species of related genera of methanotrophs (Blom et al., 2009; Dam et al., 2013). Genome sequencing and comparative genomics provided new insights into the metabolic pathways of methanotrophic bacteria. For instance, genomic analyses of the verrucomicrobial methanotroph *Methylacidiphilum fumariolicum* SolV indicated that, in contrast to proteobacterial methanotrophs, it possesses a complete Calvin cycle. Thus, it uses carbon dioxide as its sole carbon source, while it oxidizes methane to generate energy (Khadem et al., 2011; Khadem et al., 2012).

Furthermore, transcriptomics has become a common approach in the field of methanotroph research. Transcriptomics provides information about the total mRNA in the cell and about the genes that are actively expressed at the time of sampling (Baldrian and Lopez-Mondejar, 2014). Transcriptomic analyses of the type IIa methanotroph *Methylocystis* sp. strain SB2 revealed the mechanisms involved in its facultative growth on ethanol (Vorobev et al., 2014). The transcriptome of *M. alcaliphilum* 20Z, grown under limited $O_2$ supply, showed its ability of a novel methane-derived formaldehyde fermentation, leading to the formation of formate, acetate, succinate, lactate, and hydroxybutyrate as end products (Kalyuzhnaya et al., 2013). Finally, *Methylocystis* sp. strain SC2 (details in section 1.4) was used in two transcriptomic studies. These revealed the molecular mechanisms of its adaptation to high-ammonium load and salt stress (Dam et al., 2014; Han et al., 2017).

While several well-described protocols for DNA and RNA extraction from methanotrophic cells have been published, there exist nearly no report on the efficient isolation of proteins. The special architecture of the methanotrophic cells, particularly the presence of ICMs, makes these organisms a challenge for proteomics studies. In fact, special approaches are required to get access to the complete proteome of methanotrophs. Therefore, no proteomics research had been conducted on

*Methylocystis* spp., while a very few studies were published for type I methanotrophs (details in section 1.3.1).

In sharp contrast to other "omics" methods, metabolomics provides a very fast and sensitive functional read-out that is directly linked to bacterial physiology. The metabolome can respond within a second or less to environmental changes (Kresnowati et al., 2006). Metabolomics was successfully applied in several methanotroph studies. Among those, global molecular analyses of methane metabolism were carried out in *Methylosinus trichosporium* OB3b using metabolomics and $^{13}C$-labeling. In contrast to genome- and transcriptome-based reconstruction of central metabolic pathways of strain OB3b, metabolomics revealed its ability to use the EMC variant of serine pathway to assimilate the $CO_2$ produced through oxidation of methane. Moreover, the study showed that a portion of the intracellular acetyl-CoA is derived from pyruvate by the activity of pyruvate dehydrogenase (Yang et al., 2013).

### *1.3.1. Proteomics as a new tool for methanotroph research*

In 1986, Mark Wilkins introduced the term "proteome", combining two words: "PROTein" and "genOME" (Wilkins et al., 1996). A proteome is a complex of proteins in a cell, including their modifications that occur over time or under the influence of any factors. In 1997, the term proteomics was introduced in analogy to genomics which studies genes and their functions (James, 1997). Proteomics investigates the total pool of the proteins synthesized by an organism / cell under a given condition and at a particular phase of their cell cycle. Proteomics is still lagging behind genomics and transcriptomics due to insufficient sensitivity of existing methods; however, an increasing number of research studies appear using proteomics techniques, including proteomics studies of different microbial groups.

For many years mass spectrometry (MS)-based methods have been widely used as a standard approach for accurate mass determination and characterization of proteins and peptides in various biological studies (Pandey and Mann, 2000; Jensen, 2006; Cox and Mann, 2011; Aebersold and Mann, 2016). For total proteome analysis, technologies based on so-called "bottom-up" proteomics are considered the most effective (Molloy and Witzmann, 2002; Monteoliva and Albar, 2004). In "bottom-up" proteomics the total mass of proteins isolated from a biological object is first

proteolytically cleaved into peptides (**Figure 10**). To achieve this, the cellular membrane is solubilized and extracted, and proteins are further denatured using detergents or chaotropic salts. Following denaturation, proteins are digested into MS-compatible peptides. In the past years, various experimental strategies have been developed for protein extraction and digestion, which were compatible with online liquid chromatography-coupled tandem mass spectrometry (LC-MS/MS) analysis. Well-known strategies incorporate strongly solubilizing detergents that are easy to remove from the sample after in-solution digestion (ISD), or filter-aided sample preparation strategies (FASP) (Wisniewski et al., 2009). This procedures result in a highly complex peptide mixture, which is subsequently analyzed by high performance LC-MS/MS.

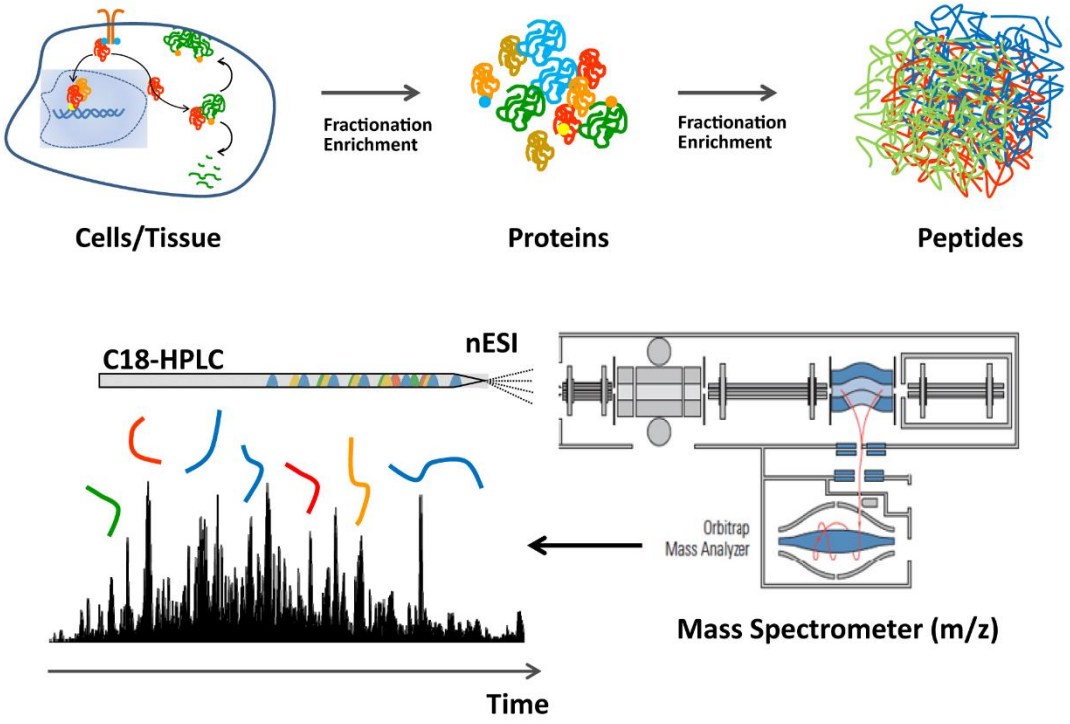



**Figure 10** | "Bottom-up" proteomics workflow. During sample preparation proteins are extracted from biological sources and digested into peptides. The complex peptide mixture is then separated on an HPLC system connected online with the mass spectrometer (MS). Here gaseous peptide ions are separated according to their mass-to-charge ratio. In a typical liquid chromatography-mass spectrometry (LC-MS) experiment thousands of MS spectra are acquired. *Adopted from Handbook of System Biology.*

The whole variety of "bottom-up" MS techniques can be divided into discovery or "shotgun" proteomics and targeted proteomics. Shotgun proteomics is intended to identify and quantify the largest possible number of proteins in the cell at a given time point and condition (Mehaffy et al., 2010). It commonly uses the data-dependent

acquisition (DDA) approach for peptide identification (**Figure 11**, *upper panel*). Here, peptides eluting from the LC column at a particular time point are first successively examined (MS1 scan) and then selected for intensity-based fragmentation to generate MS/MS (MS2) spectra (Bailey et al., 2014; Tsou et al., 2015). Further, the identification of the peptides from complex MS2 spectra is conducted based on a so-called "spectrum-centric" approach, using available public database searches (Venable et al., 2004; Hoopmann et al., 2009). The disadvantage of the DDA approach is that usually only the top most intense peptide ions are selected from MS1 scan for further fragmentation.

Targeted MS approaches offer the potential to overcome some of the challenges that scientists face with discovery proteomics. In contrast to shotgun proteomics, targeted proteomics is designed to track a relatively small set of proteins and peptides with the highest available sensitivity, accuracy, and reproducibility throughout the workflow. Once established, targeted proteomics methods are relatively fast to perform and highly reproducible. In addition, targeted proteomics can be combined with the application of isotope-labeled peptides that allow for absolute quantification (Marx, 2013; Borras and Sabido, 2017).

Selected or multiple reaction monitoring (SRM/MRM) is a well-established method of targeted proteomics (**Figure 11**, *middle panel*). One of the advantages of SRM/MRM technology is the absence of non-target detection. Thus, SRM/MRM significantly reduces the noise signal that is present in discovery proteomics and, therefore, improves the detection sensitivity (Lange et al., 2008; Ronsein et al., 2015). Nonetheless, there also exist drawbacks in SRM/PRM technology. For instance, the intensities of individual fragment product ions derived from a single precursor ion can differ substantially. To obtain a sensitive assay, it is essential to select the most intense product ions, which can be challenging and time-consuming. Moreover, because of the low-resolution of quadrupole instruments, the concomitant detection of interfering ions is a major issue in SRM experiments, especially when applied to complex biological samples (Gallien et al., 2013; Ronsein et al., 2015).

Next to SRM/MRM, parallel reaction monitoring (PRM) was developed (**Figure 11**, *middle panel*), which allows quantification of multiple peptides with increased sensitivity and specificity (Schiffmann et al., 2014; Ronsein et al., 2015). Usually, PRM is performed on high-resolution Orbitrap or time-of-flight instruments and simultaneously

monitors product ions of a targeted peptide with high resolution and mass accuracy (Gallien et al., 2012; Peterson et al., 2012). Thereby, PRM offers higher specificity than SRM on quadrupole instruments. In addition, PRM tracks the product ions with higher resolution than SRM and is therefore less likely to be affected by interfering ions. Thus, PRM is considered more advanced technology of targeted proteomics, which requires much less effort than the traditional SRM assay (Gallien et al., 2012).

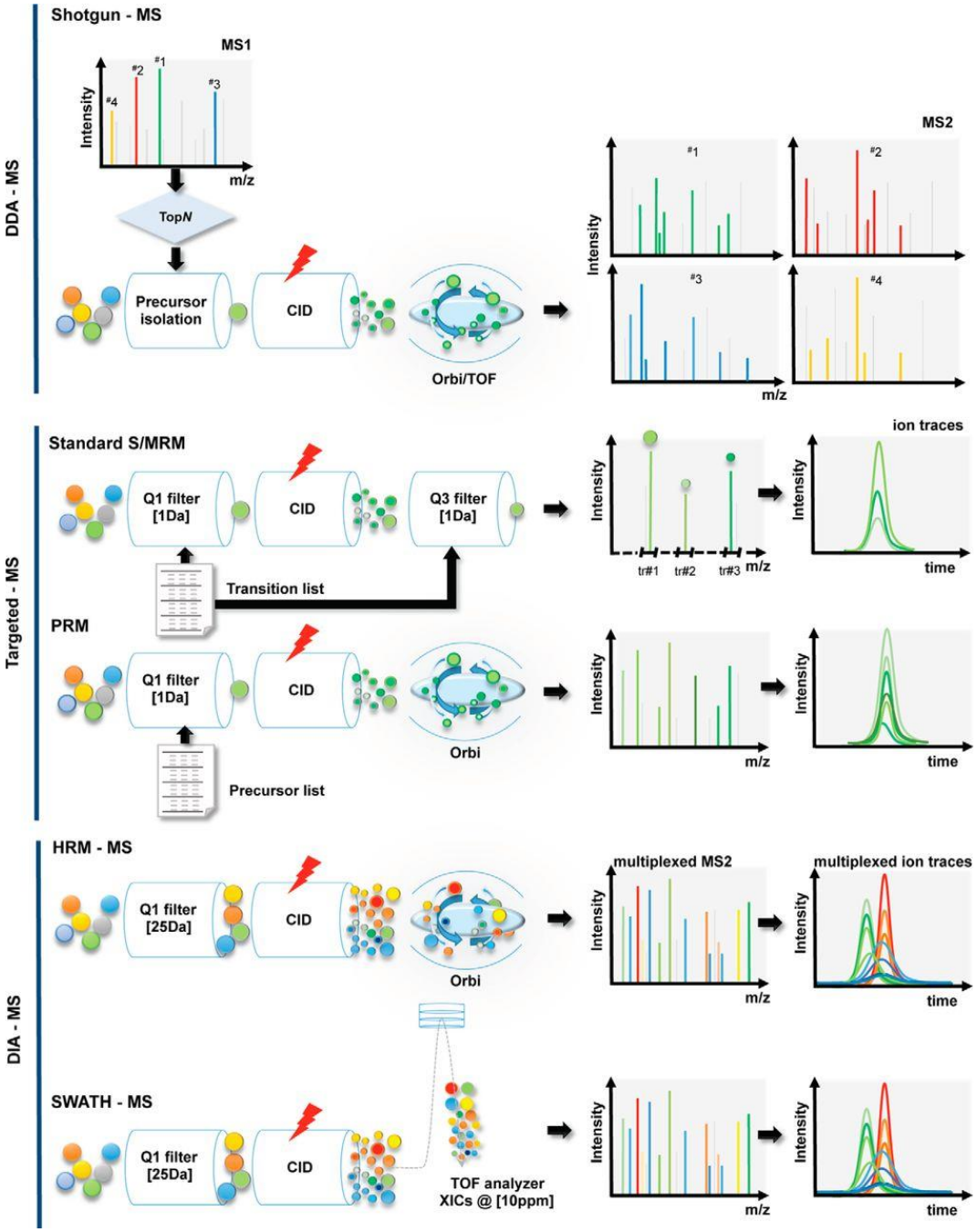



**Figure 11** | Mass spectrometry-based techniques in proteomics. All settings comprise isolation and fragmentation of precursor ions and recording of the mass-to-charge ratio of MS2 fragment ions in a mass analyzer. (*Upper panel*) In DDA MS, only the most abundant precursor ions (Top*N*) per MS1 scan are

selected for consecutive isolation and fragmentation. A linear ion trap, an Orbitrap, or a TOF analyzer is used for the high-throughput acquisition of MS1 and MS2 spectra. (*Middle panel*) In targeted MS, SRM/MRM exploits the capability of a triple quadrupole to screen a list of transitions (tr), *i.e.* precursor-fragment ion pairs. The transitions are monitored over time to yield ion traces corresponding to the peptides of interest. PRM experiments are conducted using an Orbitrap-type high-resolution and accurate mass spectrometer (*e.g.* Q-Exactive). All MS2 product ions derived from predefined peptides (MS1 precursors) are recorded over time to generate the ion chromatographic traces. (*Lower panel*) HRM and SWATH MS are two fundamentally similar data-independent acquisition methods employing an Orbitrap-type and a quadrupole TOF high-resolution accurate mass spectrometer, respectively. In DIA mode, multiplexed ion traces are acquired by repeatedly cycling through predefined consecutive precursor isolation windows (originally 32 × 25 Th) and by monitoring all fragment ions. *Adopted from (Caron et al., 2015).*

The technological development in recent years (Bin Goh et al., 2015) has led to the establishment of data-independent acquisition methods (DIA), which represent a combination of discovery and targeted proteomics approaches (**Figure 11**, *lower panel*). Here, the precursors are selected and fragmented regardless of their abundance in the sample, thereby offering a more consistent quantification of protein samples. Among those methods, SWATH (sequential window acquisition of all theoretical mass spectra) are the most well-known and widely used (Gillet et al., 2012). SWATH-MS does not require prior selection of single precursor ions, but is rather based on the acquisition of fragment-ion information by repeatedly cycling through relatively wide precursor isolation windows (Ludwig et al., 2018). The SWATH windows (e.g. 25 Da) have predefined broad range of m/z and usually cover the entire mass range of most MS-measurable precursors (Gillet et al., 2012). Thus, the SWATH strategy creates a comprehensive multi-window SWATH map for each sample with a single injection (Bin Goh et al., 2015). Similar to targeted proteomics, SWATH-DIA data can be assessed in targeted manner using spectral libraries that contain retention time and fragmentation information for the selected peptide precursors. Thus, SWATH allows for the identification and quantification of a large set of proteins, similar to classical discovery (DDA) approaches, but with the accuracy and reproducibility of SRM/PRM for a larger number of samples (Ludwig et al., 2018). However, the highly complex fragment spectra derived from multiple precursor ions makes the data-analysis quite challenging (Bin Goh et al., 2015). Moreover, SWATH-DIA data are producing more noise than most DDA data due to parallel fragmentation of an increased number of precursors (Gillet et al., 2012). Nonetheless, recent improvements in software tools and bioinformatics now enable accurate analysis of SWATH-DIA data.

It is worth to mention that when discovery approaches were implemented into the proteomics workflow, the most accurate quantification results among a large number of samples were achieved by stable isotopic labeling techniques. Common labeling strategies include protein- and peptide-labeling strategies such as stable isotope labeling of amino acids (SILAC), labeled synthetic peptides (AQUA), or multiplexed isobaric tags (iTRAQ) (**Figure 12**, *right panel*). In the field of quantitative proteomics, labeling approaches have been considered the "gold standard" for relative proteome comparisons for many years (Graumann et al., 2008).

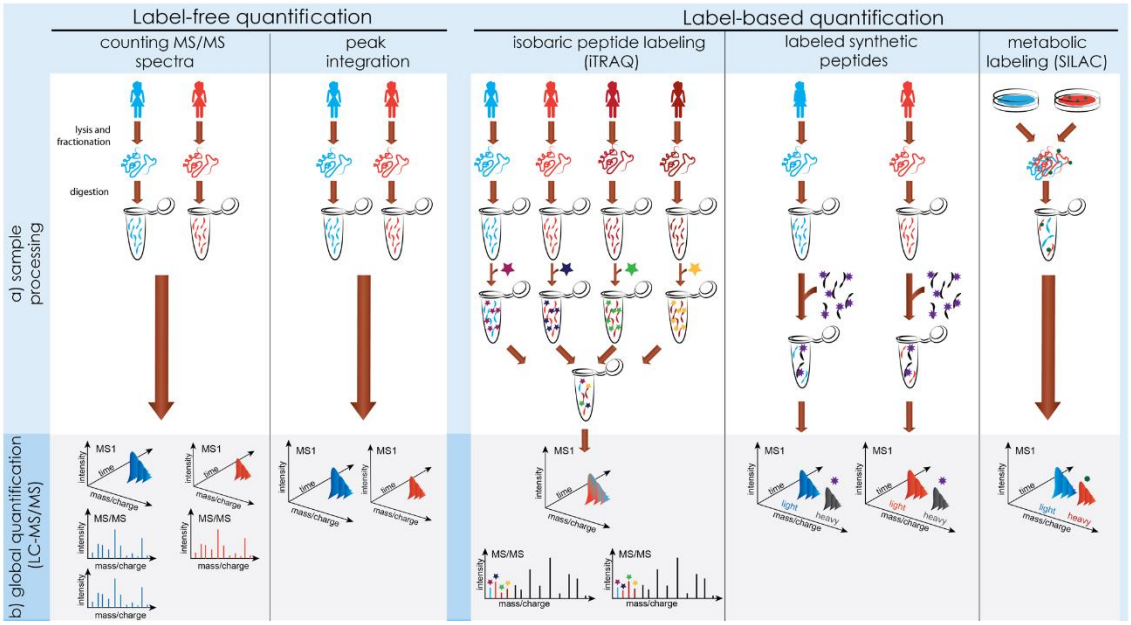



**Figure 12** | Label-free and label-based techniques of shotgun proteomics. Label-free quantification requires minimal sample manipulation, but acquires spectra from each sample in a separate mass spectrometry run. Label-based quantification varies in the timing and type of the labeling steps, but always simultaneously profiles two or more biological samples within a single run. Global label-free workflows achieve relative quantification by comparing (i) counts of MS/MS spectra or (ii) the intensities of MS peaks between runs. Global label-based workflows compare intensities of reporter MS/MS fragments (iTRAQ) or MS peaks (SILAC, synthetic peptides). *Adopted from (Kall and Vitek, 2011).*

Although labeling approaches remain efficient, label-free quantification (LFQ) approaches (**Figure 12***, left panel*) are nowadays more commonly used as robust technological platforms for relative quantification (Domon and Aebersold, 2010). LFQ aims to determine the relative amount of proteins in two or more biological samples, without incorporating stable isotope labelling (Nikolov et al., 2012). The data analysis in label-free approaches is based on the detection of peptides, match of the corresponding peptides across LC-MS data of all analyzed samples, and selection of discriminative peptides (Asara et al., 2008). Precursor signal intensity or spectral counting are usually used to determine the relative abundance changes of proteins under different

conditions upon LFQ. The differences in precursor signal intensity of the same protein between multiple samples accurately reflect relative differences in its abundance (Bantscheff et al., 2012). Typically, precursor signals are detected at the MS1 level and distinguished from chemical noise through their characteristic isotopic pattern. Following, those patterns are tracked across the retention time dimension and used to reconstruct a chromatographic elution profile of the mono-isotopic peptide mass.

In general, microbes are excellent model organisms for verification of new proteomics technologies, since microbial genomes "only" comprise 500 to 5000 genes. Therefore, a complete proteome dataset can be achieved faster for microorganisms than for eukaryotes. Membrane-associated proteins comprise the crucial part of microbial proteome. About 20% to 30% of all genes in bacteria are coding for membrane-associated proteins (Poetsch and Wolters, 2008). These proteins perform essential physiological functions, such as the import or export of metabolites, the homeostasis of metal ions, the extrusion of toxic substances or antibiotics, and the generation or conversion of energy. Large membrane-associated proteins with extra-membrane domains in the periplasmic space usually provide sufficient peptide information for protein identification even without efficient membrane solubilization (Wu et al., 2003; Vit and Petrak, 2017). However, a major portion of the membrane-associated proteins is usually embedded in the phospholipid bilayer and, therefore, inaccessible for proteolytic cleavage. Therefore, solubilization of the membrane material is one of the most important steps for the complete and unbiased proteome analysis of microbial cells. Chaotropes, organic solvents and, in particular, detergents have been used in different membrane proteomics studies. Among all known detergents, sodium deoxycolate (SDC) showed the highest efficiency in digesting hydrophobic membrane-associated proteins through their efficient denaturation and solubilization (Leon et al., 2013; Moore et al., 2016; Vit and Petrak, 2017). Albeit less powerful than detergents, chaotropes, like urea, are sometimes used in membrane proteomics to achieve disruption of protein-protein interactions, denaturation and/or to maintain the unfolded state of proteins (Vit and Petrak, 2017). To obtain a comprehensive coverage of membrane proteins, the next crucial step is the proteolytic digestion of these proteins.

Trypsin has long been the gold standard in proteomics. However, trypsin is usually very sensitive to the denaturation environment and is not completely suitable for the cleavage of membrane-associated hydrophobic proteins. In several studies, bacterial serine endoprotease Lys-C, together with trypsin, was used to increase the digestion efficiency of membrane proteins in urea environments (Nielsen et al., 2005; Dormeyer et al., 2008; Wisniewski et al., 2009; Wisniewski and Mann, 2012). The major advantage of Lys-C is its ability to remain active in high concentrations of urea, thereby enabling the specific pre-digestion of hydrophobic proteins prior to trypsin digestion (**Figure 13**). This tandem strategy has been employed and became almost a standard approach in membrane proteomics (Vit and Petrak, 2017). In addition, Lys-C/trypsin tandem digestion was shown to be more efficient in yielding fully cleaved peptides while reducing the abundance of miscleaved peptides, both in discovery and targeted proteomics approaches (Glatter et al., 2012).

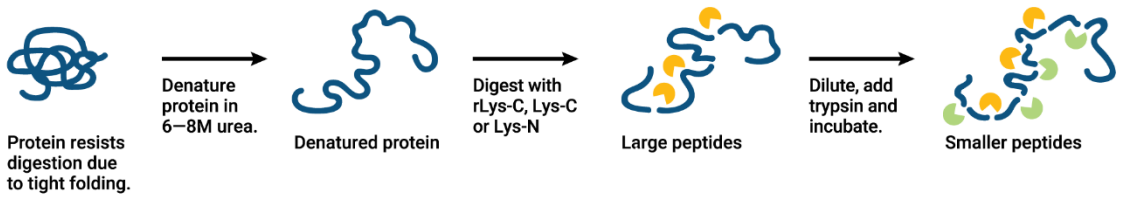



**Figure 13** | Two-step or tandem protein digestion using strong denaturing conditions for the initial digestion with LysC, followed by trypsin digestion in diluted denaturant. *Adopted from Promega Corporation.*

Compared to all other microbial guilds, methanotrophs possess a very special cell architecture characterized by an abundant membrane fraction (ICMs). This makes the proteomics of methanotrophs highly challenging. The major challenge is the high amount of membrane-associated proteins that need to be efficiently solubilized and digested for downstream analysis. Therefore, while the field of membrane proteomics has changed dramatically over the last years, only very few proteomics studies have been conducted on methanotrophs. The few existing studies focused on differential expression proteomics of the well-studied *Methylococcus capsulatus* Bath (type Ib methanotroph (Kao et al., 2004; Berven et al., 2006)) and *Methylocella silvestris* (type IIb methanotroph (Patel et al., 2009; Patel et al., 2012)).

Proteome studies on *M. capsulatus* Bath focused on the impact of copper concentration on its cellular metabolism. Copper is known to induce a switch between

the expression of sMMO and pMMO (Semrau et al., 2010). Out of the 682 proteins identified, 60 proteins were up-regulated and 68 proteins were down-regulated by at least 2-fold or more in response to the changes in copper concentration. Among up-regulated proteins, the proteins of methane and carbohydrate metabolisms were identified (Kao et al., 2004).

The proteome of *M. silvestris* has been characterized in response to growth on two different substrates (methane vs. acetate). The key enzymes of a methanotrophic lifestyle, sMMO and methanol dehydrogenase, were significantly down-regulated during the growth on acetate (Patel et al., 2009). However, the proteomics approaches applied to *M. silvestris* cannot be generalized for type II methanotrophs. In contrast to type IIa methanotrophs (family *Methylocystaceae*) and *Methylocapsa* spp., *M. silvestris* (family *Beijerinckiaceae*) possesses only sMMO but lacks ICMs and pMMO (Dunfield et al., 2003; Chen et al., 2010).

## 1.4. Our model organism *Methylocystis* sp. strain SC2

Among the well-described type I and type II methanotrophs, *Methylocystis* spp. are known to be the most oligotrophic. They represent the ecologically most relevant methanotrophs in terrestrial environments (Dedysh et al., 2003; Knief et al., 2003; Chen et al., 2007; Nauer et al., 2012). Our model organism, *Methylocystis sp.* strain SC2, is a member of type IIa methanotrophs. The strain was isolated from a polluted aquifer and identified based on morphology and 16S rRNA gene phylogeny (Dunfield et al., 2002). It is a Gram-negative, non-motile, rod-shaped bacterium, with the ability to form cysts.

Strain SC2 is relatively well-characterized based on modern molecular techniques. The complete genome of strain SC2 was obtained by whole-genome shotgun sequencing (Dam et al., 2012a). It contains a circular chromosome of 3.7 Mbp and two plasmids of 229,614 bp (pBSC2-1) and 143,536 bp (pBSC2-2) (**Figure 14A**). All genes required for a methanotrophic lifestyle were identified. It was shown to possess all the genes encoding pMMO, but not sMMO. In addition to two copies of the *pmoCAB1* gene cluster encoding “low-affinity” pMMO1 enzyme, the genome of strain SC2 contains a single *pmoCAB2* gene cluster, encoding “high-affinity” pMMO2 enzyme. Interestingly, the *pmoCAB1* and *pmoCAB2* gene clusters share only a low degree of identity on both the nucleotide (67.4–70.9%) and derived amino acid (59.3–65.6%)

sequence levels (Dunfield et al., 2002; Dam et al., 2012a). In addition, three monocistronic *pmoC* paralogs were identified, with one present on plasmid pBSC2-2 (*pmoC3*$_{Ps}$) (**Figure 14A**). The exact function of *pmoC3*$_{Ps}$ is not yet known (Dam et al., 2013).

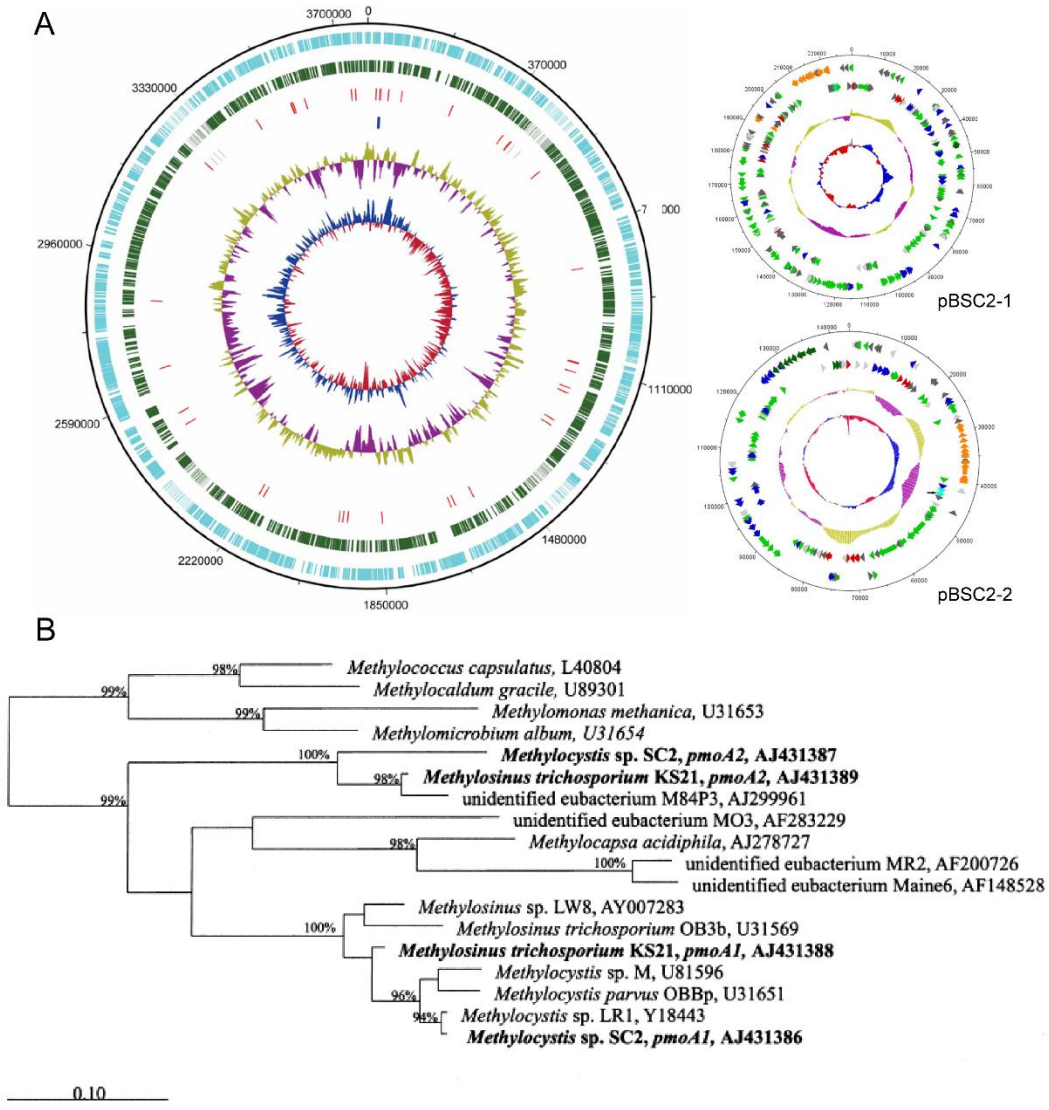



**Figure 14** | (**A**) Circular representation of the gap-closed genome of strain SC2 with two plasmids, pBSC2-1 and pBSC2-2. Open reading frames (ORFs) are colored according to their function: red, plasmid replication and maintenance; orange, conjugation; light green, ORFs having homology to some functional proteins; dark green, nitrogen metabolism; blue, transposase-like proteins; light gray, hypothetical proteins; dark gray, conserved hypothetical proteins. The singleton *pmoC* (marked with an arrow) and its upstream regulatory gene in pBSC2-2 are colored sky blue. *Adopted from (Dam et al., 2012b)* **(B)** Phylogenetic tree showing the relationship of the derived amino acid sequences of the two partial *pmoA* fragments of *Methylocystis* sp. SC2 to those of selected type II and type I methanotrophs. The tree was constructed using a neighbour-joining algorithm with a Kimura correction. Bootstrap values >80% are indicated. The scale bar represents 0.1 change per amino acid position. *Adopted from (Dunfield et al., 2002).*

To study the methane oxidation kinetics of the pMMO1 and pMMO2 isozymes, a suicide vector was introduced to strain SC2 targeting either *pmoCAB1* or *pmoCAB2* (**Figure 14B**) (Baani and Liesack, 2008). The results showed that the expression of pMMO1 takes place at methane concentrations >600 ppmv. In contrast, pMMO2 was constitutively expressed, capable of oxidizing $CH_4$ even at atmospheric concentrations of 1.75 ppmv. In addition, pMMO2 allowed strain SC2 to maintain growth at methane mixing ratios between 10 and 100 ppmv (Baani and Liesack, 2008).

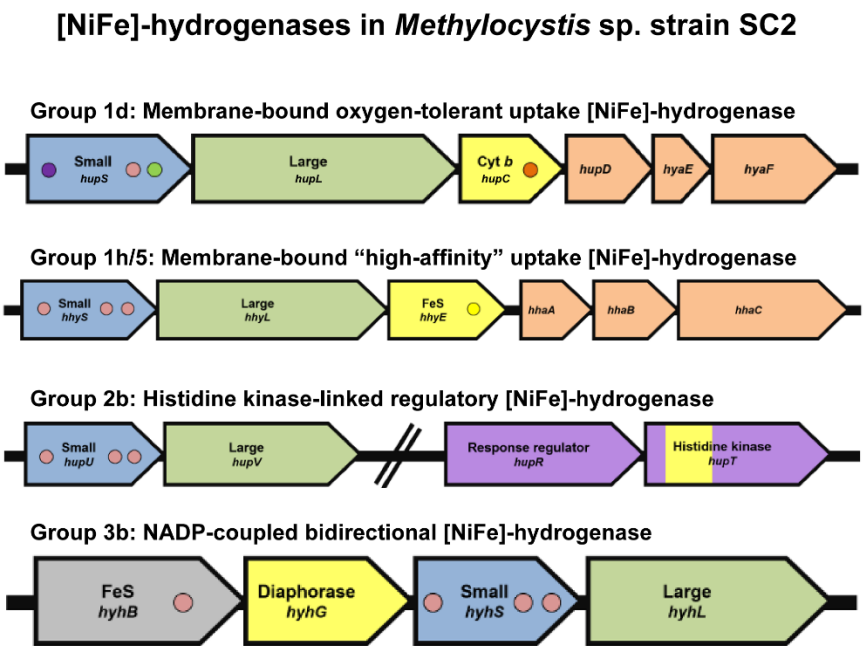



**Figure 15** | Genetic organization of four [NiFe]-hydrogenases in *Methylocystis* sp. strain SC2. Genes/ domains are color-coded as follows: green = catalytic site; blue = small subunit; yellow = electron acceptor or donor; red = redox subunit; light orange = maturation factor; dark orange = ion-translocation module; purple = regulatory module; grey = conserved hypothetical. Redox-active centers are shown in circles, where: orange = heme; red = [4Fe4S] cluster; yellow = [2Fe2S] cluster; green = [3Fe4S] cluster; purple = [4Fe3S] cluster. Genes are named according to nomenclature if previously defined. *Adopted from (Greening et al., 2016)*

Similar to other type IIa alphaproteobacterial methanotrophs, strain SC2 possesses genes involved in the serine pathway of formaldehyde assimilation. In addition, genes encoding complete TCA cycle and poly-beta-hydroxybutyrate (PHB) metabolism were identified (Dam et al., 2013). Interestingly, strain SC2 was shown to produce the maximum amount of PHB among five tested *Methylocystis* members and the third highest among all alphaproteobacterial methanotrophs (Pieja et al., 2011).

Furthermore, comparative genomics revealed a large inventory of genes involved in nitrogen metabolism. This included $N_2$ fixation, ammonium transport, assimilatory nitrate/nitrite reduction, hydroxylamine detoxification, and denitrification (Dam et al., 2013). Strain SC2 possesses a complete denitrification pathway, encoded by plasmid-borne *nosRZDFYX* operon (Dam et al., 2013). The genome sequence of strain SC2 provides evidence for its ability to thrive in environments with varying methane or nitrogen availability.

In addition, the genome of *Methylocystis* sp. strain SC2 was shown to possess four different groups of [NiFe]-hydrogenases: Group 1d, Group 1h/5, Group 2b, and Group 3b (Carere et al., 2017). Thus, strain SC2 has a well-established genetic inventory to utilize $H_2$ as an alternative energy source for aerobic respiration (**Figure 15**).

Based on the complete genome sequence of strain SC2, two comprehensive transcriptomics studies were conducted. These provided first insights into the regulation of gene expression in strain SC2. The first study focused on the differential expression of genes involved in central methane and nitrogen metabolisms under high-ammonium exposure (Dam et al., 2014). Here, RNA-Seq was used to identify strain SC2 genes that were differentially expressed in response to standard (10 mM) and high (30 mM) ammonia concentrations in the medium. Interestingly, upon high-ammonium load, *pmoCAB2* was significantly down-regulated, while the expression of *pmoCAB1* remained unaffected. Given that *Methylocystis* spp. contribute to atmospheric methane oxidation in upland soils, the differential expression of *pmoCAB2* explains, at least to some extent, the strong inhibitory effect of ammonium fertilizers on this activity. Among nitrogen metabolism-related pathways, genes involved in hydroxylamine detoxification (*haoAB*) were highly up-regulated, while those for assimilatory nitrate/nitrite reduction, high-affinity ammonium uptake and nitrogen regulatory protein PII were down-regulated (Dam et al., 2014).

The second, more recent study revealed the adaptation mechanisms of strain SC2 to salt-induced osmotic stress (Han et al., 2017). When strain SC2 was challenged with 0.75% NaCl, 301 genes were differetially expressed, with sigma factor $\sigma^{32}$ being a major controllling factor of the transcriptional stress response. The transcript levels of nearly all the genes involved in methane oxidation remained unaffected. Nonetheless, chromosomal genes involved in stress reponse (stress-induced protein, DNA-binding

protein from starved cells, and CsbD family protein) were significantly up-regulated. Amino acid profiling showed that glutamate was accumulated upon salt stress, acting as an osmoprotectant. In summary, the results revealed that strain SC2 is able to cope with salt stress, but only in a narrow range between 0.5% to 1.0 NaCl (Han et al., 2017).

Given the prevalence of *Methylocystis* spp. in both lowland and upland soils and the whole-genome transcriptome research already conducted on strain SC2, it represents an excellent model organism to study the molecular mechanisms involved in methane oxidation under a broad range of environmental factors.

## 1.5. Objectives of the study

Over the last decades, comprehensive research on the diversity, physiology, and ecology of methane-oxidizing bacteria has been published. The great interest in methanotrophic bacteria is primarily due to their ability to mitigate the greenhouse effect of methane by acting either as a biofilter for methane in methanogenic environments or as a sink for atmospheric methane in upland soils. However, methane is also a low-cost natural resource with great potential for applications to produce biopolymers, liquid value-added products, and high-quality feed protein. Despite the fact that methanotrophs have been studied for decades, there still exist major gaps in our knowledge of this important bacterial group; in particular with regard to their molecular biology and, as a consequence, their use in biotechnological applications. Primary reason is the lack of time- and cost-efficient molecular and metabolic engineering techniques for this group of organisms. Our model organism, *Methylocystis* sp. strain SC2, is one of the well-characterized members of the alphaproteobacterial type II methanotrophs, including a finished genome sequence. Therefore, it represents a suitable model organism to address questions related to the molecular biology of this bacterial group; including those with a potential for metabolic engineering and biotechnology.

### *Chapter 2. Crude-MS strategy for in-depth proteome analysis of the methane-oxidizing Methylocystis sp. strain SC2.*

To date, very few proteomics studies have been conducted on methanotrophs (details in 1.3.1). The major challenge for proteomics studies of type II methanotrophs is the

high amount of membrane-associated proteins that need to be efficiently solubilized and digested for downstream analysis. Therefore, we aimed to build on recent developments in sample preparation strategies and improvement in MS instrumentation to establish an efficient analytical proteomics workflow for the environmentally highly important members of the family *Methylocystaceae*, with *Methylocystis* sp. strain SC2 as our model organism.

**Our research questions**: *Could the optimized proteomics workflow increase the protein quantification accuracy and proteome coverage of strain SC2? In particular, will the SC2-specific proteomics workflow increase both identification and coverage of the membrane sub-proteome essential for methanotrophic lifestyle?*

### *Chapter 3. Efficient tandem LysC/Trypsin digestion in detergent conditions.*

In proteomics workflows, the efficient tryptic cleavage is characterized by the generation of the maximum number of most intense fully cleaved peptides, while keeping the miscleavage events on a minimal level. Previous studies disclosed that LysC/trypsin tandem digestion applied in a differential denaturation environment of chaotropes (e.g., urea), significantly increases the efficiency of proteolysis. Nonetheless, detergents used in in-solution digestion proteomics workflows resulted in less solubilization bias and protein-modifying properties than urea.

**Our research questions:** *Is the advantageous efficiency of LysC/trypsin tandem digestion in MS-compatible detergent environments similarly high as in chaotropic environments? Are the beneficial effects of tandem digestion caused by different target specificities of LysC and trypsin or by the digestion environment?*

### *Chapter 4. Hydrogen utilization by Methylocystis sp. strain SC2 expands the known metabolic versatility of Type IIa methanotrophs.*

Contrary to the long-held paradigm that methanotrophs obligately utilize methane as the only source of carbon and energy, recent pure culture studies have provided evidence that alphapoteobacterial type II methanotrophs possess a facultative metabolism. Several *Methylocystis* spp. have been shown to grow on simple organic acids, alcohols, and short-chain alkane gases. More recently, alphaproteobacterial

methanotrophs, including strain SC2, have also been shown to contain genes encoding different types of [NiFe]-hydrogenases.

**Our research questions:** *Is strain SC2 able to use hydrogen as an alternative energy source to optimize its biomass yield by mixed utilization of $CH_4$ and $H_2$ relative to the utilization of only $CH_4$ as an energy source? Is thereby $CH_4$ primarily used for synthesis of cell carbon and increased biomass/protein yield? Which $CH_4/O_2$ ratios may maximize the effect of hydrogen addition on the biomass yield and proteome response of strain SC2?*

### *Chapter 5. Genome scale metabolic modeling reveals the metabolic potential of three Type II methanotrophs of the genus Methylocystis.*

Methanotrophs are attracting a growing attention as potential cell factories, due to their ability to use methane as a carbon and energy source. Methane is a greenhouse gas and its utilization, as a feedstock would constitute a carbon sink. The recognition of the methane oxidation mechanism and the C1 assimilation pathways may give the information about the steps in the carbon metabolism that can be potential targets for metabolic engineering and biotechnological applications

**Our research questions:** *Is it possible to reveal the mechanisms of methane oxidation in well-studied Methylocystis species based on the Genome Scale Metabolic Models (GSMMs)? Could GSMMs predict the target metabolic pathways and culture conditions for effective metabolic engineering applications of these methanotrophs?*

## 1.6. References


Aebersold, R., and Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. *Nature* 537(7620), 347-355.

Allen, G. (2016). Biogeochemistry: Rebalancing the global methane budget. *Nature* 538(7623), 46-48.

AlSayed, A., Fergala, A., and Eldyasti, A. (2018). Sustainable biogas mitigation and value-added resources recovery using methanotrophs intergrated into wastewater treatment plants. *Rev Environ Sci Bio* 17(2), 351-393.

Anthony, C., and Williams, P. (2003). The structure and mechanism of methanol dehydrogenase. *Biochim Biophys Acta* 1647(1-2), 18-23.

Asara, J.M., Christofk, H.R., Freimark, L.M., and Cantley, L.C. (2008). A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. *Proteomics* 8(5), 994-999.

Baani, M., and Liesack, W. (2008). Two isozymes of particulate methane monooxygenase with different methane oxidation kinetics are found in *Methylocystis* sp strain SC2. *Proc Natl Acad Sci USA* 105(29), 10203-10208.

Bailey, D.J., McDevitt, M.T., Westphall, M.S., Pagliarini, D.J., and Coon, J.J. (2014). Intelligent data acquisition blends targeted and discovery methods. *J Proteome Res* 13(4), 2152-2161.

Baldrian, P., and Lopez-Mondejar, R. (2014). Microbial genomics, transcriptomics and proteomics: new discoveries in decomposition research using complementary methods. *Appl Microbiol Biotechnol* 98(4), 1531-1537.

Bantscheff, M., Lemeer, S., Savitski, M.M., and Kuster, B. (2012). Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal Bioanal Chem* 404(4), 939-965.

Bastien, C., Machlin, S., Zhang, Y., Donaldson, K., and Hanson, R.S. (1989). Organization of genes required for the oxidation of methanol to formaldehyde in three type II methylotrophs. *Appl Environ Microbiol* 55(12), 3124-3130.

Belova, S.E., Baani, M., Suzina, N.E., Bodelier, P.L.E., Liesack, W., and Dedysh, S.N. (2011). Acetate utilization as a survival strategy of peat-inhabiting *Methylocystis* spp. *Environ Microbiol Rep* 3(1), 36-46.

Berven, F.S., Karlsen, O.A., Straume, A.H., Flikka, K., Murrell, J.C., Fjellbirkeland, A., et al. (2006). Analysing the outer membrane subproteome of *Methylococcus capsulatus* (Bath) using proteomics and novel biocomputing tools. *Arch Microbiol* 184(6), 362-377.

Bin Goh, W.W., Guo, T., Aebersold, R., and Wong, L. (2015). Quantitative proteomics signature profiling based on network contextualization. *Biol Direct* 10, 71.

Blees, J., Niemann, H., Wenk, C.B., Zopfi, J., Schubert, C.J., Kirf, M.K., et al. (2014). Micro-aerobic bacterial methane oxidation in the chemocline and anoxic water column of deep south-Alpine Lake Lugano (Switzerland). *Limnol Oceanogr* 59(2), 311-324.

Blom, J., Albaum, S.P., Doppmeier, D., Puhler, A., Vorholter, F.J., Zakrzewski, M., et al. (2009). EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinf* 10, 154.

Boetius, A., Ravenschlag, K., Schubert, C.J., Rickert, D., Widdel, F., Gieseke, A., et al. (2000). A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* 407(6804), 623-626.

Borras, E., and Sabido, E. (2017). What is targeted proteomics? A concise revision of targeted acquisition and targeted data analysis in mass spectrometry. *Proteomics* 17(17-18).

Bowman, J. (2006). "The Methanotrophs—The Families *Methylococcaceae* and *Methylocystaceae*," in *The Prokaryotes: A Handbook on the Biology of Bacteria,* ed. M. Dworkin. (New York, USA: Springer Science+Business Media, LLC), 282-283.

Bowman, J.P., Sly, L.I., Nichols, P.D., and Hayward, A.C. (1993). Revised taxonomy of the methanotrophs - description of *Methylobacter* gen. nov., emendation of *Methylococcus*, validation of *Methylosinus* and *Methylocystis* species, and a proposal that the family *Methylococcaceae* includes only the group I methanotrophs. *Int J Syst Bacteriol* 43(4), 735-753.

Cai, Y., Zheng, Y., Bodelier, P.L., Conrad, R., and Jia, Z. (2016). Conventional methanotrophs are responsible for atmospheric methane oxidation in paddy soils. *Nat Commun* 7, 11728.

Carere, C.R., Hards, K., Houghton, K.M., Power, J.F., McDonald, B., Collet, C., et al. (2017). Mixotrophy drives niche expansion of verrucomicrobial methanotrophs. *ISME J* 11(11), 2599-2610.

Caron, E., Kowalewski, D.J., Koh, C.C., Sturm, T., Schuster, H., and Aebersold, R. (2015). Analysis of major histocompatibility complex (MHC) immunopeptidomes using mass spectrometry. *Mol Cell Proteomics* 14(12), 3105-3117.

Chen, Y., Crombie, A., Rahman, M.T., Dedysh, S.N., Liesack, W., Stott, M.B., et al. (2010). Complete genome sequence of the aerobic facultative methanotroph *Methylocella silvestris* BL2. *J Bacteriol* 192(14)**,** 3840-3841.

Chen, Y., Dumont, M.G., Cebron, A., and Murrell, J.C. (2007). Identification of active methanotrophs in a landfill cover soil through detection of expression of 16S rRNA and functional genes. *Environ Microbiol* 9(11)**,** 2855-2869.

Chen, Y.P., and Yoch, D.C. (1987). Regulation of two nickel-requiring (inducible and constitutive) hydrogenases and their coupling to nitrogenase in *Methylosinus trichosporium* OB3b. *J Bacteriol* 169(10)**,** 4778-4783.

Chistoserdova, L. (2015). Methylotrophs in natural habitats: current insights through metagenomics. *Appl Microbiol Biotechnol* 99(14)**,** 5763-5779.

Cicerone, R.J., and Oremland, R.S. (1988). Biogeochemical aspects of atmospheric methane. *Global Biogeochem Cycles* 2(4)**,** 299-327.

Conrad, R. (1996). Soil microorganisms as controllers of atmospheric trace gases ($H_2$, CO, $CH_4$, OCS, $N_2O$, and NO). *Microbiol Rev* 60(4)**,** 609-640.

Costello, A.M., Auman, A.J., Macalady, J.L., Scow, K.M., and Lidstrom, M.E. (2002). Estimation of methanotroph abundance in a freshwater lake sediment. *Environ Microbiol* 4(8)**,** 443-450.

Cox, J., and Mann, M. (2011). Quantitative, high-resolution proteomics for data-driven systems biology. *Annu Rev Biochem* 80**,** 273-299.

Cubasch, U., Wuebbles, D., Chen, D.L., Facchini, M.C., Frame, D., Mahowald, N., et al. (2014). Climate Change 2013; The Physical Science Basis Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Introduction. *Climate Change 2013: The Physical Science Basis***,** 119-158.

Culpepper, M.A., and Rosenzweig, A.C. (2012). Architecture and active site of particulate methane monooxygenase. *Crit Rev Biochem Mol Biol* 47(6)**,** 483-492.

Dam, B., Dam, S., Blom, J., and Liesack, W. (2013). Genome analysis coupled with physiological studies reveals a diverse nitrogen metabolism in *Methylocystis* sp. strain SC2. *PLoS One* 8(10)**,** e74767.

Dam, B., Dam, S., Kim, Y., and Liesack, W. (2014). Ammonium induces differential expression of methane and nitrogen metabolism-related genes in *Methylocystis* sp. strain SC2. *Environ Microbiol* 16(10)**,** 3115-3127.

Dam, B., Dam, S., Kube, M., Reinhardt, R., and Liesack, W. (2012a). Complete genome sequence of *Methylocystis* sp. strain SC2, an aerobic methanotroph with high-affinity methane oxidation potential. *J Bacteriol* 194(21)**,** 6008-6009.

Dam, B., Kube, M., Dam, S., Reinhardt, R., and Liesack, W. (2012b). Complete sequence analysis of two methanotroph-specific *repABC*-containing plasmids from *Methylocystis* sp. strain SC2. *Appl Environ Microbiol* 78(12)**,** 4373-4379.

Danilova, O.V., Suzina, N.E., Van De Kamp, J., Svenning, M.M., Bodrossy, L., and Dedysh, S.N. (2016). A new cell morphotype among methane oxidizers: a spiral-shaped obligately microaerophilic methanotroph from northern low-oxygen environments. *ISME J* 10(11)**,** 2734-2743.

Dedysh, S.N., and Dunfield, P.F. (2011). Facultative and obligate methanotrophs how to identify and differentiate them. *Methods Enzymol* 495**,** 31-44.

Dedysh, S.N., Dunfield, P.F., Derakshani, M., Stubner, S., Heyer, J., and Liesack, W. (2003). Differential detection of type II methanotrophic bacteria in acidic peatlands using newly developed 16S rRNA-targeted fluorescent oligonucleotide probes. *FEMS Microbiol Ecol* 43(3)**,** 299-308.

Dedysh, S.N., Khmelenina, V.N., Suzina, N.E., Trotsenko, Y.A., Semrau, J.D., Liesack, W., et al. (2002). *Methylocapsa acidiphila* gen. nov., sp. nov., a novel methane-oxidizing and dinitrogen-fixing acidophilic bacterium from Sphagnum bog. *Int J Syst Evol Microbiol* 52(1)**,** 251-261.

Dedysh, S.N., and Knief, C. (2018). "Diversity and Phylogeny of Described Aerobic Methanotrophs," in *Methane Biocatalysis: Paving the Way to Sustainability,* eds. M.G. Kalyuzhnaya & X.-H. Xing. (Cham: Springer International Publishing), 17-42.

Dedysh, S.N., Knief, C., and Dunfield, P.F. (2005). *Methylocella* species are facultatively methanotrophic. *J Bacteriol* 187(13)**,** 4665-4670.

Dedysh, S.N., Liesack, W., Khmelenina, V.N., Suzina, N.E., Trotsenko, Y.A., Semrau, J.D., et al. (2000). *Methylocella palustris* gen. nov., sp. nov., a new methane-oxidizing acidophilic bacterium from peat bogs, representing a novel subtype of serine-pathway methanotrophs. *Int J Syst Evol Microbiol* 50 (3)**,** 955-969.

del Cerro, C., Garcia, J.M., Rojas, A., Tortajada, M., Ramon, D., Galan, B., et al. (2012). Genome sequence of the methanotrophic poly-beta-hydroxybutyrate producer *Methylocystis parvus* OBBP. *J Bacteriol* 194(20), 5709-5 710.

Deng, Y., Cui, X., Luke, C., and Dumont, M.G. (2013). Aerobic methanotroph diversity in Riganqiao peatlands on the Qinghai-Tibetan Plateau. *Environ Microbiol Rep* 5(4), 566-574.

Dlugokencky, E.J., Nisbet, E.G., Fisher, R., and Lowry, D. (2011). Global atmospheric methane: budget, changes and dangers. *Philos Trans A Math Phys Eng Sci* 369(1943), 2058-2072.

Domon, B., and Aebersold, R. (2010). Options and considerations when selecting a quantitative proteomics strategy. *Nat Biotechnol* 28(7), 710-721.

Dormeyer, W., van Hoof, D., Braam, S.R., Heck, A.J.R., Mummery, C.L., and Krijgsveld, J. (2008). Plasma membrane proteomics of human embryonic stem cells and human embryonal carcinoma cells. *J Proteome Res* 7(7), 2936-2951.

Dumont, M.G., Luke, C., Deng, Y.C., and Frenzel, P. (2014). Classification of *pmoA* amplicon pyrosequences using BLAST and the lowest common ancestor method in MEGAN. *Front Microbiol* 5.

Dunfield, P.F. (2007). The Soil Methane Sink. *Greenhouse Gas Sinks*, 152-170.

Dunfield, P.F., Belova, S.E., Vorob'ev, A.V., Cornish, S.L., and Dedysh, S.N. (2010). *Methylocapsa aurea* sp. nov., a facultative methanotroph possessing a particulate methane monooxygenase, and emended description of the genus *Methylocapsa*. *Int J Syst Evol Microbiol* 60, 2659-2664.

Dunfield, P.F., Khmelenina, V.N., Suzina, N.E., Trotsenko, Y.A., and Dedysh, S.N. (2003). *Methylocella silvestris* sp. nov., a novel methanotroph isolated from an acidic forest cambisol. *Int J Syst Evol Microbiol* 53(5), 1231-1239.

Dunfield, P.F., Yimga, M.T., Dedysh, S.N., Berger, U., Liesack, W., and Heyer, J. (2002). Isolation of a *Methylocystis* strain containing a novel *pmoA*-like gene. *FEMS Microbiol Ecol* 41(1), 17-26.

Dunfield, P.F., Yuryev, A., Senin, P., Smirnova, A.V., Stott, M.B., Hou, S., et al. (2007). Methane oxidation by an extremely acidophilic bacterium of the phylum *Verrucomicrobia*. *Nature* 450(7171), 879-882.

Ettwig, K.F., Butler, M.K., Le Paslier, D., Pelletier, E., Mangenot, S., Kuypers, M.M., et al. (2010). Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* 464(7288), 543-548.

Ettwig, K.F., Zhu, B., Speth, D., Keltjens, J.T., Jetten, M.S.M., and Kartal, B. (2016). Archaea catalyze iron-dependent anaerobic oxidation of methane. *Proc Natl Acad Sci USA* 113(45), 12792-12796.

Fu, L., Li, S.W., Ding, Z.W., Ding, J., Lu, Y.Z., and Zeng, R.J. (2016). Iron reduction in the DAMO/*Shewanella oneidensis* MR-1 coculture system and the fate of Fe(II). *Water Res* 88, 808-815.

Fujino, J., Nair, R., Kainuma, M., Masui, T., and Matsuoka, Y. (2006). Multi-gas mitigation analysis on stabilization scenarios using aim global model. *Energ J*, 343-353.

Gallien, S., Duriez, E., Crone, C., Kellmann, M., Moehring, T., and Domon, B. (2012). Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Mol Cell Proteomics* 11(12), 1709-1723.

Gallien, S., Duriez, E., Demeure, K., and Domon, B. (2013). Selectivity of LC-MS/MS analysis: implication for proteomics experiments. *J Proteomics* 81, 148-158.

Gillet, L.C., Navarro, P., Tate, S., Rost, H., Selevsek, N., Reiter, L., et al. (2012). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* 11(6), O111 016717.

Glatter, T., Ludwig, C., Ahrne, E., Aebersold, R., Heck, A.J., and Schmidt, A. (2012). Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion. *J Proteome Res* 11(11), 5145-5156.

Graumann, J., Hubner, N.C., Kim, J.B., Ko, K., Moser, M., Kumar, C., et al. (2008). Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol Cell Proteomics* 7(4), 672-683.

Greening, C., Biswas, A., Carere, C.R., Jackson, C.J., Taylor, M.C., Stott, M.B., et al. (2016). Genomic and metagenomic surveys of hydrogenase distribution indicate $H_2$ is a widely utilised energy source for microbial growth and survival. *ISME J* 10(3), 761-777.

Hakobyan, A., Liesack, W., and Glatter, T. (2018). Crude-MS strategy for in-depth proteome analysis of the methane-oxidizing *Methylocystis* sp. strain SC2. *J Proteome Res* 17(9), 3086-3103.

Han, D., Link, H., and Liesack, W. (2017). Response of *Methylocystis* sp. strain SC2 to salt stress: physiology, global transcriptome, and amino acid profiles. *Appl Environ Microbiol* 83(20), e00866-17.

Hanczar, T., Csaki, R., Bodrossy, L., Murrell, J.C., and Kovacs, K.L. (2002). Detection and localization of two hydrogenases in *Methylococcus capsulatus* (Bath) and their potential role in methane metabolism. *Arch Microbiol* 177(2)**,** 167-172.

Hanson, R.S., and Hanson, T.E. (1996). Methanotrophic bacteria. *Microbiol Rev* 60(2)**,** 439-471.

Haroon, M.F., Hu, S., Shi, Y., Imelfort, M., Keller, J., Hugenholtz, P., et al. (2013). Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature* 500(7464)**,** 567-570.

Hirayama, H., Abe, M., Miyazaki, M., Nunoura, T., Furushima, Y., Yamamoto, H., et al. (2014). *Methylomarinovum caldicuralii* gen. nov., sp. nov., a moderately thermophilic methanotroph isolated from a shallow submarine hydrothermal system, and proposal of the family *Methylothermaceae* fam. nov. *Int J Syst Evol Microbiol* 64(3)**,** 989-999.

Hoehler, T.M., Alperin, M.J., Albert, D.B., and Martens, C.S. (1994). Field and laboratory studies of methane oxidation in an anoxic marine sediment - evidence for a methanogen-sulfate reducer consortium. *Global Biogeochem Cycles* 8(4)**,** 451-463.

Hoehler, T.M., and Jorgensen, B.B. (2013). Microbial life under extreme energy limitation. *Nat Rev Microbiol* 11(2)**,** 83-94.

Hoopmann, M.R., Merrihew, G.E., von Haller, P.D., and MacCoss, M.J. (2009). Post analysis data acquisition for the iterative MS/MS sampling of proteomics mixtures. *J Proteome Res* 8(4)**,** 1870-1875.

Hou, S., Makarova, K.S., Saw, J.H., Senin, P., Ly, B.V., Zhou, Z., et al. (2008). Complete genome sequence of the extremely acidophilic methanotroph isolate V4, *Methylacidiphilum infernorum*, a representative of the bacterial phylum *Verrucomicrobia*. *Biol Direct* 3**,** 26.

Hwang, I.Y., Lee, S.H., Choi, Y.S., Park, S.J., Na, J.G., Chang, I.S., et al. (2014). Biocatalytic conversion of methane to methanol as a key step for development of methane-based biorefineries. *J Microbiol Biotechn* 24(12)**,** 1597-1605.

Hyman, M.R., and Wood, P.M. (1983). Methane oxidation by *Nitrosomonas europaea*. *Biochem J* 212(1)**,** 31-37.

Im, J., Lee, S.W., Yoon, S., Dispirito, A.A., and Semrau, J.D. (2011). Characterization of a novel facultative *Methylocystis* species capable of growth on methane, acetate and ethanol. *Environ Microbiol Rep* 3(2)**,** 174-181.

Intergovernmental Panel Climate Change, W.G.I.I.I. (2014). Climate Change 2014: Mitigation of Climate Change. *Climate Change 2014: Mitigation of Climate Change***,** 1-1435.

Islam, T., Jensen, S., Reigstad, L.J., Larsen, O., and Birkeland, N.K. (2008). Methane oxidation at 55 ◦C and pH 2 by a thermoacidophilic bacterium belonging to the *Verrucomicrobia* phylum. *Proc Natl Acad Sci USA* 105(1)**,** 300-304.

James, P. (1997). Protein identification in the post-genome era: the rapid rise of proteomics. *Q Rev Biophys* 30(4)**,** 279-331.

Jensen, O.N. (2006). Interpreting the protein language using proteomics. *Nat Rev Mol Cell Bio* 7(6)**,** 391-403.

Jones, R.D., and Morita, R.Y. (1983). Methane oxidation by *Nitrosococcus oceanus* and *Nitrosomonas europaea*. *Appl Environ Microbiol* 45(2)**,** 401-410.

Kall, L., and Vitek, O. (2011). Computational mass spectrometry-based proteomics. *PLoS Comput Biol* 7(12)**,** e1002277.

Kalyuzhnaya, M.G., Puri, A.W., and Lidstrom, M.E. (2015). Metabolic engineering in methanotrophic bacteria. *Metab Eng* 29**,** 142-152.

Kalyuzhnaya, M.G., Yang, S., Rozova, O.N., Smalley, N.E., Clubb, J., Lamb, A., et al. (2013). Highly efficient methane biocatalysis revealed in a methanotrophic bacterium. *Nat Commun* 4**,** 2785.

Kao, W.C., Chen, Y.R., Yi, E.C., Lee, H., Tian, Q., Wu, K.M., et al. (2004). Quantitative proteomic analysis of metabolic regulation by copper ions in *Methylococcus capsulatus* (Bath). *J Biol Chem* 279(49)**,** 51554-51560.

Khadem, A.F., Pol, A., Wieczorek, A., Mohammadi, S.S., Francoijs, K.J., Stunnenberg, H.G., et al. (2011). Autotrophic methanotrophy in verrucomicrobia: *Methylacidiphilum fumariolicum* SolV uses the Calvin-Benson-Bassham cycle for carbon dioxide fixation. *J Bacteriol* 193(17)**,** 4438-4446.

Khadem, A.F., Wieczorek, A.S., Pol, A., Vuilleumier, S., Harhangi, H.R., Dunfield, P.F., et al. (2012). Draft genome sequence of the volcano-inhabiting thermoacidophilic methanotroph *Methylacidiphilum fumariolicum* strain SolV. *J Bacteriol* 194(14)**,** 3729-3730.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J.G., Dlugokencky, E.J., et al. (2013). Three decades of global methane sources and sinks. *Nat Geosci* 6(10)**,** 813-823.

Kits, K.D., Kalyuzhnaya, M.G., Klotz, M.G., Jetten, M.S., Op den Camp, H.J., Vuilleumier, S., et al. (2013). Genome sequence of the obligate gammaproteobacterial methanotroph *Methylomicrobium album* strain BG8. *Genome Announc* 1(2), e0017013.

Knief, C. (2015). Diversity and habitat preferences of cultivated and uncultivated aerobic methanotrophic bacteria evaluated based on *pmoA* as molecular marker. *Front Microbiol* 6, 1346.

Knief, C., and Dunfield, P.F. (2005). Response and adaptation of different methanotrophic bacteria to low methane mixing ratios. *Environ Microbiol* 7(9), 1307-1317.

Knief, C., Lipski, A., and Dunfield, P.F. (2003). Diversity and activity of methanotrophic bacteria in different upland soils. *Appl Environ Microbiol* 69(11), 6703-6714.

Knittel, K., and Boetius, A. (2009). Anaerobic oxidation of methane: progress with an unknown process. *Annu Rev Microbiol* 63, 311-334.

Kolb, S., Knief, C., Dunfield, P.F., and Conrad, R. (2005). Abundance and activity of uncultured methanotrophic bacteria involved in the consumption of atmospheric methane in two forest soils. *Environ Microbiol* 7(8), 1150-1161.

Kresnowati, M.T.A.P., van Winden, W.A., Almering, M.J.H., ten Pierick, A., Ras, C., Knijnenburg, T.A., et al. (2006). When transcriptome meets metabolome: fast cellular responses of yeast to sudden relief of glucose limitation. *Mol Syst Biol* 2.

Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* 4, 222.

Lau, M.C., Stackhouse, B.T., Layton, A.C., Chauhan, A., Vishnivetskaya, T.A., Chourey, K., et al. (2015). An active atmospheric methane sink in high Arctic mineral cryosols. *ISME J* 9(8), 1904.

Leon, I.R., Schwammle, V., Jensen, O.N., and Sprenger, R.R. (2013). Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol Cell Proteomics* 12(10), 2992-3005.

Lieberman, R.L., and Rosenzweig, A.C. (2005). Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature* 434(7030), 177-182.

Ludwig, C., Gillet, L., Rosenberger, G., Amon, S., Collins, B.C., and Aebersold, R. (2018). Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol* 14(8), e8126.

Marx, V. (2013). Targeted proteomics. *Nat Methods* 10(1), 19-22.

McDonald, I.R., and Murrell, J.C. (1997). The particulate methane monooxygenase gene *pmoA* and its use as a functional gene probe for methanotrophs. *FEMS Microbiol Lett* 156(2), 205-210.

Mehaffy, C., Hess, A., Prenni, J.E., Mathema, B., Kreiswirth, B., and Dobos, K.M. (2010). Descriptive proteomic analysis shows protein variability between closely related clinical isolates of *Mycobacterium tuberculosis*. *Proteomics* 10(10), 1966-1984.

Meinshausen, M., Smith, S.J., Calvin, K., Daniel, J.S., Kainuma, M.L.T., Lamarque, J.F., et al. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change* 109(1-2), 213-241.

Michaelis, W., Seifert, R., Nauhaus, K., Treude, T., Thiel, V., Blumenberg, M., et al. (2002). Microbial reefs in the Black Sea fueled by anaerobic oxidation of methane. *Science* 297(5583), 1013-1015.

Milucka, J., Kirf, M., Lu, L., Krupke, A., Lam, P., Littmann, S., et al. (2015). Methane oxidation coupled to oxygenic photosynthesis in anoxic waters. *ISME J* 9(9), 1991-2002.

Molloy, M.P., and Witzmann, F.A. (2002). Proteomics: technologies and applications. *Brief Funct Genomic Proteomic* 1(1), 23-39.

Monteoliva, L., and Albar, J.P. (2004). Differential proteomics: an overview of gel and non-gel based approaches. *Brief Funct Genomic Proteomic* 3(3), 220-239.

Moore, S.M., Hess, S.M., and Jorgenson, J.W. (2016). Extraction, enrichment, solubilization, and digestion techniques for membrane proteomics. *J Proteome Research* 15(4), 1243-1252.

Morita, R.Y. (1999). Is $H_2$ the universal energy source for long-term survival? *Microb Ecol* 38(4), 307-320.

Murrell, J.C., Gilbert, B., and McDonald, I.R. (2000a). Molecular biology and regulation of methane monooxygenase. *Arch Microbiol* 173(5-6), 325-332.

Murrell, J.C., McDonald, I.R., and Gilbert, B. (2000b). Regulation of expression of methane monooxygenases by copper ions. *Trends Microbiol* 8(5), 221-225.

Myhre, G., Shindell, D., Breon, F.M., Collins, W., Fuglestvedt, J., Huang, J.P., et al. (2014). Anthropogenic and Natural Radiative Forcing. *Climate Change 2013: The Physical Science Basis*, 659-740.

Nauer, P.A., Dam, B., Liesack, W., Zeyer, J., and Schroth, M.H. (2012). Activity and diversity of methane-oxidizing bacteria in glacier forefields on siliceous and calcareous bedrock. *Biogeosciences* 9(6), 2259-2274.

Nielsen, P.A., Olsen, J.V., Podtelejnikov, A.V., Andersen, J.R., Mann, M., and Wisniewski, J.R. (2005). Proteomic mapping of brain plasma membrane proteins. *Mol Cell Proteomics* 4(4)**,** 402-408.

Nikolov, M., Schmidt, C., and Urlaub, H. (2012). Quantitative mass spectrometry-based proteomics: an overview. *Methods Mol Biol* 893**,** 85-100.

Op den Camp, H.J., Islam, T., Stott, M.B., Harhangi, H.R., Hynes, A., Schouten, S., et al. (2009). Environmental, genomic and taxonomic perspectives on methanotrophic *Verrucomicrobia*. *Environ Microbiol Rep* 1(5)**,** 293-306.

Oshkin, I.Y., Beck, D.A.C., Lamb, A.E., Tchesnokova, V., Benuska, G., McTaggart, T.L., et al. (2015). Methane-fed microbial microcosms show differential community dynamics and pinpoint taxa involved in communal response. *ISME J* 9(5)**,** 1119-1129.

Oswald, K., Milucka, J., Brand, A., Hach, P., Littmann, S., Wehrli, B., et al. (2016). Aerobic gammaproteobacterial methanotrophs mitigate methane emissions from oxic and anoxic lake waters. *Limnol Oceanogr* 61**,** S101-S118.

Pandey, A., and Mann, M. (2000). Proteomics to study genes and genomes. *Nature* 405(6788)**,** 837-846.

Patel, N.A., Crombie, A., Slade, S.E., Thalassinos, K., Hughes, C., Connolly, J.B., et al. (2012). Comparison of one- and two-dimensional liquid chromatography approaches in the label-free quantitative analysis of *Methylocella silvestris*. *J Proteome Res* 11(9)**,** 4755-4763.

Patel, V.J., Thalassinos, K., Slade, S.E., Connolly, J.B., Crombie, A., Murrell, J.C., et al. (2009). A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *J Proteome Res* 8(7)**,** 3752-3759.

Peterson, A.C., Russell, J.D., Bailey, D.J., Westphall, M.S., and Coon, J.J. (2012). Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics* 11(11)**,** 1475-1488.

Piche-Choquette, S., and Constant, P. (2019). Molecular hydrogen, a neglected key driver of soil biogeochemical processes. *Appl Environ Microbiol* 85(6).

Pieja, A.J., Rostkowski, K.H., and Criddle, C.S. (2011). Distribution and selection of poly-3-hydroxybutyrate production capacity in methanotrophic proteobacteria. *Microb Ecol* 62(3)**,** 564-573.

Pimenov, N.V., Kalmychkov, G.V., Veryasov, M.B., Sigalevich, P.A., and Zemskaya, T.I. (2014). Microbial oxidation of methane in the sediments of central and southern Baikal. *Microbiology* 83(6)**,** 773-781.

Poetsch, A., and Wolters, D. (2008). Bacterial membrane proteomics. *Proteomics* 8(19)**,** 4100-4122.

Pol, A., Heijmans, K., Harhangi, H.R., Tedesco, D., Jetten, M.S.M., and den Camp, H.J.M.O. (2007). Methanotrophy below pH 1 by a new *Verrucomicrobia* species. *Nature* 450(7171)**,** 874-U817.

Reeburgh, W.S. (1976). Methane consumption in Cariaco Trench waters and sediments. *Earth Planet Sc Lett* 28(3)**,** 337-344.

Riahi, K., Grubler, A., and Nakicenovic, N. (2007). Scenarios of long-term socio-economic and environmental development under climate stabilization. *Technol Forecast Soc* 74(7)**,** 887-935.

Ricke, P., Erkel, C., Kube, M., Reinhardt, R., and Liesack, W. (2004). Comparative analysis of the conventional and novel *pmo* (particulate methane monooxygenase) operons from *Methylocystis* strain SC2. *Appl Environ Microbiol* 70(5)**,** 3055-3063.

Rigby, M., Prinn, R.G., Fraser, P.J., Simmonds, P.G., Langenfelds, R.L., Huang, J., et al. (2008). Renewed growth of atmospheric methane. *Geophys Res Lett* 35(22).

Ronsein, G.E., Pamir, N., von Haller, P.D., Kim, D.S., Oda, M.N., Jarvik, G.P., et al. (2015). Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. *J Proteomics* 113**,** 388-399.

Ross, M.O., and Rosenzweig, A.C. (2017). A tale of two methane monooxygenases. *J Biol Inorg Chem* 22(2-3)**,** 307-319.

Saunois, M., Jackson, R.B., Bousquet, P., Poulter, B., and Canadell, J.G. (2016). The growing role of methane in anthropogenic climate change. *Environ Res Lett* 11(12).

Savvichev, A.S., Kokryatskaya, N.M., Zabelina, S.A., Rusanov, II, Zakharova, E.E., Veslopolova, E.F., et al. (2017). Microbial processes of the carbon and sulfur cycles in an ice-covered, iron-rich meromictic lake Svetloe (Arkhangelsk region, Russia). *Environ Microbiol* 19(2)**,** 659-672.

Schafer, C., Friedrich, B., and Lenz, O. (2013). Novel, oxygen-insensitive group 5 [NiFe]-hydrogenase in *Ralstonia eutropha*. *Appl Environ Microbiol* 79(17)**,** 5137-5145.

Schiffmann, C., Hansen, R., Baumann, S., Kublik, A., Nielsen, P.H., Adrian, L., et al. (2014). Comparison of targeted peptide quantification assays for reductive dehalogenases by selective reaction monitoring (SRM) and precursor reaction monitoring (PRM). *Anal Bioanal Chem* 406(1)**,** 283-291.

Semrau, J.D., DiSpirito, A.A., and Yoon, S. (2010). Methanotrophs and copper. *FEMS Microbiol Rev* 34(4), 496-531.

Shah, N.N., Hanna, M.L., Jackson, K.J., and Taylor, R.T. (1995). Batch cultivation of *Methylosinus trichosporium* OB3B: IV. Production of hydrogen-driven soluble or particulate methane monooxygenase activity. *Biotechnol Bioeng* 45(3), 229-238.

Sharp, C.E., Smirnova, A.V., Graham, J.M., Stott, M.B., Khadka, R., Moore, T.R., et al. (2014). Distribution and diversity of *Verrucomicrobia* methanotrophs in geothermal and acidic environments. *Environ Microbiol* 16(6), 1867-1878.

Shen, L.D., He, Z.F., Wu, H.S., and Gao, Z.Q. (2015). Nitrite-dependent anaerobic methane-oxidising bacteria: unique microorganisms with special properties. *Curr Microbiol* 70(4), 562-570.

Sirajuddin, S., and Rosenzweig, A.C. (2015). Enzymatic oxidation of methane. *Biochemistry* 54(14), 2283-2294.

Sohngen, N.L. (1907). Methan as carbon-food and source of energy for bacteria. *P K Akad Wet-Amsterd* 8, 327-331.

Stein, L.Y., Bringel, F., DiSpirito, A.A., Han, S., Jetten, M.S., Kalyuzhnaya, M.G., et al. (2011). Genome sequence of the methanotrophic alphaproteobacterium *Methylocystis* sp. strain Rockwell (ATCC 49242). *J Bacteriol* 193(10), 2668-2669.

Stein, L.Y., Yoon, S., Semrau, J.D., Dispirito, A.A., Crombie, A., Murrell, J.C., et al. (2010). Genome sequence of the obligate methanotroph *Methylosinus trichosporium* strain OB3b. *J Bacteriol* 192(24), 6497-6498.

Stephenson, M., and Stickland, L.H. (1931). Hydrogenase: a bacterial enzyme activating molecular hydrogen: The properties of the enzyme. *Biochem J* 25(1), 205-214.

Stoecker, K., Bendinger, B., Schoning, B., Nielsen, P.H., Nielsen, J.L., Baranyi, C., et al. (2006). Cohn's *Crenothrix* is a filamentous methane oxidizer with an unusual methane monooxygenase. *Proc Natl Acad Sci USA* 103(7), 2363-2367.

Thauer, R.K. (2011). Hydrogenases and the global $H_2$ cycle. *Eur J Inorg Chem* (7), 919-921.

Timmers, P.H., Suarez-Zuluaga, D.A., van Rossem, M., Diender, M., Stams, A.J., and Plugge, C.M. (2016). Anaerobic oxidation of methane associated with sulfate reduction in a natural freshwater gas source. *ISME J* 10(6), 1400-1412.

Timmers, P.H., Welte, C.U., Koehorst, J.J., Plugge, C.M., Jetten, M.S., and Stams, A.J. (2017). Reverse methanogenesis and respiration in methanotrophic archaea. *Archaea* 2017, 1654237.

Trotsenko, Y.A., and Murrell, J.C. (2008). Metabolic aspects of aerobic obligate methanotrophy. *Adv Appl Microbiol* 63, 183-229.

Tsou, C.C., Avtonomov, D., Larsen, B., Tucholska, M., Choi, H., Gingras, A.C., et al. (2015). DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods* 12(3), 258-264.

Turner, A.J., Frankenberg, C., and Kort, E.A. (2019). Interpreting contemporary trends in atmospheric methane. *Proc Natl Acad Sci USA* 116(8), 2805-2813.

Tveit, A.T., Hestnes, A.G., Robinson, S.L., Schintlmeister, A., Dedysh, S.N., Jehmlich, N., et al. (2019). Widespread soil bacterium that oxidizes atmospheric methane. *Proc Natl Acad Sci USA* 116(17), 8515-8524.

Valentine, D.L., and Reeburgh, W.S. (2000). New perspectives on anaerobic methane oxidation. *Environ Microbiol* 2(5), 477-484.

van Teeseling, M.C., Pol, A., Harhangi, H.R., van der Zwart, S., Jetten, M.S., Op den Camp, H.J., et al. (2014). Expanding the verrucomicrobial methanotrophic world: description of three novel species of *Methylacidimicrobium* gen. nov. *Appl Environ Microbiol* 80(21), 6782-6791.

van Vuuren, D.P., Den Elzen, M.G.J., Lucas, P.L., Eickhout, B., Strengers, B.J., van Ruijven, B., et al. (2007). Stabilizing greenhouse gas concentrations at low levels: an assessment of reduction strategies and costs. *Climatic Change* 81(2), 119-159.

Venable, J.D., Dong, M.Q., Wohlschlegel, J., Dillin, A., and Yates, J.R. (2004). Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat Methods* 1(1), 39-45.

Vigliotta, G., Nutricati, E., Carata, E., Tredici, S.M., De Stefano, M., Pontieri, P., et al. (2007). *Clonothrix fusca* Roze 1896, a filamentous, sheathed, methanotrophic gamma-proteobacterium. *Appl Environ Microbiol* 73(11), 3556-3565.

Vignais, P.M., and Billoud, B. (2007). Occurrence, classification, and biological function of hydrogenases: an overview. *Chem Rev* 107(10), 4206-4272.

Vit, O., and Petrak, J. (2017). Integral membrane proteins in proteomics. How to break open the black box? *J Proteomics* 153**,** 8-20.

Vorobev, A., Jagadevan, S., Jain, S., Anantharaman, K., Dick, G.J., Vuilleumier, S., et al. (2014). Genomic and transcriptomic analyses of the facultative methanotroph *Methylocystis* sp. strain SB2 grown on methane or ethanol. *Appl Environ Microbiol* 80(10)**,** 3044-3052.

Vorobev, A.V., Baani, M., Doronina, N.V., Brady, A.L., Liesack, W., Dunfield, P.F., et al. (2011). *Methyloferula stellata* gen. nov., sp. nov., an acidophilic, obligately methanotrophic bacterium that possesses only a soluble methane monooxygenase. *Int J Syst Evol Microbiol* 61(10)**,** 2456-2463.

Vuilleumier, S., Khmelenina, V.N., Bringel, F., Reshetnikov, A.S., Lajus, A., Mangenot, S., et al. (2012). Genome sequence of the haloalkaliphilic methanotrophic bacterium *Methylomicrobium alcaliphilum* 20Z. *J Bacteriol* 194(2)**,** 551-552.

Wahlen, M. (1993). The global methane cycle. *Annu Rev Earth Pl Sc* 21**,** 407-426.

Ward, N., Larsen, O., Sakwa, J., Bruseth, L., Khouri, H., Durkin, A.S., et al. (2004). Genomic insights into methanotrophy: The complete genome sequence of *Methylococcus capsulatus* (Bath). *PLoS Biol* 2(10)**,** 1616-1628.

Whittenbury, R., Phillips, K.C., and Wilkinso.Jf (1970). Enrichment, isolation and some properties of methane-utilizing bacteria. *J Gen Microbiol* 61**,** 205-18.

Wilkins, M.R., Pasquali, C., Appel, R.D., Ou, K., Golaz, O., Sanchez, J.C., et al. (1996). From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology* 14(1)**,** 61-65.

Wisniewski, J.R., and Mann, M. (2012). Consecutive proteolytic digestion in an enzyme reactor increases depth of proteomic and phosphoproteomic analysis. *Anal Chem* 84(6)**,** 2631-2637.

Wisniewski, J.R., Zougman, A., and Mann, M. (2009). Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. *J Proteome Res* 8(12)**,** 5674-5678.

Woese, C.R. (1987). Bacterial evolution. *Microbiol Rev* 51(2)**,** 221-271.

Wu, C.C., MacCoss, M.J., Howell, K.E., and Yates, J.R., 3rd (2003). A method for the comprehensive proteomic analysis of membrane proteins. *Nat Biotechnol* 21(5)**,** 532-538.

Yang, S., Matsen, J.B., Konopka, M., Green-Saxena, A., Clubb, J., Sadilek, M., et al. (2013). Global molecular analyses of methane metabolism in methanotrophic alphaproteobacterium, *Methylosinus trichosporium* OB3b. Part II. Metabolomics and $^{13}$C-Labeling Study. *Front Microbiol* 4**,** 70.

Yimga, M.T., Dunfield, P.F., Ricke, P., Heyer, H., and Liesack, W. (2003). Wide distribution of a novel *pmoA*-like gene copy among type II methanotrophs, and its expression in *Methylocystis* strain SC2. *Appl Environ Microbiol* 69(9)**,** 5593-5602.

Zehnder, A.J., and Brock, T.D. (1979). Methane formation and methane oxidation by methanogenic bacteria. *J Bacteriol* 137(1)**,** 420-432.

# Chapter 2:

## Crude-MS Strategy for in-Depth Proteome Analysis of the Methane - Oxidizing *Methylocystis* sp. strain SC2

*J. Proteome Res.* 2018, 17, 9, 3086-3103
https://doi.org/10.1021/acs.jproteome.8b00216

**Anna Hakobyan**[1], Werner Liesack[1, 3]*, Timo Glatter[2]*

*Corresponding authors

Chapter 2 is written in research manuscript style. It was published as research article in *Journal of Proteome Research* in July, 2018. My contribution to the chapter 2 involved the experimental design, laboratory experiments, the proteomics data analysis, the writing and the revision of corresponding parts of the manuscript, including the design of the figures.

[1]Research Group of Methanotrophic Bacteria and Environmental Genomics/Transcriptomics, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, D-35043 Marburg, Germany

[2]Core Facility for Mass-Spectrometry and Proteomics, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, D-35043 Marburg, Germany

[3]Center for Synthetic Microbiology (SYNMIKRO), Philipps-Universitaẗ Marburg, Karl-von-Frisch-Str. 16, D-35043 Marburg, Germany

# 2 Crude-MS Strategy for in-Depth Proteome Analysis of the Methane- Oxidizing *Methylocystis* sp. strain SC2

## 2.1. Abstract

*Methylocystis* sp. strain SC2 is a representative of the alphaproteobacterial methane oxidizers or type IIa methanotrophs. These microorganisms play a crucial role in methane cycling. Here, we developed an efficient analytical proteomics workflow for strain SC2. It tackles the major challenges related to the high amount of integral membrane proteins that need to be efficiently solubilized and digested for downstream analysis. Each step of the workflow, including cell lysis, protein solubilization and digestion, and MS peptide quantification, was assessed and optimized. Our new crude-lysate-MS approach proved to increase protein quantification accuracy and proteome coverage of strain SC2. It captured 62% of the predicted SC2 proteome, with up to 10-fold increase in membrane-associated proteins relative to less effective conditions. The use of crude cell lysate for downstream analysis showed to be highly efficient for SC2 and other members of the family *Methylocystaceae*. Using two contrasting nitrogen conditions, we further validated our workflow efficiency by analyzing the SC2 proteome for differentially expressed proteins involved in methane and nitrogen metabolism. Our crude-MS approach may be applied to a variety of proteomic workflows incorporating cell types with challenging solubilization properties. Data are available via ProteomeXchange with identifier PXD009027.

## 2.2. Introduction

Methane is known to be the second most significant greenhouse gas in the atmosphere with global warming potential ($GWP_{100}$) 28–36-times greater than $CO_2$ (Myhre et al., 2014; Lau et al., 2015). Microbial methane oxidation is one of the fundamental processes in global methane cycle (Urmann et al., 2009). In terrestrial systems, methane-oxidizing bacteria or methanotrophs are the only biological sink of methane from natural sources. They are widespread both in aerobic interfaces of methanogenic environments and in upland soils. Inhabiting oxic/anoxic soil interfaces, methanotrophs

reoxidize 20% to 90% of the methane produced in natural wetlands or submerged rice paddies by methanogenic archaea before its emission to the atmosphere. In upland soils, methanotrophs act as a sink for 30 Tg $year^{-1}$ of atmospheric methane (Conrad, 1996; Knief and Dunfield, 2005; Denman and Brasseur, 2007; Allen, 2016). They utilize methane as their sole source of energy (Hanson and Hanson, 1996). Historically, proteobacterial methane-oxidizing bacteria have been classified into type I and type II methanotrophs based on intracytoplasmic membranes (ICM) arranged in the cell, biochemical pathways of carbon fixation, the capability of nitrogen fixation, the formation of resting stages, and the phospholipid fatty acid composition (Whittenbury and Dalton, 1981). These bacteria oxidize methane to methanol via particulate methane monooxygenase (pMMO) and/or soluble methane monooxygenase (sMMO) enzymes, which are unique to methanotrophs (Murrell et al., 2000a; Murrell and Jetten, 2009). While sMMO is localized in the cytoplasm of methanotrophic cells, pMMO is a copper-containing enzyme that is associated with the ICM of methanotrophs (Murrell et al., 2000b; Semrau et al., 2010; Lombardi, 2015). Phylogenetic analyses of 16S rRNA gene sequences confirmed the initial classification into type I and type II methanotrophs. They were grouped into *Gammaproteobacteria* (type I) and *Alphaproteobacteria* (type II). More recent research, however, turned this initial concept of type I and type II methanotrophs into question. Besides phylogeny, the carbon fixation pathway remained the only major feature of the above-mentioned criteria that validly differentiates between type I and type II methanotrophs (Knief, 2015). As suggested by Knief (Knief, 2015), we therefore use these terms only as synonyms for the phylogenetic groups of *Gamma*- and *Alphaproteobacteria*. The methanotrophic *Alphaproteobacteria* were recently divided into type IIa (*Methylocystaceae*) and type IIb (*Beijerinckiaceae*) methanotrophs.

Among the well-described methanotrophs, *Methylocystis* spp. are known to be both most oligotrophic and most widespread in terrestrial environments (Dedysh et al., 2003; Knief et al., 2003; Chen et al., 2007; Nauer et al., 2012). Our model organism, *Methylocystis* sp. strain SC2 (further referred to as SC2), contains two pMMO isozymes (pMMO1, pMMO2) with different methane oxidation kinetics (Dunfield et al., 2002; Ricke et al., 2004; Baani and Liesack, 2008). The conventional pMMO1 enzyme is active and oxidizes only elevated levels of methane as they occur in the aerobic interfaces of

methanogenic environments. By contrast, the high-affinity pMMO2 allows SC2 to oxidize methane even at the atmospheric trace level of 1.8 ppm of $CH_4$. Given that *Methylocystis* spp. are prevalent in both lowland and upland soils, SC2 represents an excellent model organism for studying the molecular mechanisms involved in methane oxidation under a broad range of methane mixing ratios. To date, all studies done with SC2 have been on transcriptome level. One of these studies focused on the differential expression of genes involved in central methane and nitrogen metabolisms under contrasting nitrogen sources (Dam et al., 2014). Another, more recent study revealed the adaptation mechanisms of SC2 to salt-induced osmotic stress (Han et al., 2017). Undoubtedly, transcriptome research increased our understanding of gene regulation in SC2. Nonetheless, it does not elucidate the structure, function, and expression of all translated proteins of this organism, which are the functional units that define the cellular phenotype.

For many years, mass spectrometry (MS)-based methods have been widely used as a standard approach for accurate mass determination and characterization of proteins and peptides in numerous biological studies (Walther and Mann, 2010; Cox and Mann, 2011; Altelaar and Heck, 2012).

The general shotgun MS proteomics workflow starts with the extraction of the total cellular protein pool. Common strategies incorporate chaotropic salts or strong solubilizing detergents that are easily removed from the sample after digestion. Alternatively, filter-aided sample preparation strategies (FASP) are being used for extraction and digestion (Manza et al., 2005; Wisniewski et al., 2009a). Both strategies facilitate MS compatibility due to the removal of the solubilizing detergent and adjustment of a fully compatible digestion environment. Thereby, the proteome is converted into peptides amenable for liquid chromatography–MS analysis (LC–MS).
With the improvement in MS instrumentation and development of sophisticated MS acquisition strategies along with efficient software solutions, proteomes can now be studied in-depth on a global level with a variety of labeling and label-free quantitative MS approaches for relative proteome comparisons (Domon and Aebersold, 2010; Rauniyar and Yates, 2014; Collins et al., 2017).

As an extension to relative sample comparisons, absolute protein quantification is being widely used in the community (Schwanhausser et al., 2011; Simpson and

Beynon, 2012; Zeiler et al., 2012; Picotti et al., 2013; Simicevic et al., 2013; Maass and Becher, 2016). Well-known strategies use isotopically labeled reference peptides or proteins as internal standards. The standards are incorporated into the workflow for quantification of individual proteins or even total proteomes (Picotti et al., 2013). Such global absolute quantification strategies are highly accurate but unfortunately very costly. Therefore, alternative strategies have been developed estimating protein abundances based on label-free quantification (LFQ) data. They assume a correlation between peptide precursor intensities and protein abundances (iBAQ or Top3) (Silva et al., 2006; Schwanhausser et al., 2011). These alternative approaches have gained much attention over recent years because such absolute abundance estimators can be readily extracted from total proteome measurements, yet absolute quantification studies may suffer from nonoptimal solubilization conditions (Glatter et al., 2015). Therefore, we anticipated that currently available sample preparation strategies, aiming at a universal solution, might have to be optimized for yielding better extraction efficiencies on specialized proteomes.

To date, very few proteomics studies have been conducted on methanotrophs. The only existing studies focused on differential expression proteomics of two well-studied methanotrophs, *Methylococcus capsulatus* Bath (type I, (Kao et al., 2004; Berven et al., 2006)) and *Methylocella silvestris* (type IIb, (Patel et al., 2009)). In contrast to type IIa methanotrophs (family *Methylocystaceae*), *M. silvestris* (family *Beijerinckiaceae*) possesses only sMMO but lacks ICM and pMMO (Dunfield et al., 2003; Chen et al., 2010). In this study, we aimed to build on recent developments in sample preparation strategies and improvement in MS instrumentation to establish an efficient analytical workflow for the environmentally highly important members of the family *Methylocystaceae*, with *Methylocystis* sp. strain SC2 as our model organism.

## 2.3. Materials and methods

### *2.3.1. Cell culture conditions*

For all steps of protocol optimization, *Methylocystis* sp. strain SC2 was cultivated in nitrate mineral salts (NMS) medium with the same basic composition as reported earlier (Heyer et al., 2002; Dam et al., 2014). As a nitrogen source, the medium was

supplemented with 1 g $KNO_3$ (~10 mM $NO_3^-$) per liter. Cells were grown to the mid log phase ($OD_{600}$ of 0.2–0.3) in 300 mL culture volume, with the initial $OD_{600}$ of 0.01 ± 0.003. The 1000 mL bottles (Schott Duran) were sealed with butyl rubber septum and plastic screw caps. The headspace was filled with filter-sterilized 20% methane (v/v) and the bottles were incubated on a rotary shaker at 130 rpm. The cultures were incubated at 25 °C for 48–72 h. They were transferred and grown to mid log phase in respective medium at least two times before the experimental treatment. Subsequently, the cells were collected by centrifugation (7000 × *g* for 20 min, 4 °C) and thoroughly washed twice with phosphate buffer (5.4 g $Na_2HPO_4 \times 7\ H_2O$ and 2.6 g $KH_2PO_4$ per liter distilled $H_2O$) to remove traces of the growth medium. To compare changes in the proteome in response to different nitrogen treatments, the cultures of SC2 were subjected to high ammonium level, in the same manner as performed previously (Dam et al., 2014). In brief, cells pregrown in NMS and washed with phosphate buffer were resuspended in an equivalent volume of medium containing 15 mM $(NH_4)_2SO_4$ (~30 mM $NH_4^+$) and then incubated for additional 10 h. Cell pellets (50 ± 10 mg) were collected in 2 mL sterile safe-lock microcentrifuge tubes (Eppendorf). At least three biological replicates were processed and used to extract and further analyze the total proteome of SC2.

*Methylocystis rosea* strain $SV97^T$ and *Methylosinus trichosporium* strain $OB3b^T$ were used as in-group references to compare the efficiency of our newly developed protein extraction method among close relatives. Both strains were incubated in NMS medium using the same culture conditions as specified in the beginning of this section. The widely used wild-type *Escherichia coli* strain K-12 substr. MG1655 was employed as an outgroup reference organism. The strain was cultivated overnight in Luria Broth medium, washed twice with phosphate buffered-saline, and pelleted as specified above. All cell pellets were stored at −80 °C until further use.

### *2.3.2. Sample preparation and protein solubilization*

Sample preparation followed previous studies (Leon et al., 2013; Glatter et al., 2015; Vit and Petrak, 2017) with some modifications. For protein solubilization, the frozen cell pellets were resuspended and sonicated (Hielscher Ultrasound Technology) in 2% sodium lauroyl sarcosinate (SLS), 2% sodium deoxycholate (SDC), 2% sodium dodecyl sulfate (SDS), or 8 M urea (all dissolved in 100 mM ammonium bicarbonate buffer

(following referred to as ABC buffer)) in the presence of 5 mM tris(2-carboxyethyl)phosphine. All urea-containing samples were incubated at 37 °C, whereas all other samples were incubated at 95 °C. To determine the optimal incubation time for protein solubilization, aliquots of 50 µL were collected from each sample after 15, 30, 60, and 90 min of incubation. The remaining samples were additionally sonicated after each aliquot collection step. After incubation, crude lysates were divided into two subsamples of equal volume. The first subsample was used to produce cleared lysate, by centrifuging for 10 min at 21 000 × *g*. The second subsample was maintained as crude lysate containing the cell debris. To test for further increase in protein solubilization, two other physical cell disruption techniques were examined. First, frozen cell pellets were resuspended in TE buffer (10 mM Tris-HCl, 1 mM $Na_2EDTA$, pH 7.4–8.0) containing protease inhibitor cocktail (cOmplete mini protease inhibitor cocktail; MERCK), followed by either bead beating or French press. Bead beating was performed for 40 s with 6.5 speed in FAST-PREP-24 (MP Biomedicals) bead beater using Lysing Matrix E tubes (MP Biomedicals) containing specialized beads. Alternatively, two rounds of French press (Thermo Electron Corporation) cell lysis were carried out under 1200 atm pressure, using cell pellets diluted in 10 mL TE buffer. Subsequently, protein extracts from both lysis procedures were concentrated with trichloroacetic acid/acetone protein precipitation procedure. Then, extracted proteins were resuspended in 2% SDC and processed as described in the beginning of this section. Regardless of the methods used for extraction and solubilization, all the protein samples were allowed to cool and were incubated with 10 mM iodoacetamide at 25 °C for 30 min. The concentration of proteins in each sample was measured with BCA protein assay kit (Thermo Fisher Scientific) following manufacturer's instructions.

### *2.3.3. Protein digestion*

For protein digestion step, 50 µg of total solubilized protein from each sample was used (modified from Glatter et al. (Glatter et al., 2012)). For in-solution digestion (ISD), the samples were diluted to 0.5% concentration of solubilizing (SDC, SLS) agent. The dilution was made using 100 mM ABC buffer. For FASP digestion, protein samples were mixed with 100 mM ABC buffer and transferred into the preconditioned spin filters (Microcon YM-30, Millipore). Following two additional wash/centrifugation cycles, protein samples

were finally suspended in 100 mM ABC for further digestion. LysC (0.5 µg, Wako Chemicals GmbH) was added directly to the protein extract and incubated for 4 h at 30 °C, followed by trypsin (1 µg, Promega) digestion overnight at 30 °C. Subsequently, the peptides were eluted by centrifugation of the spin filters. The filters were further washed twice with 15% acetonitrile. All eluates were collected into the same tube and diluted to an acetonitrile concentration of 1%. Before LC–MS analysis, traces of SDC or SLS were precipitated using 2% trifluoroacetic acid, and all protein digests were desalted using C18 microspin columns (Harvard Apparatus) according to the manufacturer's instructions.

### *2.3.4. LC–MS/MS analyses, peptide/protein identification, and LFQ*

LC–MS/MS analysis of protein digests was performed on Q-Exactive Plus mass spectrometer connected to an electrospray ion source (Thermo Fisher Scientific). Peptide separation was carried out using Ultimate 3000 nanoLC-system (Thermo Fisher Scientific), equipped with in-house packed C18 resin column (Magic C18 AQ 2.4 µm, Dr. Maisch). The peptides were first loaded onto a C18 precolumn (preconcentration setup) and then eluted in backflush mode with a gradient from 98% solvent A (0.15% formic acid) and 2% solvent B (99.85% acetonitrile, 0.15% formic acid) to 25% solvent B over 105 min, continued from 25% to 35% of solvent B up to 135 min. The flow rate was set to 300 nL/min. The data acquisition mode for the initial LFQ study was set to obtain one high-resolution MS scan at a resolution of 60 000 (*m*/*z* 200) with scanning range from 375 to 1500 *m*/*z* followed by MS/MS scans of the 10 most intense ions. To increase the efficiency of MS/MS shots, the charged state screening modus was adjusted to exclude unassigned and singly charged ions. The dynamic exclusion duration was set to 30 s. The ion accumulation time was set to 50 ms (both MS and MS/MS). The automatic gain control (AGC) was set to $3 \times 10^6$ for MS survey scans and $1 \times 10^5$ for MS/MS scans.

Using MASCOT (v.2.5, Matrix Science), the MS raw data of all identification (ID)-based experiments was searched against the UniProt *Methylocystis* sp. strain SC2 protein database (downloaded 12/02/2016, 4428 proteins) or *E. coli* strain K-12 protein database (downloaded 03/09/2016, 4306 proteins) containing 386 common contaminant/background proteins that were manually added. Searches were subjected to MASCOT using Proteome Discoverer (v.1.4, Thermo Fisher Scientific) and the

following parameters: full tryptic specificity required (cleavage after lysine or arginine residues); two missed cleavages allowed; carbamidomethylation (C) set as a fixed modification; and oxidation (M) set as a variable modification. The mass tolerance was set to 10 ppm for precursor ions and 0.02 Da for fragment ions generated by high energy-collision dissociation (HCD). ID-based data were tested for Gene Ontology (GO) term enrichment using the open source software LAGO (see below). With this approach, we obtained first evidence which categories of proteins were particularly enriched in crude lysates relative to clear lysates.

Discovery-LFQ was done using Progenesis QI software (Nonlinear Dynamics, version 2.0). MS raw files were imported into Progenesis and the output data (MS/MS spectra) were exported in mgf format. MS/MS spectra were then searched using MASCOT against a decoy database of the predicted proteomes from *Methylocystis* sp. strain SC2 and *E. coli* strain K-12. The search criteria were set as for ID-based evaluation of the data. Results from the database search were imported back to Progenesis to map peptide identifications to MS1 features. The peak heights of all MS1 features annotated with the same peptide sequence were summed, and protein abundance was calculated per LC–MS run. Next, the data obtained from Progenesis were evaluated using SafeQuant R-package version 2.2.2 (Glatter et al., 2012). Hereby, 1% FDR of identification and quantification as well as intensity-based absolute quantification (iBAQ) values were calculated. MS raw files derived from each heating time point and detergent condition were analyzed by Progenesis and SafeQuant to test for quantification biases between each experimental condition set up for crude and clear lysates. Significantly enriched proteins ($\log_2$ fold change ≥1, *q*-value ≤0.01) were then loaded into LAGO (https://go.princeton.edu/cgi-bin/LAGO) to test for GO term enrichment (Boyle et al., 2004).

### *2.3.5. Development of PRM assays for targeted MS*

For parallel reaction monitoring (PRM), 50 µg of crude or cleared protein sample solubilized in SDC was used for digestion. PRM was performed in three biological replicates with Lys-C/trypsin combinatorial digestion as described above (see Protein digestion section). Subsequently, protein digests were C18 purified using spin columns (Harvard Apparatus) according to the manufacturer's instructions. The samples were

then combined with equal volumes of heavy labeled peptide standards (TQL, JPT Peptide Technologies) of PmoB1 protein. The PmoB1 proteotypic reference peptides contained an arginine residue at the C-terminus. This residue was labeled with $^{13}C_6/^{15}N_4$, resulting in a +10 Da mass shift compared to the unlabeled counterpeptide. Proteotypicity was assessed using Scaffold. The peptide standards were prepared according to the manufacturer's instructions. The spectral library for PRM was generated within Skyline (version 3.6.0) using the data derived from the shotgun LC–MS/MS of the reference peptide mix. Up to five transitions per peptide were traced on the Q-Exactive Plus mass spectrometer. LC settings were the same as described before. For PRM, fragment spectra were obtained with 35000 resolution, with the ion accumulation time set to 100 ms and the AGC set to $2 \times 10^5$. Normalized collision energy (NCE) was set to 27. The MS data was imported into Skyline, which was used for visualization, transition detection, and calculation of transition ratios.

### *2.3.6. Additional data processing and visualization*

Protter version 1.0 was used to predict the number and structure of transmembrane domains of pMMO1 and pMMO2 (Omasits et al., 2014). Using this web-based tool, interactive protein data analysis and hypothesis generation was performed by visualizing the sequences of the selected proteins with regards to their localization and protein topology.

The visualization and analysis of the pathways involved in methane and nitrogen metabolism were done by mapping the identified proteins from LFQ output to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways database of SC2.

A multiple alignment of 16S rRNA gene sequences was made using the Clustal W alignment tool. The sequences were obtained from the GenBank database. The alignment was constructed to compare the 16S rRNA gene sequence of SC2 with those of a representative set of type II methanotrophs. Phylogenetic tree construction was done with maximum-likelihood method using Kimura's two-parameter calculation model available in Mega 7 (Kumar et al., 2016). Confidence values of branches of the phylogenetic tree were determined by using bootstrap analyses based on 1000 resampled data sets.

Visualization of the final graphs and figures was accomplished with SigmaPlot version 2.5, TIBCO Spotfire and Adobe Illustrator CS5.

### *2.3.7. Raw files and associated data deposition*

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the data set identifier PXD009027.

## 2.4. Results and discussion

Our model organism, *Methylocystis* sp. strain SC2, is a representative of one of the important methanotroph groups in terrestrial environments. As all molecular studies hitherto made with SC2 were on the transcriptome level (Dam et al., 2014; Han et al., 2017), we performed the first in-depth proteomic investigation. By first profiling the solubilization preferences, we obtained critical parameters for the ICM-containing SC2. Given the strong difference in protein concentration between crude and clear lysates of SC2, we decided to examine whether solubilization and lysate effects can also be observed for related type IIa methanotrophs of the family *Methylocystaceae* (*Methylocystis/Methylosinus* group). Using the crude-MS strategy, we showed how sample preparation strategies contributed to quantification biases, in particular for membrane-associated subproteomes within whole-cell proteomics. Our results suggest that the crude-MS strategy may also be efficiently applied to other cell types with complex membrane architecture. As a proof of concept, we finally applied our new protocol to assess the differential expression of a targeted set of proteins under high ammonium levels in comparison with standard growth condition (NMS). The proteome results were related to transcriptome data obtained previously using the same experimental setup.

### *2.4.1. Elucidating efficient protein solubilization conditions for strain SC2*

Because of the limited biochemical characterization of SC2 and the lack of proteomics studies, we started our research by establishing an efficient workflow for the analysis of SC2 proteome. It was recently shown how solubilization preferences of individual cell

systems may introduce biases in absolute proteome quantification (Glatter et al., 2015). Therefore, we thoroughly examined our workflow performance on proteome coverage and specific abundance biases of the methane oxidation pathway proteins of SC2.

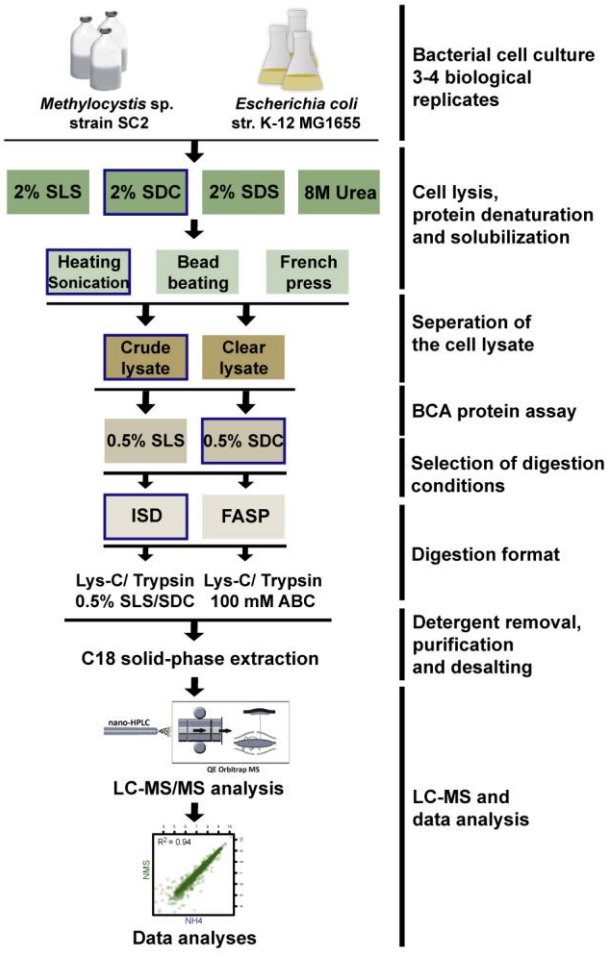



**Figure 1** | Flowchart illustrating the different solubilization and digestion conditions, and procedural steps used to obtain an optimized proteomics workflow for *Methylocystis* sp. strain SC2. Cell pellets of strain SC2, reference methanotrophs, and *E. coli* were reconstituted in different buffer systems. Different cell disruption techniques were applied. Cell debris were either removed or maintained in the lysate to obtain clear and crude lysates, respectively. The protein extraction efficiency was assessed by BCA protein assay. After determining the optimal solubilization conditions, crude-MS performance was tested under in-solution or FASP digestion conditions. All peptides derived from the individual assessment steps were purified using C18 solid phase extraction and analyzed via LC–MS. All experiments were performed in at least three biological replicates. For details, see section 2.3. The best performing conditions of each step are framed in blue.

On the basis of previous observations and general usage in proteome research, we chose the following solubilization buffers to test their solubilization properties on SC2: (i) 8 M urea, (ii) 2% SDS, (iii) 2% SLS, and (iv) 2% SDC (**Figure 1**).While urea, SDS, and SDC are

widely used for ISD or FASP (Wisniewski et al., 2009b; Proc et al., 2010; Wisniewski et al., 2011; Leon et al., 2013; Erde et al., 2014; Wisniewski et al., 2014; Glatter et al., 2015; Schmidt et al., 2016), SLS-based workflows are not yet commonly used in whole-shotgun proteomics. However, SLS has been used as a reconstitution and digestion reagent in the past (Masuda et al., 2009; Ottman et al., 2016). Nevertheless, because of its structural similarity, it represents an attractive alternative to SDS. In addition, like SDC, it is considered as MS-compatible due to its acid-labile properties. Hence, there is no need to deplete the detergent by filtration or protein precipitation techniques before conducting LC–MS.

After cell pellet resuspension in the detergent or urea-containing buffers, we performed solubilization and measured the protein amount of the lysate over incubation time (details in section 2.3.2). To avoid possible loss of proteins, we did not clear the lysate from the cell debris at this stage.

The use of SDS, urea, or SLS had no detectable effect on the protein concentration (**Figure 2A**). In the presence of SDC, the protein amount, however, gradually increased over time up to two-fold and was the highest after 60 min of heating. Bead beating was frequently used as an efficient and rapid technique for disruption of methanotrophic cells in many ecological studies (Kuske et al., 1998; Steinkamp et al., 2001; Radajewski et al., 2002; Kolb et al., 2003; Bussmann et al., 2004; Kolb et al., 2005). However, in the present study, SDC-heating/sonication outperformed alternative cell lysis strategies like bead beating and French press (**Figure S-1**). Furthermore, the overall protein amount strongly decreased after removal of the cell debris from the crude lysate in all tested conditions (**Figure 2A**), suggesting that SC2 cells contain a large pool of insoluble proteins. Most likely, these are membrane-associated. Consequently, the formation of a clear lysate will lead to strong depletion of a substantial fraction from the proteome pool.

A previous study had already shown that solubilization preferences might differ between organisms due to differences in their cell membrane integrity and cell structure (Glatter et al., 2015). Therefore, we applied our different solubilization conditions to *Methylocystis rosea* strain SV97 (*M. rosea*) and *Methylosinus trichosporium* strain OB3b (*M. trichosporium*), but also included *Escherichia coli* in our research as outgroup reference. The BCA assays revealed that the protein extraction profile of *M. rosea* was

highly similar to that of SC2 (**Figure 2B**). This result was anticipated because *M. rosea* is, among type IIa methanotrophs, one of the closest relatives of SC2. Their 16S rRNA gene sequence identity is 99% (**Figure S-2**). Moreover, both strains share the same membrane architecture specific to *Methylocystis* genus (Wartiainen et al., 2006).

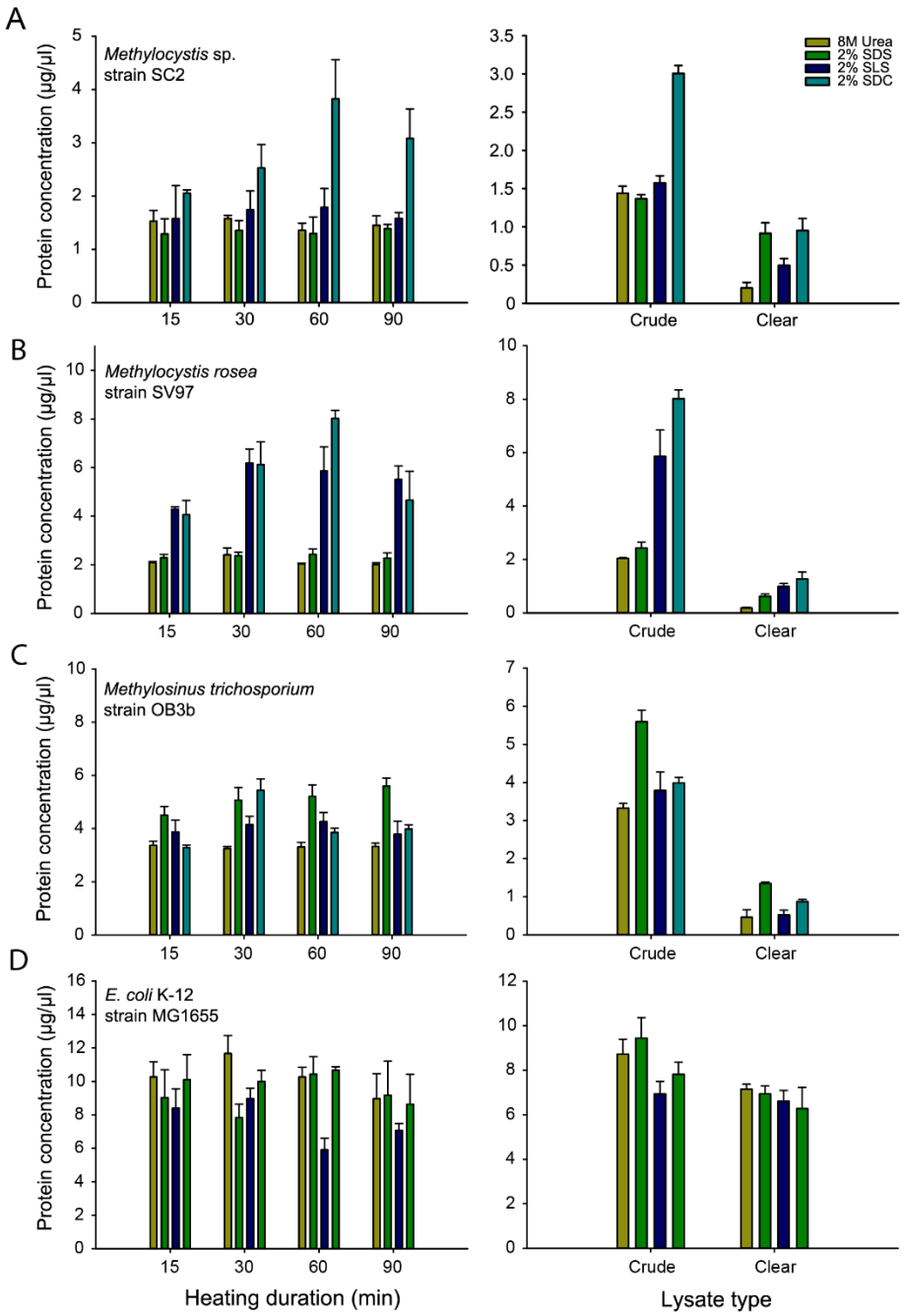



**Figure 2** | Assessment of protein extraction efficiencies using different solubilization buffers and lysate types. Different buffer systems were tested for protein solubilization efficiency by BCA protein assay. The protein content of crude lysates (lysates containing cell debris) was measured over a time course applying heating/sonication to lysates of *Methylocystis* sp. strain SC2, *Methylocystis rosea* strain SV97$^T$, *Methylosinus trichosporium* strain OB3b$^T$, and *E. coli* strain K-12 substr. MG1655 (*left graph* **A–D**). Additional conditions (bead beating and French press) were tested, with results being shown in **Figure S-**

**1**. The impact of the lysate type (crude versus clear) was assessed for cell pellets of all tested organisms after 60 min of incubation (*right graph* **A–D**). The error bars indicate standard deviation. SDC, sodium deoxycholate; SLS, sodium lauroyl sarcosinate; SDS, sodium dodecyl sulfate.

It is, however, worth mentioning that, in contrast to SC2, both *M. rosea* and *M. trichosporium* possess only the conventional pMMO (pMMO1), encoded by *pmoCAB1* (Yimga et al., 2003; Wartiainen et al., 2006). This may be an explanation why SLS and SDC conditions showed nearly the same degree of increase in solubilized proteins for cells of *M. rosea*. In contrast, in SC2, the amount of solubilized proteins was highly increased only for SDC condition. Contrary to SC2 and *M. rosea*, *M. trichosporium* and *E. coli* showed no selective detergent preference (**Figure 2C and 2D**). However, regardless of the lysis technique used, the protein amount of crude lysates consistently outperformed clear lysates by at least three-fold among the type IIa methanotrophs (**Figure 2B and 2C**, right graph). This considerable difference between crude and clear lysates was not detected for *E. coli* (**Figure 2D**, right graph). Our results clearly indicate that the impact of detergents, lysis technique, and lysate type must be thoroughly examined for each microorganism prior to its first-time proteomics analysis. At the same time, we conclude that closely related organisms have similar solubilization requirements due to their similar cell and membrane structures.

### *2.4.2. Proteome coverage and quantification accuracy in clear-MS and crude-MS*

Because of the remarkable difference in the amount of extracted proteins between crude and clear lysates, we decided to further elucidate general aspects of our crude lysate-MS strategy. Employing crude lysates in shotgun proteomics is somewhat unusual, as the most common workflows incorporate removal of cell debris before adding the proteases for digestion (Wisniewski et al., 2009b; Nagaraj et al., 2011; Wisniewski et al., 2011; Hughes et al., 2014; Wisniewski et al., 2014; Glatter et al., 2015; Schmidt et al., 2016). As we did not find any detailed information in previous reports on the performance of crude lysate-MS, we were specifically interested in addressing whether or not such strategy performs well and, with regard to protein identification sensitivity, quantification accuracy, and reproducibility, could be superior to the more common workflows. In particular, we aimed to thoroughly investigate if the lysate formation would have a strong impact on the methane oxidation pathway recovery. For

a more detailed assessment of the workflow, we chose the two best solubilizing agents (SDC and SLS) along with the two most prominent determinants of lysate efficiency, namely lysate type (crude or clear) and heating duration (15 or 60 min).

At first, we monitored the ID-based data of the different workflows. To enzymatically digest our prepared SC2 extracts, we adjusted the protein amount to 50 µg and performed a LysC/trypsin tandem digestion scheme in 0.5% detergent (as detailed in section 2.3.3). Subsequent LC–MS/MS analysis showed us that the detergent in crude-MS format outperformed clear-MS with regard to both spectrum counts and protein identification (**Table S-1**). We observed the following differences in total spectral counts between crude and clear lysates: 30 and 22% in SDC buffer, 17 and 13% in SLS buffer, after 15 and 60 min of incubation, respectively (**Figure S-3A and S-3B**). Obviously, the centrifugation step depletes a significant fraction of the proteome from the initial extract and thus largely contributes to the difference in proteome coverage (**Figure S-3C**). In addition, crude-MS allowed us to detect increased number of proteins with higher average spectrum counts than clear-MS (**Figure S-4A**). Interestingly, when testing for “Cellular component” GO terms, “Membrane”-related terms were significantly enriched among those proteins only identified by crude-MS (**Figure S-4B** and **Table S-2**). Thus, we concluded that crude-MS is essential for sensitive analysis of the SC2 proteome. Though the same amount of protein extract (50 µg) was used for digestion, the difference in proteome coverage between crude-MS and clear-MS could not be eliminated.

To further investigate advantages and pitfalls of our workflows, we performed LFQ of the different conditions (details in section 2.3.4). As studies using crude-MS strategies have rarely been done, we first assessed the variation of crude-LFQ on a global level. Following LFQ and SafeQuant analysis, we plotted the median coefficient of variation (CV) for the different conditions (**Table S-3**).

Global median CVs were generally higher for cleared SC2 samples than for crude lysates (**Figure 3A**). Nonetheless, even CVs obtained from the least efficient condition were among the best CV values hitherto reported for shotgun-MS-LFQ workflows on cultivated microorganisms (Bruderer et al., 2017; Sielaff et al., 2017). After 60 min of incubation, crude lysate in SDC buffer (SDC-60-CR) exhibited the lowest median CV among all eight experimental conditions (3.2%). This corroborates that the depletion of

protein material without exhaustive solubilization affects negatively the protein quantification accuracy and proteome coverage.

Finally, we compared the impact of the digestion scheme and found that ISD and FASP show excellent LFQ correlation ($R^2$ = 0.975, **Figure S-5 and Table S-4**). We concluded that lysis and solubilization conditions have a much greater impact on the protein/peptide recovery and quantification biases than the digestion scheme. Therefore, because of the easy handling and reduced number of procedural steps, we preferred the use of ISD conditions for our further workflow assessment. One explanation for the low impact of the digestion strategy on the proteome coverage may be our high protein input. As shown in a previous study (Sielaff et al., 2017), the impact of the digestion format increases with the input protein decreasing to the lower microgram range.

### *2.4.3. Enrichment of membrane-related proteins in all crude-MS conditions*

Next, we decided to explore whether the lysate type selectively enriched for individual protein classes using GO term analysis. Therefore, we performed a separate LFQ comparison for each heating time point and detergent. We tested whether the proteins that significantly differed between clear-MS and crude-MS were enriched in particular "Cellular component" GO terms. Examining the normalized results for SDC-60 revealed a specific enrichment of "Membrane"-related terms for crude-MS. By contrast, clear-MS showed biases toward ribosomal proteins (**Figure 3B**, *left plot*, **and Figure 3C**, *upper plot*). We, however, got the impression that data normalization (protein sum normalization) may have an impact on the number of significantly enriched proteins. Therefore, we used SafeQuant to compare the statistical parameters before and after normalization. In LFQ-SDC-60, we indeed observed that before and after data normalization, the number of significantly changed proteins differed between crude-MS and clear-MS (**Figure 3B**, right plot**, Table S-5**). However, "Membrane"-related proteins were not affected by normalization bias (**Table S-6**). Regardless of whether normalized or not, they were consistently enriched to a similar level in crude-MS, but not in clear-MS.

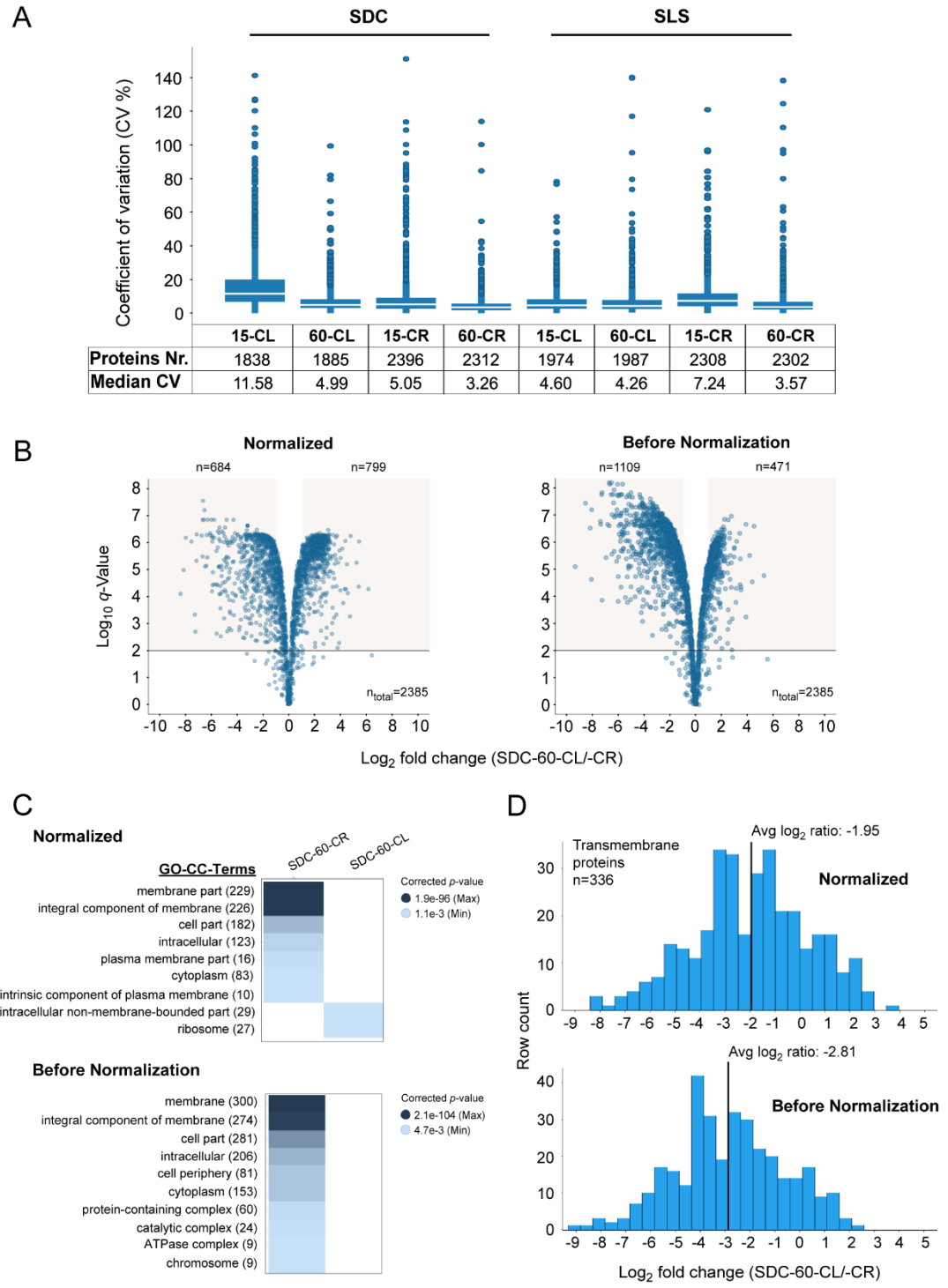



**Figure 3** | **(A)** Median coefficient of variation (CV) for clear-MS and crude-MS samples after label-free quantification (LFQ). Three biological replicates of SC2 were analyzed using different clear-MS and crude-MS strategies. The median CVs were calculated for each conditional group upon LC–MS and LFQ analysis (see section 2.3). **CR**, crude lysate; **CL**, cleared lysate. **(B)** Volcano plot displaying crude-clear LFQ comparisons for SDC-60 before (*right*) and after (*left*) data normalization. The x-axis shows the $\log_2$-fold change of protein abundance, while the y-axis indicates the negative $\log_{10}q$-value. Numbers of both totally quantified proteins and significantly enriched proteins are given. Significance areas are marked (min $\log_2$-fold change ±1, $\log_{10}q$-value ≥ 2). **(C)** Heat map showing enriched "Cellular component" GO terms before (lower) and after (upper) data normalization. Significantly enriched proteins were loaded into LAGO (Boyle

et al., 2004) and selected GO terms were plotted. All *p*-values are Bonferroni corrected for multiple tests. The heat maps of selected GO terms of crude-clear comparisons can be assessed in **Figure S-6**. **(D)** Histogram of $log_2$-fold change ratios of all detected transmembrane proteins in the SC2 proteome. The histogram for the data set before (*lower*) and after (*upper*) data normalization with corresponding average $log_2$ fold change ratios is shown. SDC-60-CR, sodium deoxycholate-60 min heating-crude lysate; SDC-60-CL, sodium deoxycholate-60 min heating-clear lysate.

By contrast, "Ribosomal proteins" were enriched by clear-MS only upon normalization. In fact, no enriched GO terms were detected in clear-MS without normalization (**Figure 3C**, lower plot, and **Table S-6**). Analysis of all experimental conditions between clear-MS and crude-MS showed that regardless of normalization, "Membrane"-related protein terms were consistently enriched in all crude-MS conditions (**Figure S-6 and Table S-6**). A histogram for extracted "Transmembrane" proteins showed that the normalization compressed the average clear/crude-log ratio from −2.8 to around −2 (**Figure 3D**). Relative to crude-MS, a slightly increased number of cytoplasmic proteins were identified in SLS-clear-MS; however, most of these were affiliated with "Ribosomal" GO terms (**Figure S-6 and Table S-6**). In conclusion, the lysate format strongly impacts proteome analysis performance, and specialized proteomes require dedicated and optimal parameter settings for accurate quantification. Here, SDC-60-crude-MS showed the most comprehensive coverage of the SC2 proteome with least quantification bias of protein groups, when compared to corresponding clear-MS conditions.

### *2.4.4. Targeted quantification and intensity-based absolute quantification of membrane-associated pMMO complexes*

Above, we showed that relative to clear-MS, crude-MS results in an increased presence of membrane-associated proteins. Therefore, we decided next to specifically investigate the abundance profile of the integral membrane pMMO protein complex. This complex catalyzes the crucial step in $CH_4$ oxidation, namely the oxidation of C–H bonds (reviewed in (Scheutz et al., 2009; Chidambarampadmavathy et al., 2015)). Strain SC2 is among the subset of type IIa methanotrophs that harbor two pMMO isozymes exhibiting low (pMMO1) and high (pMMO2) affinity to methane (Dunfield et al., 2002; Yimga et al., 2003; Baani and Liesack, 2008).

The enzyme itself is constituted of three subunits, known as PmoB, PmoA, and PmoC (α, β, and γ respectively) and organized as $\alpha_3\beta_3\gamma_3$ trimer (Zahn and DiSpirito, 1996; Lieberman and Rosenzweig, 2004). On the basis of crystallographically modeled pMMO protein, it was proposed that the active catalytic site of the enzyme is localized in the dicopper center of the PmoB subunit (Rosenzweig, 2008; Balasubramanian et al., 2010; Culpepper and Rosenzweig, 2012). This subunit plays not only the central catalytic role in methane oxidation, but also generated a larger number of tryptic peptides than the A and C subunits (**Figure S-7 and Tables S-3**). Therefore, we specifically traced the abundance of PmoB1 and PmoB2 subunits for quantifying the pMMO expression of the two SC2 isozymes.

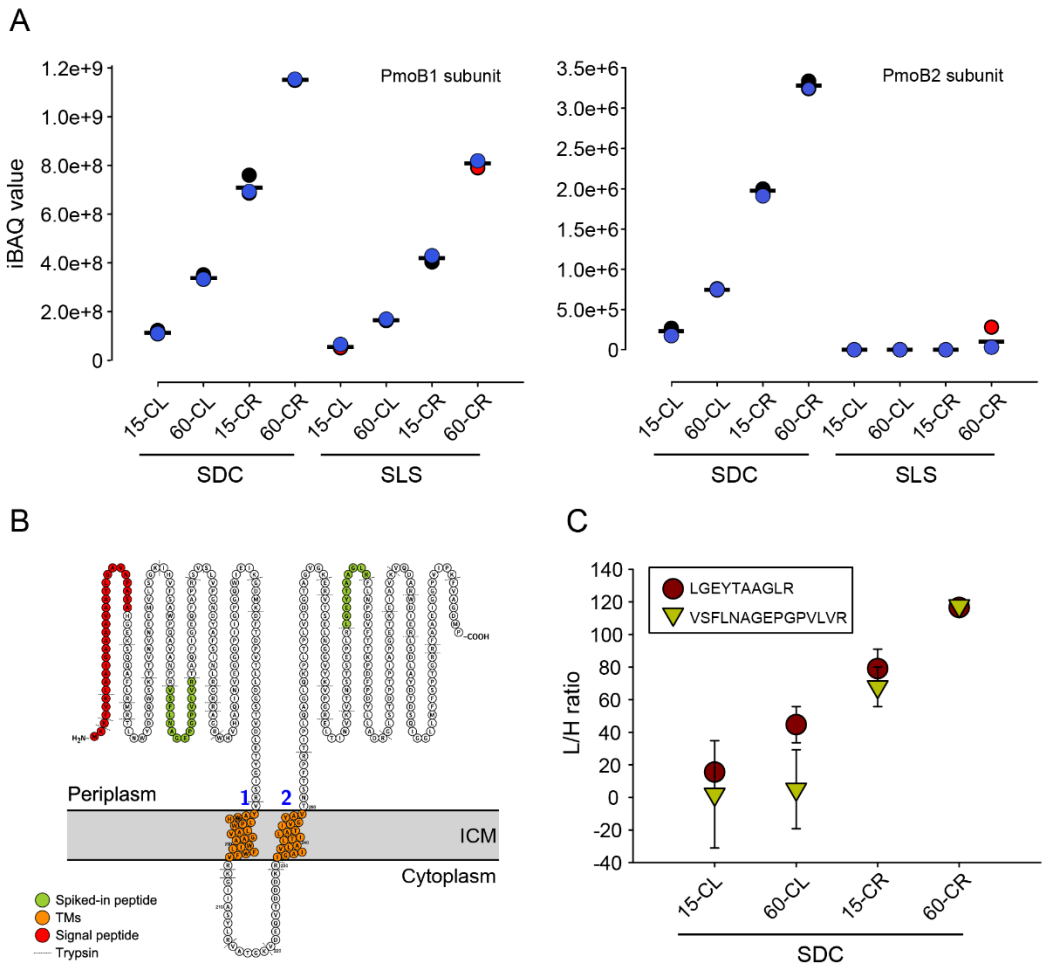



**Figure 4 | (A)** Abundance profile of the SC2 PmoB1 and PmoB2 proteins. iBAQ values for different clear-MS and crude-MS conditions are shown for PmoB1 (*left*) and PmoB2 (*right*). **(B)** Predicted structure of PmoB1. Prediction was performed using Protter (Omasits et al., 2014). The location of the peptides used as isotopically labeled reference standards is indicated in green. Each peptide matches to one of the two periplasmic tails of PmoB1. **(C)** Light/heavy (L/H) ratio of the two PmoB1 peptides across different SDC-clear-MS and SDC-crude-MS conditions. Intensities for the endogenous light peptides and heavy labeled

reference peptides were determined using Skyline following targeted MS measurements (for details, see section 2.3). SDC, sodium deoxycholate; SLS, sodium lauroyl sarcosinate.

In recent years, the iBAQ strategy has been widely used as an absolute quantification index in LFQ workflows.(Nagaraj et al., 2011; Arike et al., 2012; Geiger et al., 2012; Ahrne et al., 2013; Carpy et al., 2014; Soufi et al., 2015; Schmidt et al., 2016) We followed this strategy and calculated iBAQ values for the two PmoB proteins (**Figure 4A**). Consistent with BCA protein assay measurements and ID-based data evaluation, crude-MS confirmed to be superior over clear-MS with the highest iBAQ values detected for PmoB1 and PmoB2 subunits. SDC-60-CR gave us outstanding results. It captured 10-fold more of the PmoB pool than the least efficient condition. In particular, the low-abundant PmoB2 protein was hardly detectable in SLS lysates. Therefore, we omitted SLS conditions from further evaluation and focused on SDC only.

To complement the pMMO-LFQ data and to further investigate the relation between lysate method and quantification accuracy, we performed targeted MS of SDC-solubilized crude and clear lysates using PRM (Gallien et al., 2012). We expected to gain a more detailed understanding of how the different conditions affect total protein recovery and the digestion process. Because of its high abundance and impact on the overall quantification of SC2 proteome, we specifically traced PmoB1 abundance in the different SDC buffer systems. We generated two heavy proteotypic peptide standards for pmoB1 to conduct targeted MS: R.LGEYTAAGLR.F (peptide 1) and R.VSFLNAGEPGPVLVR.T (peptide 2). Each of the two peptides matched one of the periplasmic tails of PmoB1. In addition, they were not located within sequence motifs that were likely to produce miscleavages (Glatter et al., 2012) (**Figure 4B**). The two heavy standards were spiked into SC2 peptide samples and the light-to-heavy (L/H) ratios were monitored in all four conditions following the PRM assays. In general, we found that the average L/H relative intensity ratios of our targeted analysis were consistently the highest in crude lysates after 60 min of incubation (60-CR) (**Figure 4C and Table S-7**). It showed around 1.6-fold and 13-fold higher ratio values than 15-CR and 15-CL, respectively (**Figure S-8**). These differences in PmoB1 recovery were in good agreement with our LFQ data. When comparing the individual L/H ratios of peptide 1 and peptide 2, we found that these two peptides had highly similar L/H ratios only in the 60-CR samples. By contrast, their L/H ratios strongly varied in 15-CR and 15-CL samples (**Figure S-8 and Table S-7**). This observation suggests that each periplasmic tail of PmoB1 may

have different solubilization and denaturation properties and optima. We cannot exclude that such a heterogeneous solubilization/denaturation within a given protein is restricted only to a limited number of individual proteins, such as PmoB1. Even so, such property will certainly increase the likelihood of more erroneous protein quantification, especially when targeted MS is performed with only one peptide standard per protein.

### *2.4.5. Recovery of methane oxidation and C1 assimilation proteins in SDC-60-CR samples*

We observed major differences in the recovery of PmoB between the different lysate types. To further elucidate the SC2 proteome coverage, we assessed whether our new procedural steps introduce biases into the recovery of downstream proteins involved in methane oxidation and carbon assimilation. Formaldehyde is a central intermediate that is generated by the consecutive oxidation of methane and methanol. It is either assimilated via the serine pathway into cell carbon or further oxidized to formate and finally carbon dioxide to gain energy for growth. Serine cycle is being used to produce acetyl-CoA, which is channeled either into polyhydroxybutyrate (PHB) synthesis or the tricarboxylic acid (TCA) cycle under nutrient-deficient and nutrient-sufficient conditions, respectively (Semrau et al., 2010; Chidambarampadmavathy et al., 2015; Karthikeyan et al., 2015). The required energy for the methane oxidation step is produced during oxidation of methanol to formaldehyde by the activity of pyrroloquinoline quinone (PQQ)-dependent methanol dehydrogenase. Additional energy, in form of reduced equivalents, is generated in the two successive oxidation steps from formaldehyde to formate (formaldehyde dehydrogenase) and formate to carbon dioxide (formate dehydrogenase). Formaldehyde oxidation step involves four cofactors: tetrahydrofolate ($H_4F$), tetrahydromethanopterin ($H_4MPT$), glutathione (GSH), and mycothiol (MSH) (Vorholt, 2002; Fei et al., 2014). Pathway reactions that include $H_4F$/$H_4MPT$ are common among type IIa methanotrophs and many methylotrophs, and usually involve formaldehyde activating enzyme (Fae) (Vorholt et al., 2000).

We identified 78 proteins that were related to pMMO complexes, methanol oxidation and formaldehyde assimilation (FAA), serine cycle, PHB, glycerate regeneration (GR), and TCA cycles (**Table S-8**). These proteins cover major metabolic pathways of SC2 and thus allow for an accurate assessment of the experimental

conditions (**Figure 5A**). We generated line charts averaged on protein group $log_2$-iBAQ intensities for each of the four experimental conditions studied (**Figure 5B**). Consistent with our previous results, abundance changes differed most between 15-CL and 60-CR. Differences were also detectable between 15-CR and 60-CR, albeit less pronounced. Importantly, effects on protein abundance were not only observed for integral membrane PmoB protein, but also for proteins involved in methanol oxidation and in PHB/GR and TCA cycles. Among other proteins, these pathways contain membrane proteins or membrane proximal proteins involved in electron transfer. Like PmoB, these proteins showed an increased iBAQ profile in 60-CR lysates relative to all other lysate types.

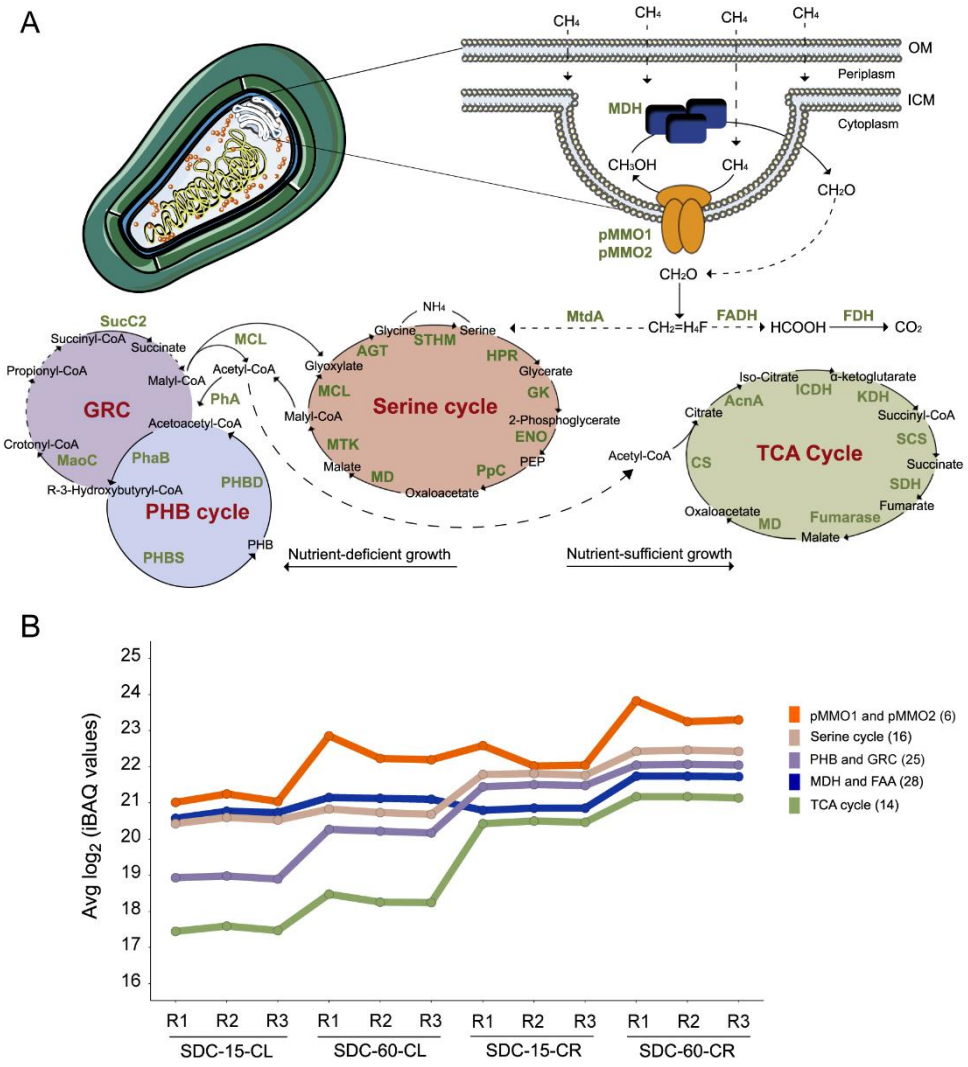



**Figure 5** | **(A)** Illustration of the SC2 cell structure and main biochemical pathways involved in methane oxidation and carbon assimilation. All identified proteins are shown in green. pMMO, particulate methane monooxygenase; MDH, methanol dehydrogenase; FADH, formaldehyde dehydrogenase; FDH, formate

dehydrogenase; MtdA, methylene tetrahydromethanopterin dehydrogenase; HPR, hydroxypyruvate reductase; GK, glycerate 2-kinase; ENO, enolase; PpC, phosphoenolpyruvate carboxylase; MD, malate dehydrogenase; MTK, malate thiokinase; MCL, malyl coenzyme A lyase; AGT, alanine-glyoxylate transaminase; STHM, serine hydroxymethyl transferase; CS, citrate synthase; AcnA, aconitase; ICDH, isocitrate dehydrogenase; KDH, α-ketoglutarate dehydrogenase; DDH, dihydrolipoyl dehydrogenase; SCS, succinyl-CoA synthase; SDH, succinic dehydrogenase; PhA, acetyl-CoA acetyltransferase; PhaB, acetoacetyl-CoA reductase; PHBS, PHB synthase; PHBD, PHB depolymerase; MaoC, 3-hydroxybutyryl-CoA dehydratase; SucC2, succinyl-CoA synthetase, GRC, glyoxylate regeneration cycle; PHB, polyhydroxybutyrate; TCA, tricarboxylic acid; FAA, formaldehyde assimilation. **(B)** Line chart linked to the different pathway groups with averaged $log_2$-iBAQ intensities for different SDC-clear-MS and SDC-crude-MS conditions. The number of quantified proteins belonging to the different pathway groups are shown in parentheses. The color of each group corresponds to the color code as used in panel A.

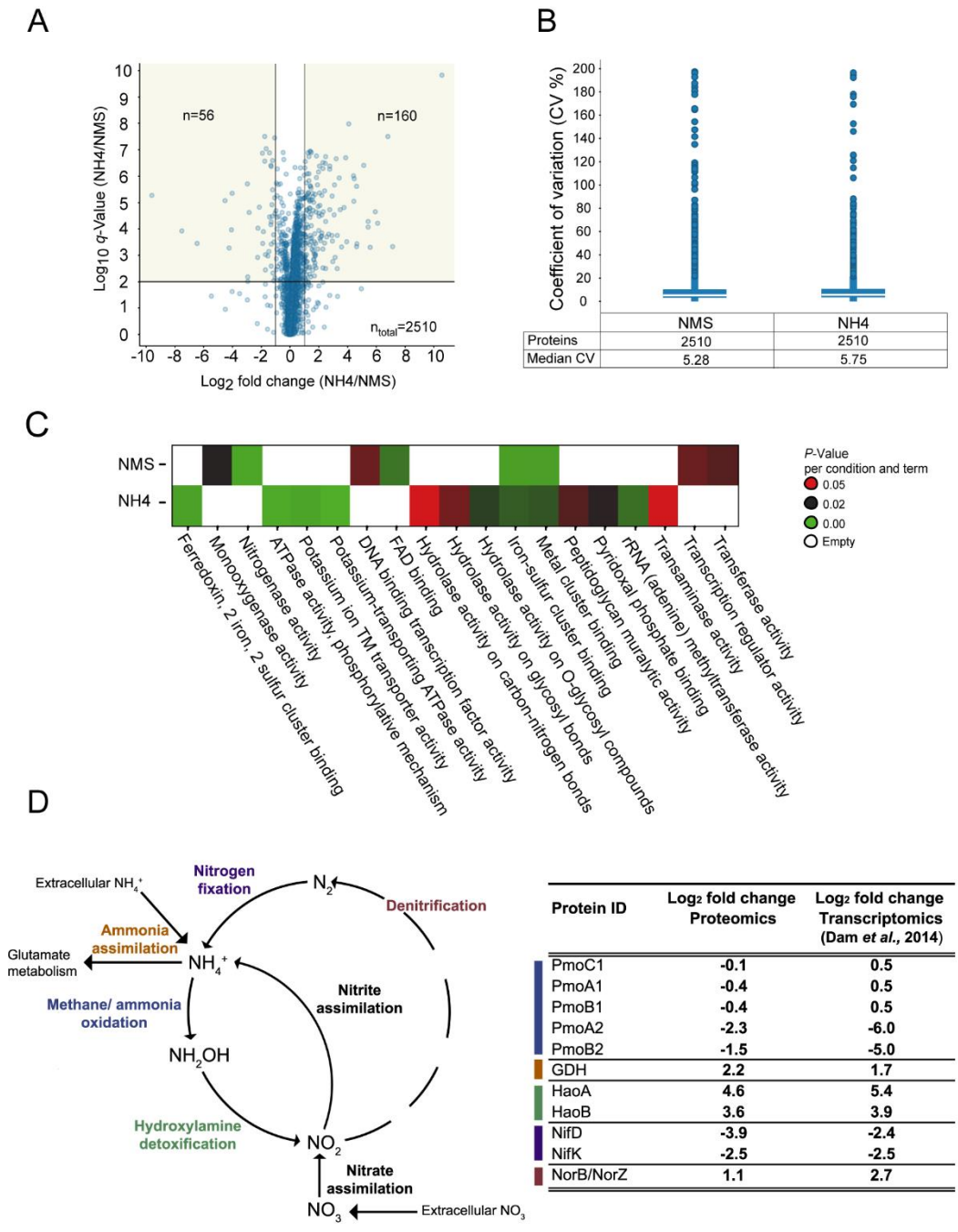
Thus, the solubilization of membrane-rich cells, such as those of *Methylocystis* spp., is a heterogeneous process that, to avoid major quantification biases, requires highly optimized conditions for comprehensive proteome data analysis. This is particularly important for studies on biochemical processes that involve highly abundant membrane proteins or membrane proximal protein groups. Therefore, it is recommended that, initially, a solubilization screening will be performed to identify whether the study organism exhibits “specializer” properties, similar as observed for SC2.

### *2.4.6. Assessment of the newly developed proteomics workflow by exposure of SC2 to high ammonium*

To further validate the efficiency of our optimized SC2 workflow, we analyzed the SC2 proteome under low and high nitrogen conditions, with a focus on the differential expression of proteins involved in methane and nitrogen metabolism. Various methanotroph studies have shown that aerobic methane oxidation is very sensitive to the addition of N-based fertilizers in upland soils (Mohanty et al., 2006). In particular, ammonium-containing fertilizers strongly inhibit the sink capacity of these soils for atmospheric methane. This inhibition effect primarily occurs at low-methane and high ammonium concentrations. It is thought to be related to a competitive interaction between methane and ammonia for the pMMO and related to the production of nitrite (Bedard and Knowles, 1989; Nyerges and Stein, 2009; Lu et al., 2012).

First insights into how high ammonium exposure affects SC2 cells were obtained by Dam et al. (Dam et al., 2014) using whole-genome transcriptomics. This study

primarily focused on the effect of increasing ammonium on the differential expression of methane and nitrogen metabolism-related genes.




| Protein ID | Log2 fold change Proteomics | Log2 fold change Transcriptomics (Dam *et al.*, 2014) |
|---|---|---|
| PmoC1 | -0.1 | 0.5 |
| PmoA1 | -0.4 | 0.5 |
| PmoB1 | -0.4 | 0.5 |
| PmoA2 | -2.3 | -6.0 |
| PmoB2 | -1.5 | -5.0 |
| GDH | 2.2 | 1.7 |
| HaoA | 4.6 | 5.4 |
| HaoB | 3.6 | 3.9 |
| NifD | -3.9 | -2.4 |
| NifK | -2.5 | -2.5 |
| NorB/NorZ | 1.1 | 2.7 |

**Figure 6** | **(A)** Volcano plot illustrating LFQ data obtained from cells grown under control (NMS) and high ammonium (NH4) conditions. The *x*-axis shows the $\log_2$-fold change of protein abundance, while the *y*-axis indicates the negative $\log_{10}q$-values. Significance areas are highlighted by color. **(B)** Box plots show total protein numbers and median CV. **(C)** Heat map showing enriched GO terms among differentially expressed proteins in control and high ammonium treated samples upon LFQ analysis (see Volcano plot). Significance level of GO term enrichment is indicated by different colors. **(D)** Comparative transcriptome/proteome analyses of strain SC2 under standard (NMS) and high ammonium conditions. Nitrogen metabolism of strain SC2 based on identified pathway proteins is shown (*left graph*). $\text{Log}_2$ fold change ratios (NH4/NMS) of methane oxidation and nitrogen metabolism-related transcripts (Dam et al.,

2014) and proteins are summarized in the table (*right graph*). NMS is specified as the control in both data sets. Pmo, particulate methane monooxygenase; Nif, nitrogenase; Nor, nitric-oxide reductase; Nas/Nir, nitrate/nitrite assimilation; Hao, hydroxylamine oxidoreductase; GDH, glutamate dehydrogenase.

We applied the same experimental setup as has been done by Dam et al. (Dam et al., 2014) in their transcriptome study. SC2 cells pregrown to mid log phase in NMS were exposed to "high ammonium" (30 mM $NH_4^+$) for 10 h (details in section 2.3.1) or, as a control, were maintained in standard NMS (10 mM $NO_3^-$) medium. After harvesting, the proteomics workflow, including LC–MS/MS and statistical analyses, was applied as described above. Next, we compared the results of our proteomics approach with the above-mentioned SC2 transcriptome data sets.

Similar as on transcriptome level, our proteomics approach was successful in identifying almost all methane oxidation and carbon assimilation proteins within the most abundant proteins in the data set. In total, 2510 proteins were identified in LFQ analyses (**Figure 6A and Table S-9**). This corresponds to about 62% of the 4040 predicted CDSs in the SC2 genome, with high reproducibility between the four replicates of each experimental condition (global median CVs 5%) (**Figure 6B**). All the proteins involved in methane oxidation, C1 assimilation, PHB/GR and TCA cycles were quantified in the LFQ analyses. On the basis of significant $\log_2$ fold change ratios of peptide intensities in NH4/NMS (>1 or < – 1; *q*-value ≤0.01) a total of 216 proteins were differentially expressed (**Figure 6A**).

On the basis of the proteins differentially expressed between the two treatments, we determined the significantly enriched "Molecular function" GO terms using LAGO (**Figure 6C and Table S-10**). The majority of GO terms enriched in high ammonium treatment were linked to the early response in ionic-osmotic stress. Among these, proteins related to "Potassium ion transmembrane activity" and "Potassium-transporting ATPase activity" terms were most significantly enriched. Therefore, we calculated the ionic strength of our nitrogen treatments to reveal the possible ionic pressure on SC2 upon exposure to high ammonium load (Solomon, 2001). The high ammonium treatment had ionic strength of 78 mM in comparison to the control with 43 mM ionic strength. Therefore, SC2 responded to the high ammonium treatment with active accumulation of potassium into the cell. It has been shown that the primary response of most bacteria to ionic-osmotic stress is the accumulation of potassium ions ($K^+$) to high cytoplasmic concentrations. This "salt in" strategy requires a specific set of

proteins including influx systems and passive transport via $K^+$ channels and porins (Oren et al., 2002; Sleator and Hill, 2002; Epstein, 2003). In most of the cases, the intracellular $K^+$ concentration is regulated by the *kdp* operon whose expression may be permanently induced in the presence of osmotic solutes like NaCl and $(NH_4)_2SO_4$ (Csonka, 1989). Parallel to $K^+$ uptake, most bacteria accumulate high cytoplasmic levels of glutamate as a counterion for potassium (Ogahara et al., 1995). Glutamate can be synthesized from ammonia and α-ketoglutarate using either glutamate dehydrogenase (GDH) or glutamine synthetase and glutamate synthase (GS/GOGAT) (Csonka, 1989). In both transcriptome and proteome studies of SC2, glutamate dehydrogenase (GDH) responded to high ammonium with a significant upregulation ($log_2$ fold change ratio (proteome) of 2.2) (Figure 6D and Table S-11). This enzyme is involved in ammonium assimilation by which ammonia is processed and incorporated into cell biomass (Pengpeng and Tan, 2013). Under high ammonium load organisms switch to GDH from GS/GOGAT system since $K_m$ values of GDH for ammonia are generally higher and its assimilation is less energy-demanding than by GS (Murrell and Dalton, 1983; Kanamori et al., 1989). In addition, we observed that "Transaminase activity" and "Pyridoxal phosphate binding" GO terms were highly enriched in high ammonium treatment (Figure 6C andTable S-10). Transaminases are pyridoxal-phosphate requiring enzymes that catalyze a transamination reaction between amino acids and production of glutamate from other amino acids (Feldman and Gunsalus, 1950). One of those enriched proteins was acetylornithine aminotransferase responsible for synthesis of glutamate from $N_2$-acetyl-l-ornithine(Albrecht and Vogel, 1964). A study on the osmotic stress response of *Rhizobium meliloti* showed that some of the accumulated glutamate was derived from transamination of amino acids (Botsford and Lewis, 1990). In addition to proteins responsible for $K^+$ transport into cell, "Peptidoglycan muralytic activity" and "Hydrolase activity" GO terms were highly enriched under high ammonium concentration. Ionic-osmotic stress may lead to changes in cellular size to maintain the turgor pressure in bacterial cells (Piuri et al., 2005). To enlarge the cells, newly synthesized peptidoglycan subunits are exported from the cytoplasm and then incorporated into the wall. Simultaneously, hydrolases assist actively in cleaving cell wall components to achieve the insertion of synthesized material necessary for enlargement of the cell and construction of new transport systems (Lee and Huang, 2013). "rRNA

methyltransferase activity" was another significantly enriched GO term under high ammonium load (**Figure 6C and Table S-10**). This housekeeping enzyme supports ribosomal activities and generally exhibits increasing protein levels under stress conditions (Kyuma et al., 2015).

The exposure of SC2 to high ammonium levels had no effect on pMMO1 expression. Its three subunits were highly expressed in both nitrogen treatments, with high reproducibility between the biological replicates. By contrast, the expression level of pMMO2 was significantly lowered in high ammonium, with a mean $\log_2$ fold change ratio of −1.9 (**Figure 6D and Table S-11**). This lowered expression level agrees well with the transcriptome response of SC2 to high ammonium, although the downregulation of pMMO2 on transcript level was much greater (mean $\log_2$ fold change ratio (mRNA) of −5.5) than on proteome level. It is assumed that the inhibitory effect of ammonia on methane oxidation is due to the structural homology between ammonia and methane (Bedard and Knowles, 1989). The pMMO2 isozyme is responsible for high-affinity methane oxidation. Therefore, high concentrations of ammonia block accessibility of the active site of pMMO2 for methane. In addition, excessive ammonia oxidation produces toxic products, like hydroxylamine (reviewed in (Stein, 2012)). The latter is further oxidized to nitrite, known to inhibit methane oxidation in cultivated and soil methanotrophs (Nyerges et al., 2010).

Like the SC2 proteins involved in methane metabolism, the majority of proteins related to nitrogen metabolism were identified to be highly expressed. In the different pathway categories, there were a number of proteins with differential expression between the two nitrogen treatments, as previously observed for their transcripts (**Figure 6D and Tables S-11**). One of these enzymes was hydroxylamine oxidoreductase (Hao) with its two subunits, displaying $\log_2$ fold change ratios of 3.6 (HaoB) and 4.6 (HaoA). The production of this enzyme was significantly increased under high ammonium condition. In a detoxification step, it converts hydroxylamine into the less-toxic nitrite. In agreement with the transcriptome study, the proteins involved in nitrogen fixation showed lower expression levels than most other SC2 proteins. Moreover, NifD and NifK proteins encoding α and β subunits of nitrogenase were significantly downregulated (**Figure 6D, Table S-9 and S-11**). Given that the active site cluster (called FeMo-cofactor) of the nitrogenase enzyme is localized within NifD subunit

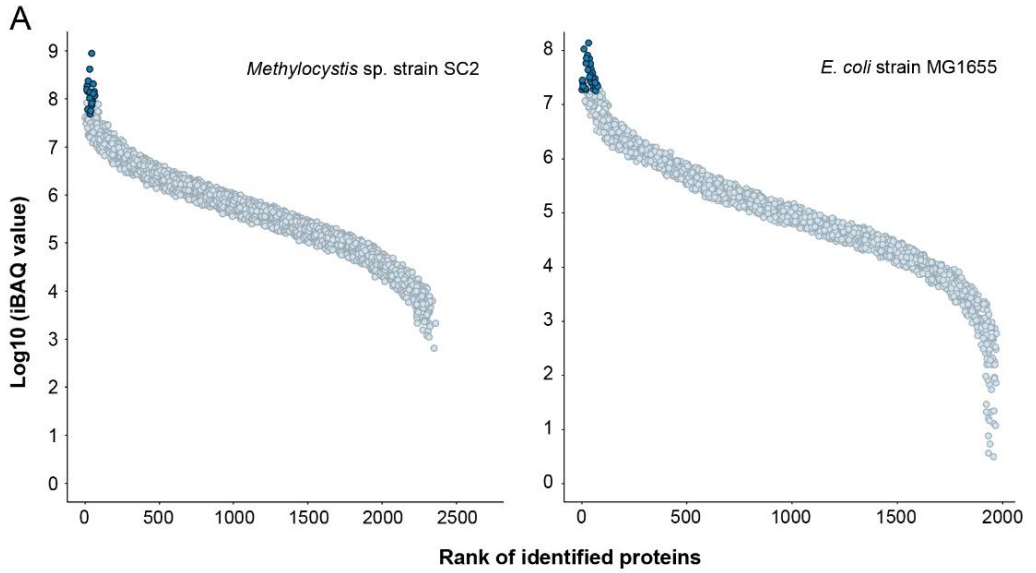
(Lovell et al., 2003; Staples et al., 2007), NifDK complex is expected to more rapidly respond to the excess of alternative nitrogen sources than other subunits of the enzyme. In summary, we successfully tested our newly developed proteomics workflow on the differential expression analysis of proteins involved in the methanotrophic lifestyle of SC2 under different nitrogen conditions. Key proteins that were highly regulated on their transcript level showed a concurrent trend of abundance change on protein level. Therefore, we conclude that our crude-lysis proteomics workflow is a fast and reliable way to accurately analyze the response of strain SC2 with high coverage to any type of environmental change.

### *2.4.7. Strain SC2: A highly specialized proteome compared to E. coli*

The unexpectedly high impact of the lysis method on the SC2 proteome coverage made us decide to also apply our newly developed workflow to *E. coli* strain K-12, which showed a different solubilization profile (**Figure 1D**). In particular, we were interested in identifying the major differences in workflow performance between SC2 and *E. coli*; two bacteria that obviously differ in their cell structure. The *E. coli* proteome has been studied in detail (Soufi et al., 2015). Most recently, its in-depth quantitative analysis showed a condition-dependent proteome-wide representation of protein copy numbers per cell (Schmidt et al., 2016).

To compare the two proteomes, we processed mid log phase cultures of SC2 and *E. coli* and calculated iBAQ values for the digested proteomes (**Table S-3 and S-12**). The limited annotation coverage of SC2 relative to the well-annotated *E. coli* proteome made it difficult to compare the two proteomes in detail. Therefore, we ranked all the iBAQ protein values (**Figure 7A**) and focused on the first 20 most expressed proteins (Top20). SC2 had a large number of methane oxidation pathway proteins among the highly expressed proteins, with the pMMO1 proteins ranking always within the Top20 (**Figure 7A and 7B**). This relates well to the results of global transcriptome studies on type IIa methanotrophs, where *pmoCAB1* genes had always been higher expressed than all other pathway genes (Matsen et al., 2013; Dam et al., 2014; Vorobev et al., 2014; Han et al., 2017). Next, we calculated the proteome fraction (PF) of individual proteins and the contribution of Top20 proteins to the total SC2 proteome. The PF coverage of Top20 proteins was 39% in SC2 compared to around 36% in *E. coli*. Although it was in a similar

range for both bacteria, it became obvious that a higher number of growth/translation-related proteins belonged to the Top20 in *E. coli* (**Figure 7B**). This agrees well with the results of a recent quantitative protein study in *E. coli* (Schmidt et al., 2016) (**Figure S-9 and Table S-13**).




B

| | Strain SC2 | *E.coli* |
|---|---|---|
| | Proteins | |
| 1 | PmoB1 | Elongation factor Tu (Tuf) |
| 2 | FAE | Major outer membrane lipoprotein (Lpp) |
| 3 | MxaF | Outer membrane protein A (ompA) |
| 4 | Uncharacterized protein | Outer membrane protein C (ompC) |
| 5 | PmoA1 | Acyl carrier protein (AcpP) |
| 6 | 60 kDa chaperonin (GroL) | 50S ribosomal protein L7/L12 (RplL) |
| 7 | Elongation factor Tu (Tuf) | 50S ribosomal protein L24 (RplX) |
| 8 | Acyl carrier protein (AcpP) | 50S ribosomal protein L29 (RpmC) |
| 9 | Cold-shock family protein | 50S ribosomal protein L33 (RpmG) |
| 10 | Uncharacterized protein | 30S ribosomal protein S16 (RpsP) |
| 11 | 50S ribosomal protein L7/L12 (RplL) | 50S ribosomal protein L31 (RpmE) |
| 12 | ATP synthase subunit beta (AtpD) | 50S ribosomal protein L22 (RplV) |
| 13 | MxaI | GapA |
| 14 | PmoC1 | DNA-binding protein HU-alpha |
| 15 | Aminotransferase class V | Phosphoglycerate kinase (Pgk) |
| 16 | Uncharacterized protein | 50S ribosomal protein L9 (RplI) |
| 17 | Citryl-CoA lyase | 30S ribosomal protein S5 (RpsE) |
| 18 | Glutamine synthetase | 30S ribosomal protein S13 (RpsM) |
| 19 | Cold-shock DNA-binding protein | 30S ribosomal protein S10 (RpsJ) |
| 20 | OmpW family protein | 50S ribosomal protein L2 (RplB) |

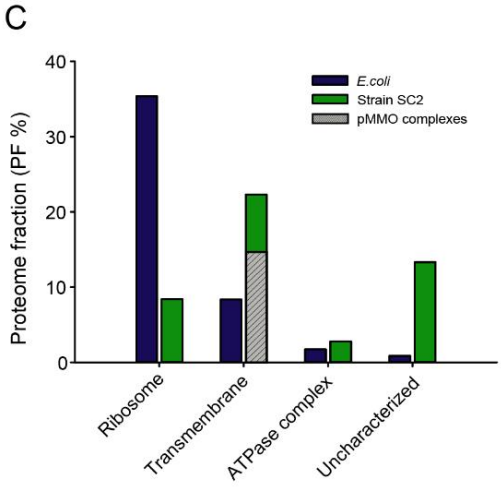



**Figure 7** | **(A)** Ranked iBAQ values for the SC2 and *E. coli* proteomes. Proteins marked in dark blue refer to the proteins with Top20 highest iBAQ values. **(B)** List of Top20 SC2 and *E. coli* proteins. Proteins linked to the methane oxidation pathway are most intense proteins among SC2-Top20 and highlighted in green. In *E. coli*, the majority of the Top20 proteins is related to cell growth (Tuf and ribosomal proteins) and to outer membrane proteins (Lpp, OmpA, OmpC). **(C)** Proteome fractions of selected protein groups. The iBAQ-proteome fraction (PF) was calculated for ribosomal proteins, transmembrane proteins (TM), the ATPase complex, and proteins defined as "uncharacterized" based on Swiss-Prot keywords entries. While *E. coli* shows a higher ribosome-PF than SC2, strain SC2 contains higher level of TM proteins (23% TM-PF in SC2, 8% TM-PF in *E. coli*). The majority of the SC2-TM consists of pMMO (15% TM-PF). On the contrary, the ATPase complex has similar PFs in both organisms. Relative to *E. coli*, a higher fraction of "uncharacterized" proteins was found in SC2.

We further calculated the PF of the methane oxidation pathway proteins and found that they comprise around 25% of the SC2 proteome. Relative to standard organisms, like *E. coli*, this is an exceptionally high level of proteome coverage by a niche-subproteome. In addition, the high proportion of membrane-related proteins on the proteome explains the previously described impact of the normalization step on the number of significantly regulated proteins in a crude-clear-lysate comparison (**Figure 3**).

Following, we compared PFs of particular functional categories between SC2 and *E. coli* (**Figure 3C**). The *E. coli* ribosomal PF was around 20% of the total proteome and thus 2.5-fold higher than in SC2. The differing PF values can be explained with the specific ecological lifestyle adaptations of these two bacteria. With the synthesis of up to 70 000 ribosomes per cell, *E. coli* grows rapidly at a rate of one generation per 20 min in carbohydrate-rich (e.g., glucose) environments (Bremer and Dennis, 2008). The necessary level of ribosomes and rapid growth rate is related to the transcriptional activity of seven rRNA operons present in the *E. coli* genome (Condon et al., 1995). On the contrary, SC2 harbors only a single rRNA operon copy (Dam et al., 2012) and grows relatively slow (doubling time around 8.5 h) with methane as the only energy and carbon source.

In opposite to ribosomal proteins, transmembrane (TM) proteins in *E. coli* covered only 8% of its proteome, while the PF in SC2 was as high as 23%. Interestingly, among that 23% of the SC2 TM proteins, the pMMO complex covered around 15% (TM-PF). It is known that the amount of ICM and the relative concentration of pMMO increase concurrently with the growth rate of methanotrophic bacteria (Bowman, 2006). The increase in ICM amount expands the surface area available for pMMO membrane integration and thus allows for increase in growth rate (Dunfield et al., 1999). Obviously, the strong discrepancy in the PF between the TM subproteomes of SC2 and *E. coli* is related to the pMMO complex, with the PF of their other TM proteins being relatively similar. This is well illustrated by ATP synthase, which has a PF of 1–2% in both SC2 and *E. coli* (**Figure 7C**).

The high content of TM proteins, but in particular the pMMO complex, is a specific proteome property of methanotrophs that contributes to the initially found lysis-type-specific solubilization preferences. The high pMMO-PF is a specific metabolic characteristic of methanotrophs and is directly linked to their ecological lifestyle.

## 2.5. Conclusions

In this study, we developed an optimized global proteome analysis workflow for a representative member of type IIa methanotrophs, *Methylocystis* sp. strain SC2. Among the methane-oxidizing bacteria, *Methylocystis* spp. are most widespread in nature (Knief, 2015). We showed that detergent-based, and in particular SDC-based, procedures are more efficient in extracting and solubilizing SC2 proteins than chaotropic strategies. The incubation time proved to be one of the critical parameters for optimal protein yields. Another unexpected finding was that the removal of cell debris from the crude cell lysate strongly reduced the fraction of membrane-associated proteins from the total SC2 proteome, including pMMO complexes. Expanding our research to other members of the *Methyolocystis/Methylosinus* group and to *E. coli* allowed us to show that the preference for particular protein solubilization conditions is a specific property of methanotrophic bacteria. Introducing our new crude-MS workflow to LFQ analyses, we obtained excellent quantification accuracy and proteome coverage for SC2, with CVs among the lowest reported for the global proteome quantification of cultured microorganisms. In particular, our SC2-crude-MS-strategy not only significantly increased the coverage of pMMO complexes, but also improved the detection of most C1 assimilation pathway proteins. These findings were further validated by assessment of the proteome response of SC2 to contrasting nitrogen conditions. Transcripts and proteins related to pMMO complexes and nitrogen metabolism were highly expressed and showed concurrent trends of differential regulation in both transcriptomics and proteomics. The results corroborate that our newly developed workflow is a fast and reliable way to comprehensively elucidate the global proteome response of SC2 to environmental change.

Finally, compared to *E. coli* strain K-12, strain SC2 contains an exceptionally high content of transmembrane proteins. This explains the lysis-specific solubilization preferences of SC2 or, more general, type IIa methanotrophs. On the basis of our results, we conclude that specialized proteomes, but in particular those with specific membrane architecture, require a dedicated workflow design, for example, in the format of our crude-MS strategy. We found that such workflows critically depend on specifically optimized lysis conditions, whereas the digestion format may be neglectable.

## 2.6. Acknowledgements

AH is a member of the International Max Planck Research School for Environmental, Cellular, and Molecular Microbiology (IMPRS-Mic). AH thanks Christiane Nüsslein-Volhard-Foundation (CNV) for CNV grant awarded to excellent women scientists with children in the field of experimental natural sciences. This study was supported by the Deutsche Forschungsgemeinschaft (DFG) through Collaborative Research Center SFB 987.

## 2.7. Supporting information

Supporting figures are provided below in section 2.7.1. The complete supporting information, including the supporting tables (xlsx), is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jproteome.8b00216. In addition, it is included in the DVD provided with the hard copy of the thesis.

### *2.7.1. Supporting figures*

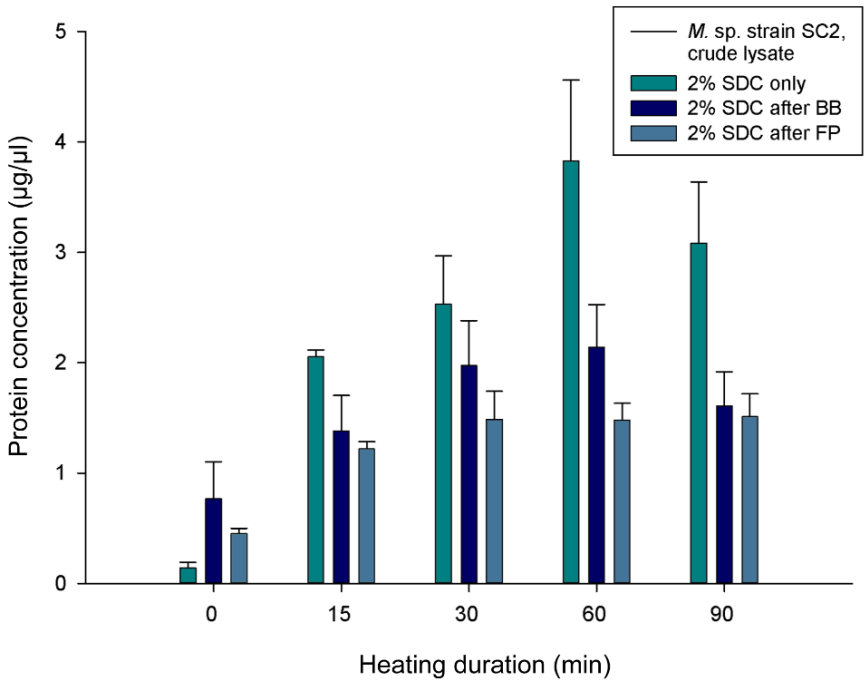



**Figure S-1 |** BCA protein assay of SC2 protein extracts. Bead beating or French press were used as an alternative physical cell lysis technique in addition to SDC-heating/sonication in *Methylocystis* sp. strain SC2. SDC-heating/sonication outperformed both of the techniques.

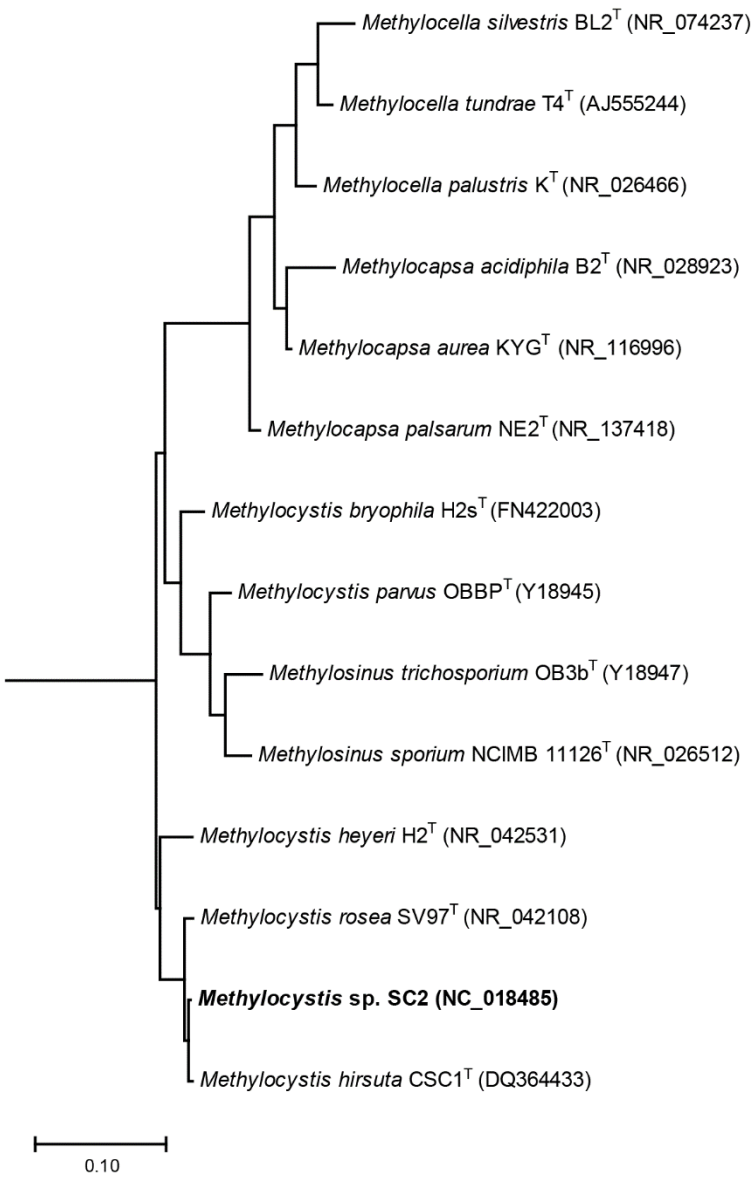



**Figure S-2** | Maximum-likelihood tree of type II methanotrophs based on 16S rRNA gene sequences. The tree shows the relationship of *Methylocystis* sp. strain SC2 to selected type II methanotrophic bacteria (T - type strains). *Escherichia coli* str. K-12 substr. MG1655 was used as the outgroup reference. The scale bar represents 0.1 change per nucleotide position. The GenBank accession numbers of the sequences are given in parentheses after each species name. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016) using Kimura-2-parameter model (Kimura, 1980).

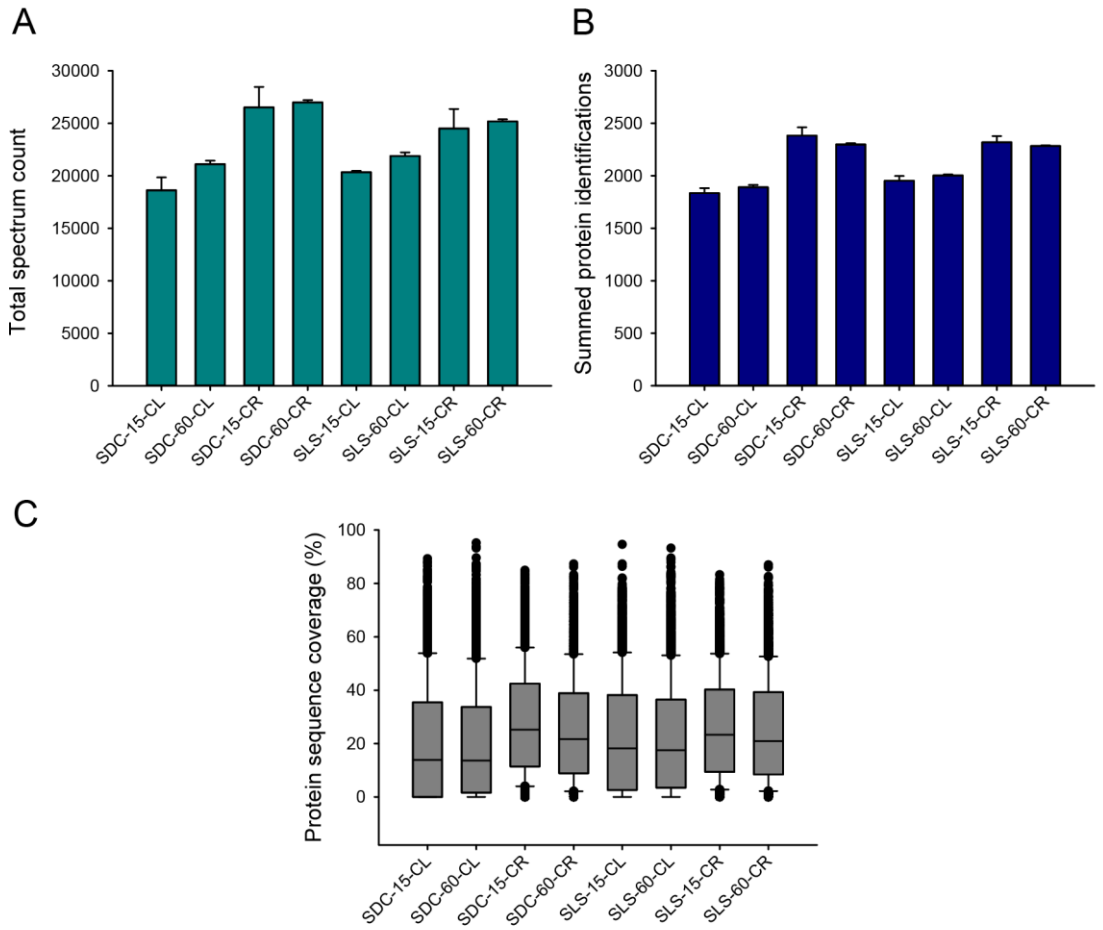
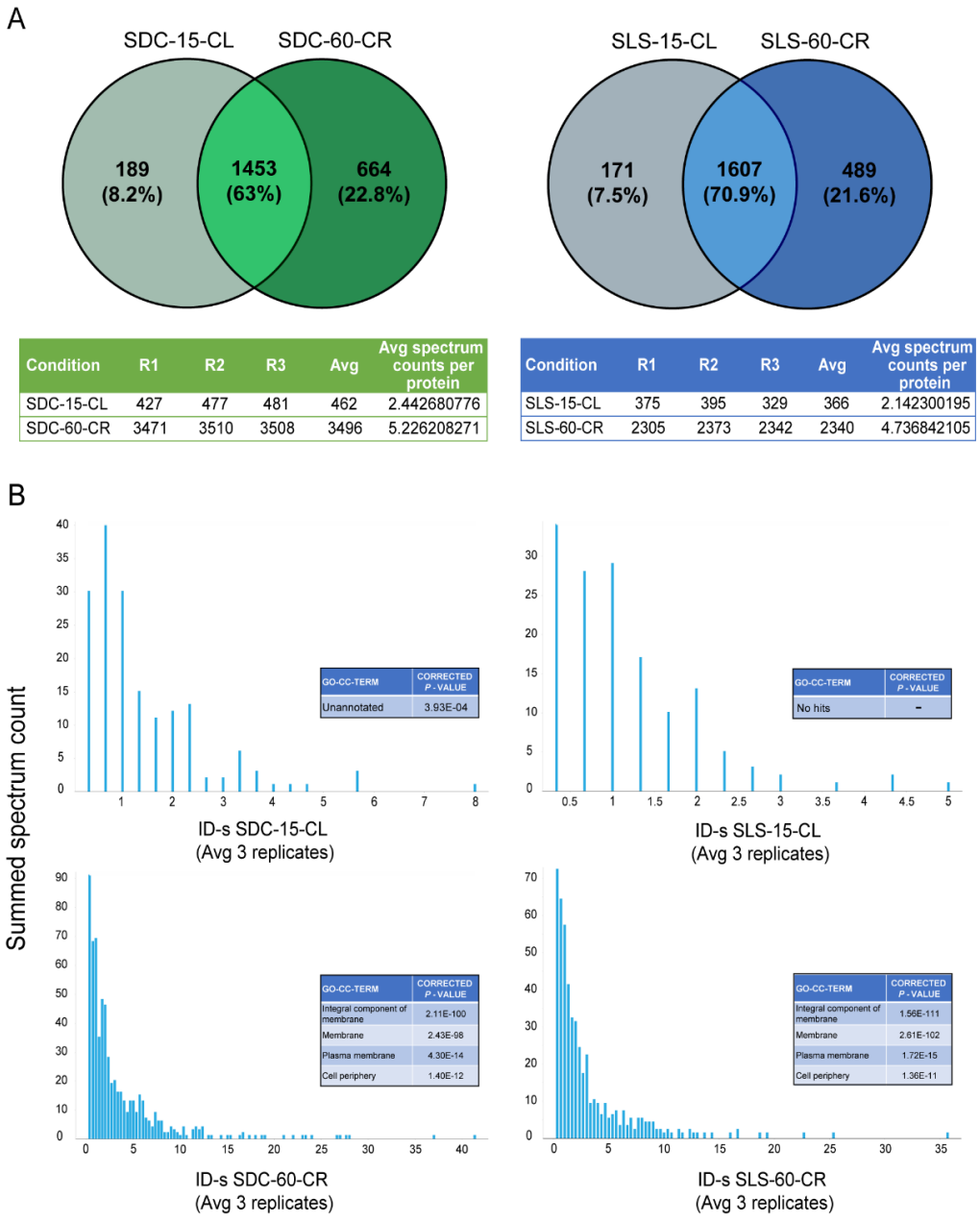



**Figure S-3** |Qualitative identification (ID) - based evaluation of the experimental conditions. The data was evaluated within Scaffold 4 (Proteome Software) and the protein/peptide false discovery rate (FDR) was set to 1% for protein/peptide identification-based assessment of the data. Within Scaffold, analysis of variance (ANOVA) was performed using Hochberg-Benjamini correction (adjusted *p*-value ≤ 0.01). **(A)** Distribution of summed up total spectral counts of eight different digestion conditions with the two best-solubilizing agents SDC and SLS (three replicates). **(B)** Total protein identifications shown as averaged result of three replicates per experimental condition. The error bars indicate the standard deviation. **(C)** Coverage of summed protein sequences in SDC buffer system. CR - crude lysate; CL - clear lysate; SDC - sodium deoxycholate; SLS - sodium lauroyl sarcosinate.




| Condition | R1 | R2 | R3 | Avg | Avg spectrum counts per protein |
|---|---|---|---|---|---|
| SDC-15-CL | 427 | 477 | 481 | 462 | 2.442680776 |
| SDC-60-CR | 3471 | 3510 | 3508 | 3496 | 5.226208271 |

| Condition | R1 | R2 | R3 | Avg | Avg spectrum counts per protein |
|---|---|---|---|---|---|
| SLS-15-CL | 375 | 395 | 329 | 366 | 2.142300195 |
| SLS-60-CR | 2305 | 2373 | 2342 | 2340 | 4.736842105 |

**Figure S-4** |ID-based comparison of identified proteins in both SDC and SLS treated crude or clear lysates. **(A)** Venn diagrams showing how many proteins were identified in total, and uniquely in crude and clear lysates in SDC and SLS treated samples. Accompanying tables represent the numbers of spectrum counts and average spectrum counts per protein for corresponding additionally identified proteins in either clear or crude-MS. **(B)** Histograms displaying summed spectrum counts and number of IDs of proteins detected in average in three replicates SDC or SLS-treated crude or clear lysates. In addition, significant enrichment of GO "Cellular component" terms is presented by small tables for the additionally identified proteins in crude-MS. The enrichment of GO terms was calculated using LAGO (Boyle et al., 2004) (https://go.princeton.edu/cgi-bin/LAGO). CR - crude lysate; CL - clear lysate; SDC - sodium deoxycholate; SLS - sodium lauroyl sarcosinate, CC-cellular component.

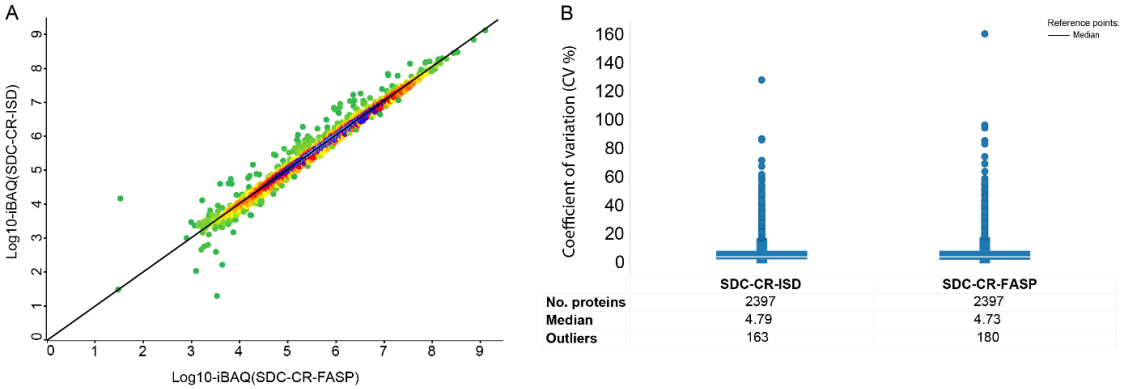



**Figure S-5** | Comparison of ISD and FASP digestion strategies in SDC buffer system; **(A)** Density plot correlation ($R^2$=0.975) between log10 iBAQ values calculated based on LFQ of the two digestion strategies presented in density plots (density descending from light blue to dark green color), **(B)** Median coefficient of variation (CV) in percentage based on three individual biological replicates of each digestion condition. SDC - sodium deoxycholate; ISD-in-solution digestion; FASP - filter-assisted sample preparation.

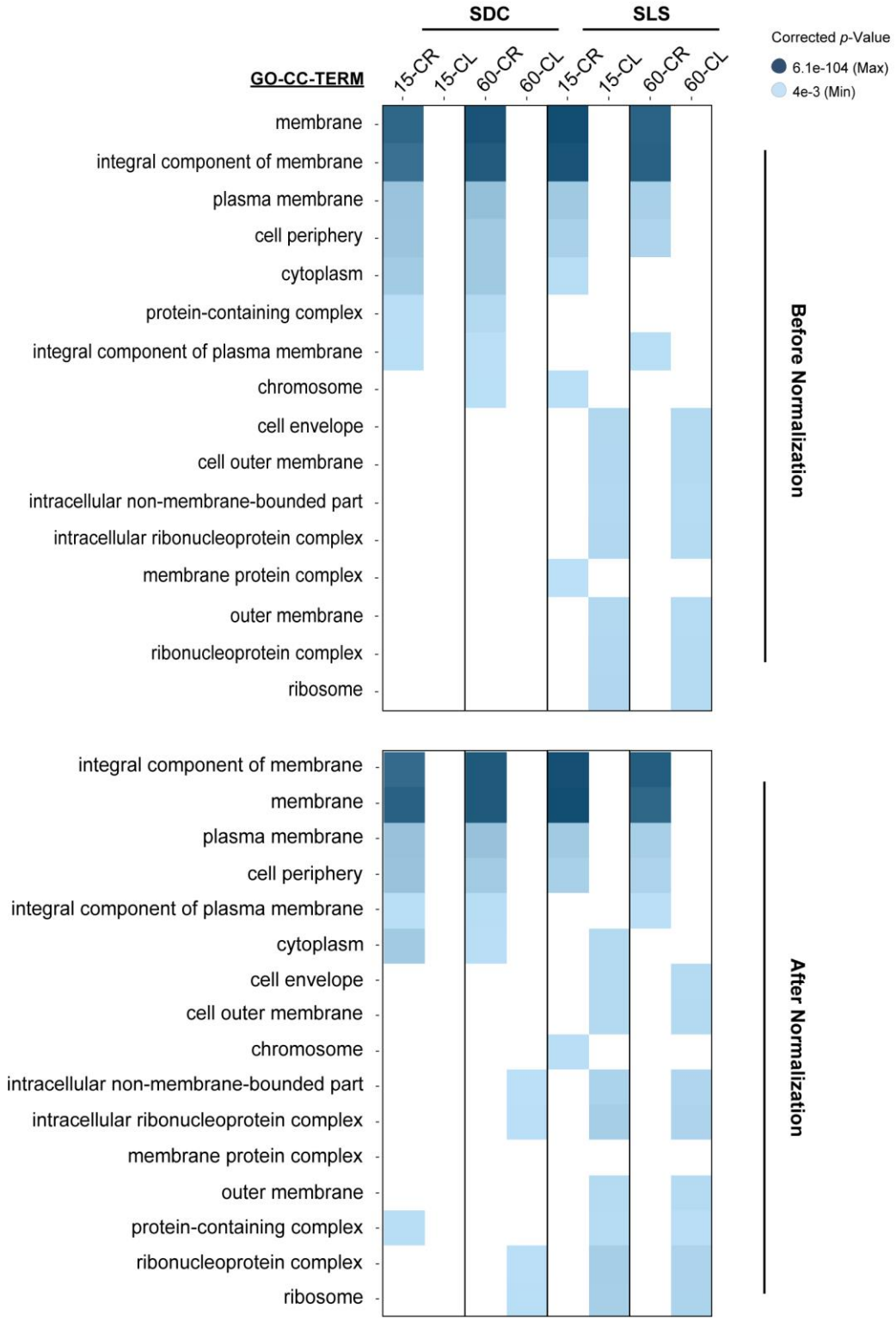



**Figure S-6 |** Heat map showing enriched "Cellular Compartment" GO terms before (upper) and after (lower) LFQ data normalization. Significantly enriched proteins from SDC and SLS-treated crude or clear lysates were loaded into LAGO (Boyle et al., 2004) and selected terms were plotted. The shown *p*-values were adjusted using Bonferroni correction.

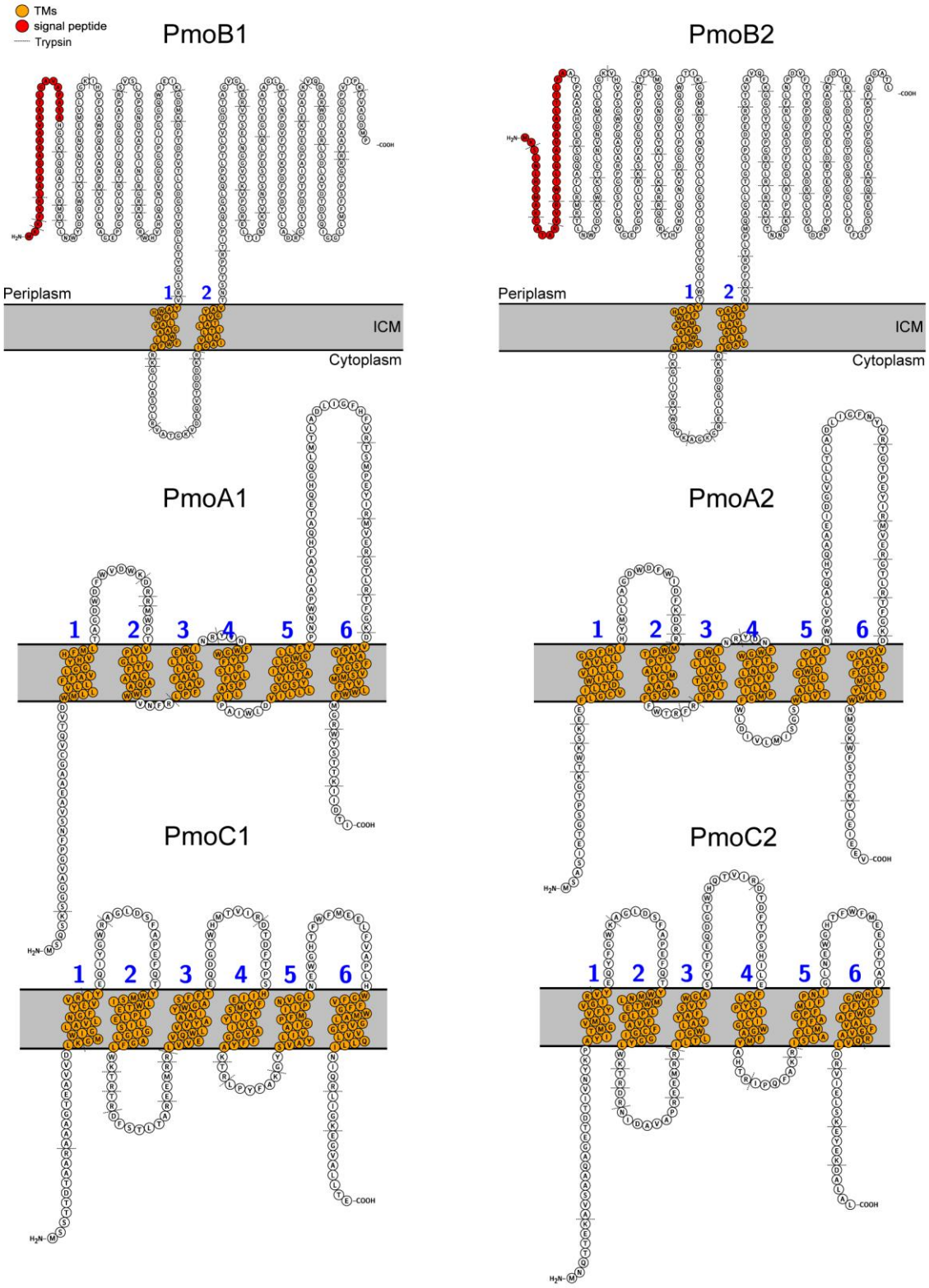
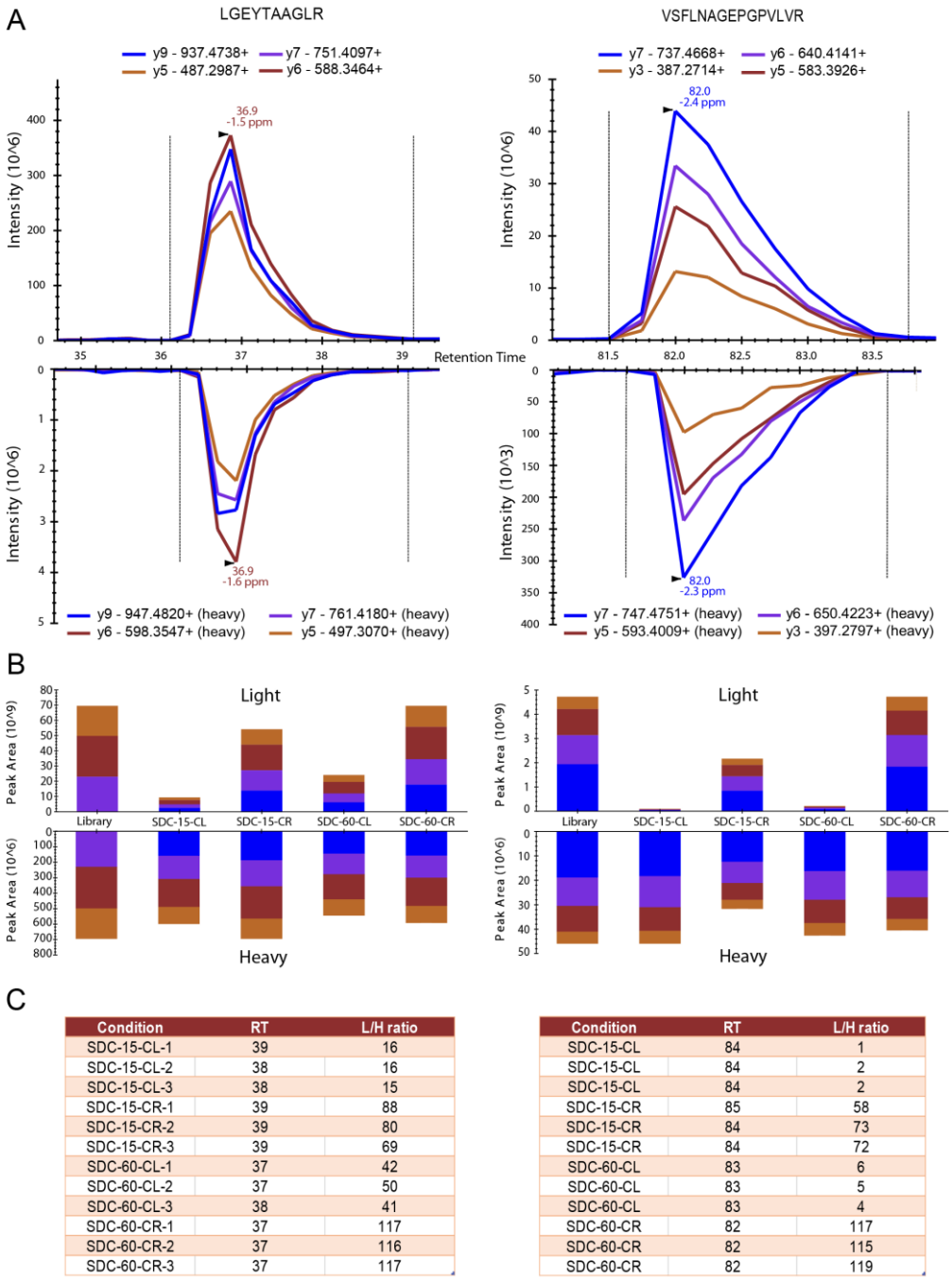



**Figure S-7** | Number and structure of transmembrane domains in pMMO1 and pMMO2 predicted with Protter version 1.0 (Omasits et al., 2014). Localization and protein topology of the pMMO subunits are visualized, with the transmembrane domains (TM) shown in orange. Number of TMs are marked in blue, tryptic digest sites are specified with black dashes, and signaling sites are highlighted in red.




C

| Condition | RT | L/H ratio |
|---|---|---|
| SDC-15-CL-1 | 39 | 16 |
| SDC-15-CL-2 | 38 | 16 |
| SDC-15-CL-3 | 38 | 15 |
| SDC-15-CR-1 | 39 | 88 |
| SDC-15-CR-2 | 39 | 80 |
| SDC-15-CR-3 | 39 | 69 |
| SDC-60-CL-1 | 37 | 42 |
| SDC-60-CL-2 | 37 | 50 |
| SDC-60-CL-3 | 38 | 41 |
| SDC-60-CR-1 | 37 | 117 |
| SDC-60-CR-2 | 37 | 116 |
| SDC-60-CR-3 | 37 | 117 |

| Condition | RT | L/H ratio |
|---|---|---|
| SDC-15-CL | 84 | 1 |
| SDC-15-CL | 84 | 2 |
| SDC-15-CL | 84 | 2 |
| SDC-15-CR | 85 | 58 |
| SDC-15-CR | 84 | 73 |
| SDC-15-CR | 84 | 72 |
| SDC-60-CL | 83 | 6 |
| SDC-60-CL | 83 | 5 |
| SDC-60-CL | 83 | 4 |
| SDC-60-CR | 82 | 117 |
| SDC-60-CR | 82 | 115 |
| SDC-60-CR | 82 | 119 |

**Figure S-8** | Targeted (PRM) quantification of PmoB1 protein analyzed by Skyline version 3.6.0. For PRM the protein digests were combined with equal volumes of heavy labeled peptide standards of PmoB1 protein. The ratio of PmoB1 endogenous peptides to their reference heavy peptides (L/H ratio) was calculated in SDC buffer system in crude and clear lysates after 15 or 60 min of incubation. **(A)** L/H ratios of peptide transition intensities of peptide 1 (R.LGEYTAAGLR.F) and peptide 2 (R.VSFLNAGEPGPVLVR.T) in SDC-60-CR condition. **(B)** L/H peak areas in SC2 crude versus clear lysates using SDC buffer, after 15 or 60 min of incubation. Samples were acquired in three replicates. (**C**) Tables presenting the retention times (RT) and L/H ratios of peptide 1 and peptide 2 in all the experimental conditions. Comparison of the individual L/H ratios of peptide 1 and peptide 2 showed that considered two peptides had highly similar L/H ratios only in 60-CR samples. CR - crude lysate; CL - clear lysate; SDC - sodium deoxycholate; SLS - sodium lauroyl sarcosinate.

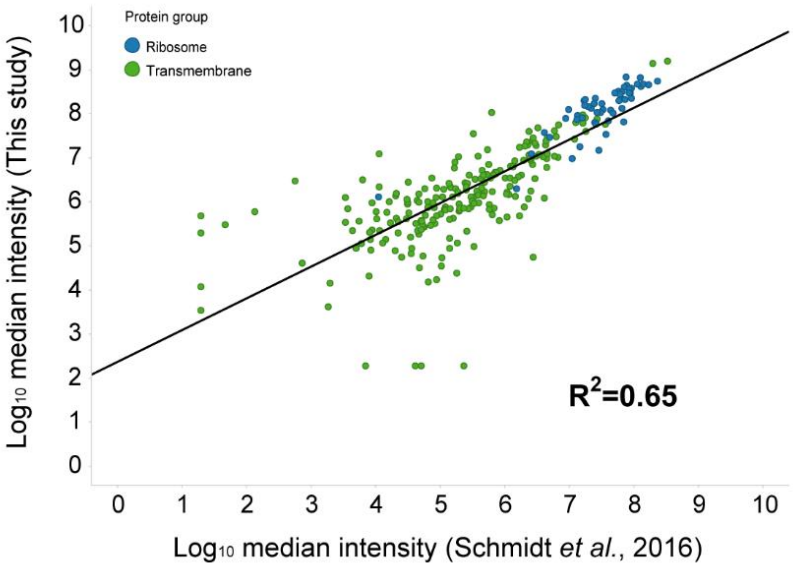



**Figure S-9** | Correlation between *E. coli* datasets from this study and Schmidt et al., 2016 (Schmidt et al., 2016). The median intensities of "transmembrane" and "ribosomal" proteins extracted from our dataset and that of Schmidt et al. showed a correlated quite well ($R^2$=0.65).

## 2.9. References


Ahrne, E., Molzahn, L., Glatter, T., and Schmidt, A. (2013). Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *Proteomics* 13(17)**,** 2567-2578.

Albrecht, A.M., and Vogel, H.J. (1964). Acetylornithine delta-transaminase. partial purification and repression behavior. *J Biol Chem* 239**,** 1872-1876.

Allen, G. (2016). Biogeochemistry: Rebalancing the global methane budget. *Nature* 538(7623)**,** 46-48.

Altelaar, A.F., and Heck, A.J. (2012). Trends in ultrasensitive proteomics. *Curr Opin Chem Biol* 16(1-2)**,** 206-213.

Arike, L., Valgepea, K., Peil, L., Nahku, R., Adamberg, K., and Vilu, R. (2012). Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli. J Proteomics* 75(17)**,** 5437-5448.

Baani, M., and Liesack, W. (2008). Two isozymes of particulate methane monooxygenase with different methane oxidation kinetics are found in *Methylocystis* sp strain SC2. *Proc Natl Acad Sci USA* 105(29)**,** 10203-10208.

Balasubramanian, R., Smith, S.M., Rawat, S., Yatsunyk, L.A., Stemmler, T.L., and Rosenzweig, A.C. (2010). Oxidation of methane by a biological dicopper centre. *Nature* 465(7294)**,** 115-119.

Bedard, C., and Knowles, R. (1989). Physiology, biochemistry, and specific inhibitors of $CH_4$, $NH_4^+$, and CO oxidation by methanotrophs and nitrifiers. *Microbiol Rev* 53(1)**,** 68-84.

Berven, F.S., Karlsen, O.A., Straume, A.H., Flikka, K., Murrell, J.C., Fjellbirkeland, A., et al. (2006). Analysing the outer membrane subproteome of *Methylococcus capsulatus* (Bath) using proteomics and novel biocomputing tools. *Arch Microbiol* 184(6)**,** 362-377.

Botsford, J.L., and Lewis, T.A. (1990). Osmoregulation in *Rhizobium meliloti*: Production of glutamic acid in response to osmotic stress. *Appl Environ Microbiol* 56(2)**,** 488-494.

Bowman, J. (2006). "The Methanotrophs—The Families Methylococcaceae and Methylocystaceae," in *The Prokaryotes: A Handbook on the Biology of Bacteria,* ed. M. Dworkin. (New York, USA: Springer Science+Business Media, LLC), 282-283.

Boyle, E.I., Weng, S.A., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., et al. (2004). GO::TermFinder - open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* 20(18)**,** 3710-3715.

Bremer, H., and Dennis, P.P. (2008). Modulation of chemical composition and other parameters of the cell at different exponential growth rates. *EcoSal Plus* 3(1).

Bruderer, R., Bernhardt, O.M., Gandhi, T., Xuan, Y., Sondermann, J., Schmidt, M., et al. (2017). Optimization of experimental parameters in data-independent mass spectrometry significantly increases depth and reproducibility of results. *Mol Cell Proteomics* 16(12)**,** 2296-2309.

Bussmann, I., Pester, M., Brune, A., and Schink, B. (2004). Preferential cultivation of type II methanotrophic bacteria from littoral sediments (Lake Constance). *FEMS Microbiol Ecol* 47(2)**,** 179-189.

Carpy, A., Krug, K., Graf, S., Koch, A., Popic, S., Hauf, S., et al. (2014). Absolute proteome and phosphoproteome dynamics during the cell cycle of *Schizosaccharomyces pombe* (Fission Yeast). *Mol Cell Proteomics* 13(8)**,** 1925-1936.

Chen, Y., Crombie, A., Rahman, M.T., Dedysh, S.N., Liesack, W., Stott, M.B., et al. (2010). Complete genome sequence of the aerobic facultative methanotroph *Methylocella silvestris* BL2. *J Bacteriol* 192(14)**,** 3840-3841.

Chen, Y., Dumont, M.G., Cebron, A., and Murrell, J.C. (2007). Identification of active methanotrophs in a landfill cover soil through detection of expression of 16S rRNA and functional genes. *Environ Microbiol* 9(11)**,** 2855-2869.

Chidambarampadmavathy, K., Obulisamy, P.K., and Heimann, K. (2015). Role of copper and iron in methane oxidation and bacterial biopolymer accumulation. *Eng Life Sci* 15(4)**,** 387-399.

Collins, B.C., Hunter, C.L., Liu, Y., Schilling, B., Rosenberger, G., Bader, S.L., et al. (2017). Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nat Commun* 8(1)**,** 291.

Condon, C., Liveris, D., Squires, C., Schwartz, I., and Squires, C.L. (1995). rRNA operon multiplicity in *Escherichia coli* and the physiological implications of rrn inactivation. *J Bacteriol* 177(14)**,** 4152-4156.

Conrad, R. (1996). Soil microorganisms as controllers of atmospheric trace gases ($H_2$, CO, $CH_4$, OCS, $N_2O$, and NO). *Microbiol Rev* 60(4)**,** 609-40.

Cox, J., and Mann, M. (2011). Quantitative, high-resolution proteomics for data-driven systems biology. *Annu Rev Biochem* 80**,** 273-299.

Csonka, L.N. (1989). Physiological and genetic responses of bacteria to osmotic stress. *Microbiol Rev* 53(1)**,** 121-147.

Culpepper, M.A., and Rosenzweig, A.C. (2012). Architecture and active site of particulate methane monooxygenase. *Crit Rev Biochem Mol Biol* 47(6)**,** 483-492.

Dam, B., Dam, S., Kim, Y., and Liesack, W. (2014). Ammonium induces differential expression of methane and nitrogen metabolism-related genes in *Methylocystis* sp. strain SC2. *Environ Microbiol* 16(10)**,** 3115-3127.

Dam, B., Dam, S., Kube, M., Reinhardt, R., and Liesack, W. (2012). Complete genome sequence of *Methylocystis* sp. strain SC2, an aerobic methanotroph with high-affinity methane oxidation potential. *J Bacteriol* 194(21)**,** 6008-6009.

Dedysh, S.N., Dunfield, P.F., Derakshani, M., Stubner, S., Heyer, J., and Liesack, W. (2003). Differential detection of type II methanotrophic bacteria in acidic peatlands using newly developed 16S rRNA-targeted fluorescent oligonucleotide probes. *FEMS Microbiol Ecol* 43(3)**,** 299-308.

Denman, K.L., and Brasseur, G. (2007). Couplings Between Changes in the Climate System and Biogeochemistry. *Climate Change 2007: The Physical Science Basis***,** 499-587.

Domon, B., and Aebersold, R. (2010). Options and considerations when selecting a quantitative proteomics strategy. *Nat Biotechnology* 28(7)**,** 710-721.

Dunfield, P.F., Khmelenina, V.N., Suzina, N.E., Trotsenko, Y.A., and Dedysh, S.N. (2003). *Methylocella silvestris* sp. nov., a novel methanotroph isolated from an acidic forest cambisol. *Int J Syst Evol Microbiol* 53(5)**,** 1231-1239.

Dunfield, P.F., Liesack, W., Henckel, T., Knowles, R., and Conrad, R. (1999). High-affinity methane oxidation by a soil enrichment culture containing a type II methanotroph. *Appl Environ Microbiol* 65(3)**,** 1009-1014.

Dunfield, P.F., Yimga, M.T., Dedysh, S.N., Berger, U., Liesack, W., and Heyer, J. (2002). Isolation of a *Methylocystis* strain containing a novel *pmoA*-like gene. *FEMS Microbiol Ecol* 41(1)**,** 17-26.

Epstein, W. (2003). The roles and regulation of potassium in bacteria. *Prog Nucleic Acid Res Mol Biol* 75**,** 293-320.

Erde, J., Loo, R.R., and Loo, J.A. (2014). Enhanced FASP (eFASP) to increase proteome coverage and sample recovery for quantitative proteomic experiments. *J Proteome Res* 13(4)**,** 1885-1895.

Fei, Q., Guarnieri, M.T., Tao, L., Laurens, L.M.L., Dowe, N., and Pienkos, P.T. (2014). Bioconversion of natural gas to liquid fuel: Opportunities and challenges. *Biotechnol Adv* 32(3)**,** 596-614.

Feldman, L.I., and Gunsalus, I.C. (1950). The occurrence of a wide variety of transaminases in bacteria. *J Biol Chem* 187(2)**,** 821-830.

Gallien, S., Duriez, E., Crone, C., Kellmann, M., Moehring, T., and Domon, B. (2012). Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Mol Cell Proteomics* 11(12)**,** 1709-1723.

Geiger, T., Wehner, A., Schaab, C., Cox, J., and Mann, M. (2012). Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol Cell Proteomics* 11(3)**,** M111.014050.

Glatter, T., Ahrne, E., and Schmidt, A. (2015). Comparison of different sample preparation protocols reveals lysis buffer-specific extraction biases in Gram-negative bacteria and human cells. *J Proteome Res* 14(11)**,** 4472-4485.

Glatter, T., Ludwig, C., Ahrne, E., Aebersold, R., Heck, A.J., and Schmidt, A. (2012). Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion. *J Proteome Res* 11(11)**,** 5145-5156.

Han, D., Link, H., and Liesack, W. (2017). Response of *Methylocystis* sp. strain SC2 to salt stress: physiology, global transcriptome, and amino acid profiles. *Appl Environ Microbiol*.

Hanson, R.S., and Hanson, T.E. (1996). Methanotrophic bacteria. *Microbiol Rev* 60(2)**,** 439-471.

Heyer, J., Galchenko, V.F., and Dunfield, P.F. (2002). Molecular phylogeny of type II methane-oxidizing bacteria isolated from various environments. *Microbiology* 148(Pt 9)**,** 2831-2846.

Hughes, C.S., Foehr, S., Garfield, D.A., Furlong, E.E., Steinmetz, L.M., and Krijgsveld, J. (2014). Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol* 10(10).

Kanamori, K., Weiss, R.L., and Roberts, J.D. (1989). Ammonia assimilation pathways in nitrogen-fixing *Clostridium kluyverii* and *Clostridium butyricum*. *J Bacteriol* 171(4)**,** 2148-2154.

Kao, W.C., Chen, Y.R., Yi, E.C., Lee, H., Tian, Q., Wu, K.M., et al. (2004). Quantitative proteomic analysis of metabolic regulation by copper ions in *Methylococcus capsulatus* (Bath). *J Biol Chem* 279(49)**,** 51554-51560.

Karthikeyan, O.P., Chidambarampadmavathy, K., Cires, S., and Heimann, K. (2015). Review of sustainable methane mitigation and biopolymer production. *Crit Rev Env Sci Tec* 45(15)**,** 1579-1610.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. *J Mol Evol* 16(2)**,** 111-120.

Knief, C. (2015). Diversity and habitat preferences of cultivated and uncultivated aerobic methanotrophic bacteria evaluated based on *pmoA* as molecular marker. *Front Microbi* 6, 1346

Knief, C., and Dunfield, P.F. (2005). Response and adaptation of different methanotrophic bacteria to low methane mixing ratios. *Environ Microbiol* 7(9)**,** 1307-1317.

Knief, C., Lipski, A., and Dunfield, P.F. (2003). Diversity and activity of methanotrophic bacteria in different upland soils. *Appl Environ Microbiol* 69(11)**,** 6703-6714.

Kolb, S., Knief, C., Dunfield, P.F., and Conrad, R. (2005). Abundance and activity of uncultured methanotrophic bacteria involved in the consumption of atmospheric methane in two forest soils. *Environ Microbiol* 7(8)**,** 1150-1161.

Kolb, S., Knief, C., Stubner, S., and Conrad, R. (2003). Quantitative detection of methanotrophs in soil by novel *pmoA*-targeted real-time PCR assays. *Appl Environ Microbiol* 69(5)**,** 2423-2429.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33(7)**,** 1870-1874.

Kuske, C.R., Banton, K.L., Adorada, D.L., Stark, P.C., Hill, K.K., and Jackson, P.J. (1998). Small-Scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. *Appl Environ Microbiol* 64(7)**,** 2463-2472.

Kyuma, T., Kimura, S., Hanada, Y., Suzuki, T., Sekimizu, K., and Kaito, C. (2015). Ribosomal RNA methyltransferases contribute to *Staphylococcus aureus* virulence. *FEBS Jl* 282(13)**,** 2570-2584.

Lau, M.C., Stackhouse, B.T., Layton, A.C., Chauhan, A., Vishnivetskaya, T.A., Chourey, K., et al. (2015). An active atmospheric methane sink in high Arctic mineral cryosols. *ISME Jl* 9(8)**,** 1904.

Lee, T.K., and Huang, K.C. (2013). The role of hydrolases in bacterial cell-wall growth. *Curr Opin Microbiol* 16(6)**,** 760-766.

Leon, I.R., Schwammle, V., Jensen, O.N., and Sprenger, R.R. (2013). Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol Cell Proteomics* 12(10)**,** 2992-3005.

Lieberman, R.L., and Rosenzweig, A.C. (2004). Biological methane oxidation: regulation, biochemistry, and active site structure of particulate methane monooxygenase. *Crit Rev Biochem Mol Biol* 39(3)**,** 147-164.

Lombardi, A. (2015). Metalloproteins: Simple structure, complex function. *Nat Chem Biol* 11(10)**,** 760-761.

Lovell, T., Liu, T., Case, D.A., and Noodleman, L. (2003). Structural, spectroscopic, and redox consequences of a central ligand in the FeMoco of nitrogenase: a density functional theoretical study. *J Am Chem Soc* 125(27)**,** 8377-8383.

Lu, F., He, P., Guo, M., Yang, N., and Shao, L. (2012). Ammonium-dependent regulation of aerobic methane-consuming bacteria in landfill cover soil by leachate irrigation. *J Environ Sci* 24(4)**,** 711-719.

Maass, S., and Becher, D. (2016). Methods and applications of absolute protein quantification in microbial systems. *J Proteomics* 136**,** 222-233.

Manza, L.L., Stamer, S.L., Ham, A.J.L., Codreanu, S.G., and Liebler, D.C. (2005). Sample preparation and digestion for proteomic analyses using spin filters. *Proteomics* 5(7)**,** 1742-1745.

Masuda, T., Saito, N., Tomita, M., and Ishihama, Y. (2009). Unbiased quantitation of *Escherichia coli* membrane proteome using phase transfer surfactants. *Mol Cell Proteomics* 8(12)**,** 2770-2777.

Matsen, J.B., Yang, S., Stein, L.Y., Beck, D., and Kalyuzhnaya, M.G. (2013). Global molecular analyses of methane metabolism in methanotrophic alphaproteobacterium, *Methylosinus trichosporium* OB3b. Part I: transcriptomic study. *Front Microbiol* 4.

Mohanty, S.R., Bodelier, P.L., Floris, V., and Conrad, R. (2006). Differential effects of nitrogenous fertilizers on methane-consuming microbes in rice field and forest soils. *Appl Environ Microbiol* 72(2)**,** 1346-1354.

Murrell, J.C., and Dalton, H. (1983). Ammonia assimilation in *Methylococcus capsulatus* (Bath) and other obligate methanotrophs. *J Gen Microbiol* 129(Apr)**,** 1197-1206.

Murrell, J.C., Gilbert, B., and McDonald, I.R. (2000a). Molecular biology and regulation of methane monooxygenase. *Arch Microbiol* 173(5-6)**,** 325-332.

Murrell, J.C., and Jetten, M.S. (2009). The microbial methane cycle. *Environ Microbiol Rep* 1(5), 279-284.
Murrell, J.C., McDonald, I.R., and Gilbert, B. (2000b). Regulation of expression of methane monooxygenases by copper ions. *Trends Microbiol* 8(5), 221-225.
Myhre, G., Shindell, D., Breon, F.M., Collins, W., Fuglestvedt, J., Huang, J.P., et al. (2014). Anthropogenic and Natural Radiative Forcing. *Climate Change 2013: The Physical Science Basis*, 659-740.
Nagaraj, N., Wisniewski, J.R., Geiger, T., Cox, J., Kircher, M., Kelso, J., et al. (2011). Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol* 7, 548.
Nauer, P.A., Dam, B., Liesack, W., Zeyer, J., and Schroth, M.H. (2012). Activity and diversity of methane-oxidizing bacteria in glacier forefields on siliceous and calcareous bedrock. *Biogeosciences* 9(6), 2259-2274.
Nyerges, G., Han, S.K., and Stein, L.Y. (2010). Effects of ammonium and nitrite on growth and competitive fitness of cultivated methanotrophic bacteria. *Appl Environ Microbiol* 76(16), 5648-5651.
Nyerges, G., and Stein, L.Y. (2009). Ammonia cometabolism and product inhibition vary considerably among species of methanotrophic bacteria. *FEMS Microbiol Lett* 297(1), 131-136.
Ogahara, T., Ohno, M., Takayama, M., Igarashi, K., and Kobayashi, H. (1995). Accumulation of glutamate by osmotically stressed *Escherichia coli* Is dependent on pH. *J Bacteriol* 177(20), 5987-5990.
Omasits, U., Ahrens, C.H., Muller, S., and Wollscheid, B. (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 30(6), 884-886.
Oren, A., Heldal, M., Norland, S., and Galinski, E.A. (2002). Intracellular ion and organic solute concentrations of the extremely halophilic bacterium *Salinibacter ruber*. *Extremophiles : life under extreme conditions* 6(6), 491-498.
Ottman, N., Huuskonen, L., Reunanen, J., Boeren, S., Klievink, J., Smidt, H., et al. (2016). Characterization of outer membrane proteome of *Akkermansia muciniphila* reveals sets of novel proteins exposed to the human intestine. *Front Microbiol* 7, 1157.
Patel, V.J., Thalassinos, K., Slade, S.E., Connolly, J.B., Crombie, A., Murrell, J.C., et al. (2009). A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *J Proteome Res* 8(7), 3752-3759.
Pengpeng, W., and Tan, Z. (2013). Ammonia assimilation in rumen bacteria: a review. *Anim Biotechnol* 24(2), 107-128.
Picotti, P., Clement-Ziza, M., Lam, H., Campbell, D.S., Schmidt, A., Deutsch, E.W., et al. (2013). A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature* 494(7436), 266-270.
Piuri, M., Sanchez-Rivas, C., and Ruzal, S.M. (2005). Cell wall modifications during osmotic stress in *Lactobacillus casei*. *J Appl Microbiol* 98(1), 84-95.
Proc, J.L., Kuzyk, M.A., Hardie, D.B., Yang, J., Smith, D.S., Jackson, A.M., et al. (2010). A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin. *J Proteome Res* 9(10), 5422-5437.
Radajewski, S., Webster, G., Reay, D.S., Morris, S.A., Ineson, P., Nedwell, D.B., et al. (2002). Identification of active methylotroph populations in an acidic forest soil by stable-isotope probing. *Microbiology* 148(Pt 8), 2331-2342.
Rauniyar, N., and Yates, J.R., 3rd (2014). Isobaric labeling-based relative quantification in shotgun proteomics. *J Proteome Res* 13(12), 5293-5309.
Ricke, P., Erkel, C., Kube, M., Reinhardt, R., and Liesack, W. (2004). Comparative analysis of the conventional and novel *pmo* (particulate methane monooxygenase) operons from *Methylocystis* strain SC2. *Appl Environ Microbiol* 70(5), 3055-3063.
Rosenzweig, A.C. (2008). The metal centres of particulate methane mono-oxygenase. *Biochem Soc Trans* 36(6), 1134-1137.
Scheutz, C., Kjeldsen, P., Bogner, J.E., De Visscher, A., Gebert, J., Hilger, H.A., et al. (2009). Microbial methane oxidation processes and technologies for mitigation of landfill gas emissions. *Waste Manage Res* 27(5), 409-455.
Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrne, E., Volkmer, B., Callipo, L., et al. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nat Biotechnol* 34(1), 104-110.
Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., et al. (2011). Global quantification of mammalian gene expression control. *Nature* 473(7347), 337-342.
Semrau, J.D., DiSpirito, A.A., and Yoon, S. (2010). Methanotrophs and copper. *FEMS Microbiol Rev* 34(4), 496-531.
Sielaff, M., Kuharev, J., Bohn, T., Hahlbrock, J., Bopp, T., Tenzer, S., et al. (2017). Evaluation of FASP, SP3, and iST protocols for proteomic sample preparation in the low microgram range. *J Proteome Res*.

Silva, J.C., Gorenstein, M.V., Li, G.Z., Vissers, J.P., and Geromanos, S.J. (2006). Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics* 5(1)**,** 144-156.

Simicevic, J., Schmid, A.W., Gilardoni, P.A., Zoller, B., Raghav, S.K., Krier, I., et al. (2013). Absolute quantification of transcription factors during cellular differentiation using multiplexed targeted proteomics. *Nat Methods* 10(6)**,** 570-576.

Simpson, D.M., and Beynon, R.J. (2012). QconCATs: design and expression of concatenated protein standards for multiplexed protein quantification. *Anal Bioanal Chem* 404(4)**,** 977-989.

Sleator, R.D., and Hill, C. (2002). Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiol Rev* 26(1)**,** 49-71.

Solomon, T. (2001). The definition and unit of ionic strength. *J Chem Educ* 78(12)**,** 1691-1692.

Soufi, B., Krug, K., Harst, A., and Macek, B. (2015). Characterization of the *E. coli* proteome and its modifications during growth and ethanol stress. *Front Microbiol* 6**,** 103.

Staples, C.R., Lahiri, S., Raymond, J., Von Herbulis, L., Mukhophadhyay, B., and Blankenship, R.E. (2007). Expression and association of group IV nitrogenase NifD and NifH homologs in the non-nitrogen-fixing archaeon *Methanocaldococcus jannaschii*. *J Bacteriol* 189(20)**,** 7392-7398.

Stein, L.Y., Real, R., and Dunfield, P. F. (2012). Aerobic Methanotrophy and Nitrification: Processes and Connections. *In eLS, John Wiley & Sons,* Ltd: Chichester.

Steinkamp, R., Zimmer, W., and Papen, H. (2001). Improved method for detection of methanotrophic bacteria in forest soils by PCR. *Curr Microbiol* 42(5)**,** 316-322.

Urmann, K., Lazzaro, A., Gandolfi, I., Schroth, M.H., and Zeyer, J. (2009). Response of methanotrophic activity and community structure to temperature changes in a diffusive $CH_4/O_2$ counter gradient in an unsaturated porous medium. *FEMS Microbiol Ecol* 69(2)**,** 202-212.

Vit, O., and Petrak, J. (2017). Integral membrane proteins in proteomics. How to break open the black box? *J Proteomics* 153**,** 8-20.

Vizcaino, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 44(22)**,** 11033.

Vorholt, J.A. (2002). Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. *Arch Microbiol* 178(4)**,** 239-249.

Vorholt, J.A., Marx, C.J., Lidstrom, M.E., and Thauer, R.K. (2000). Novel formaldehyde-activating enzyme in *Methylobacterium extorquens* AM1 required for growth on methanol. *J Bacteriol* 182(23)**,** 6645-6650.

Vorobev, A., Jagadevan, S., Jain, S., Anantharaman, K., Dick, G.J., Vuilleumier, S., et al. (2014). Genomic and transcriptomic analyses of the facultative methanotroph *Methylocystis* sp. strain SB2 grown on methane or ethanol. *Appl Environ Microbiol* 80(10)**,** 3044-3052.

Walther, T.C., and Mann, M. (2010). Mass spectrometry-based proteomics in cell biology. *J Cell Biol* 190(4)**,** 491-500.

Wartiainen, I., Hestnes, A.G., McDonald, I.R., and Svenning, M.M. (2006). *Methylocystis rosea* sp. nov., a novel methanotrophic bacterium from Arctic wetland soil, Svalbard, Norway (78 °N). *Int J Syst Evol Microbiol* 56(3)**,** 541-547.

Whittenbury, R., and Dalton, H. (1981). "The Methylotrophic Bacteria," in *The Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria,* eds. M.P. Starr, H. Stolp, H.G. Trüper, A. Balows & H.G. Schlegel. (Berlin, Heidelberg: Springer Berlin Heidelberg), 894-902.

Wisniewski, J.R., Hein, M.Y., Cox, J., and Mann, M. (2014). A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. *Mol Cell Proteomics* 13(12)**,** 3497-3506.

Wisniewski, J.R., Ostasiewicz, P., and Mann, M. (2011). High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers. *J Proteome Res* 10(7)**,** 3040-3049.

Wisniewski, J.R., Zougman, A., and Mann, M. (2009a). Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. *J Proteome Res* 8(12)**,** 5674-5678.

Wisniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009b). Universal sample preparation method for proteome analysis. *Nat Methods* 6(5)**,** 359-362.

Yimga, M.T., Dunfield, P.F., Ricke, P., Heyer, H., and Liesack, W. (2003). Wide distribution of a novel *pmoA*-like gene copy among type II methanotrophs, and its expression in *Methylocystis* strain SC2. *Appl Environ Microbiol* 69(9)**,** 5593-5602.

Zahn, J.A., and DiSpirito, A.A. (1996). Membrane-associated methane monooxygenase from *Methylococcus capsulatus* (Bath). *J Bacteriol* 178(4)**,** 1018-1029.

Zeiler, M., Straube, W.L., Lundberg, E., Uhlen, M., and Mann, M. (2012). A protein epitope signature tag (PrEST) library allows SILAC-based absolute quantification and multiplexed determination of protein copy numbers in cell lines. *Mol Cell Proteomics* 11(3).

# Chapter 3:

## Efficient Tandem LysC/Trypsin Digestion in Detergent Conditions

*Proteomics* 2019, 19(20), e1900136
https://doi.org/10.1002/pmic.201900136

**Anna Hakobyan**[1], Martin Bernd Schneider[2, 4], Werner Liesack[1, 3], Timo Glatter[2]*

*Corresponding author

Chapter 3 is written in research manuscript style. It was published as research article in *Proteomics* in September, 2019. My contribution to the chapter 3 involved the laboratory experiments, proteomics data analysis and the writing and revision of corresponding parts of the manuscript, including the design of the figures.

[1]Research Group of Methanotrophic Bacteria and Environmental Genomics/Transcriptomics, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, D-35043 Marburg, Germany

[2]Core Facility for Mass-Spectrometry and Proteomics, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, D-35043 Marburg, Germany

[3]Center for Synthetic Microbiology (SYNMIKRO), Philipps-Universitaẗ Marburg, Karl-von-Frisch-Str. 16, D-35043 Marburg, Germany

[4]Genomics and Proteomics Core Facility, German Cancer Research Center, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

# 3 Efficient Tandem LysC/Trypsin Digestion in Detergent Conditions

## 3.1. Abstract

All shotgun proteomics experiments rely on efficient proteolysis steps for sensitive peptide/protein identification and quantification. Previous reports suggest that the sequential tandem LysC/trypsin digest yields higher recovery of fully tryptic peptides than single-tryptic proteolysis. Based on the previous studies, it is assumed that the advantageous effect of tandem proteolysis requires a high sample denaturation state for the initial LysC digest. Therefore, to date, all systematic assessments of LysC/trypsin proteolysis are done in chaotropic environments such as urea. Here, sole trypsin is compared with LysC/trypsin and it is shown that tandem digestion can be carried with high efficiency in Mass Spectrometry-compatible detergents, thereby resulting in higher quantitative yields of fully cleaved peptides. It is further demonstrated that higher cleavage efficiency of tandem digests has a positive impact on absolute protein quantification using intensity-based absolute quantification (iBAQ) values. The results of the examination of divergent urea tandem conditions imply that beneficial effects of the initial LysC digest do not depend on the sample denaturation state, but, are mainly caused by different target specificities of LysC and trypsin. The observed detergent compatibility enables tandem digestion schemes to be implemented in efficient cellular solubilization proteomics procedures without the need for buffer exchange to chaotropic environments.

## 3.2. Introduction

The endopeptidase trypsin is the most widely used protease in shotgun MS experiments (Hervey et al., 2007; Chen et al., 2008; Wisniewski et al., 2009; Glatter et al., 2012). This is due to the high efficiency and specificity of trypsin in cleaving the peptide-bonds on the C-terminal side of lysine and arginine residues, thereby generating peptides with suitable properties for MS analysis. Efficient tryptic cleavage is characterized by the generation of the maximum number of most intense fully cleaved peptides (FCPs), while keeping the miscleavage events minimal. One suggested strategy to improve digestion

efficiency is to include a lysine-specific endopeptidase in a sequential or tandem digest together with trypsin to obtain a tandem proteolysis workflow (further referred to as LT) (Glatter et al., 2012; Wisniewski and Mann, 2012; Chiva et al., 2014; Betancourt et al., 2018). Due to its higher stability than trypsin, LysC is utilized in LT strategies as the first protease in a high denaturation environment, followed by trypsin digest in diluted denaturant concentrations (Glatter et al., 2012). Urea is often used to create such differential denaturation state across the LT digestion procedure, which can result in superior cleavage efficiency compared to sole trypsinization (Glatter et al., 2012). However, it is well known that the use of urea may have negative side effects due to an increased protein carbamylation at higher temperatures (Poulsen et al., 2013). Furthermore, urea creates biases when extracting proteins from intact cells, especially in the recovery of membrane proteins (Glatter et al., 2015), and when heat exposure is a critical parameter for protein extraction (Hakobyan et al., 2018). Alternatively, different detergents are utilized in in-solution digestion (ISD) schemes with less solubilization bias (Glatter et al., 2015) and protein-modifying properties than urea. However, the assessment of LT digests in detergent buffers using state-of-the-art quantitative proteomics approaches is still missing.

The general advantage of detergents in one-step ISD routines prompted us to compare the efficiency of tandem LT and single-trypsin digest in MS-compatible detergent buffer systems. We decided to focus our comparative study on the MS-compatible and budget-friendly products, like sodium deoxycholate (DOC) and sodium lauroyl sarcosinate (SLS). While DOC is broadly used for cellular proteomics workflows (Kulak et al., 2014; Glatter et al., 2015; Schmidt et al., 2016), SLS is not frequently applied. Recent studies, however, have shown that SLS can be efficiently implemented in the ISD strategies of comparative proteomics studies (Hakobyan et al., 2018; Sander et al., 2019).

## 3.3. Results and discussion

Our research was strictly focused on the digestion effect, with the aim to minimize solubilization-mediated differences between the detergents. Therefore, we first extracted proteins from frozen *Escherichia coli* cell pellets using SLS, followed by acetone

protein precipitation and reconstitution of the protein pellets in the detergent buffers. Then protein digestion in LT and single-trypsin formats (for more details see Supporting Information) was carried out in constant detergent concentration of 0.5% (**Figure 1A**); thereby anticipating that potential differences between LT and single-trypsin digest could not be the results of a differential denaturation state of the sample as assumed for urea buffers.

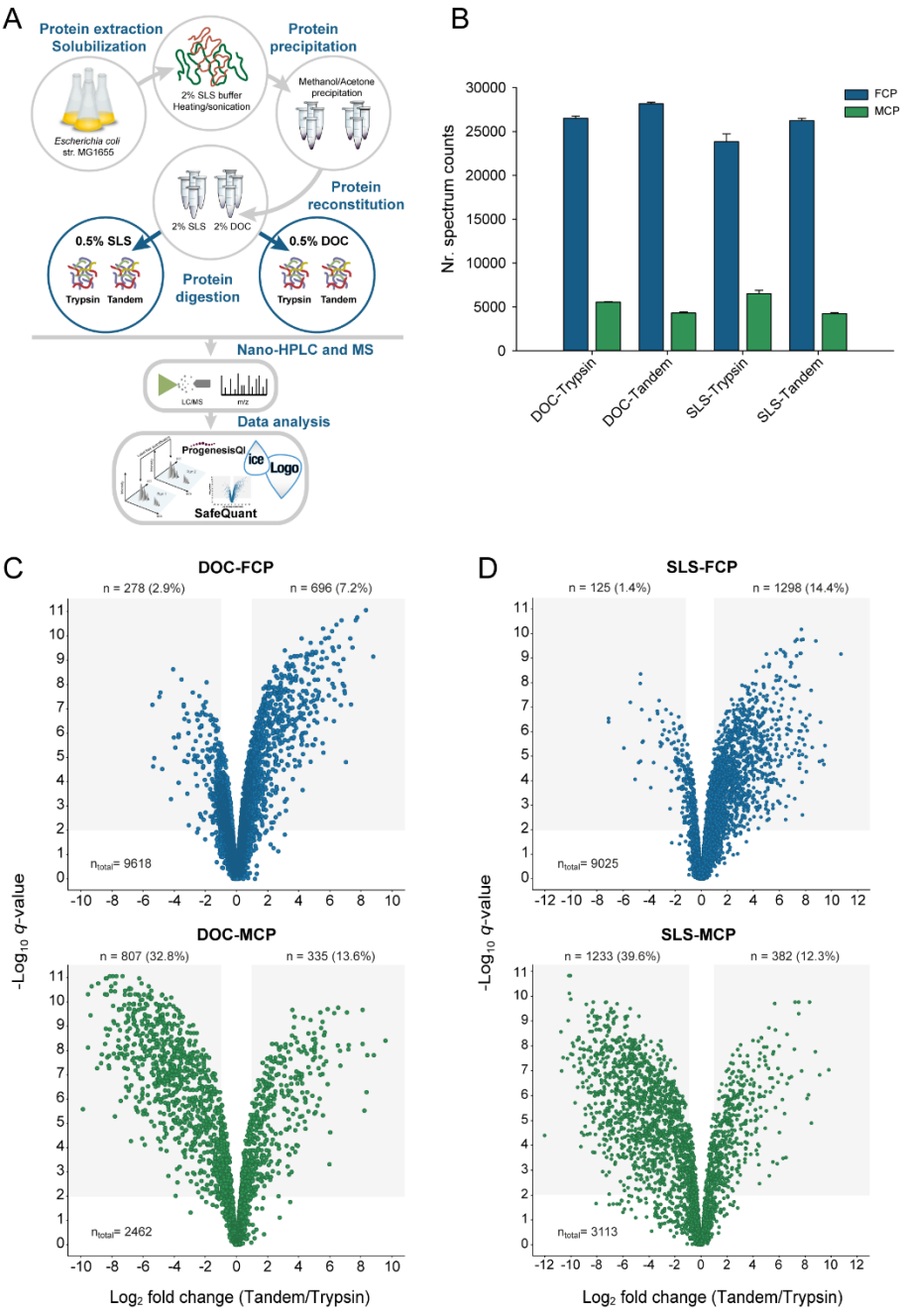



**Figure 1** | **(A)** Experimental design to assess the efficiency of tandem LysC/trypsin (LT) versus single-trypsin digest in sodium lauroyl sarcosinate (SLS) and sodium deoxycholate (DOC). *E. coli* strain MG1655 cell pellets were lysed using SLS buffer, acetone/methanol precipitated and reconstituted in DOC or SLS.

Proteins were then digested by LysC/trypsin (LT) or sole trypsin in 0.5% detergent. Following digestion, the peptide samples were analyzed using LC–MS with subsequent label-free quantification (LFQ). **(B)** Bar chart representing the number of fully (FCP) and miscleaved peptide (MCP) spectra detected upon applying the different digestion workflows. **(C)** Volcano plot comparing the relative enrichment of FCPs and MCPs as a result of the LT and single-trypsin workflows in DOC**. (D)** The same as in (C) but in SLS buffer.

In a first attempt to investigate the differences between LT and single-trypsin digest in detergent buffers, we extracted total spectrum counts from the data set, grouped the identified spectra into fully, and miscleaved (**Figure 1B and Table S-1**, *Supporting Information*). As expected, the number of fully cleaved peptide (FCP) spectra were the most prominently identified. They exceeded the number of miscleaved peptide (MCP) spectrum counts by at least 3.5-fold. Both LT workflows identified more FCP spectra than single-trypsin digestion. Thus, in both LT workflows, the MCP counts concurrently decreased with the increase in FCP spectra. Nonetheless, we did not observe a major influence of the digestion scheme on the protein identification rate in the moderately complex *E. coli* lysate (**Figure S-1**, *Supporting Information*).

In order to compare the workflow efficiencies in greater detail, we performed label-free quantification (LFQ) of the different digestion strategies. Out of a total of 9618 FCPs quantified in DOC (**Table S-2**, *Supporting Information*) and 9025 FCPs quantified in SLS (**Table S-3**, *Supporting Information*), FCP levels were consistently higher in all LT strategies than in single-trypsin workflows (**Figure 1C and 1D**). In contrast to SLS-trypsin, 1298 FCPs were significantly enriched in SLS-LT (**Table S-3**, *Supporting Information*). This particular enrichment was the most striking difference between LT and single-trypsin digests and corresponded to 14.4% of all quantified FCPs. On the contrary, the MCPs ($n$ = 2462 for DOC and $n$ = 3113 for SLS) showed reversed asymmetric distribution, with MCPs being more abundantly represented after single-trypsin digest (**Figure 1D**).

Next, we assessed how the digestion strategies affect quantification reproducibility and variation. The box plots in **Figure 2A and 2B** illustrate median FCP-coefficient of variation (CV) ranging from 8.9% to 13%, which are consistently lower than the CVs determined for MCPs (10.7–18%). In SLS-LT, CVs were lower than in the corresponding single-trypsin workflow. In DOC, however, we noted that LT exhibited a higher signal variation than single trypsin. This may be caused by the influence of DOC on LysC activity, for which critical digestion parameters are less well studied than for trypsin.

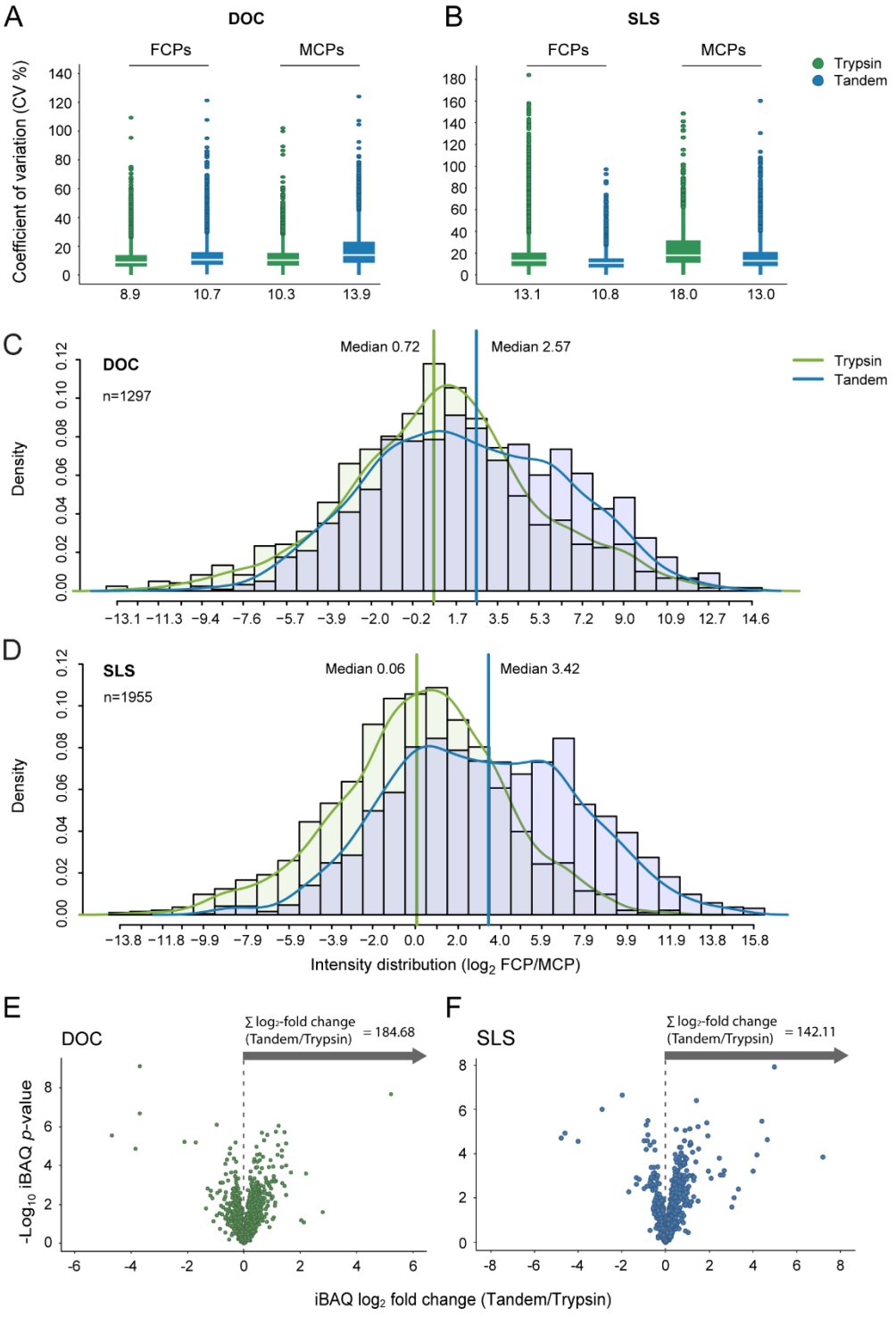



**Figure 2** | **(A)** and **(B)** Box plots representing the coefficient of variation (CV) of FCPs and MCPs for LT (tandem) and single-trypsin digestion in DOC **(A)** and in SLS **(B)**. Median CVs are shown. **(C)** and **(D)** $Log_2$-ratio distribution of FCPs with detected MCP counterpart in DOC **(C)** and in SLS **(D).** Median FCP/MCP intensity ratios are indicated in the graph. **(E)** and **(F)** Volcano plot illustrating the $log_2$-fold change ratio of iBAQ values upon LFQ analysis of LT and single-trypsin digestion data sets in DOC **(E)** and in SLS **(F)**. The gray arrows indicate the direction of the sum $log_2$-fold change in iBAQ values.

To further monitor the cleavage efficiency of the proteolytic workflows, we identified all MCPs with a corresponding FCP counterpart and calculated the LFQ intensity ratio of the FCP/MCP pairs (**Figure 2C and 2D**). The distribution of the FCP/MCP ratios showed that the FCP/MCP LT median in both DOC and SLS is shifted to higher values than for single-trypsin digestion.

Based on our observations, we hypothesize that a high FCP rate has an effect on absolute protein quantification. Accordingly, absolute quantification strategies that exclusively incorporate FCP intensities to calculate protein quantities, such as the intensity-based absolute quantification index (iBAQ (Schwanhausser et al., 2011)), may get compressed when a digestion strategy generates reduced FCP intensities. In order to test our assumption, we generated iBAQ values from LT and single-trypsin data sets and found that indeed LT generates higher iBAQ values than single-trypsin digest (**Figure S-2, Tables S-4 and S-5**, *Supporting Information*). This finding is further illustrated by the LT-directed shape of the iBAQ volcano plots and the associated sum of all $log_2$-fold change ratios of iBAQ values biased toward LT in both SLS and DOC (**Figure 2E and 2F**). Therefore, the more comprehensive cleavage makes LT to the potentially more reliable procedure in representing absolute protein abundances via indices like iBAQ than single-trypsin strategies.

Analysis of the sequence context of occurring cleavage and miscleavage events provides valuable information on workflow specificities. For this analysis, we first assessed the global MCP target residues and created iceLogo plots (Colaert et al., 2009) for all detected MCPs. In agreement with other studies (Glatter et al., 2012) we observed that the pool of totally detected MCPs is characterized by the frequent occurrence of acidic residues adjacent to scissile lysine and arginine (**Figure S-3**, *Supporting Information*). In addition, proline in +1 position and basic residues at both sides of cleavage occur. This cleavage consensus was observed for both detergent buffers.

Interestingly, we found a highly similar cleavage context in FCPs originating from tandem digests (**Figure 3A and 3B**). Here LT-FCPs clearly showed a particular enrichment of surrounding acidic amino acids as well as proline in +1, but also leucine in –1 positions next to cleaved lysine residues. This enrichment pattern was observed in both DOC and SLS (**Figure 3A**). Compared to LT-FCPs, no obvious cleavage preference was observed for

trypsin-FCPs, except for an expected overrepresentation of arginine and a decrease in lysine.

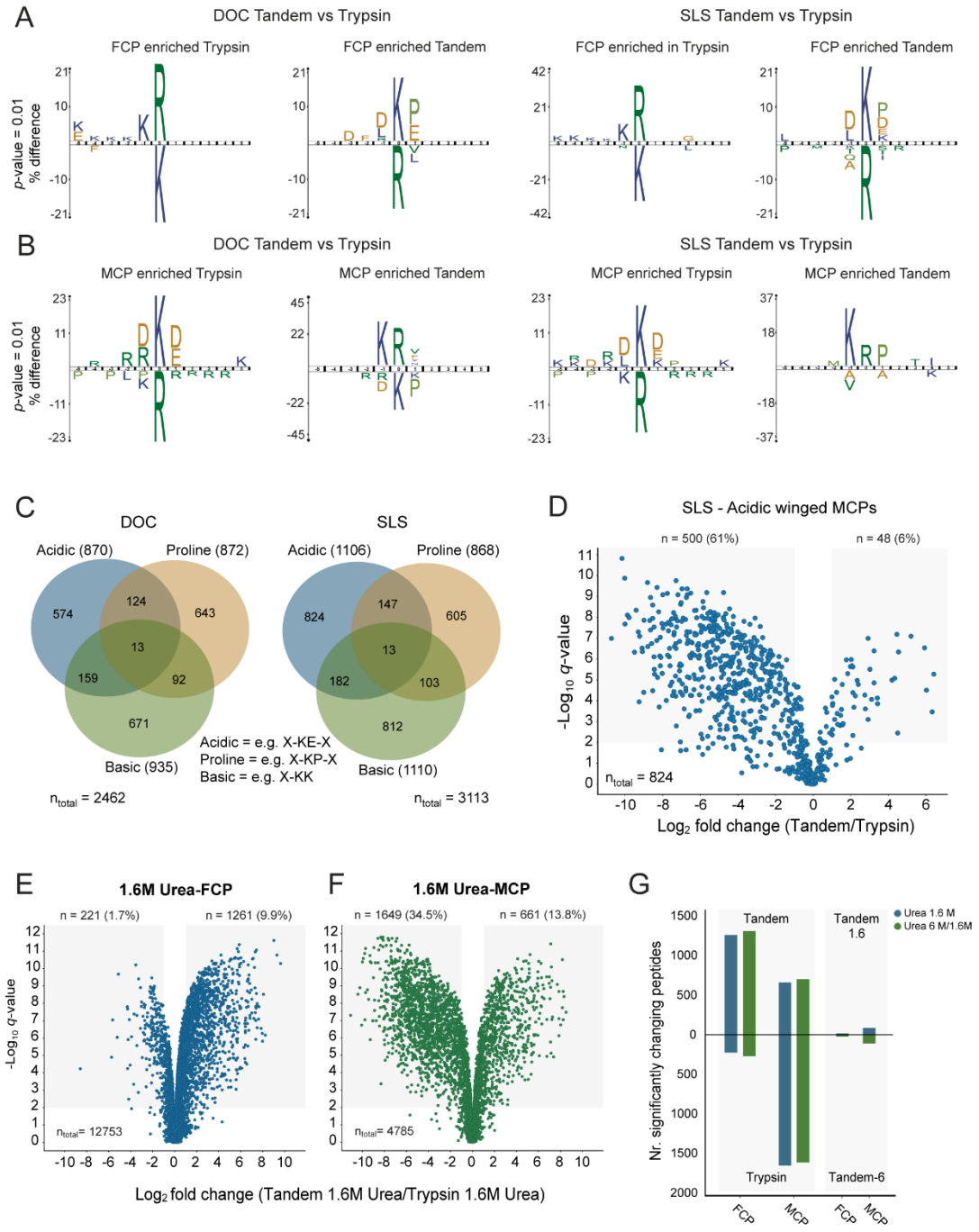



**Figure 3** | IceLogo plots showing enriched sequences at **(A)** fully cleaved and **(B)** miscleaved peptide stretches. Plots were generated using significantly enriched peptides from LT versus trypsin LFQ comparison in DOC and SLS. **(C)** Venn diagrams showing the overlap between main classes of MCPs. **(D)** Volcano plot illustrating the $\log_2$-fold change in all MCPs with acidic amino acids surrounding the miscleavage side, which resulted from LT versus single-trypsin digests in SLS. **(E)** and **(F)** Volcano plots of all FCPs **(E)** or MCPs **(F)**, resulting from a LT versus trypsin digestion with all tandem steps including LysC digest carried out in 1.6 m urea. **(G)** Bar chart representing the number of significantly enriched FCPs and MCPs from tandem 1.6 m urea, tandem 6 m urea, and single-trypsin 1.6 m urea digestion comparisons.

This finding was anticipated due to the well-known target selectivity toward arginine residues in single-trypsin digests relative to tandem digests with an initial LysC treatment (Raijmakers et al., 2010; Glatter et al., 2012; Giansanti et al., 2016). Nonetheless, consistent with the enriched LT-FCP motifs, trypsin-MCPs (**Figure 3B**) exhibited more acidic and basic amino acids surrounding the target side. This corresponds well to results obtained previously for urea–LT (Glatter et al., 2012) and LysC cleavage (Giansanti et al., 2016).

Following the extraction of peptides matching to all major MCP classes, we found that such peptides are highly enriched in the total MCP pools of the DOC (92%) and SLS (86%) experiments (**Figure 3C**). The Venn diagram shows that the numbers for each class are homogenously distributed, but with partial overlaps. In particular, we found that acidic-MCPs were considerably enriched in all trypsin conditions. This is illustrated in **Figure 3D** showing enriched "acidic" MCPs after single-trypsin digest in SLS as the most affected MCP class (**Figure S-4A**, *Supporting Information*). In addition, we observed an increased proteolytic cleavage after proline in LT digests (**Figure S-4**, *Supporting Information*), which agrees with a previous report (Glatter et al., 2012).

Our data demonstrate that LT-tandem digests perform efficiently in constant detergent concentrations, and do not require a high chaotropic state for the initial LysC digest. To further support our conclusion, we performed another LT digest experiment using 6 or 1.6 m urea for the LysC step and compared it to single-trypsin digest in 1.6 m urea. We observed that tandem proteolysis was consistently more efficient in increasing the quantity of FCPs (**Figure 3E–G**). The comparison of the two tandem digestion strategies, however, showed that the different urea concentrations (1.6 or 6 m) used in LysC digest had no major differential impact on the proteolytic result (**Figure 3G**; **Figure S-5 and Table S-6**, *Supporting Information*). We additionally confirmed that the sequence cleavage context for FCPs and MCPs was preserved among all comparisons, including the LT urea (**Figure S-6**, *Supporting Information*) and LT detergent workflows (**Figure 3A and 3B**).

## 3.4. Conclusions

Taking together, we aimed to investigate the digestion efficiency of tandem LysC/trypsin digestion in DOC and SLS. As concluded from the increased number of FCPs, increased

cleavage efficiency and conserved cleavage motifs, we show that LT digests in these two MS-compatible detergents are highly efficient. Furthermore, the increased FCP yield has a positive impact on quantitative approaches building on them. Finally, we suggest that the main positive attributes of LT strategies are not an immediate effect of the sample denaturation state, but rather a result of the different target specificities between LysC and trypsin.

## 3.5. Raw files and associated data deposition

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD013273.

## 3.6. Supporting information

Supporting text and supporting figures are provided below in sections 3.6.1 and 3.6.2. The complete supporting Information, including the supporting tables (xlsx), is available from the Wiley Online Library at DOI: 10.1002/pmic.201900136. In addition, it is included in the DVD provided with the hard copy of the thesis.

### *3.6.1. Supporting text. Materials and methods*

**Cell culture conditions**

*Escherichia coli* strain K-12 substr. MG1655 was used as a model organism. *E.coli* cultures were cultivated overnight in Luria Broth medium, washed twice with phosphate buffered-saline, to remove traces of the growth medium. Cell pellets (50 ± 10 mg) were collected in 2 mL sterile safe-lock microcentrifuge tubes (Eppendorf). Four biological replicates were processed and used for the protein extraction. All cell pellets were stored at -80 °C until further use.

**Extraction and precipitation of proteins**

For initial cell lysis and protein solubilization, the frozen cell pellets were resuspended and sonicated (Hielscher Ultrasound Technology) in 2% sodium lauroyl sarcosinate ((SLS, Sigma Aldrich), dissolved in 100 mM ammonium bicarbonate buffer (ABC buffer)). Further samples were incubated at 95 °C for 30 min in the presence of 5 mM TCEP (Pierce/Thermo Fisher Scientific) as a reducing agent. Following, all samples were allowed to cool and were alkylated with 10 mM iodoacetamide (Sigma Aldrich) at 25 °C for 30 min. The concentration of proteins in each sample was measured with BCA protein assay kit (Pierce/Thermo Fisher Scientific) following manufacturer's instructions. For protein precipitation, 1.5 mg of total solubilized proteins was precipitated with acetone/methanol precipitation. Eight sample volumes of freezer cold acetone and one sample volume of freezer cold methanol was mixed with the each sample in chemically resistant tubes. Next, the samples were let to precipitate for 2 hours at -80 °C. Afterwards, the precipitated proteins were pelleted (21,120 × g for 15 min), washed twice with freezer cold methanol and air-dried pellets were stored at -80 °C until further use.

**Protein digestion**

The frozen protein pellets were resuspended in 2% SLS or 2% DOC (Sigma-Aldrich) in ABC buffer, sonicated for 20 sec and incubated at 95 °C for 5 min. The concentration of proteins in each sample after precipitation was measured with BCA protein assay kit (Thermo Fisher Scientific) following manufacturer's instructions.

For protein digestion step, 50 µg of total precipitated protein from each sample was used. For in-solution digestion (ISD), the detergents were diluted to 0.5% using ABC buffer. The digestion was carried out in either tandem or trypsin-only digestion formats as follows: (1) For the tandem (LT) strategy, proteins were digested by 0.5 µg LysC (Wako Chemicals GmbH) and incubated for 4 h at 30 °C, followed by trypsin (1 µg, Promega) digestion overnight at 30 °C. (2) The single-trypsin proteolysis was carried out only by trypsin (1 µg, Promega) digestion and incubated overnight at 30 °C. Before LC-MS analysis, SDC or SLS were precipitated using 2% trifluoroacetic acid, and all protein digests were desalted using C18 microspin columns (Harvard Apparatus) according to the manufacturer's instructions.

**LC-MS/MS analyses, peptide/protein identification**

LC-MS/MS analysis of protein digests was performed on Q-Exactive Plus mass spectrometer connected to an electrospray ion source (Thermo Fisher Scientific). Peptide separation was carried out using Ultimate 3000 nanoLC-system (Thermo Fisher Scientific), equipped with packed in-house C18 resin column (Magic C18 AQ 2.4 µm, Dr. Maisch). The peptides were first loaded onto a C18 precolumn (preconcentration setup) and then eluted in backflush mode with a gradient from 96% solvent A (0.15% formic acid) and 4% solvent B (99.85% acetonitrile, 0.15% formic acid) to 20% solvent B over 70 min, continued from 25% to 35% of solvent B for additional 20 min. The flow rate was set to 300 nL/ min. The data acquisition mode for the initial LFQ study was set to obtain one high-resolution MS scan at a resolution of 60000 (*m/z* 200) with scanning range from 375 to 1500 *m/z* followed by MS/MS scans of the 10 most intense ions. To increase the efficiency of MS/MS shots, the charged state screening modus was adjusted to exclude unassigned and singly charged ions. The dynamic exclusion duration was set to 30 sec. The ion accumulation time was set to 50 ms (both MS and MS/MS). The automatic gain control (AGC) was set to $3 \times 10^6$ for MS survey scans and $1 \times 10^5$ for MS/MS scans.

For all identification-based (ID) experiments, the MS raw data was converted into MASCOT-compatible mgf files using Thermo Proteome Discoverer Daemon (Thermo Fisher Scientific). Further assessments were performed by searching raw data using MASCOT (v.2.5, Matrix Science) against the *E. coli* strain K-12 protein database

(downloaded 03/09/2016, 4306 proteins) containing 386 common contaminant/background proteins that were manually added. The following search parameters were used: full tryptic specificity required (cleavage after lysine or arginine residues); two missed cleavages allowed; carbamidomethylation (C) set as a fixed modification; and oxidation (M) set as a variable modification. The mass tolerance was set to 10 ppm for precursor ions and 0.02 Da for fragment ions for high energy-collision dissociation (HCD). The data was further evaluated within Scaffold 4 (Proteome Software) and peptide false discovery rate (FDR) was set to 1% for ID-based assessment of the data.

**Label-free quantification**

Discovery-LFQ was done using Progenesis QI software (Nonlinear Dynamics, version 2.0). MS raw files were imported into Progenesis and the output data (MS/MS spectra) were exported in mgf format. MS/MS spectra were then searched using MASCOT against a decoy database of the predicted proteomes from *E. coli* strain K-12 downloaded from UniProt database. The search criteria were set as for ID-based evaluation of the data. Results from the database search were imported back to Progenesis in order to map peptide identifications to MS1 features. The peak heights of all MS1 features annotated with the same peptide sequence were summed and peptide abundance ratios across different digestion conditions and accompanying q-values (i.e., p-values adjusted for multiple testing) were calculated using SafeQuant R-package version 2.2.2, which is part of the BioConductor open development software project. MS raw files derived from each detergent digest condition were analyzed by LFQ separately to avoid potential quantification biases between each experimental condition.

### *3.6.2. Supporting figures*

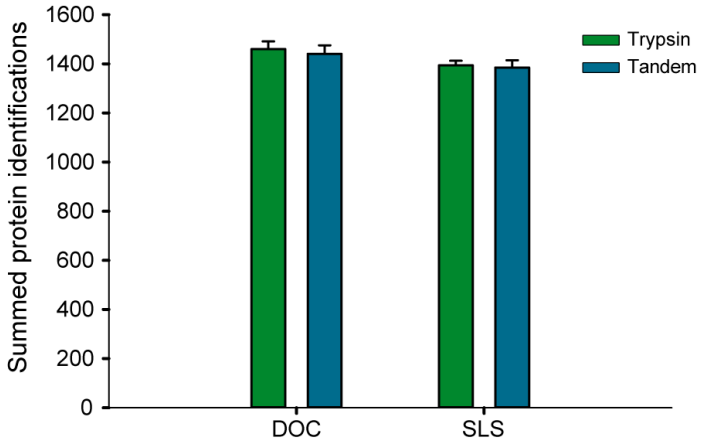



**Figure S-1 |** Total protein identifications shown as averaged result of four replicates per experimental condition; tandem or trypsin only digestion. The error bars indicate the standard deviation. DOC - sodium deoxycholate; SLS - sodium lauroyl sarcosinate.

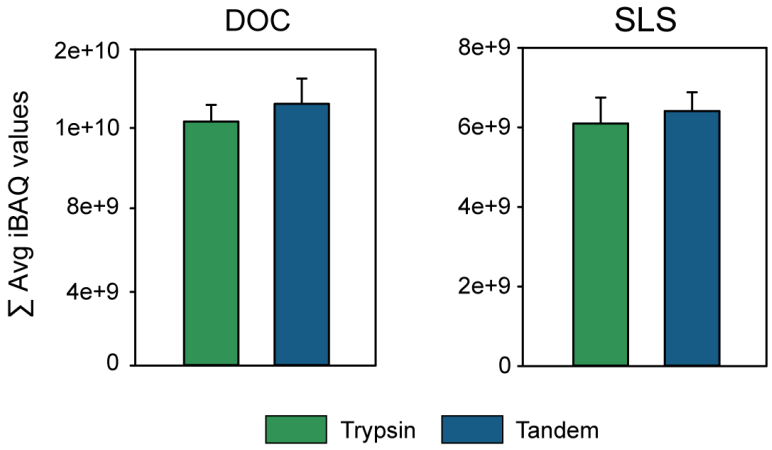



**Figure S-2 |** Sum of average iBAQ values from tandem and trypsin digestions. The iBAQ values were extracted using SafeQuant 2.2.2 R package. The average iBAQ was calculated based on data of four biological replicates. The error bars represent sum of standard deviation between four replicates. DOC - sodium deoxycholate; SLS - sodium lauroyl sarcosinate.

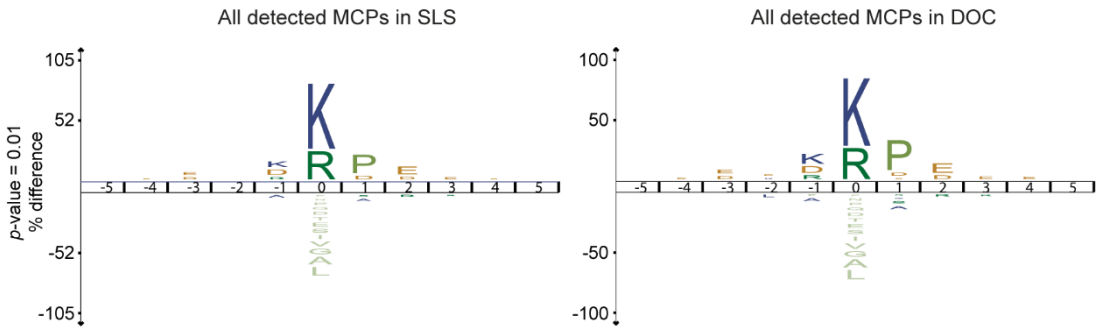



**Figure S-3 |** IceLogo plots showing enriched sequence context of all detected miscleaved peptide stretches. Plots were generated using significantly enriched peptides from tandem versus trypsin label-free quantification (LFQ) comparison in SLS (left graph) and DOC (right graph). DOC - sodium deoxycholate; SLS - sodium lauroyl sarcosinate.

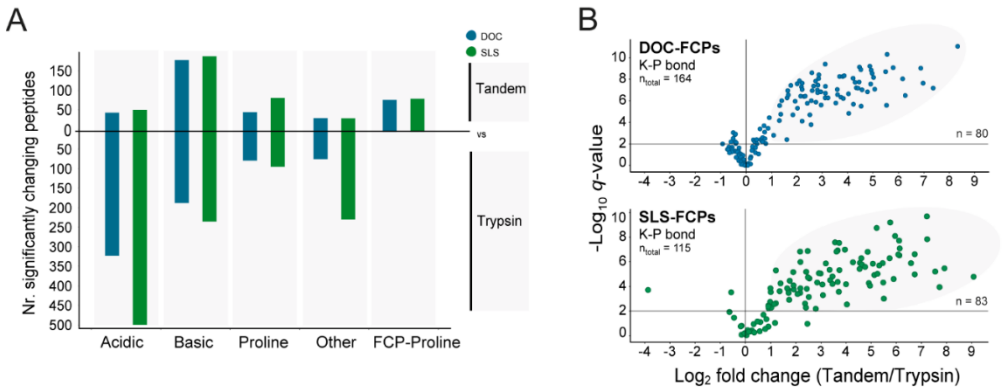



**Figure S-4 | (A)** Number of all significantly enriched peptides obtained from LFQ comparisons in tandem versus trypsin digest. **(B)** Volcano plots of all FCPs with cleavage at a proline (e.g. K-P) residue. DOC - sodium deoxycholate; SLS - sodium lauroyl sarcosinate.

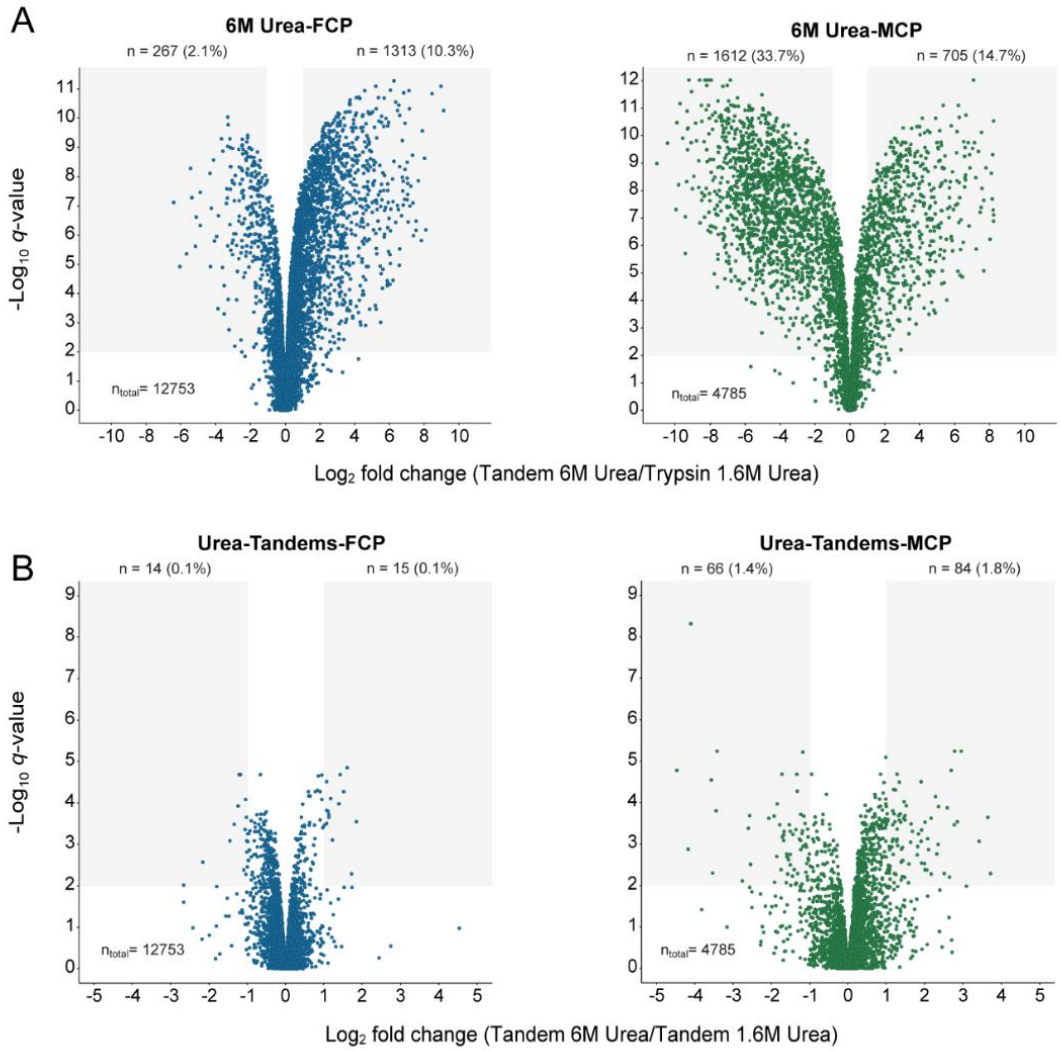



**Figure S-5 | (A)** Volcano plots of relative enrichment of FCPs and MCPs after tandem digest using LysC in 6M Urea following with trypsin in 1.6 Urea environment in comparison to trypsin-only digestion in 1.6 M Urea. **(B)** Volcano plots of relative enrichment of FCPs and MCPs after tandem digest using LysC in 1.6M Urea following with trypsin in 1.6 Urea environment in comparison to tandem digest using LysC in 6M Urea following with trypsin in 1.6 Urea environment.

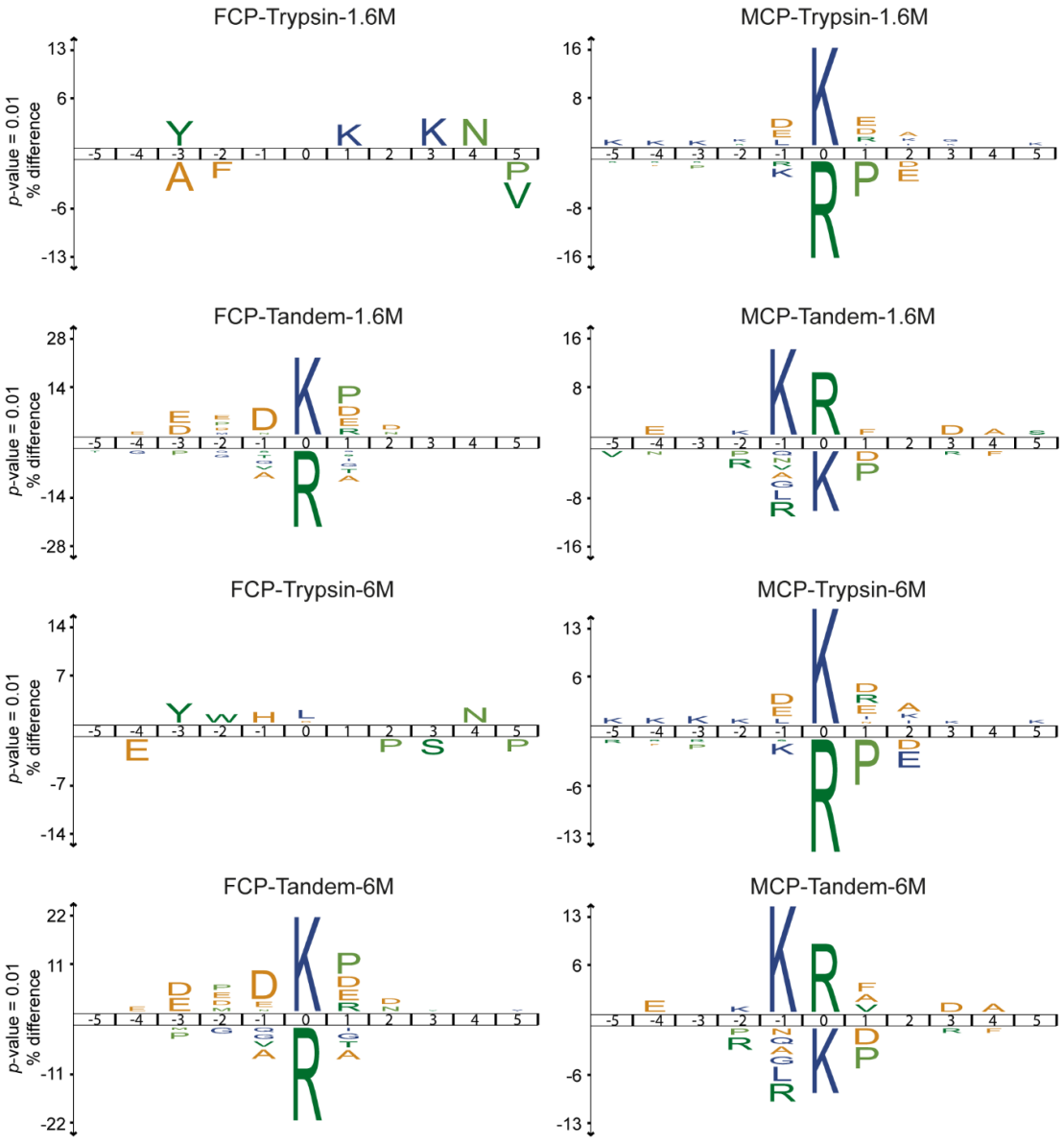



**Figure S-6 |** IceLogo plots showing enriched sequences at fully (FCPs) and miscleaved (MCPs) peptide stretches. Plots were generated using significantly enriched peptides from tandem (LT) versus trypsin label-free quantification (LFQ) comparison in 1.6 M or 6 M Urea

## 3.7. References

Betancourt, L.H., Sanchez, A., Pla, I., Kuras, M., Zhou, Q., Andersson, R., et al. (2018). Quantitative assessment of urea in-solution Lys-C/trypsin digestions reveals superior performance at room temperature over traditional proteolysis at 37 °C. *J Proteome Res* 17(7)**,** 2556-2561.

Chen, E.I., McClatchy, D., Park, S.K., and Yates, J.R., 3rd (2008). Comparisons of mass spectrometry compatible surfactants for global analysis of the mammalian brain proteome. *Ana Chem* 80(22)**,** 8694-8701.

Chiva, C., Ortega, M., and Sabido, E. (2014). Influence of the digestion technique, protease, and missed cleavage peptides in protein quantitation. *J Proteome Res* 13(9)**,** 3979-3986.

Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009). Improved visualization of protein consensus sequences by iceLogo. *Nat Methods* 6(11)**,** 786-787.

Giansanti, P., Tsiatsiani, L., Low, T.Y., and Heck, A.J.R. (2016). Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat Protoc* 11(5)**,** 993-1006.

Glatter, T., Ahrne, E., and Schmidt, A. (2015). Comparison of different sample preparation protocols reveals lysis buffer-specific extraction biases in Gram-negative bacteria and human cells. *J Proteome Res* 14(11)**,** 4472-4485.

Glatter, T., Ludwig, C., Ahrne, E., Aebersold, R., Heck, A.J., and Schmidt, A. (2012). Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion. *J Proteome Res* 11(11)**,** 5145-5156.

Hakobyan, A., Liesack, W., and Glatter, T. (2018). Crude-MS strategy for in-depth proteome analysis of the methane-oxidizing *Methylocystis* sp. strain SC2. *J Proteome Res* 17(9)**,** 3086-3103.

Hervey, W.J., Strader, M.B., and Hurst, G.B. (2007). Comparison of digestion protocols for microgram quantities of enriched protein samples. *J Proteome Res* 6(8)**,** 3054-3061.

Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* 11(3)**,** 319-324.

Poulsen, J.W., Madsen, C.T., Young, C., Poulsen, F.M., and Nielsen, M.L. (2013). Using guanidine-hydrochloride for fast and efficient protein digestion and single-step affinity-purification mass spectrometry. *J Proteome Res* 12(2)**,** 1020-1030.

Raijmakers, R., Neerincx, P., Mohammed, S., and Heck, A.J. (2010). Cleavage specificities of the brother and sister proteases Lys-C and Lys-N. *Chem Commun* 46(46)**,** 8827-8829.

Sander, T., Farke, N., Diehl, C., Kuntz, M., Glatter, T., and Link, H. (2019). Allosteric feedback inhibition enables robust amino acid biosynthesis in *E. coli* by enforcing enzyme overabundance. *Cell Syst* 8(1)**,** 66-75 e68.

Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrne, E., Volkmer, B., Callipo, L., et al. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. *Nat Biotechnol* 34(1)**,** 104-110.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., et al. (2011). Global quantification of mammalian gene expression control. *Nature* 473(7347)**,** 337-342.

Wisniewski, J.R., and Mann, M. (2012). Consecutive proteolytic digestion in an enzyme reactor increases depth of proteomic and phosphoproteomic analysis. *Anal Chem* 84(6)**,** 2631-2637.

Wisniewski, J.R., Zougman, A., and Mann, M. (2009). Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. *J Proteome Res* 8(12)**,** 5674-5678.

# Chapter 4:

## Hydrogen Utilization by *Methylocyctis* sp. Strain SC2 Expands the Known Metabolic Versatility of Type IIa Methanotrophs

*Submitted* 2019

**Anna Hakobyan**[1 #], Jing Zhu[1, 2 #], Timo Glatter[3], Werner Liesack[1, 4*]

*Corresponding author

#Equal contribution

Chapter 4 is written in research manuscript style. It was submitted as research article in November, 2019. My contribution to the chapter 4 involved the experimental design, thermodynamic calculations, laboratory experiments, proteomics data analysis, and the writing of the manuscript, including the design of the figures.

[1]Research Group of Methanotrophic Bacteria and Environmental Genomics/Transcriptomics, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, D-35043 Marburg, Germany

[2]Institute of Environmental Science and Technology, Zhejiang University, Hangzhou, China

[3]Core Facility for Mass-Spectrometry and Proteomics, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Str. 10, D-35043 Marburg, Germany

[4]Center for Synthetic Microbiology (SYNMIKRO), Philipps-Universitaẗ Marburg, Karl-von-Frisch-Str. 16, D-35043 Marburg, German

# 4 Hydrogen Utilization by *Methylocyctis* sp. Strain SC2 Expands the Known Metabolic Versatility of Type IIa Methanotrophs

## 4.1. Abstract

Methane, a non-expensive natural substrate, is used by *Methylocystis* spp. as a single source of carbon and energy. Therefore, these bacteria attract increasing attention as an engineering platform for biotechnological applications and the production of high-quality feed protein. Here, we assessed whether *Methylocystis* sp. strain SC2 is able to also utilize hydrogen as an energy source. The addition of 2% $H_2$ to the culture headspace had a significant positive effect on both growth yield and $CH_4$ consumption rate under $CH_4$ (6%) and $O_2$ (3%) limited conditions. The SC2 biomass yield almost doubled from 5.57 (±0.54) to 10.3 (±1.04) mg cell dry weight per mmol $CH_4$, while the consumption of $CH_4$ significantly decreased. These changes in physiological growth parameters agreed well with theoretical max biomass yield calculations. Hydrogen utilization was linked to significant changes in the SC2 proteome. In addition to hydrogenase accessory proteins, the production of Group 1d and Group 2b hydrogenases was significantly increased in both short- and long-term incubations. Changes in the expression level of the $H_2$-responsive proteins showed a good correspondence to those of their transcripts. Long-term $H_2$ treatments (37 d) induced a switch in the electron flow from NADH dehydrogenase I to a flavin-based electron bifurcating FixABCX complex. Apparently, strain SC2 has the metabolic capacity to channel hydrogen-derived electrons into the quinone pool, which provides a link between hydrogen oxidation and energy production. In summary, $H_2$ can serve in biotechnologically oriented methanotroph projects as an alternative energy source to significantly increase biomass yield from $CH_4$.

## 4.2. Introduction

Methane is the most abundant reduced compound in the atmosphere and plays an important role in the carbon cycle (Dean et al., 2018). In addition, methane is a powerful greenhouse gas trapping 28 times more heat per mass unit than carbon dioxide (Myhre

et al., 2014). Aerobic methane-oxidizing bacteria, or methanotrophs, are crucial players in the global cycle of methane. They can attenuate methane emissions from major sources and constitute the only biological sink for atmospheric methane from terrestrial environments (Hanson and Hanson, 1996). Methanotrophs can be found in three bacterial phyla: *Proteobacteria*, *Verrucomicrobia*, and NC10. The traditionally known methanotrophs belong to the phylum *Proteobacteria*. Proteobacterial methanotrophs have been detected in a wide range of environments such as rice paddies, different upland and hydromorphic soils, landfills, and peatlands (Conrad, 2007; Hanson and Hanson, 1996; Knief, 2015; Semrau et al., 2010).

Based on phylogenetic, morphological and physiological characteristics, the proteobacterial methanotrophs are historically divided into type I and type II methanotrophs (*Gamma*- and *Alphaproteobacteria*, respectively (Bowman, 2006; Trotsenko and Murrell, 2008; Whittenbury and Dalton, 1981)). The methanotrophic *Alphaproteobacteria* were recently divided into type IIa (*Methylocystaceae*) and type IIb (*Beijerinckiaceae*) (Knief, 2015).

Proteobacterial methanotrophs utilize methane as their sole source of carbon and energy. The primary step that differentiates methanotrophs from other methylotrophs is the oxidation of methane by the particulate and/or soluble methane monooxygenases (Murrell et al., 2000a; Murrell et al., 2000b). Subsequently, formaldehyde is generated as the central intermediate by the consecutive oxidation of methane and methanol. Formaldehyde is either assimilated via the serine or ribulose monophosphate pathway into cell carbon or further oxidized to formate and carbon dioxide to gain energy for growth (Chidambarampadmavathy et al., 2015; Karthikeyan et al., 2015; Semrau et al., 2010).

Even though methanotrophs have been studied for many years, major gaps still exist in the fundamental knowledge of this important microbial group, making them underrepresented in metabolic engineering strategies (Kalyuzhnaya et al., 2015). It is widely assumed that utilization of methane is the primary growth strategy of the methanotrophs; however, little is known about how methanotrophic bacteria adapt to changes in environmental factors such as $CH_4$ and/or $O_2$ limitation or the availability of putative alternative electron donors (e.g., methanol, hydrogen, or multicarbon organic compounds). In particular, this is true on the molecular level, including their

transcriptome and proteome. These gaps in our knowledge make it difficult to create useful metabolic models or predict key targets for metabolic engineering of methanotrophs (Kalyuzhnaya et al., 2015). Nonetheless, the biotechnological potential of type IIa methanotrophs has been broadly discussed (Dalton, 1983; Fei et al., 2014; Murrell and Smith, 2009; Trotsenko and Murrell, 2008), but to date only a few methane-based products have been made at pilot- and/or commercial scales. Among those were the production of methanotrophic single cell protein (SCP), which has been explored in UK ("BioProtein" (Nofferm/Calysta)), Norway and Denmark ("Uniprotein" (Unibio A/S)), and the former USSR ("Gaprin"). However, these trademarks did not involve metabolically engineered strains (Kalyuzhnaya et al., 2015).

Recent pure culture studies have provided evidence that alphapoteobacterial methanotrophs are indeed metabolically more versatile than previously thought. Thus, these bacteria offer a number of metabolic properties for specific biotechnological applications. Contrary to the long-held paradigm that methanotrophs are obligate methylotrophs, species from three alphaproteobacterial genera have been shown to grow on simple organic acids, alcohols and short-chain alkane gases (Belova et al., 2011; Crombie and Murrell, 2014; Dedysh et al., 2005; Im et al., 2011).

More recently, alphaproteobacterial methanotrophs have also been shown to contain genes encoding different types of [NiFe]-hydrogenases (Carere et al., 2017; Greening et al., 2016). These organisms are widely distributed in diverse habitats (in particular *Methylocystis* spp.), but most frequently in oxic-anoxic interfaces of methanogenic environments with counter gradients of oxygen, methane and hydrogen (Brune et al., 2000; Piche-Choquette and Constant, 2019; Piche-Choquette et al., 2018). Thus, it was assumed that in addition to organic compounds, they are capable of using hydrogen ($H_2$) as an alternative source of energy in the so-called "*knallgas"* reaction (Eberhardt, 1969). The key enzymes involved hydrogen metabolism are hydrogenases that channel electrons from hydrogen into the quinone pool, thereby providing the link between hydrogen oxidation and energy production (Peterson et al., 2012). The most abundant hydrogenases, including those responsible for aerobic hydrogen oxidation, contain a [NiFe] active site (Shafaat et al., 2013). Utilization of $H_2$ as an alternative energy source may be an effective strategy to improve the biomass yield for metabolic engineering using type IIa methanotrophs.

Nowadays, most efforts to metabolically engineer type IIa methanotrophs are focused on well characterized species such as *Methylosinus trichosporium* OB3b and on *Methylocystis* spp. The latter organisms are known to be both most oligotrophic and most widespread in terrestrial environments (Chen et al., 2007; Dedysh et al., 2003; Knief et al., 2003; Nauer et al., 2012). Our model organism, strain SC2, is one of the very few *Methylocystis* spp. with a well-annotated gap-closured genome sequence (Dam et al., 2012; Dunfield et al., 2002). Among other genes of interest, *Methylocystis sp. strain SC2 was shown to* possess genes encoding four different groups of [NiFe]-hydrogenases: Group 1d, Group 1h/5, Group 2b, and Group 3b. Therefore, it is a suitable model organism to study the multifaceted role of $H_2$ metabolism for the improvement of metabolic engineering strategies of type IIa methanotrophs.

The objective of our study was to assess whether strain SC2, as a representative for *Methylocystis* spp., can utilize hydrogen to optimize its biomass yield by mixed utilization of $CH_4$ and $H_2$, rather than $CH_4$ as the sole source of energy. Thus, we aimed to show that, in the presence of $H_2$, $CH_4$ will primarily be used for synthesis of cell carbon and increased biomass/protein yield. In particular, we intended to explore those $CH_4/O_2$ ratios, which maximize the effect of hydrogen addition on the biomass yield and proteome reconstruction of strain SC2. To achieve our goals, we combined hydrogen-based growth experiments with our recently optimized proteomics workflow (Hakobyan et al., 2018).

## 4.3. Materials and methods

### *4.3.1. Culture conditions for hydrogen treatments*

Cells of *Methylocystis* sp. strain SC2 were pre-grown to the exponential phase ($OD_{600}$ = 0.2 ± 0.05) in 120-ml serum bottles containing 30 ml of nitrate-based mineral salts medium (NMS) with the same basic composition as reported earlier (Dam et al., 2014; Heyer et al., 2002). One gram of $KNO_3$ (~10 mM $NO_3^-$) per liter was added to the medium as the nitrogen source. The headspace was filled with air and filter-sterilized methane (air/$CH_4$ ratio of 80/20 %, v/v). The serum bottles were incubated on a rotary shaker at 130 rpm and 25°C for 48–72 h. Next, the SC2 cells were collected by centrifugation (7000 × *g* for 20 min, 4 °C) and thoroughly washed twice with phosphate buffer saline (1 × PBS,

pH 7.4). The washed cells were used to investigate the influence of $H_2$ on biomass yield and $CH_4$ consumption under $CH_4/O_2$-replete and $CH_4/O_2$-limiting conditions.

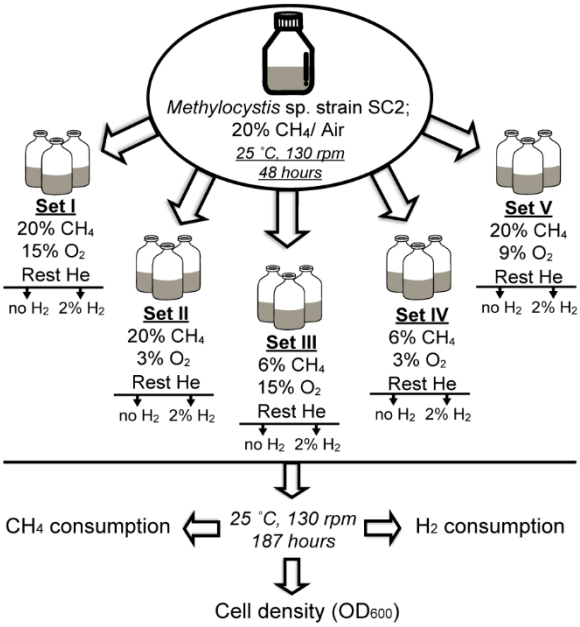



**Figure 1** | The physiological growth experiments involved five experimental sets (I to V) with different $CH_4/O_2$ ratios, with and without the addition of 2% $H_2$. Cells of *Methylocystis* sp. strain SC2 were pre-grown to the exponential phase ($OD_{600}$ = 0.2 ± 0.05) on nitrate-based mineral salts medium (NMS). The pre-cultured cells were used to investigate the influence of $H_2$ on biomass yield and $CH_4$ consumption of strain SC2 under $CH_4/O_2$-replete and $CH_4/O_2$-limiting conditions. For that purpose, the headspace of the batch cultures was filled with five different $O_2/CH_4$ ratios and balanced with helium (He). For $H_2$ treatment, 2% $H_2$ was added to the headspace. The cultures without $H_2$ addition acted as a control. Both cell density and changes in the headspace $CH_4$ and $H_2$ concentrations were repeatedly monitored during the cultivation period. Measurements were done in triplicate cultures.

The hydrogen treatments were done using the same cultural conditions as described above, but with adjusted headspace. The headspace was filled with five different $O_2/CH_4$ ratios (Experimental Sets I to V in **Figure 1**), with or without the addition of 2% $H_2$, and finally balanced with helium (He). The cell density was monitored by measuring the turbidity at 600 nm ($OD_{600}$) using Eppendorf Biophotometer UV/Vis spectrophotometer (Eppendorf, Germany). Changes in the headspace $CH_4$ and $CO_2$ concentrations were monitored over time using a gas chromatograph equipped with a flame ionization detector (GC-FID, SRI Instruments, Earl St. Torrance, CA). The utilization of $H_2$ by strain SC2 was measured using $H_2$ microsensor with piercing-needle (Unisens A/S, Denmark).

### *4.3.2. Determination of cell dry weight per $OD_{600}$*

Biomass expressed as cell dry weight (CDW) was used to calculate the growth rate and specific gas consumption rate of strain SC2. In order to determine the CDW per $OD_{600}$, SC2 cells were pre-incubated in NMS under standard conditions and harvested in the exponential growth phase. After incubation, the culture was subjected to centrifugation (7000 × *g* for 30 min, 4 °C) and the pelleted SC2 cells were washed twice with PBS. The washed cells were dried for 24 h at 100 °C, and the CDW was determined. An $OD_{600}$ of 1 corresponded to 0.26 (± 0.01) g CDW/L of strain SC2 culture.

Biomass yield was expressed as mg CDW per mmol $CH_4$ consumed. Methane consumption rate was indicated as mmol $CH_4$ consumed per g CDW per day. Hydrogen consumption rate was shown as mmol $H_2$ consumed per g CDW per day.

### *4.3.3. Culture conditions for proteomics and RNA extraction*

In order to collect biomass for proteomics and RNA extraction, cells of strain SC2 were pre-grown to mid-exponential phase using standard conditions as described above (air/$CH_4$ ratio of 80/20 %, v/v). Following, cells in the logarithmic growth phase were collected, washed and used for the $H_2$ treatments. The washed cells were incubated under a headspace of 6% $CH_4$, 3% $O_2$, 2% $H_2$, and 89% He. SC2 cultures without the addition of $H_2$ were used as a control. Biomass for proteomics and RNA extraction was collected after an incubation period of 19 h, 10 d, and 37 d. The headspace of the 10-day and 37-day cultures were readjusted to the starting concentrations of $CH_4/O_2/H_2$ every three to four days. Finally, the cells of all the $H_2$ and control treatments were collected by centrifugation (7000 × *g* for 20 min, 4 °C) and thoroughly washed twice with PBS buffer to remove traces of the growth medium. The cells (50 ± 10 mg) were pelleted in 2 mL sterile safe-lock microcentrifuge tubes (Eppendorf) and stored at -80 °C for further use.

### *4.3.4. Extraction, solubilization and digestion of proteins*

The proteomics sample preparation was done according to the previous study (Hakobyan et al., 2018). In brief, the frozen cell pellets were resuspended and sonicated (Hielscher Ultrasound Technology) in 2 % sodium deoxycholate buffer (SDC, dissolved

in 100 mM ammonium bicarbonate) in the presence of 5 mM tris[2-carboxyethyl]phosphine and subsequently incubated at 95 °C for 60 min. The samples were subjected to additional 20 sec rounds of sonication after 15 and 30 min of incubation. Upon the SDC treatment, all the protein samples were allowed to cool followed by an incubation with 10 mM iodoacetamide at 25 °C for 30 min. The resulting crude lysates were used for further digestion as described below. The concentration of proteins in each sample was measured with BCA protein assay kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

For protein digestion step, 50 µg of total solubilized protein from crude lysate samples was used. For in-solution digestion, the samples were diluted to 0.5 % of solubilizing (SDC) agent in 100 mM ammonium bicarbonate buffer. LysC (0.5 µg, Wako Chemicals GmbH) was added directly to the protein extract and incubated for 4 h at 30 °C, followed by trypsin (1 µg, Promega) digestion overnight at 30 °C. Before LC-MS analysis, traces of SDC were precipitated using 2 % trifluoroacetic acid, and all protein digests were desalted using C18 microspin columns (Harvard Apparatus) according to the manufacturer's instructions.

### *4.3.5. LC-MS/MS analyses, peptide/protein identification and LFQ quantification*

The LC-MS/MS analysis of protein digests was performed on a Q-Exactive Plus mass spectrometer connected to an electrospray ion source (Thermo Fisher Scientific). Peptide separation was carried out using the Ultimate 3000 nanoLC-system (Thermo Fisher Scientific), equipped with an in-house packed C18 resin column (Magic C18 AQ 2.4 µm, Dr. Maisch). The peptides were first loaded onto a C18 precolumn (preconcentration set-up) and then eluted in backflush mode using a gradient from 96 % solvent A (0.15 % formic acid) and 4 % solvent B (99.85 % acetonitrile, 0.15 % formic acid) to 30 % solvent B over 115 min. The flow rate was set to 300 nL/min. The data acquisition mode for the initial LFQ study was set to obtain one high-resolution MS scan at a resolution of 60,000 (*m/z* 200) with a scanning range from 375 to 1500 *m/z* followed by MS/MS scans of the 10 most intense ions. To increase the efficiency of MS/MS acquisition, the charged state screening modus was activated to exclude unassigned and singly charged ions. The dynamic exclusion duration was set to 30 sec. The ion

accumulation time was set to 50 ms (both MS and MS/MS). The automatic gain control (AGC) was set to $3 \times 10^6$ for MS survey scans and $1 \times 10^5$ for MS/MS scans.

Discovery-LFQ was done using Progenesis QI software (Nonlinear Dynamics, version 2.0) as discussed before (for details see (Hakobyan et al., 2018)). Next, the data obtained from Progenesis were evaluated using SafeQuant R package, version 2.2.2 (Glatter et al., 2012). Hereby, 1% identification and quantification false discovery rate (FDR) were calculated.

### *4.3.6. Prediction of protein-protein interactions*

Prediction of protein-protein interactions was conducted with the online analysis tool STRING version 11 (Szklarczyk et al., 2019). The predicted network was grouped into five clusters by "kmeans clustering" filter. The network edges were analyzed based on molecular action, where the line shape indicates the predicted mode of action. The thickness of each edge indicates the confidence of interaction (0–1.0).

### *4.3.7. RNA extraction*

Total RNA was extracted from frozen cell pellets of both $H_2$ and control treatments using the RiboPure™-Bacteria Kit (Ambion, Life Technologies, TX, USA) according to the manufacturer`s instructions, but with slight modifications. In order to increase the yield of total RNA, the RiboPure™-Bacteria Kit reagents used for binding and elution of the RNA extract were replaced with those of the RNA Clean and Concentrator™-5 Kit (Zymo Research, CA, USA). The purified RNA was treated with RNase-free DNase I (Ambion, Austin, TX, USA) and quantified by Qubit fluorometer (Life Technologies, CA, USA).

### *4.3.8. Real-time quantitative PCR and primer design*

A total of eight genes were selected for real-time quantitative PCR (RT-qPCR), in order to validate the differential regulation of methane, nitrogen and hydrogen metabolism-related proteins between the experimental treatments on the transcriptome level. These genes encode the large subunits of Group 1d (*hupB*), Group 1h/5 (*hhyL*), Group 2b (*hupV*) and Group 3b (*hyhL*) hydrogenases and *nifH* (nitrogenase), *pmoA1* (pMMO1) and *pmoA2* (pMMO2) genes. In addition, the *rplI* gene encoding 50S ribosomal protein

L9 was used as a reference in RT-qPCR. Total RNA was reverse transcribed using the GoScript™ Reverse Transcription System (Promega, WI, USA) according to the manufacturer's instructions. Random primers (0.5 µg/reaction) provided by the kit were used to produce single-stranded cDNA. A serial dilution series of cDNA (1:10 and 1:50) was used to quantify the mRNA transcript copy numbers of each target gene. Total DNA of strain SC2 was extracted using the GeneElute™ Bacterial Genomic DNA Kit (Merck, Germany). Standard curves were constructed using genomic DNA from strain SC2 as the template (calibration range from 1 to $10^7$ copies, based on Applied Biosystems protocols).

The primers used for RT-qPCR (**Table S-1**) were designed with the NCBI Primer design tool using Primer 3 settings as suggested previously (Thornton, 2011). Specificity and secondary structures of the newly designed primers were checked by Beacon DesignerTM free edition (Premier Biosoft International). RT-qPCR was performed using the SYBR® Green JumpStart™ Taq ReadyMix™ (Merck, Germany) on a Bio-Rad CFX96 Real-Time PCR system (Bio-Rad Laboratories Inc., USA). All reactions were carried out in three biological and three technical replicates for each $H_2$ and control treatment (total of nine RT-qPCR reactions for each treatment and time point). The PCR efficiency was at least 80% ($R^2 > 0.90$). The presence of unspecific products was checked by melt curve analysis.

Statistical analysis of the RT-qPCR data was conducted in SigmaPlot 14. Significance (*p*-value) for the differential regulation of transcripts between $H_2$-treated and control samples was calculated. Each data set was checked for normal distribution using Shapiro-Wilk normality test. For the data sets that passed the normality test ($p>0.050$), Student's t-test was computed. While for the data sets that failed the normality test ($p<0.050$), Mann-Whitney Rank Sum Test was applied. The differential expression of transcripts with $\log_2$-fold change of >1 or <-1 and *p*-value of ≤ 0.01 was considered significant.

### *4.3.9. Raw files and associated data deposition*

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the dataset identifier PXD015700 and 10.6019/PXD015700.

## 4.4. Theory/calculations

### *4.4.1. Thermodynamic calculations for cell carbon synthesis under $H_2$ availability*

In order to reveal the theoretical impact of $H_2$ addition on maximum cell carbon synthesis, we applied thermodynamic electron equivalents (eeq) model two (TEEM2) with some modifications (McCarty, 2007). This model was created for aerobic heterotrophic reactions of C1 compounds that include oxygenases. Theoretically, the energy reactions for aerobic $CH_4$ and $H_2$ oxidation can be displayed by the stoichiometric Equations (1) and (2), where reaction free energy is based on one electron equivalents. The free energy of each reaction was calculated based on the half-reaction reduction potential for the electron donors ($CH_4$, $H_2$) and acceptor ($O_2$).

$$CH_4 + 2O_2 = CO_2 + 2H_2O \qquad \Delta G^{0'} = -102.25\ kJ/eeq \tag{1}$$

$$2H_2 + O_2 = 2H_2O \qquad \Delta G^{0'} = -237.34\ kJ/eeq \tag{2}$$

Based on TEEM2 model we were able to calculate the true or maximum cell yield ($Y_{C/C}$) using either $CH_4$-only or $CH_4$/$H_2$-combined as electron donors for strain SC2. The calculations of $Y_{C/C}$ where carried out based on Equation (3):

$$Y_{C/C} = \frac{y_d}{y_x} f_s^0 \tag{3}$$

Here, $y_d$ and $y_x$ are showing the degree of reduction of C1 electron donor and degree of reduction of cells. In case of $CH_4$, $y_d$ equals to 8 and $y_x$ was estimated based on the assumed relative composition of cells ($C_5H_7O_2N$) and equals to 4. Following, $f_s^0$ which represents the true yield expressed as a fraction of electron donor converted for the synthesis (eeq cells/eeq donor), was calculated based on Equations (4) and (5), for the control and $H_2$ treatments, respectively.

$$f_s^0 = \frac{\Delta G_a(O_2) - \Delta G_d(CH_4) - (q/p)\Delta G_{xy}}{(\Delta G_a(O_2) - \Delta G_d(CH_4) - (q/p)\Delta G_{xy}) - \Delta G_s/\varepsilon} \tag{4}$$

$$f_s^0 = \frac{\Delta G_{Eq.(2)} + (\Delta G_a(O_2) - \Delta G_d(CH_4) - (q/p)\Delta G_{xy})}{(\Delta G_{Eq.(2)} + (\Delta G_a(O_2) - \Delta G_d(CH_4 \pm H_2) - (q/p)\Delta G_{xy})) - \Delta G_s/\varepsilon} \tag{5}$$

Here, $\Delta G_a(O_2)$ and $\Delta G_d(CH_4)$ indicate the half-reaction reduction potential for electron acceptor ($O_2$) and C1 electron donor ($CH_4$). The half reaction reduction potential for $O_2$

is -78.72 kJ/eeq and for $CH_4$ is 23.53 kJ/eeq (McCarty, 2007). $\Delta G_{xy}$ is the energy loss involved in oxygenase usage and is represented by the standard energy associated with NADH oxidation and equals to -219.2 kJ/mol. Next, *p* shows the number of electron equivalents per mol of substrate from half-reaction reduction equation and equals to 8 for $CH_4$ (McCarty, 2007; Thauer et al., 1977). Value *q* indicates the number of oxygenase reactions per mole $CH_4$. In aerobic methanotrophs, the oxygenase, namely pMMO, is used once during the oxidation of methane to $CO_2$ and therefore, *q*=1. $\varepsilon$ represents the energy transfer efficiency with a value of 0.37. The $\Delta G_s$ value represents the Gibbs free energy for cell synthesis reactions and is calculated based on Equation (6).

$$\Delta G_s(CH_4) = \frac{(\Delta G_{fa} - \Delta G_{CH_4})}{\varepsilon^m} + \frac{(\Delta G_{in} - \Delta G_{fa})}{\varepsilon^n} + \frac{\Delta G_{pc}}{\varepsilon} \quad (6)$$

Aerobic methanotrophs oxidize $CH_4$ to formaldehyde, which is assimilated into cell biomass, either through the ribulose monophosphate pathway or the serine pathway (details in Section 1). Therefore, the energy loss ($\Delta G_{fa}$=46.53 kJ/eeq) through the oxidation of $CH_4$ to formaldehyde is included in the calculation of $\Delta Gs$. $\Delta G_{in}$ is the half-reaction reduction potential for acetyl-CoA (30.9 kJ/eeq) and $\Delta G_{pc}$ is the Gibbs free energy for the incorporation of acetyl-CoA into cell biomass (McCarty, 2007). $\Delta G_{pc}$ equals to 13.45 kJ/eeq, when cells are incubated with nitrate as a nitrogen source (NMS for strain SC2) and with cell biomass formulation of $C_5H_7O_2N$. Next, m=n and equals to 1 given that $\Delta G_{in}$-$\Delta G_d$ ($CH_4$)>0. In summary, $\Delta Gs$ for strain SC2 with $CH_4$ as C1 substrate equals to 56.27 kJ/eeq.

## 4.5. Results

### *4.5.1. Impact of the $CH_4/O_2$ ratio on $H_2$ utilization by strain SC2*

To determine the impact of 2% $H_2$ addition on the growth of strain SC2 under $CH_4$ and/or $O_2$ limited conditions, SC2 cells were incubated with five different $CH_4/O_2$ headspace ratios under nitrogen-free conditions (**Figure 1**). Cell density ($OD_{600}$) and the concentrations of $CH_4$ and $H_2$ were regularly measured during the one-week (187 h) incubation period.

Hydrogen-treated cultures of all five experimental sets showed the ability to oxidize $H_2$ that was added to the headspace (**Figure 2A and Figure S-1**). However, $H_2$ utilization was much faster in the cultures with limited $O_2$ (3%, Sets II and IV) than with elevated $O_2$ concentrations. Hydrogen had almost been fully consumed after 70 h and 160 h in cultures of Set IV (**Figure 2A**) and Set II (**Figure S-1**), respectively. In Set V, $H_2$ was almost fully consumed after 180 h of incubation (**Figure S-1**). In Sets I and III, $H_2$ had not been fully consumed even after the complete incubation period of 187 h (**Figure S-1**).

The results of our $H_2$ consumption measurements prompted us to calculate cell dry weight, biomass yield, and $CH_4$ and $H_2$ consumption rates after 46 h (**Table S-2**), 70 h (**Figure 3 and Table S-3**), and the complete incubation period of 187 h (**Table S-4**). These time points correspond, respectively, to the exponential, early stationary, and late stationary growth phases of strain SC2.

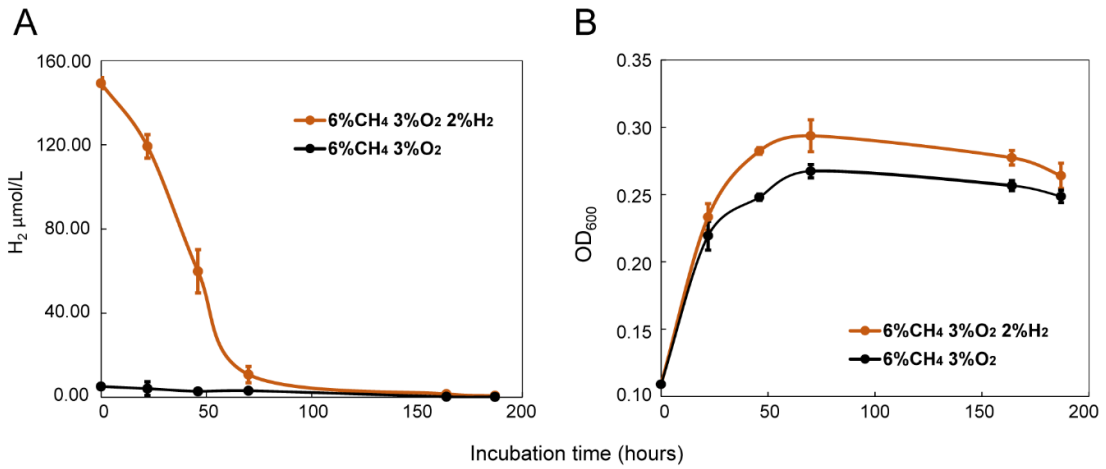



**Figure 2** | Hydrogen utilization by *Methylocystis* sp. strain SC2. **(A)** Growth of strain SC2 under an atmosphere of 6% $CH_4$/ 3% $O_2$ / He (Set IV), with (orange) and without (black) the addition of 2% $H_2$. **(B)** Hydrogen consumption by strain SC2 as determined by a $H_2$ microsensor equipped with a piercing-needle (Unisens A/S). $H_2$-treated and control cultures were incubated for a total of 187 h. Error bars show standard deviations of triplicate cultures.

SC2 cultures with $H_2$ in the headspace demonstrated significantly higher growth during the first 70 h of incubation under $O_2$-limited (3%) conditions than those without $H_2$ **(Figure 2B and Figure S-1)**. This time period includes both the exponential growth phase and the transition to the early stationary phase. The positive growth effect of $H_2$ under $O_2$ limited conditions was significant regardless of whether 6% $CH_4$ (Set IV) or 20% $CH_4$

(Set II) had been added to the headspace (**Figure 3 and Table S-3**). The addition of $H_2$ under $O_2$ limited conditions significantly increased the biomass yield (**Figure 3A**) but decreased the $CH_4$ consumption rate of strain SC2 (**Figure 3B**). The differences in biomass yield and $CH_4$ consumption rate between control and $H_2$ treatment were clearly more pronounced in cultures of Set IV (6% $CH_4$) than in those of Set II (20% $CH_4$). Relative to the controls, biomass yield in SC2 cultures of Set IV was increased by 85% after 70 h of incubation (**Figure 3A**), while the $CH_4$ consumption rate was decreased by 40% (**Figure 3B and Table S-3**). Moreover, the difference in biomass yield between $H_2$-treated and control cultures was clearly more significant after 46 h (**Table S-2**) than after 70 h (**Table S-3**) of incubation.

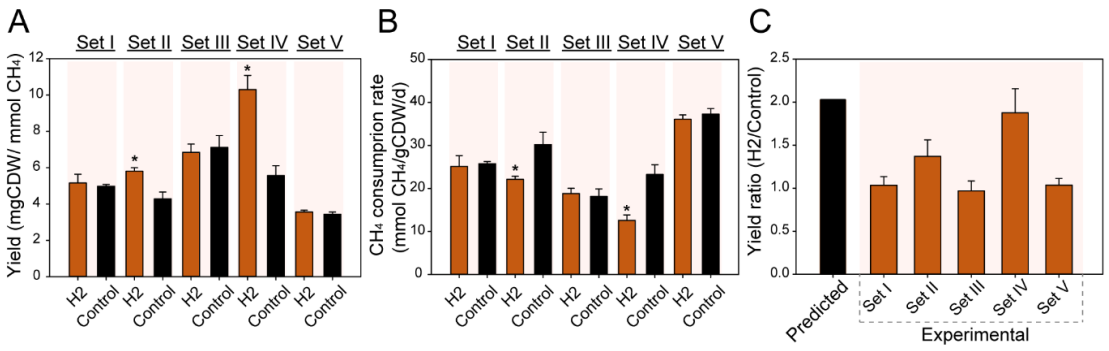



**Figure 3 |** Physiological growth parameters of early stationary SC2 cultures determined for experimental sets I to V and their comparison to the predicted theoretical yield ratio between $CH_4$-only and $CH_4/H_2$-combined treatments (see details in section 3.1). **(A)** Biomass yield was expressed as mg cell dry weight (CDW) per mmol $CH_4$ consumed. **(B)** Methane consumption rate was calculated as mmol $CH_4$ consumed per g CDW per day. An $OD_{600}$ of 1 corresponded to 0.26 (± 0.01) g CDW/L of strain SC2 culture. All the measurements and calculations were made in triplicate cultures. The asterisk (*) indicates significant difference ($p$-value ≤ 0.05) between $H_2$ and control treatments. **(C)** Bar-chart comparing the theoretical max yield (black bar) with the experimental biomass yields obtained for sets I to V (orange bars), calculated as the ratio of the biomass yields between $H_2$ and control treatments (details in section 3.1). Error bars show the observed standard deviation among the triplicate cultures of each treatment.

When calculations were made for the physiological growth parameters over the complete incubation period of 187 h, the positive growth effect of $H_2$ was significant only in the $O_2$ limited SC2 cultures of Set IV (6% $CH_4$), but not in those of Set II (20% $CH_4$) (**Table S-4**). No effects of $H_2$ addition on the growth activity of strain SC2 were observed under $O_2$ replete (9% and 15% $O_2$) conditions (**Figure 3, Table S-2, Table S-3 and Table S-4**).

Furthermore, based on thermodynamic TEEM2 calculations (details in section 3.1), the theoretical maximum cell yield ($Y_{C/C}$) for strain SC2 using $CH_4$ as the only energy source was 0.66 mol cell C/mol $CH_4$ and 1.34 mol cell C/mol $CH_4$ when $CH_4$/$H_2$-combined was considered as energy source. Thus, the ratio between $H_2$ and control treatments equaled 2.03, indicating that $H_2$ addition increased the theoretical max yield of strain SC2 by two-fold. Among all experimental sets tested (**Figure 1**), the predicted max yield ratio was highly similar to the measured yield ratio in experimental Set IV (**Figure 3C**).

### *4.5.2. Impact of $H_2$ on SC2 growth during long-term incubation*

The following experiments were performed using those culture conditions (6% $CH_4$/ 3% $O_2$, Set IV) that had the greatest $H_2$ effect on strain SC2 in our physiological growth experiments (**Figure 3**). The cultures were incubated for a total of 37 d, with and without the addition of 2% $H_2$. To maintain the SC2 cultures in the exponential growth phase and thus to maximize the effects of $H_2$ treatment, the headspace of the 10-day and 37-day cultures was regularly readjusted to the starting concentrations of $CH_4$/$O_2$/$H_2$ (Set IV, **Figure 1**). The cell dry weight was significantly increased in $H_2$-treated cultures after short-term incubation (19 h) (**Figure S-2A**). However, the positive impact of $H_2$ availability was clearly more pronounced after long-term incubation (37 d), with significant effects across all SC2 growth parameters monitored over incubation time (**Figure S-2C**). In particular, cell dry weight and biomass yield were increased by 25% and 52% in $H_2$-treated cultures, respectively. Concomitantly, their $CH_4$ consumption rate was decreased by 34%.

### *4.5.3. Effects of $H_2$ availability on the SC2 proteome*

SC2 biomass was collected for proteomics after short-term (19 h) and long-term (10 d and 37 d) incubations, with or without the addition of $H_2$ (**Figure 4**). The experimental approach is detailed in section 2.4. Overall, about 63% (2526 proteins) of the SC2 reference proteome (4040 proteins, Uniprot Proteomes Database) was covered with high reproducibility among three biological replicates, in both $H_2$ and control treatments after 19 h, 10 d, and 37 d of incubation (**Figure 4, Figure S-3 and Table S-5**).

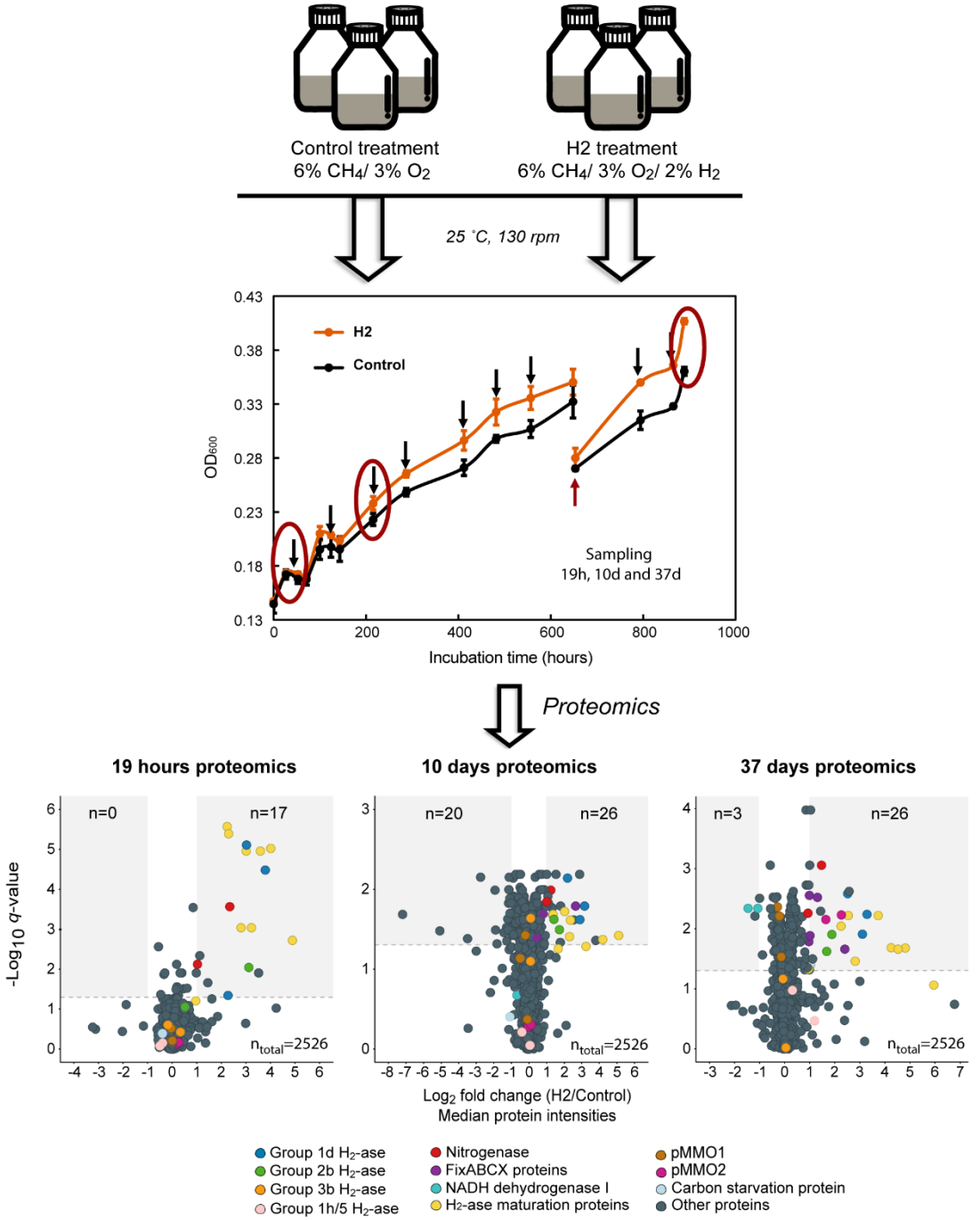



**Figure 4 |** Flow chart showing the proteome analysis of strain SC2 upon incubation under an atmosphere of 6% $CH_4$ and 3% $O_2$, with and without the addition of 2% $H_2$. **(A)** In order to collect biomass for proteomics, cells of strain SC2 were cultured for a total of 37 d (details in Section 2.3). Biomass for proteomics was collected after an incubation period of 19 h, 10 d, and 37 d (red circles). Black arrows in the growth curve indicate the addition of fresh gas mixture to the headspace. The red arrow indicates the time point when SC2 cultures were washed and supplemented with fresh medium. **(B)** The Volcano plots illustrate the differential regulation of proteins between $H_2$-treated and control cells after 19 h, 10 d, and 37 d of incubation. The x-axis shows the $\log_2$-fold change of protein abundance, while the y-axis indicates the negative $\log_{10}$ *q*-values. Significance areas are highlighted by color (details in Table S-6 and Table S-7). The complete experiment is based on the independent analysis of triplicate cultures for both $H_2$ and control treatments.

In the short-term (19 h) incubation, $H_2$ availability induced a significant upregulation of 17 proteins (*q*-Value < 0.05, $log_2$-fold change ratio ≥ 1) (**Figure 4 and Table S-6**). This involved the concurrent upregulation of all three subunits of the low-affinity Group 1d hydrogenase and small subunit of regulatory Group 2b hydrogenase ($log_2$ change ratio 3.1) (**Figure 4 and Table S-7**). Simultaneously, nine members of the proteins responsible for expression, formation, and maturation of the hydrogenases were significantly upregulated (**Figure 4 and Table S-6**). Notably, we also observed a significant upregulation of NifH and NifX subunits of the nitrogenase enzyme complex, with a $log_2$ fold change ratio of 2.3 and 1.0, respectively (**Table S-6**). In addition, ferredoxin protein was differentially regulated with high significance ($log_2$ change ratio of 3.5).

In mid-term (10 d) incubation, the number of differentially expressed proteins was greater than in short-term (19 h) incubation. A total of 26 and 20 proteins were significantly upregulated and downregulated, respectively (**Table S-6**). The expression level of Group 1d hydrogenase, hydrogenase accessory proteins (maturation/formation/expression) and the nitrogenase NifH and NifX subunits maintained significantly increased in the $H_2$-treated cells (**Table S-7**). In addition, the small and large subunits of Group 2b regulatory hydrogenase and FixX were significantly upregulated (**Figure 4 and Table S-6).** Serine protease inhibitors (Kazal domain proteins), ABC transporter proteins, cytochrome P450, NADH-quinone oxidoreductase subunit H and several other proteins were among the downregulated proteins (**Table S-6**).

In long-term (37 d) incubation, the majority of the above-mentioned proteins maintained upregulated on a constantly high level (**Table S-5 and Table S-6**). These included the Group 1d and Group 2b regulatory hydrogenases, all nine hydrogenase accessory proteins and NifH protein (**Figure 4 and Table S-7**). In addition to FixX, all other subunits of the FixABCX protein complex got upregulated **(Table S-7)**. By contrast, two other hydrogenases of strain SC2, namely the high-affinity Group 1h/5 hydrogenase and Group 3b bidirectional hydrogenase, did not respond to the $H_2$ treatments (**Figure 4 and Table S-7**). Two subunits of pMMO2 were significantly upregulated ($log_2$ fold change ratio of 1.6 [PmoA2] and 2.3 [PmoB2]), while there was no differential regulation of pMMO1 (**Table S-7**). In addition, the carbon starvation protein CstA was significantly

upregulated (**Table S-7**). Finally, the NADH dehydrogenase I (complex I) with two subunits (20 kDa and 49 kDa) along with cytochrome C oxidase subunit III were significantly downregulated in response to $H_2$ availability, but only after 37 days of incubation (**Figure 4 and Table S-7**).

### *4.5.4. Predicted protein-protein interactions in response to the long-term $H_2$ treatment*

In order to display the relationship between differentially regulated proteins after long-term incubation (37 d), protein-protein interactions were predicted *in silico* for selected proteins (based on the proteins list in **Table S-7**). A total of 35 proteins were blasted against the *Methylocystis* sp. strain SC2 protein database available in STRING.

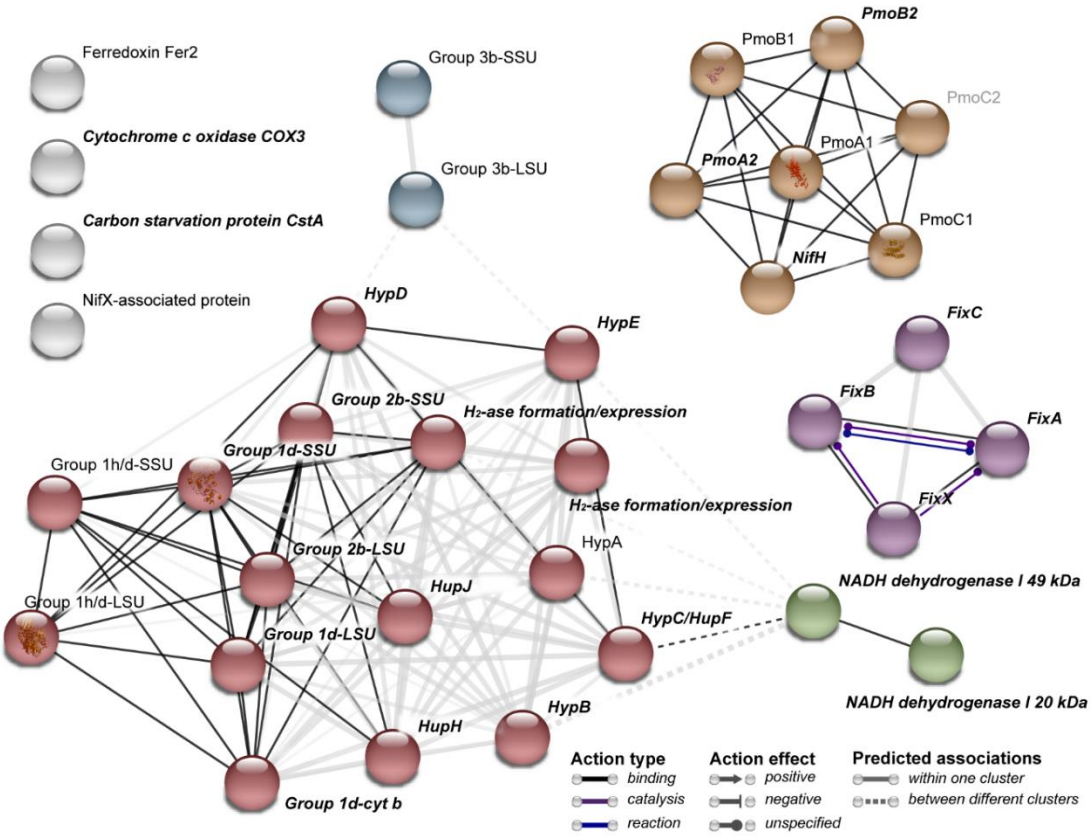



**Figure 5 |** STRING network analysis of the proteins differentially regulated after 37 d of $H_2$ treatment. Each node in the network denotes a protein while edges between nodes indicate interactions between corresponding proteins. The network edges were analyzed based on molecular action, where the line shape indicates the predicted mode of action. The thickness of each edge indicates the confidence of interaction (0–1.0). In the current network, this confidence ranges from 0.4 to 0.9. The color of the edges indicate different action modes (binding, catalysis, and biochemical reaction). Binding was the only action

mode within each of the interaction clusters, except for FixABCX which involved all three action types. The dashed edges show the interaction between different clusters. Nodes are grouped and colored based on interaction clusters.

STRING "kmeans clustering" revealed five functionally distinct protein clusters (**Figure 5**). The largest cluster included uptake Group 1d, regulatory Group 2b and high-affinity Group 1h/5 hydrogenases and all hydrogenase accessory proteins; however, the Group 3b with its two identified subunits clustered separately from the main hydrogenase cluster. Complex I with its two subunits also grouped separately from the hydrogenase cluster. Though, these three clusters were highly interrelated. The next large cluster included the three subunits of pMMO1 and pMMO2, and NifH protein. The FixABCX complex formed another cluster (**Figure 5**). In both the main hydrogenase and pMMO clusters, most of the proteins were predicted to have strong protein-protein "binding" interactions. In addition to "binding" actions, there occurred a strong "catalysis" interaction between FixA, FixB and FixX proteins of the FixABCX complex. Moreover, FixA and FixB exhibited a "biochemical reaction" interaction linked to the electron transport reaction between these two proteins (**Figure 5**).

Four differentially regulated proteins were not assigned to any of the five clusters; namely ferredoxin (Fer2), cytochrome c oxidase subunit III (COX3), carbon starvation protein (CstA), and nitrogenase-associated NifX protein.

### *4.5.5. Differential SC2 gene expression in response to $H_2$ availability*

In order to relate our proteomics results to mRNA expression, we selected eight genes involved in methane, nitrogen, and hydrogen metabolisms to quantify their differential expression in response to $H_2$ availability by RT-qPCR. As for proteomics, samples for RNA extraction were collected after 19 h, 10 d, and 37 d of culture incubation under 6% $CH_4$/ 3% $O_2$/ He atmosphere, with and without the addition of 2% $H_2$. The eight genes encoded the large subunits (LSU) of Group 1d, Group 1h/5, Group 2b and Group 3b hydrogenases, *nifH* (nitrogenase), *pmoA1* (pMMO1) and *pmoA2* (pMMO2), and *rplI* (**Table S-1**). The latter was used as a reference in RT-qPCR. Overall, the transcript expression patterns of these genes agreed well with their proteomics results. In the presence of $H_2$, *the transcript levels* of Group 1d and Group 2b hydrogenases were significantly increased already after short-term (19 h) incubation (*p*-value ≤ 0.01 of their

$log_2$-fold change ratio values) and remained constantly high during mid-term (10 d) and long-term (37 d) incubations (**Figure 6 and Table S-8**). In none of the three incubation periods, the genes encoding LSU of the Group 1h/5 and Group 3b hydrogenases showed a differential transcript expression in response to $H_2$ availability.

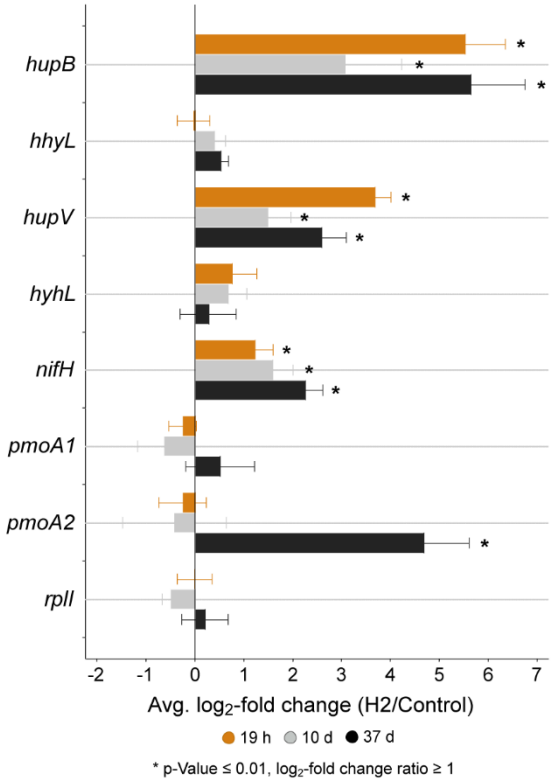



**Figure 6** | Validation of proteomics results by real-time quantitative reverse-transcription PCR (RT-qPCR). The changes in transcript numbers are shown as $log_2$-fold ratio values between $H_2$-treated and control samples. A total of eight genes were selected for RT-qPCR, encoding the large subunits of Group 1d (*hupB*), Group 1h/5 (*hhyL*), Group 2b (*hupV*) and Group 3b (*hyhL*) hydrogenases, *nifH* (nitrogenase), and *pmoA1* (pMMO1) and *pmoA2* (pMMO2) genes. In addition, the *rplI* gene encoding 50S ribosomal protein L9 was used as a reference. All RT-qPCR reactions were carried out in three biological replicates and three technical replicates for both $H_2$-treated and control cultures (total of nine RT-qPCR reactions for each treatment and time point). The PCR efficiency was at least 80% ($R^2$ > 0.90). The values with $log_2$-fold changes ≥ 1 and $p$ ≤ 0.01 were considered significant.

The expression of *pmoA1* gene was not affected at all by $H_2$ availability. The expression of *pmoA2* gene was not affected by $H_2$ availability after short-term (19 h) and mid-term (10 d) incubations, but its expression was significantly increased after long-term (37 d) incubation ($log_2$ fold change ratio of 1.4, $p$-value ≤ 0.01) (**Figure 6 and Table S-8**). Thus, the effect of $H_2$ availability on the differential expression of *pmoA2* corresponds perfectly to the differential regulation of PmoA2 and PmoB2 subunits (**Figure 6**). The

transcript level of *nifH* was increased in response to $H_2$ availability. This increase was significant in both short-term (19 h) and long-term (37 d) incubations. As expected, the *rplI* gene showed no differential transcript expression in response to $H_2$ availability (**Figure 6 and Table S-8**).

## 4.6. Discussion

In recent years, aerobic $H_2$ oxidizing bacteria were considered a major target for metabolic engineering to improve their ability to act as a source of renewable fuel molecules and high quality feed proteins (Nybo et al., 2015). Various proteobacterial and verrucomicrobial methanotrophs were shown to oxidize $CH_4$ and $H_2$ under aerobic conditions (Chen and Yoch, 1987; Mohammadi et al., 2017; Piche-Choquette and Constant, 2019). In particular, it was demonstrated that the biomass yield of the verrucomicrobial Methylacidiphilum sp. RTK17.1 significantly increased when $CH_4$ and $H_2$ were simultaneously supplied under $O_2$-limited conditions (3.5 %) (Carere et al., 2017). Recently, Greening and colleagues (Greening et al., 2016) showed that hydrogenase-encoding genes are widely distributed among alphaproteobacterial methanotrophs, including *Methylocystis* spp. A genomic survey of our model organism revealed us that strain SC2 possesses the complete genetic machinery for $H_2$ oxidation. Therefore, we decided to test the impact of $H_2$ supplementation on the growth of strain SC2.

### *4.6.1. Metabolic adaptation of strain SC2 to H2 availability under CH4/O2 limited conditions*

The results obtained for Methylacidiphilum sp. RTK17.1 prompted us to test the effect of a similar $H_2$ (2%) concentration on the growth activity of strain SC2 in combination with different $O_2/CH_4$ mixing ratios (Figure 1). We found that $H_2$ addition did not affect the $CH_4$ oxidation rates under elevated $O_2$ (15%) conditions. However, $O_2$ limitation (3%) in combination with carbon source limitation (6% $CH_4$) resulted in the greatest $H_2$ consumption rate among all tested conditions (Set IV, Figure 1 and Table S-3). In fact, SC2 cultures of Set IV showed twice higher $H_2$ oxidation rates than those (Set II) grown under elevated $CH_4$ concentration (Table S-2, Table S-3 and Table S-4). In addition, the

biomass yield was increased by more than two-fold during the exponential growth phase in Set IV cultures (Table S-2). Evidently, $H_2$ availability stimulates the growth of strain SC2 under $CH_4$ and $O_2$ limiting conditions. Moreover, the $CH_4$ oxidation rate under $O_2$-limiting conditions was decreased by 40% (Figure 3B). We assume that, similar to the verrucomicrobial strain RTK17.1 (Carere et al., 2017), Methylocystis sp. strain SC2 is able to regulate the rates of $CH_4$ and $H_2$ consumption in response to $O_2$ availability; in order to balance between energy generation and biomass yield. Furthermore, the thermodynamic model calculations (details in section 3.1) revealed that the theoretical max yield of cell carbon for $CH_4$/$H_2$-combined is two-fold higher than when $CH_4$-only is supplied. The ratio of the theoretical max yield between $H_2$ treatment and control corresponds perfectly to the experimental SC2 biomass yield ratio between $H_2$ treatment and control in Set IV (**Figure 3C**). Thus, the $H_2$ treatment in Set IV provided optimal growth conditions for strain SC2, thereby resulting in maximum biomass yield with less $CH_4$ and $O_2$ consumed (**Figure 3**).

To elucidate the physiological adaptation of SC2 cells to repeated $H_2$ supply, cultures were incubated under Set IV (6% $CH_4$/ 3% $O_2$) with or without the addition of 2% $H_2$. The SC2 cultures were continuously maintained in the exponential growth phase to maximize the $H_2$ effect and thus biomass yield (**Figure S-2**). The experiment revealed that the hydrogen effect was more substantial after long-term incubation (37 d) than after short-term incubation (19 h) (**Figure S-2**). Thus, strain SC2 adapts to the continuous supply of $H_2$ as an alternative energy source, thereby leading to an increased biomass yield and concomitant decrease in $CH_4$ consumption.

### *4.6.2. Differential regulation of key enzymes in response to $H_2$ availability*

Hydrogen availability induced a specific proteome response of strain SC2. The production of low-affinity pMMO1 was not affected by the $H_2$ treatment throughout the whole incubation period. However, pMMO2 known to oxidize $CH_4$ with high affinity (Baani and Liesack, 2008) was upregulated toward the end of the incubation period (37 d) (**Figure 4 and Table S-7**). Based on our physiological growth measurements (data not shown) we suggest that this effect was due to the misbalance between SC2 biomass and available $CH_4$. The supplied 3% $CH_4$ was rapidly oxidized by the cells. In the end of our

experimental treatment, the cell density was thus too high relative to $CH_4$ availability, triggering the upregulation of high-affinity pMMO2.

Regardless of whether or not hydrogen was added, all four known hydrogenases were well detectable in the SC2 proteome. However, $H_2$ addition was correlated with a significant increase in the enzyme levels of the low-affinity Group 1d hydrogenase and Group 2b regulatory hydrogenase in both short-term (19 h) and long-term (37 d) treatments. These two hydrogenases are known to be $O_2$-tolerant due to the presence of a [4Fe3S] cluster that protects the $O_2$-sensitive active site from oxidative damage (Greening et al., 2016). As discussed previously (Carere et al., 2017), these enzymes can therefore maintain their activity at relatively high $O_2$ tensions (3-15 %) as applied in our cultivation study. All three subunits of the low-affinity Group 1d hydrogenase were highly expressed and significantly upregulated. This enzyme catalyzes hydrogenotrophic respiration using $O_2$ as the terminal electron acceptor. It is a heterodimeric periplasmic [NiFe]-hydrogenase that is attached to the membrane by a b-type cytochrome and is primarily involved in the generation of a proton gradient through the cytoplasmic membrane (Friedrich et al., 2005). This enzyme is directly linked to the quinone pool of the aerobic respiratory chain, thereby supplying electrons for either the methane monooxygenase reaction or the terminal oxidase complex (Carere et al., 2017).

The significantly increased production of Group 2b regulatory hydrogenase enzyme may be explained by the requirement of aerobic $H_2$-oxidizing bacteria to respond rapidly to changes in $H_2$ supply, given that aerobic hydrogen oxidizers receive $H_2$ only occasionally and in limited amounts. Therefore, the regulatory hydrogenase that forms a complex with a histidine protein kinase rapidly recognizes $H_2$ in the environment and transmits the signal to a response regulator, which in turn controls transcription of the hydrogenase genes (Greening and Cook, 2014). Furthermore, it was shown that two hydrogenase operons (regulatory and uptake hydrogenases) are coordinately regulated and transcribed in response to $H_2$ and the amount of available carbon (Friedrich et al., 2005). Our RT-qPCR results suggest that such a coordinated regulation may also occur in strain SC2. The expression of genes encoding Group 2b hydrogenase (*hupV*) and Group 1d hydrogenase (*hupB*) were both significantly increased during the whole $H_2$ treatment (**Figure 6**). Thus, hydrogenase transcription is repressed under the elevated concentration of carbon sources (Schwartz et al., 1999), but the uptake hydrogenase will

be transcribed at high level under carbon limitation due to the inhibition of the histidine kinase of regulatory hydrogenase (Greening and Cook, 2014). In fact, the $H_2$ oxidation rate was very low when strain SC2 was grown under elevated $CH_4$ concentrations. We observed no growth rate difference between control and $H_2$-treated cultures (**Figure S-1 and Table S-3**). By contrast, SC2 cells rapidly responded to the addition of $H_2$ under limiting concentrations of $CH_4$ (6%) and $O_2$ (3%) in the headspace, resulting in the highest $H_2$ oxidation rate among all four treatments.

In addition to Group 1d uptake hydrogenase and Group 2b regulatory hydrogenase, nine proteins involved in the expression, formation, and maturation of the hydrogenase were significantly upregulated (**Figure 4 and Table S-7**). The genes (*hup* genes) encoding HupJ and HupH proteins are usually localized downstream from the structural genes (Vignais et al., 2001). The Hup proteins are involved in the maturation of the heterodimeric hydrogenase enzyme. Another set of proteins, encoded by the *hyp* ('p' for pleiotropic) genes, is involved in the insertion of Ni, Fe, CO and CN into the active site of both uptake and regulatory hydrogenases. Accordingly, all Hup and Hyp proteins clustered together with uptake and regulatory hydrogenases in the predicted protein-protein interaction network (**Figure 5**). Moreover, the network indicated a strong "binding" interaction between the two hydrogenases (Group 1d and 2b) and Hup maturation proteins, presumably due to the fact that most hydrogenases exhibit oxidoreductase activity and are involved in the electron transport chain. A highly specific binding between proteins of the electron transport chain is a prerequisite for ensuring efficient electron transfer (Le and Ou, 2016).

In contrast to the Group 1d and Group 2b hydrogenases, the addition of 2% $H_2$ had no effect on the low expression level of Group 1h/5 high-affinity hydrogenase (**Figure 4 and Table S-7**). This hydrogenase catalyzes aerobic hydrogen oxidation under very low, even atmospheric (ca. 0.53 ppmv), mixing ratios of $H_2$ (Constant et al., 2010; Greening et al., 2014; Greening et al., 2016; Schafer et al., 2013). Since we aimed to investigate the $H_2$ effect on the physiology and proteome of exponentially growing SC2 cells, we refilled the culture headspace during the long-term incubation (37 d) every 3 to 4 days with a fresh gas mixture (**Figure 4**). Apparently, this $H_2$ level did not trigger differential regulation of Group 1h/5 high-affinity hydrogenase.

Likewise, the addition of 2% $H_2$ had no effect on the expression level of Group 3b soluble hydrogenase under limiting $CH_4$ (6%) and $O_2$ (3%) conditions (**Table S-7**). The Group 3b hydrogenases are known to be bidirectional hydrogenases (Piche-Choquette and Constant, 2019), meaning that they either directly couple oxidation of NAD(P)H to the evolution of $H_2$ or oxidize $H_2$ by the reverse reaction (Ma et al., 1993). In addition, Group 3b hydrogenase clustered separately from the main hydrogenase cluster in our predicted protein-protein interactions, (**Figure 5**). Overall, our results confirm that the expression of Group 3b hydrogenase is controlled by factors other than those shown to regulate the production of Group 1d / Group 1h/5 hydrogenases. Despite the fact that Group 3b hydrogenases are known for more than 20 years (Ma et al., 1993), the factors controlling their expression remained elusive (Vignais et al., 2001). Thus, these factors need to be elucidated in future studies.

Another important finding of our proteomics experiments was the upregulation of NifH and NifX nitrogenase proteins, even though our experiments were conducted under a nitrogen-free headspace. Strain SC2 contains the complete set of nitrogenase genes and is able to fix atmosphere $N_2$, when alternative nitrogen sources are unavailable (Dam et al., 2013). Many diazotrophic (nitrogen-fixing) *Proteobacteria* and *Cyanobacteria* possess uptake hydrogenases encoded by *hup* genes, which are usually co-expressed with the nitrogenase complex (Annan et al., 2012; Brito et al., 2005; Vignais et al., 2001). These uptake hydrogenases are known to reutilize $H_2$ produced by nitrogenase activity, thereby reconstituting some of the ATP consumed during nitrogen fixation (Rey et al., 2006). Moreover, in proteobacterial diazotrophic bacteria, special NtrC-like transcription factors, namely HoxA (Durmowicz and Maier, 1997; Van Soom et al., 1997), control the transcription of genes involved in nitrogen metabolism and, in addition, *hupSL* genes encoding uptake hydrogenase (Weyman et al., 2008). In the genome of strain SC2, *hoxA* (gene ID BN69_2342) is located upstream of the *hupSL* (Group 1d) and *hupVU* (Group 2b) hydrogenase genes (Dam et al., 2012). In our study, the transcription factor encoded by *hoxA* was higher expressed in SC2 cells in $H_2$ treatment than in control (gene ID BN69_2342, **Table S-5**). Thus, like in other diazotrophic bacteria, *hoxA* may induce not only the upregulation of uptake and regulatory hydrogenases in strain SC2, but also particular structural subunits of nitrogenase.

### *4.6.3. Proposed model of electron flow and energy conservation in strain SC2 under mixed utilization of $CH_4$ and $H_2$*

Recently, Bordel and colleagues showed by genome-scale metabolic modeling and experimental research of three *Methylocystis* spp., including strain SC2, that the electron supply for methane oxidation functions as a so-called "redox arm" mechanism (Bordel et al., 2019). Under non-limiting culture conditions, complex I (NADH dehydrogenase I) catalyzes the transfer of electrons from NADH to ubiquinone upon oxidative phosphorylation. Eventually, ubiquinone further serves as the electron donor for pMMO.

Our proteomics results revealed that upon long-term incubation with $H_2$ and constant upregulation of Group 1d hydrogenase, complex I was significantly downregulated. Concurrently, FixABCX complex was significantly upregulated (**Figure 4 and Table S-7**). Thus, our proteomics results suggest that during our long-term (37 d) incubation experiment, the functional role of complex I was taken over by Group 1d hydrogenase and FixABCX complex. The latter has been shown to oxidize NADH by flavin-based electron bifurcation (Ledbetter et al., 2017). In general, during electron bifurcation, pairs of electrons are transferred to two different one-electron acceptors; one with higher reduction potential and the other with a lower reduction potential (Buckel and Thauer, 2018). Upon NADH oxidation by the FixABCX complex, the pair of electrons initiate first the bifurcation of FixA, with one electron being transferred to the quinone pool via the flavins in FixB and FixC and the other to ferredoxin via the low-potential FixX (Edgren and Nordlund, 2004; Ledbetter et al., 2017). Genes encoding FixABCX proteins usually are organized within the *nif* cluster. Under nitrogen-fixing conditions, the Fix proteins further transfer electrons to the nitrogenase complex via ferredoxin, thereby providing energy for nitrogenase activity (Carvalho et al., 2010). As already discussed above, it is well known that in most diazotrophic bacteria, production and activity of uptake hydrogenase and nitrogenase are linked to each other. Because our experiments were carried out under a nitrogen-free atmosphere, we propose that not only one but both NADH-derived electrons were transferred to ubiquinone. Thus, ubiquinone may have served as both high-potential electron acceptor and low-potential electron acceptor in strain SC2 under $H_2$-utilizing growth conditions. Due to its water solubility ubiquinone can interact with both FixC and FixX (Buckel and Thauer, 2018). In

summary, it may be reasonable to speculate that long-term incubation of strain SC2 under $H_2$ atmosphere induces a switch from complex I to flavin-based electron bifurcation to provide additional electrons to the electron transport chain.

## 4.7. Conclusions

Our results of both proteomics and RT-qPCR prompted us to propose a model (**Figure 7**) of how strain SC2 responds to $H_2$ availability under $CH_4$ and $O_2$ limitation. Our results showed that under long-term incubation (37d) with 2% $H_2$, strain SC2 reconfigures the electron transport chain. It uses the membrane-integrated Group 1d uptake hydrogenase to provide electrons to the electron transport chain through the quinone pool. Reduced quinone then transfers the electrons either to pMMO for $CH_4$ oxidation or to cytochrome c and terminal oxidase for ATP generation. Furthermore, production and activity of uptake hydrogenase is continuously controlled by Group 2b regulatory hydrogenase and accessory Hup and Hyp proteins. Moreover, NADH oxidation may provide additional electrons through activity of the FixABCX bifurcation complex.

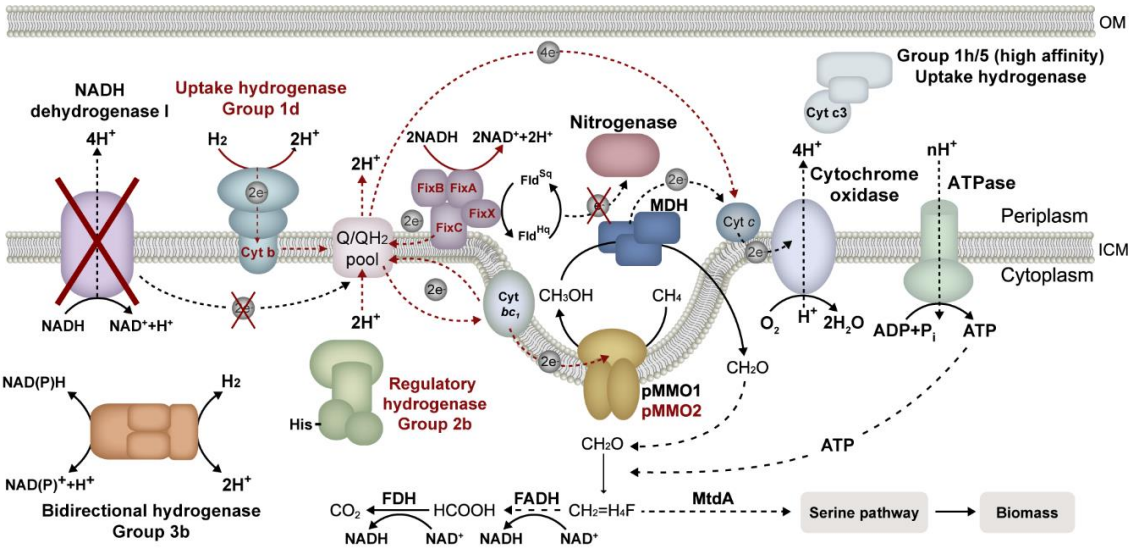



**Figure 7 |** Proposed model of metabolic response of *Methylocystis* sp. strain SC2 to the addition of $H_2$ under limited $CH_4$ and $O_2$ conditions. During the growth of strain SC2, $H_2$ is oxidized by Group 1d uptake hydrogenase, yielding reduced quinones ($QH_2$) and, in consequence, a large proton-motive force for production of ATP. In addition, some of the $QH_2$ provides electrons for $CH_4$ oxidation by pMMO. Formaldehyde oxidation via the activity of formaldehyde dehydrogenase (FADH) and formate dehydrogenase (FDH) leads to the production of reducing equivalents (NADH) that feed the respiratory chain through the FixABCX complex, resulting in additional ATP available for growth. All the proteins and pathways that were differentially regulated are highlighted in red.

In summary, strain SC2 can utilize $H_2$ as an additional energy source to optimize biomass yield under $CH_4$ and $O_2$ limitation. Hydrogen for use as an electron donor to support methanotroph growth can be produced by non-expensive green energy sources such as wind and solar energy powered water electrolysis (Matassa et al., 2015). Given that most *Methylocystis* and *Methylosinus* spp. have been shown to possess the genetic potential to produce Group 1d uptake hydrogenases and Group 2b regulatory hydrogenases, the metabolic capacity to utilize $H_2$ is most likely not only limited to strain SC2 but widely distributed among type IIa methanotrophs.

## 4.8. Acknowledgements

AH is a member of the International Max Planck Research School for Environmental, Cellular, and Molecular Microbiology (IMPRS-Mic). AH thanks Christiane Nüsslein-Vollhard-Foundation (CNV) for CNV grant awarded to excellent women scientists with children in the field of experimental sciences. JZ was supported by a Ph.D. scholarship of the Chinese Scholarship Council. This study was supported by the Deutsche Forschungsgemeinschaft (DFG) through Collaborative Research Center SFB987. Prof. Dr. Rudolf K. Thauer is acknowledged for insightful discussion. Authors thank the in-house research group of Prof. Dr. Andreas Brune for providing Unisense Multimeter instrument to conduct $H_2$ oxidation measurements.

## 4.9. Supporting information

Supporting figures are provided below in section 4.9.1. In addition, Tables S-1 to S-4 and S-7 are provided in section 4.9.2. The extended supporting Information, including all supporting tables (Table S-1 to S-8), is included in the DVD provided with the hard copy of the thesis.

### *4.9.1. Supporting figures*

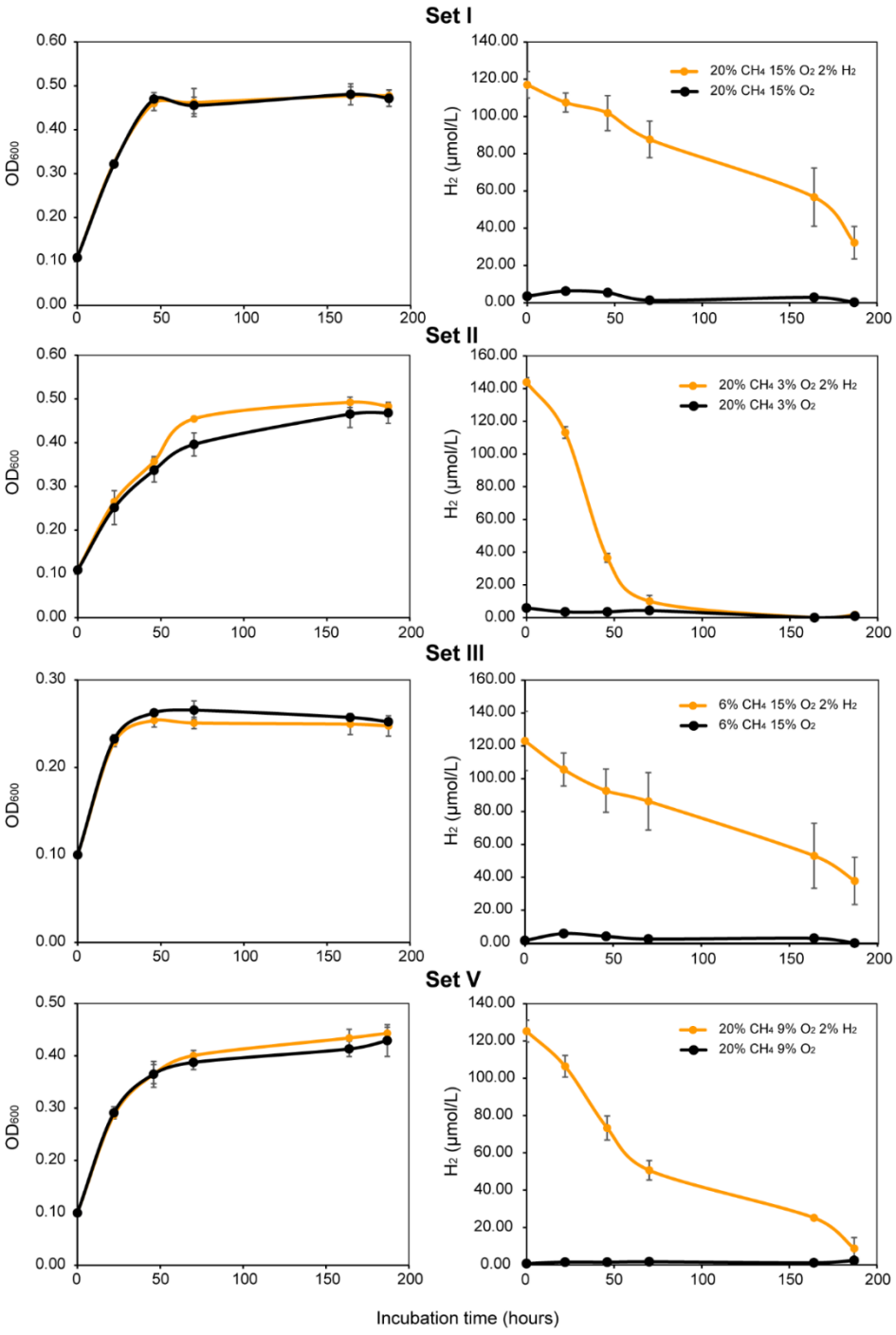
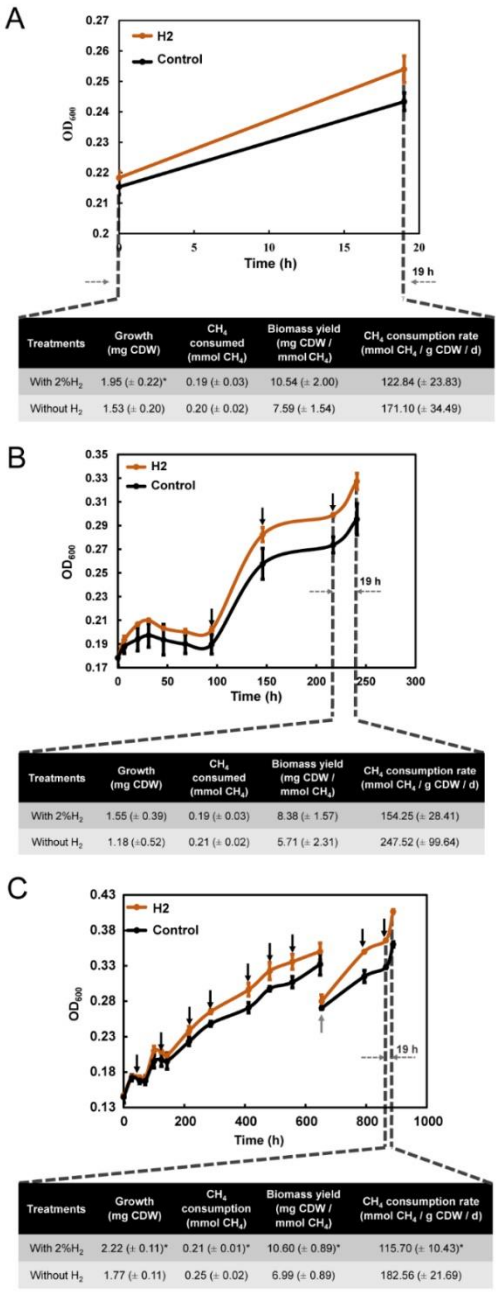



**Figure S-1 |** Physiological growth experiments of *Methylocystis* sp. strain SC2 with or without addition of $H_2$. The headspace was filled with five different $O_2/CH_4$ ratios (Experimental Sets I to V in Table 1), with or without the addition of 2% $H_2$, and finally balanced with helium (He). On the left graphs growth curves of strain SC2 are shown. As an initial biomass the mid-log phase grown cells ($OD_{600}$=0.109) of strain SC2 were used. The $H_2$-treated and control cultures were incubated in total for 187 h. On the right graphs oxidation of the headspace $H_2$ is shown. Hydrogen oxidation was measured over 187 h using $H_2$ microsensor with piercing-needle from Unisens A/S. The error bars show the standard deviation calculated from triplicate measurements.




| Treatments | Growth (mg CDW) | $CH_4$ consumed (mmol $CH_4$) | Biomass yield (mg CDW / mmol $CH_4$) | $CH_4$ consumption rate (mmol $CH_4$ / g CDW / d) |
|---|---|---|---|---|
| With 2%$H_2$ | 1.95 (± 0.22)* | 0.19 (± 0.03) | 10.54 (± 2.00) | 122.84 (± 23.83) |
| Without $H_2$ | 1.53 (± 0.20) | 0.20 (± 0.02) | 7.59 (± 1.54) | 171.10 (± 34.49) |

| Treatments | Growth (mg CDW) | $CH_4$ consumed (mmol $CH_4$) | Biomass yield (mg CDW / mmol $CH_4$) | $CH_4$ consumption rate (mmol $CH_4$ / g CDW / d) |
|---|---|---|---|---|
| With 2%$H_2$ | 1.55 (± 0.39) | 0.19 (± 0.03) | 8.38 (± 1.57) | 154.25 (± 28.41) |
| Without $H_2$ | 1.18 (±0.52) | 0.21 (± 0.02) | 5.71 (± 2.31) | 247.52 (± 99.64) |

| Treatments | Growth (mg CDW) | $CH_4$ consumption (mmol $CH_4$) | Biomass yield (mg CDW / mmol $CH_4$) | $CH_4$ consumption rate (mmol $CH_4$ / g CDW / d) |
|---|---|---|---|---|
| With 2%$H_2$ | 2.22 (± 0.11)* | 0.21 (± 0.01)* | 10.60 (± 0.89)* | 115.70 (± 10.43)* |
| Without $H_2$ | 1.77 (± 0.11) | 0.25 (± 0.02) | 6.99 (± 0.89) | 182.56 (± 21.69) |

**Figure S-2 |** Short- and long-term incubations of strain SC2 with and without addition of $H_2$. The experiments were performed using 6% $CH_4$/ 3% $O_2$ headspace. The cultures were incubated for a total of 37 d, with and without the addition of 2% $H_2$. Samples for physiological and proteomics analysis were taken after an incubation period of 19 h, 10 d, and 37 d. The headspace of the 10-day and 37-day cultures was readjusted to the starting concentrations of $CH_4$/$O_2$/$H_2$ every three to four days. Black arrows in the growth curve show the addition of fresh gaseous headspace. The gray arrow shows where cultures were washed and supplemented with fresh NMS medium. The tables illustrate the biomass yield, growth rate and $CH_4$ consumption rate in the control and $H_2$ treatments. As an initial biomass the mid-log phase grown cells of strain SC2 were used. Gram cell dry weight (gCDW) measurements were determined by correlating $OD_{600}$ measurements to an existing standard curve ($OD_{600}$ = 1 corresponds to 0.26 (± 0.01) gCDW $l^{-1}$). For each calculated parameter the average value of three replicates is shown. The standard deviation is presented in parentheses. The difference between control and treated cultures with $p < 0.05$ (paired t-test, one-tailed) was considered as significant.

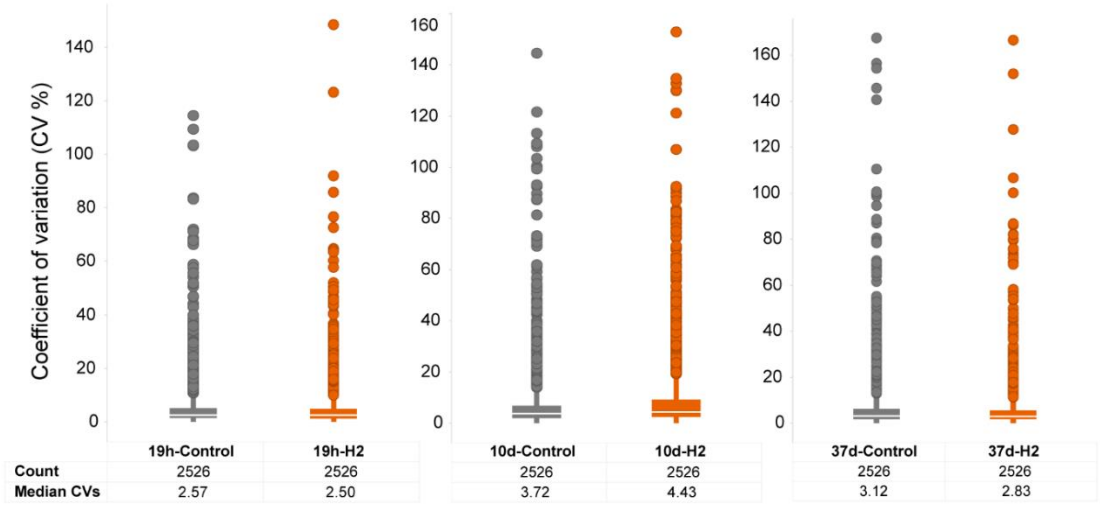



**Figure S-3 |** Median coefficient of variation (CV) after label-free quantification (LFQ) for the treatments with and without 2% H2 after 19h, 10d and 37d of incubation. The median CVs were calculated for each conditional group upon LC-MS and LFQ analysis (see Experimental section). Median CVs in percentage and total count of identified proteins are shown. The experiments were conducted in triplicates.

## *4.9.2. Supporting tables (Tables S-1, S-2, S-3, S-4, S-7)*

**Table S-1** | Primers used in the current study for RT-qPCR.

| Protein description | Genes | Primers | Annealing temperature | Sequences (5'-3') | Reference |
|---|---|---|---|---|---|
| **Group 1d H2-ase LSU** | *hupB* | hupB-617F | 54 °C | GCCCCTTCAAGAACGGCTAT | This study |
| | | hupB-617R | | AGGTTCAGGCGTTCCATGTT | This study |
| **Group 1h/5 H2-ase LSU** | *hhyL* | hhyL-1625F | 57 °C | ACGAGGACGCGGTTCAGAATAC | This study |
| | | hhyL-1625R | | TCGGCGAATGCGTGGTCTTA | This study |
| **Group 2b H2-ase LSU** | *hupV* | hupV-580F | 55 °C | CGCTTCCTTGAGACGACGGT | This study |
| | | hupV-669R | | CCCTTCGCAATAACGCTCCA | This study |
| **Group 3b H2-ase LSU** | *hyhL* | hyhL-590F | 49 °C | CGGCTGGTTTTGATTTTCC | This study |
| | | hyhL-673R | | CCTGGTTGAAGGGATATTCG | This study |
| **PmoA1 (pMMO1 complex)** | *pmoA1* | pmoA1-349F | 55 °C | ACCTACTTCCCGATCAGCCT | This study |
| | | pmoA1-442R | | TCACATAGGAGCCCGACAGA | This study |
| **PmoA2 (pMMO2 complex)** | *pmoA2* | pmoA2-96F | 52 °C | TCTTCTGTTCTTCGCCGTTC | This study |
| | | pmoA2-193R | | ACATACGACGATCCTTCCAGTC | This study |
| **Nitrogenase** | *nifH* | nifH-481F | 55 °C | GCGGCCAACAACATCTCCAA | This study |
| | | nifH-481R | | CTGGGACTCCGCTGCATATT | This study |
| **Ribosomal 50S-L9** | *rplI* | 50SL9-F | 51 °C | GCAATCTCGAGGCTAGGAA | Han et al., 2017 |
| | | 50SL9-R | | ATCGTCCATCGACGTCTCTT | Han et al., 2017 |

**Table S-2 |** Physiological growth parameters of exponentially growing SC2 cultures determined for the experimental sets I to V.

| Treatments | Growth (mgCDW) | Growth rate (mgCDW/d) | Yield (mgCDW/ mmol $CH_4$) | $CH_4$ consumed (mmol) | $H_2$ consumed (mmol) | $CH_4$ consumprion rate (mmol $CH_4$/gCDW/d) | $H_2$ consumprio n rate (mmol $H_2$/gCDW/d) |
|---|---|---|---|---|---|---|---|
| **Set I-H2** | 5.50 (± 0.27) | 0.71 (± 0.03) | 6.06 (± 0.12) | 0.91 (± 0.03) | 0.02 (± 0.00) | 21.19 (± 0.42) | 0.36 (± 0.07) |
| **Set I-Control** | 5.64 (± 0.23) | 0.72 (± 0.03) | 6.10 (± 0.28) | 0.93 (± 0.03) | – | 21.10 (± 0.96) | – |
| **Set II-H2** | 3.89 (± 0.17) | 0.50 (± 0.02) | **5.63 (± 0.81)*** | **0.70 (± 0.07)*** | 0.11 (± 0.00) | **23.27 (± 3.25)*** | 3.55 (± 0.17) |
| **Set II-Control** | 3.57 (± 0.42) | 0.46 (± 0.05) | 4.03 (± 0.29) | 0.88 (± 0.05) | – | 32.05 (± 2.40) | – |
| **Set III-H2** | 2.40 (± 0.12) | 0.31 (± 0.02) | 8.16 (± 0.32) | 0.29 (± 0.00) | 0.03 (± 0.02) | 15.76 (± 0.63) | 1.60 (± 0.84) |
| **Set III-Control** | 2.54 (± 0.04) | 0.33 (± 0.01) | 8.05 (± 0.40) | 0.32 (± 0.01) | – | 15.98 (± 0.77) | – |
| **Set IV-H2** | **2.71 (± 0.04)**** | **0.35 (± 0.01)**** | **12.53 (± 0.79)**** | **0.22 (± 0.01)*** | 0.09 (± 0.01) | **10.28 (± 0.63)**** | 4.24 (± 0.52) |
| **Set IV-Control** | 2.17 (± 0.04) | 0.28 (± 0.00) | 5.88 (± 0.52) | 0.37 (± 0.04) | – | 22.00 (± 2.05) | – |
| **Set V-H2** | 4.14 (± 0.39) | 0.53 (± 0.05) | 3.16 (± 0.31) | 1.31 (± 0.04) | 0.05 (± 0.00) | 40.93 (± 3.72) | 1.62 (± 0.13) |
| **Set V-Control** | 4.15 (± 0.28) | 0.53 (± 0.04) | 3.21 (± 0.21) | 1.29 (± 0.01) | – | 40.19 (± 2.63) | – |

**Table S-3 |** Physiological growth parameters of exponentially growing SC2 cultures determined for the experimental sets I to V.

| Treatments | Growth (mgCDW) | Growth rate (mgCDW/d) | Yield (mgCDW/ mmol $CH_4$) | $CH_4$ consumed (mmol) | $H_2$ consumed (mmol) | $CH_4$ consumprion rate (mmol $CH_4$/gCDW/d) | $H_2$ consumprion rate (mmol $H_2$/gCDW/d) |
|---|---|---|---|---|---|---|---|
| **Set I-H2** | 5.52 (± 0.50) | 0.71 (± 0.06) | 5.16 (± 0.48) | 1.07 (± 0.02) | 0.03 (± 0.01) | 25.13 (± 2.52) | 0.67 (± 0.23) |
| **Set I-Control** | 5.42 (± 0.29) | 0.70 (± 0.04) | 4.98 (± 0.10) | 1.09 (± 0.04) | – | 25.77 (± 0.51) | – |
| **Set II-H2** | **5.41 (± 0.03)*** | **0.69 (± 0.00)*** | **5.81 (± 0.19)*** | **0.93 (± 0.03)*** | 0.13 (± 0.00) | **22.13 (± 0.72)*** | 3.18 (± 0.04) |
| **Set II-Control** | 4.49 (± 0.41) | 0.58 (± 0.05) | 4.28 (± 0.38) | 1.05 (± 0.02) | – | 30.25 (± 2.83) | – |
| **Set III-H2** | 2.36 (± 0.10) | 0.30 (± 0.01) | 6.85 (± 0.45) | 0.34 (± 0.01) | 0.03 (± 0.02) | 18.81 (± 1.22) | 1.81 (± 0.98) |
| **Set III-Control** | 2.59 (± 0.17) | 0.33 (± 0.02) | 7.12 (± 0.65) | 0.36 (± 0.02) | – | 18.17 (± 1.72) | – |
| **Set IV-H2** | **2.89 (± 0.19)**** | **0.37 (± 0.02)**** | **10.30 (± 1.04)*** | **0.28 (± 0.02)*** | 0.14 (± 0.00) | **12.58 (± 1.27)*** | 6.33 (± 0.42) |
| **Set IV-Control** | 2.37 (± 0.08) | 0.30 (± 0.01) | 5.57 (± 0.54) | 0.43 (± 0.04) | – | 23.26 (± 2.26) | – |
| **Set V-H2** | 4.70 (± 0.15) | 0.60 (± 0.02) | 3.56 (± 0.10) | 1.32 (± 0.03) | 0.06 (± 0.01) | 36.10 (± 1.02) | 1.77 (± 0.37) |
| **Set V-Control** | 4.50 (± 0.21) | 0.58 (± 0.03) | 3.44 (± 0.12) | 1.30 (± 0.02) | – | 37.31 (± 1.28) | – |

**Table S-4 |** Physiological growth parameters of late stationary SC2 cultures determined for the experimental sets I to V.

| Treatments | Growth (mgCDW) | Growth rate (mgCDW/d) | Yield (mgCDW/ mmol $CH_4$) | $CH_4$ consumed (mmol) | $H_2$ consumed (mmol) | $CH_4$ consumprion rate (mmol $CH_4$/gCDW/d) | $H_2$ consumprion rate (mmol $H_2$/gCDW/d) |
|---|---|---|---|---|---|---|---|
| **Set I-H2** | 5.76 (± 0.22) | 0.74 (± 0.03) | 4.61 (± 0.29) | 1.25 (± 0.04) | 0.08 (± 0.02) | 27.94 (± 1.67) | 1.89 (± 0.38) |
| **Set I-Control** | 5.67 (± 0.29) | 0.73 (± 0.04) | 4.52 (± 0.10) | 1.25 (± 0.05) | – | 28.39 (± 0.65) | – |
| **Set II-H2** | 5.85 (± 0.13) | 0.54 (± 0.08) | 5.02 (± 0.33) | 1.17 (± 0.09) | 0.14 (± 0.00) | 25.67 (± 1.74) | 3.12 (± 0.06) |
| **Set II-Control** | 5.62 (± 0.38) | 0.51 (± 0.05) | 4.80 (± 0.43) | 1.17 (± 0.05) | – | 26.94 (± 2.46) | – |
| **Set III-H2** | 2.30 (± 0.18) | 0.30 (± 0.02) | 5.50 (± 0.46) | 0.42 (± 0.00) | 0.09 (± 0.02) | 23.51 (± 2.11) | 4.74 (± 1.16) |
| **Set III-Control** | 2.38 (± 0.03) | 0.31 (± 0.00) | 5.60 (± 0.37) | 0.43 (± 0.02) | – | 23.04 (± 1.47) | – |
| **Set IV-H2** | 2.42 (± 0.15) | 0.31 (± 0.02) | **6.73 (± 0.95)*** | **0.37 (± 0.04)*** | 0.15 (± 0.00) | **19.47 (± 2.82)*** | 7.89 (± 0.46) |
| **Set IV-Control** | 2.19 (± 0.07) | 0.28 (± 0.01) | 3.96 (± 0.20) | 0.55 (± 0.05) | – | 32.46 (± 1.65) | – |
| **Set V-H2** | 5.37 (± 0.19) | 0.69 (± 0.02) | 4.03 (± 0.04) | 1.33 (± 0.03) | 0.12 (± 0.01) | 31.85 (± 0.33) | 2.79 (± 0.19) |
| **Set V-Control** | 5.15 (± 0.47) | 0.66 (± 0.06) | 3.92 (± 0.30) | 1.31 (± 0.02) | – | 32.92 (± 2.43) | – |

**Table to S-2 to S-4**; gram cell dry weight (g CDW) was determined by correlating OD600 measurements to predetermined OD600 of 1, (OD600 = 1 corresponds to 0.26 (± 0.01) g CDW l-1).

Biomass yield; mg CDW per mmol CH4 consumed.

Methane consumption rate; mmol CH4 consumed per g CDW per day.
Hydrogen consumption rate; mmol H2 consumed per g CDW per day.

*Significant increase (p-value ≤ 0.05) in H2 treatment compared to control, paired t-test (one-tailed).

** Significant increase (p-value ≤ 0.01) in H2 treatment compared to control, paired t-test (one-tailed).

**Table S-7 |** List of selected significantly regulated proteins in response to H2 treatment after 19 h, 10 d, and 37 d of incubation (corresponds to volcano plots in Fig. 4).

| Protein | Protein ID | $Log_2$ fold change (H2/ Control) | *q*-value (H2/ Control) | $Log_2$ fold change (H2/ Control) | *q*-value (H2/ Control) | $Log_2$ fold change (H2/ Control) | *q*-value (H2/ Control) |
|---|---|---|---|---|---|---|---|
| | | 19 hours | | 10 days | | 37 days | |
| **Group 1d-LSU** | **J7QQM9** | 3.8 | 0.0000 | 2.9 | 0.0239 | 3.1 | 0.0123 |
| **Group 1d-SSU** | **J7Q4R0** | 3.0 | 0.0000 | 3.2 | 0.0162 | 3.3 | 0.0058 |
| **Group 1d-cytochrome b** | **J7Q9B4** | 2.3 | 0.0456 | 2.2 | 0.0072 | 2.5 | 0.0026 |
| **Group 1h/5 LSU** | **J7QAK4** | 0.0 | 0.9231 | -0.2 | 0.5613 | 0.4 | 0.1047 |
| **Group 1h/5 SSU** | **J7QJ71** | 0.3 | 0.7611 | 0.1 | 0.5852 | 1.2 | 0.3402 |
| **Group 2b-LSU** | **J7QUF4** | 0.5 | 0.0876 | 1.4 | 0.0237 | 1.7 | 0.0240 |
| **Group 2b SSU** | **J7QHN0** | 3.1 | 0.0090 | 1.7 | 0.0320 | 1.9 | 0.0125 |
| **Group 3b LSU** | **J7QFF8** | -0.1 | 0.3808 | 0.4 | 0.0364 | 0.1 | 0.0651 |
| **Group 3b SSU** | **J7QM94** | 0.3 | 0.4487 | -0.2 | 0.0878 | -0.1 | 0.9532 |
| **HypA/H2-ase maturation** | **J7Q4R2** | 4.9 | 0.0019 | 3.2 | 0.0518 | 6.0 | 0.0865 |
| **HypB/H2-ase maturation** | **J7QQP2** | 2.2 | 0.0000 | 1.4 | 0.0206 | 2.3 | 0.0091 |
| **HypC/HupF/H2-ase maturation** | **J7Q4Q9** | 2.8 | 0.0009 | 5.1 | 0.0379 | 4.3 | 0.0208 |
| **HypD/H2-ase maturation** | **J7QUF7** | 3.2 | 0.0009 | 1.7 | 0.0556 | 2.5 | 0.0061 |
| **HypE/H2-ase maturation** | **J7Q4R1** | 2.3 | 0.0000 | 2.3 | 0.0390 | 2.8 | 0.0347 |
| **HupJ/H2-ase maturation** | **J7QHN4** | 1.0 | 0.0621 | 0.5 | 0.1389 | 0.9 | 0.0469 |
| **HupH/H2-ase maturation** | **J7QQP7** | 4.0 | 0.0000 | 2.3 | 0.0244 | 4.8 | 0.0209 |
| **H2-ase expression/formation** | **J7QHM7** | 3.6 | 0.0000 | 4.2 | 0.0426 | 4.5 | 0.0220 |
| **H2-ase expression/formation** | **J7Q4R3** | 3.0 | 0.0000 | 2.0 | 0.0189 | 3.7 | 0.0061 |
| **PmoA1** | **Q70EF3** | 0.0 | 0.6206 | -0.2 | 0.0375 | -0.2 | 0.0061 |
| **PmoB1** | **Q70EF2** | 0.0 | 0.3014 | -0.1 | 0.4244 | -0.1 | 0.0294 |
| **PmoC1** | **J7QM98** | -0.1 | 0.7547 | -0.5 | 0.0729 | -0.3 | 0.0043 |
| **PmoA2** | **Q6MZ17** | 0.3 | 0.6768 | 0.1 | 0.4969 | 1.6 | 0.0070 |
| **PmoB2** | **Q6MZ16** | 0.1 | 0.7484 | 0.0 | 0.5224 | 2.3 | 0.0059 |
| **Carbon starvation protein CstA** | **J7QQF5** | -0.4 | 0.4067 | -1.1 | 0.3932 | 1.0 | 0.0395 |
| **NifH/Nitrogenase** | **J7Q4W7** | 2.3 | 0.0003 | 1.2 | 0.0102 | 1.5 | 0.0009 |
| **NIfX/Nitrogenase-associated** | **J7QV42** | 1.0 | 0.0075 | 1.0 | 0.0143 | 0.9 | 0.0055 |
| **FixC/FixABC complex** | **J7Q4X7** | 0.3 | 0.5259 | 0.6 | 0.0239 | 1.0 | 0.0164 |
| **FixA/FixABC complex** | **J7Q9P3** | 0.5 | 0.6033 | 0.8 | 0.0202 | 1.3 | 0.0030 |
| **FixB/FixABC complex** | **J7QRH1** | 0.0 | 0.9666 | 0.5 | 0.0402 | 1.0 | 0.0028 |
| **FixX/Ferredoxin** | **J7QV82** | 0.0 | 1.0000 | 2.6 | 0.0162 | 2.4 | 0.0220 |
| **Ferredoxin** | **J7Q9P9** | 3.5 | 0.0124 | -3.5 | 0.5515 | -0.1 | 0.5892 |
| **NADH dehydrogenase I 20kDa** | **J7Q516** | -0.1 | 0.5517 | -0.4 | 0.1228 | -1.5 | 0.0046 |
| **NADH dehydrogenase I 49kDa** | **J7QVP4** | 0.0 | 0.8534 | -0.7 | 0.2116 | -1.1 | 0.0046 |
| **Cytochrome c oxidase III COX3** | **J7Q4T0** | -0.3 | 0.6615 | -0.3 | 0.1688 | -1.2 | 0.0061 |

*Gray-highlighted rows show the significant ($log_2$ fold change ratio < -1 and > 1, *q*-value < 0.05) upregulation or downregulation of the proteins.

## 4.10. References


Annan, H., Golding, A. L., Zhao, Y., Dong, Z. (2012). Choice of hydrogen uptake (Hup) status in legume-rhizobia symbioses. *Ecol Evol* 2**,** 2285-90.

Baani, M., Liesack, W. (2008). Two isozymes of particulate methane monooxygenase with different methane oxidation kinetics are found in *Methylocystis* sp. strain SC2. *Proc Natl Acad Sci USA* 105**,** 10203-10208.

Belova, S. E., Baani, M., Suzina, N. E., Bodelier, P. L. E., Liesack, W., Dedysh, S. N. (2011). Acetate utilization as a survival strategy of peat-inhabiting *Methylocystis* spp. *Environ Microbiol Rep* 3**,** 36-46.

Bordel, S., Rodriguez, Y., Hakobyan, A., Rodriguez, E., Lebrero, R., Munoz, R. (2019). Genome scale metabolic modeling reveals the metabolic potential of three type II methanotrophs of the genus *Methylocystis*. *Metab Eng* 54**,** 191-199.

Bowman, J. (2006). The Methanotrophs - The Families *Methylococcaceae* and *Methylocystaceae*. In: Dworkin, M. (Ed.), *The Prokaryotes: A Handbook on the Biology of Bacteria. V. 5. Springer Science+Business Media*, LLC, New York, USA, pp. 282-283.

Brito, B., Baginsky, C., Palacios, J. M., Cabrera, E., Ruiz-Argueso, T., Imperial, J. (2005). Biodiversity of uptake hydrogenase systems from legume endosymbiotic bacteria. *Biochem Soc Trans* 33**,** 33-5.

Brune, A., Frenzel, P., Cypionka, H. (2000). Life at the oxic-anoxic interface: microbial activities and adaptations. *FEMS Microbiol Rev* 24**,** 691-710.

Buckel, W., Thauer, R. K. (2018). Flavin-based electron bifurcation, A new mechanism of biological energy coupling. *Chem Rev* 118**,** 3862-3886.

Carere, C. R., Hards, K., Houghton, K. M., Power, J. F., McDonald, B., Collet, C., Gapes, D. J., Sparling, R., Boyd, E. S., Cook, G. M., Greening, C., Stott, M. B., (2017). Mixotrophy drives niche expansion of verrucomicrobial methanotrophs. *ISME J* 11**,** 2599-2610.

Carvalho, F. M., Souza, R. C., Barcellos, F. G., Hungria, M., Vasconcelos, A. T. (2010). Genomic and evolutionary comparisons of diazotrophic and pathogenic bacteria of the order *Rhizobiales*. *BMC Microbiol* 10**,** 37.

Chen, Y., Dumont, M. G., Cebron, A., Murrell, J. C. (2007). Identification of active methanotrophs in a landfill cover soil through detection of expression of 16S rRNA and functional genes. *Environ Microbiol* 9**,** 2855-2869.

Chen, Y. P., Yoch, D. C. (1987). Regulation of two nickel-requiring (inducible and constitutive) hydrogenases and their coupling to nitrogenase in *Methylosinus trichosporium* Ob3b. *J Bacteriol* 169**,** 4778-4783.

Chidambarampadmavathy, K., Obulisamy, P. K., Heimann, K. (2015). Role of copper and iron in methane oxidation and bacterial biopolymer accumulation. *Eng Life Sci* 15**,** 387-399.

Conrad, R., (2007). Microbial ecology of methanogens and methanotrophs. *Adv Agron* 96**,** 1-63.

Constant, P., Chowdhury, S. P., Pratscher, J., Conrad, R. (2010). Streptomycetes contributing to atmospheric molecular hydrogen soil uptake are widespread and encode a putative high-affinity [NiFe]-hydrogenase. *Environ Microbiol* 12**,** 821-9.

Crombie, A. T., Murrell, J. C. (2014). Trace-gas metabolic versatility of the facultative methanotroph *Methylocella silvestris*. *Nature* 510**,** 148-51.

Dalton, H., (1983). The Biochemistry of methylotrophs - Anthony,C. Trends. *Biochem Sci* 8**,** 342-343.

Dam, B., Dam, S., Blom, J., Liesack, W. (2013). Genome analysis coupled with physiological studies reveals a diverse nitrogen metabolism in *Methylocystis* sp. strain SC2. *PLoS One*. 8**,** e74767.

Dam, B., Dam, S., Kim, Y., Liesack, W. (2014). Ammonium induces differential expression of methane and nitrogen metabolism-related genes in *Methylocystis* sp. strain SC2. *Environ Microbiol* 16**,** 3115-27.

Dam, B., Dam, S., Kube, M., Reinhardt, R., Liesack, W. (2012). Complete genome sequence of *Methylocystis* sp. strain SC2, an aerobic methanotroph with high-affinity methane oxidation potential. *J Bacteriol* 194**,** 6008-9.

Dean, J. F., Middelburg, J. J., Rockmann, T., Aerts, R., Blauw, L. G., Egger, M., Jetten, M. S. M., de Jong, A. E. E., Meisel, O. H., Rasigraf, O., Slomp, C. P., in't Zandt, M. H., Dolman, A. J. (2018). Methane feedbacks to the global climate system in a warmer world. *Rev Geophys* 56**,** 207-250.

Dedysh, S. N., Dunfield, P. F., Derakshani, M., Stubner, S., Heyer, J., Liesack, W. (2003). Differential detection of type II methanotrophic bacteria in acidic peatlands using newly developed 16S rRNA-targeted fluorescent oligonucleotide probes. *FEMS Microbiol Ecol* 43**,** 299-308.

Dedysh, S. N., Knief, C., Dunfield, P. F. (2005). *Methylocella* species are facultatively methanotrophic. *J Bacteriol* 187**,** 4665-70.

Dunfield, P. F., Yimga, M. T., Dedysh, S. N., Berger, U., Liesack, W., Heyer, J. (2002). Isolation of a *Methylocystis* strain containing a novel *pmoA*-like gene. *FEMS Microbiol Ecol* 41**,** 17-26.

Durmowicz, M. C., Maier, R. J. (1997). Roles of HoxX and HoxA in biosynthesis of hydrogenase in *Bradyrhizobium japonicum*. *J Bacteriol* 179**,** 3676-3682.

Eberhardt, U. (1969). On chemolithotrophy and hydrogenase of a gram-positive knallgas bacterium. *Arch Mikrobiol* 66**,** 91-104.

Edgren, T., Nordlund, S. (2004). The *fixABCX* genes in *Rhodospirillum rubrum* encode a putative membrane complex participating in electron transfer to nitrogenase. *J Bacteriol* 186**,** 2052-60.

Fei, Q., Guarnieri, M. T., Tao, L., Laurens, L. M. L., Dowe, N., Pienkos, P. T. (2014). Bioconversion of natural gas to liquid fuel: Opportunities and challenges. *Biotechnol Adv* 32**,** 596-614.

Friedrich, B., Buhrke, T., Burgdorf, T., Lenz, O. (2005). A hydrogen-sensing multiprotein complex controls aerobic hydrogen metabolism in *Ralstonia eutropha*. *Biochem Soc Trans* 33**,** 97-101.

Glatter, T., Ludwig, C., Ahrne, E., Aebersold, R., Heck, A. J., Schmidt, A. (2012). Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion. *J Proteome Res* 11**,** 5145-56.

Greening, C., Berney, M., Hards, K., Cook, G. M., Conrad, R. (2014). A soil actinobacterium scavenges atmospheric $H_2$ using two membrane-associated, oxygen-dependent [NiFe] hydrogenases. *Proc Natl Acad Sci USA* 111**,** 4257-4261.

Greening, C., Biswas, A., Carere, C. R., Jackson, C. J., Taylor, M. C., Stott, M. B., Cook, G. M., Morales, S. E. (2016). Genomic and metagenomic surveys of hydrogenase distribution indicate $H_2$ is a widely utilised energy source for microbial growth and survival. *ISME J* 10**,** 761-77.

Greening, C., Cook, G. M. (2014). Integration of hydrogenase expression and hydrogen sensing in bacterial cell physiology. *Curr Opin Microbiol* 18**,** 30-8.

Hakobyan, A., Liesack, W., Glatter, T. (2018). Crude-MS strategy for in-depth proteome analysis of the methane-oxidizing *Methylocystis* sp. strain SC2. *J Proteome Res* 17**,** 3086-3103.

Hanson, R. S., Hanson, T. E. (1996). Methanotrophic bacteria. *Microbiol Rev* 60**,** 439-71.

Heyer, J., Galchenko, V. F., Dunfield, P. F. (2002). Molecular phylogeny of type II methane-oxidizing bacteria isolated from various environments. *Microbiology* 148**,** 2831-46.

Im, J., Lee, S. W., Yoon, S., Dispirito, A. A., Semrau, J. D. (2011). Characterization of a novel facultative *Methylocystis* species capable of growth on methane, acetate and ethanol. *Environ Microbiol* Rep. 3**,** 174-81.

Kalyuzhnaya, M. G., Puri, A. W., Lidstrom, M. E. (2015). Metabolic engineering in methanotrophic bacteria. *Metabol Eng* 29**,** 142-152.

Karthikeyan, O. P., Chidambarampadmavathy, K., Cires, S., Heimann, K. (2015). Review of sustainable methane mitigation and biopolymer production. *Crit Rev Env Sci Tec* 45**,** 1579-1610.

Knief, C. (2015). Diversity and habitat preferences of cultivated and uncultivated aerobic methanotrophic bacteria evaluated based on *pmoA* as molecular marker. *Front Microbiol* 6, 1346.

Knief, C., Lipski, A., Dunfield, P. F. (2003). Diversity and activity of methanotrophic bacteria in different upland soils. *Appl Environ Microbiol* 69**,** 6703-6714.

Le, N. Q., Ou, Y. Y. (2016). Prediction of FAD binding sites in electron transport proteins according to efficient radial basis function networks and significant amino acid pairs. *BMC Bioinf* 17**,** 298.

Ledbetter, R. N., Garcia Costas, A. M., Lubner, C. E., Mulder, D. W., Tokmina-Lukaszewska, M., et al. (2017). The electron bifurcating FixABCX protein complex from *Azotobacter vinelandii*: generation of low-potential reducing equivalents for nitrogenase catalysis. *Biochemistry* 56**,** 4177-4190.

Ma, K., Schicho, R. N., Kelly, R. M., Adams, M. W. (1993). Hydrogenase of the hyperthermophile *Pyrococcus furiosus* is an elemental sulfur reductase or sulfhydrogenase: evidence for a sulfur-reducing hydrogenase ancestor. *Proc Natl Acad Sci USA.* 90**,** 5341-4.

Matassa, S., Boon, N., Verstraete, W. (2015). Resource recovery from used water: The manufacturing abilities of hydrogen-oxidizing bacteria. *Water Res* 68**,** 467-478.

McCarty, P. L. (2007). Thermodynamic electron equivalents model for bacterial yield prediction: modifications and comparative evaluations. *Biotechnol Bioeng* 97**,** 377-88.

Mohammadi, S., Pol, A., van Alen, T. A., Jetten, M. S., Op den Camp, H. J. (2017). *Methylacidiphilum fumariolicum* SolV, a thermoacidophilic 'Knallgas' methanotroph with both an oxygen-sensitive and -insensitive hydrogenase. *ISME J* 11**,** 945-958.

Murrell, J. C., Gilbert, B., McDonald, I. R. (2000a). Molecular biology and regulation of methane monooxygenase. *Arch Microbiol* 173**,** 325-32.

Murrell, J. C., McDonald, I. R., Gilbert, B. (2000b). Regulation of expression of methane monooxygenases by copper ions. *Trends Microbiol* 8**,** 221-5.

Murrell, J. C., Smith, T. J. (2009). Microbial biotechnology meets environmental microbiology. *Microb Biotechnol* 2**,** 142-143.

Myhre, G., Shindell, D., Breon, F. M., Collins, W., Fuglestvedt, J., Huang, J. P., et al. (2014). Anthropogenic and Natural Radiative Forcing. *Climate Change 2013: The Physical Science Basis.* 659-740.

Nauer, P. A., Dam, B., Liesack, W., Zeyer, J., Schroth, M. H. (2012). Activity and diversity of methane-oxidizing bacteria in glacier forefields on siliceous and calcareous bedrock. *Biogeosciences* 9**,** 2259-2274.

Nybo, S. E., Khan, N. E., Woolston, B. M., Curtis, W. R. (2015). Metabolic engineering in chemolithoautotrophic hosts for the production of fuels and chemicals. *Metab Eng* 30**,** 105-120.

Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., Coon, J. J. (2012). Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics* 11**,** 1475-88.

Piche-Choquette, S., Constant, P. (2019). Molecular hydrogen, a neglected key driver of soil biogeochemical processes. *Appl Eviron Microbiol* 85.

Piche-Choquette, S., Khdhiri, M., Constant, P. (2018). Dose-response relationships between environmentally-relevant $H_2$ concentrations and the biological sinks of $H_2$, $CH_4$ and CO in soil. *Soil Biol Biochem* 123**,** 190-199.

Rey, F. E., Oda, Y., Harwood, C. S. (2006). Regulation of uptake hydrogenase and effects of hydrogen utilization on gene expression in *Rhodopseudomonas palustris*. *J Bacteriol* 188**,** 6143-52.

Schafer, C., Friedrich, B., Lenz, O., 2013. Novel, oxygen-insensitive group 5 [NiFe]-hydrogenase in *Ralstonia eutropha*. *Appl Environ Microbiol* 79**,** 5137-45.

Schwartz, E., Buhrke, T., Gerischer, U., Friedrich, B. (1999). Positive transcriptional feedback controls hydrogenase expression in *Alcaligenes eutrophus* H16. *J Bacteriol* 181**,** 5684-92.

Semrau, J. D., DiSpirito, A. A., Yoon, S. (2010). Methanotrophs and copper. *FEMS Microbiol Rev* 34**,** 496-531.

Shafaat, H. S., Rudiger, O., Ogata, H., Lubitz, W. (2013). [NiFe] hydrogenases: a common active site for hydrogen metabolism under diverse conditions. *Biochim Biophys Acta* 1827**,** 986-1002.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47**,** D607-D613.

Thauer, R. K., Jungermann, K., Decker, K. (1977). Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol Rev* 41**,** 100-80.

Thornton, B. B., C. (2011). Real-time PCR (qPCR) primer design using free online software. *Biochem Mol Biol Educ* 39**,** 145-154.

Trotsenko, Y. A., Murrell, J. C. (2008). Metabolic aspects of aerobic obligate methanotrophy. *Adv Appl Microbiol* 63**,** 183-229.

Van Soom, C., de Wilde, P., Vanderleyden, J. (1997). HoxA is a transcriptional regulator for expression of the hup structural genes in free-living *Bradyrhizobium japonicum*. *Mol Microbiol* 23**,** 967-977.

Vignais, P. M., Billoud, B., Meyer, J. (2001). Classification and phylogeny of hydrogenases. *FEMS Microbiol Rev* 25**,** 455-501.

Vizcaino, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 44**,** 11033.

Weyman, P. D., Pratte, B., Thiel, T. (2008). Transcription of *hupSL* in *Anabaena variabilis* ATCC 29413 is regulated by NtcA and not by hydrogen. *Appl Environ Microbiol* 74**,** 2103-2110.

Whittenbury, R., Dalton, H. (1981). The methylotrophic bacteria. In: Starr, M. P., Stolp, H., Trüper, H. G., Balows, A., Schlegel, H. G., Eds.). *The Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria. Springer Berlin Heidelberg, Berlin, Heidelberg*, 894-902.

# Chapter 5:

## Genome Scale Metabolic Modeling Reveals the Metabolic Potential of Three Type II Methanotrophs of the Genus *Methylocystis*

*Met Eng* 2019, 54, 191-199
https://doi.org/10.1016/j.ymben.2019.04.001

Sergio Bordel[1,2,*], Yadira Rodríguez[1,2], **Anna Hakobyan**[3], Elisa Rodríguez[1,2], Raquel Lebrero[1,2], Raúl Muñoz[1,2]

*Corresponding author

Chapter 5 is written in research manuscript style. It was published as research article in *Metabolic Engineering* in July, 2019. My contribution to the chapter 5 involved the laboratory experiments for validation of metabolic models, the assistance in writing and revision of the corresponding parts of the manuscript.

[1]Departamento de Ingeniería Química y Tecnología del medio ambiente, Escuela de Ingenierías Industriales, Universidad de Valladolid, Spain

[2]Institute of Sustainable Processes, Universidad de Valladolid, Spain

[3]Research Group of Methanotrophic Bacteria and Environmental Genomics/Transcriptomics, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany

# 5 Genome Scale Metabolic Modeling Reveals the Metabolic Potential of Three Type II Methanotrophs of the Genus *Methylocystis*

## 5.1. Abstract

Genome Scale Metabolic Models (GSMMs) of the recently sequenced *Methylocystis hirsuta* and two other methanotrophs from the genus *Methylocystis* have been reconstructed. These organisms are type II methanotrophs with the ability of accumulating Polyhydroxyalkanoates under nutrient limiting conditions. For the first time, GSMMs have been reconstructed for type II methanotrophs. These models, combined with experimental biomass and PHB yields of *Methylocystis hirsuta*, allowed elucidating the methane oxidation mechanism by the enzyme pMMO (particulate methane monooxygenase) in these organisms. In contrast to type I methanotrophs, which use the "direct coupling mechanism", type II methanotrophs appear to use the so called "redox arm mechanism". The utilization of the "redox arm mechanism", which involves the coupling between methane oxidation and complex I of the respiratory chain, was confirmed by inhibition of complex I with catechol. Utilization of the "redox arm" mechanism leads to lower biomass yields on methane compared to type I methanotrophs. However, the ability of type II metanotrophs to redirect high metabolic carbon fluxes towards acetoacetyl-CoA under nitrogen limiting conditions makes these organisms promising platforms for metabolic engineering.

## 5.2. Introduction

Methanotrophic organisms, able to use methane as a carbon and energy source, are attracting a growing attention as potential cell factories (Strong et. al., 2016). Methane has been recently identified as a cost effective feedstock for microbial fermentations as a result of the low price of natural gas and the residual nature of some $CH_4$-laden emissions (Comer et. al., 2017). Indeed, this gas has the potential of becoming a virtually free substrate based on the fact that methane is originated in most waste treatment processes and more than 70 million tonnes of $CH_4$ are being yearly released into the

atmosphere worldwide (Abbasi et. al., 2012). The replacement of glucose by methane would also have positive environmental and societal impacts. Methane is a greenhouse gas and its utilization as a feedstock would constitute a carbon sink. However, to this date, only a few methane-based industrial bioproducts have reached the market, and none of them involves metabolically engineered strains.

Nevertheless, some interesting works on metabolic engineering of methanotrophs have been already conducted as a result of the development of adapted genetic tools and the reconstruction of Genome Scale Metabolic Models (GSMMs) for these organisms (Kalyuzhnaya et. al., 2015). Most of these examples have been carried out in type I methanotrophs, which assimilate methane by oxidizing it to formaldehyde and condensing formaldehyde with ribulose monophosphate, resulting in fructose-6-phosphate production. The first published GSMM of a methanotroph was that of *Methylomicrobium buriatense* (de la Torre et. al., 2015). *M. buriatense* was also the first methanotroph to be metabolically engineered to produce lactate (Henard et. al., 2016), thanks to the development of techniques for its genetic manipulation (Puri et. al., 2015). A GSMM of *Methylomicrobium alcaliphilum* has been also developed (Akberdin et. al., 2018) and used to optimize the production of 2,3-butanediol (Nguyen et. al., 2018) by applying the OptGene algorithm (Patil et. al., 2005). The two previously mentioned GSMMs allowed to elucidate the methane oxidation mechanism of the two mentioned *Methylomicrobium* species, which in both cases appeared to be the so called “direct coupling”. The first step in methane degradation is the oxidation of methane to methanol by a membrane bound methane monooxygenase (pMMO), with the consumption of a molecule of oxygen and the reduction of an electron donor. “Direct coupling” is the mechanism by which the mentioned electron donor is reduced again in connection to the oxidation of methanol to formaldehyde by the enzyme Methanol Dehydrogenase (MeDH). The identification of the electron donor to pMMO has been for a long time an unsolved problem in the study of methanotrophs. A summary of the different theories explaining the mechanisms of methane oxidation can be found in the review by Kalyuzhnaya and coworkers (Kalyuzhnaya et. al., 2015). The identification of the correct methane oxidation mechanism is essential for the design of metabolic engineering strategies, as it has a deep impact on the demand for redox cofactors, which interplay with all the metabolic pathways of the cell.

This study presents, to the best of our knowledge, the three first GSMMs of type II methanotrophs reported in literature. These organisms use the serine cycle for methane assimilation, which results in a relatively high flux through Acetyl-CoA (Kalyuzhnaya et. al., 2015) and might make them particularly suitable for the production of lipid derivatives such as biofuels or carotenoids. Type II methanotrophs from the genera *Methylocystis*, *Methylosinus* and *Methylocella* are also able to synthesize polyhydroxyalkanoates (PHAs) under conditions of nutrient limitation (Pieja et al., 2017). Among them, *Methylocystis hirsuta* has been found to accumulate polyhydroxybutyrate (PHB) up to 45% of its total biomass (García-Pérez et. al., 2018). In addition to the commercial interest of PHAs as biodegradable plastics, the high flux towards acetoacetyl-CoA of these organisms could potentially be deviated towards butanol or mevalonate derivatives (valencene, geraniol etc.), which makes them promising future cell factories. Type II methanotrophs exhibit lower methane yields than those of type I methanotrophs, this phenomenon has been attributed to a "relatively expensive route for C1-assimilation", the serine pathway (Kalyuzhnaya et. al., 2015). However, this statement cannot be fully proven without using GSMMs. In fact, the GSMMs reconstructed in this work, revealed that the most likely cause for this lower yield is a different mechanism of methane oxidation (the so called "redox arm" instead of "direct coupling").

This study aimed at constructing GSMMs of the recently sequenced *Methylocystis hirsuta* CSC1 (Bordel et. al., 2018), as well as of its close relative species *Methylocystis* sp. SC2 and *Methylocystis sp.* SB2. In this context, *Methylocystis* sp. SC2 was the first type II methanotroph genetically engineered (Baani and Liesack, 2008) and has shown a versatile metabolism, including a complete denitrification pathway (Dam et. al., 2013), which is truncated in *M. hirsuta* (Bordel et. al., 2018). On the other hand, *Methylocystis* sp. SB2 is a facultative methanotroph able to grow in 2-carbon substrates such as acetate and ethanol (Im et. al., 2011). In this work, the GSMM of *Methylocystis hirsuta* has been validated using experimental measurements of growth yields, methane and oxygen consumption, and PHB production. The similarities and differences of the metabolic network of *Methylocystis hirsuta* with its two close relatives are also discussed.

## 5.3. Materials and methods

### *5.3.1. Reconstruction of GSMMs*

The genome of *Methylocystis hirsuta* CSC1 was previously sequenced, assembled, and annotated (Bordel et. al., 2018), and has been deposited in GenBank with the accession number QWDD00000000 (BioProject: PRNJ487728). The genome of *Methylocystis* sp. SC2 (Dam et. al., 2012) was obtained from GenBank with accession number NC_018485 (BioProject: PRJNA224116), while that of *Methylocystis* sp. SB2 was also obtained from GenBank, with accession number GCA_000499825 (BioProject: PRJNA219752). Draft models were generated using the SEED server (Overbeek et. al., 2005) based on a previous annotation with RAST (Overbeek et. al., 2014). Each draft was manually curated by checking all the gene-reaction associations and correcting dubious annotations. For example, the SEED reaction with identifier rxn00470_c0, that corresponds to ornithine decarboxylation, was associated to the gene with identifier 369798.4.peg.465, which was annotated by RAST as diaminopimelate decarboxylase (EC 4.1.1.20), the corresponding EC entry was checked in KEGG and it was found not to be associated to ornithine decarboxylation, therefore the annotation of this reaction was changed manually to the gene 369798.4.peg.3994, annotated by RAST as ornithine decarboxylase (EC 4.1.1.17). Another example of annotation corrected manually, was nitrate reductase, with SEED identifier rxn10121_c0, which was associated to the gene 369798.4.peg.730, annotated by RAST as a nitrite reductase subunit (EC 1.7.1.4), this association was changed manually to the genes 369798.4.peg.732 and 369798.4.peg.733, which were annotated by RAST as nitrate reductases (EC 1.7.99.4). The models were manipulated in SBML format using the Python library COBRApy (Ebrahim et. al., 2013). Reaction directionality was modified in order to avoid unfeasible ATP production, so that the model only simulates ATP production in the respiratory chain and in the lower glycolysis. Directionality of the reactions involving proton transport was also modified in order to guarantee that no proton flux towards the exterior of the cell is allowed if it is not mediated by the components of the respiratory chain or associated to some energy cost. This was done in order to avoid proton flux against its concentration gradient. The biomass equations used were those generated by SEED.

### *5.3.2. Strain, chemicals and culture conditions*

The strain *Methylocystis hirsuta* CSC1 was obtained from the DSMZ culture collection (DSM-18500). The strain was cultured in Whittenbury nitrate mineral medium (pH 6.8) (Whittenbury et. al., 1970) or alternatively Whittenbury ammonium mineral medium or nitrogen free Whittenbury mineral medium for the corresponding experiments. *M. hirsuta* CSC1 growth and PHB accumulation experiments were carried out in 2.15 L serum bottles closed with butyl-rubber stoppers and aluminum crimp seals (incubation at 25 ºC and 300 rpm). Each flask contained 400 mL of mineral medium and a headspace containing mixtures of synthetic biogas (70% $CH_4$, 30% $CO_2$) and $O_2$ from Abelló Linde S.A. (Barcelona, Spain). Catechol was purchased from Sigma Aldrich®, St. Louis, USA. Headspace mixtures with 2:1, 1.5:1 and 1:1 $O_2$:$CH_4$ molar ratios were prepared in 25 L Tedlar bags (Sigma Aldrich®, St. Louis, USA) and subsequently pumped into the bottles to replace the original air headspace. The time course of the headspace composition and biomass concentration were monitored by GC-TCD and optical density (after performing a calibration curve) respectively. The growth rates were obtained from the slopes of the logarithms of the biomass concentration with respect to time, while biomass yields and oxygen-methane consumption ratios were obtained from the slope of the biomass with respect to the consumed methane and from the slope of the consumed oxygen with respect to the consumed methane. Methane specific consumption rate was obtained by dividing the specific growth rate by the yield of biomass on methane. Total amounts of oxygen and methane were calculated from their gas concentrations taking into account the gas and liquid volumes in the flasks and the solubility of each gas at 25 °C. Total $CO_2$ accumulation was calculated based on the pH, the $CO_2$ concentration in the headspace, the Henry's law constant and the equilibrium constant with $HCO_3^-$ and $CO_3^{2-}$.

Experiments assessing *M. hirsuta* growth on acetate and the inhibition of its respiratory chain by catechol were carried out at 25 °C in 125 mL serum bottles crimp sealed under sterile conditions with 50 mL of NMS and 20 mL NMS, respectively, under a $O_2$:$CH_4$ atmosphere (2:1 molar ratio).

*Methylomicrobium alcaliphilum* 20Z, a type I methanotroph, was purchased from DSMZ (DSM-19304). *M. alcaliphilum* was grown on a high alkalinity (pH 9) as described previously (Kalyuzhnaya et. al. 2008). The catechol inhibition tests were carried out under the same conditions as for *M. hirsuta*.

Control experiments with *Methylocystis* sp. SC2 were carried out at Prof. Liesack's laboratory (Max Planck Institute for Terrestrial Microbiology) under culture conditions previously reported for the mentioned microorganism (Dam et. al., 2012)

### *5.3.3. Analytical methods*

The headspace gas concentrations ($CH_4$, $O_2$ and $CO_2$) were measured by gas chromatography in a Bruker 430 GC-TCD (Bruker, Palo Alto, USA) with two columns: a CP-Molsieve 5A (15 m × 0.53 mm × 10 mm) and a CP-PoraBOND Q (25 m × 0.53 mm × 10 mm). Biomass concentration was estimated using $OD_{600}$, which was previously correlated with total suspended solids (TSS) as described previously (López et. al., 2014). Nitrate concentrations were measured by HPLC-IC using a Waters 432 conductivity detector (Waters, USA). The specific nitrate production rate for the experiment in which ammonium was used as nitrogen source was calculated multiplying the slope of nitrate concentration versus biomass by the specific growth rate.

The PHB content was extracted from the biomass as described by Zuniga and coworkers (Zuniga et. al., 2011). After extraction and hydrolysis, hydroxybutyrate was measured using an Agilent 7820A GC coupled with a 5977E MSD (Agilent technologies, Santa Clara, USA) equipped with a DB-wax column (30 m × 250 µm × 0.25 µm). The detector and injector temperatures were kept constant at 250 °C and the oven temperature was increased from 40 °C to 200 °C at 10 °C per minute and maintained at 200 °C for 2 minutes, before being increased again at 5 °C per minute until reaching 240 °C. Acetaldehyde was measured using an Agilent 1200 HPLC with 355 nm UV detector and a C18 column (250×4.6 mm) at 30 °C.

## 5.4. Results

### *5.4.1. Genome scale metabolic models*

After the manual curation of the SEED models, three GSMMs whose statistics are represented in **Table 1** were obtained. Essential reactions are defined as those whose absence makes the model unable to produce biomass using methane as a carbon and energy source.

**Table 1** | Sizes of the reconstructed GSMMs.

| | Genes | Reactions | Metabolites | Essential reactions |
|---|---|---|---|---|
| *M. hirsuta* | 2748 | 1399 | 1460 | 383 |
| *M.* sp. SC2 | 2251 | 1449 | 1434 | 381 |
| *M.* sp. SB2 | 2281 | 1380 | 1453 | 374 |

An ATP maintenance reaction was added to all the models, as well as two alternative methane monooxygenase reactions with identifiers pMMO1 and pMMO2, respectively. Reaction pMMO1 uses cytochrome-c as co-factor and is aimed to represent the direct coupling mechanism for methane oxidation, while pMMO2 uses ubiquinone as redox cofactor and is aimed to represent the redox-arm mechanism. Serine-glyoxylate aminotransferase (with EC 2.6.1.45), which is present in the three genomes and belongs to the serine cycle, was added manually. A PHB synthesis reaction, representing the polymerization of (S)-3-hydroxybutyril-CoA to form PHB, was also added in order to account for the PHB accumulation observed in the members of the genus *Methylocystis*, even if no clear annotation was found in the genome. The reaction with identifier rxn10113_c0, representing complex III in the respiratory chain was modified to represent an electron transfer between ubiquinol and cytochrome-c (in the draft model this reaction was written lumped to the final electron transfer to $O_2$). If this reaction is allowed to proceed backwards, this mimics the so called uphill electron transfer hypothesis (Kalyuzhnaya et. al., 2015). The reconstructed models in SBML format and in tab separated format, as well as the connectivity information for the most connected metabolites and the lists of essential reactions, have been made publicly available at https://github.com/SergioBordel/ModelsMethanotrophs. The models in SBML format have also been deposited in BioModels (Chelliah et. al., 2015) with identification numbers: MODEL1812120001, MODEL1812120002 and MODEL1812120003.

**Figure 1** represents the overlap between the metabolic networks of the three species herein considered. The three species shared more than 90% of their metabolic reactions and metabolites, and the theoretical biomass and PHB yields on methane were identical.

Sorbitol was identified among the metabolites with potential commercial

interest specific to a single species. This metabolite was found only in *Methylocystis hirsuta* due to the presence of a sorbitol dehydrogenase (EC 1.1.1.14), which was absent in the two other species and could catalyze the transformation of D-fructose into sorbitol with the consumption of a NADH molecule. Other metabolites specific to *M. hirsuta* were cyromazine (an insecticide) and its synthesis intermediates such as cyanuric acid. However, cyromazine synthesis was annotated as being the result of the promiscuous transaminase activity of a single dCMP deaminase (EC 3.5.4.45), which makes the presence of this pathway uncertain.

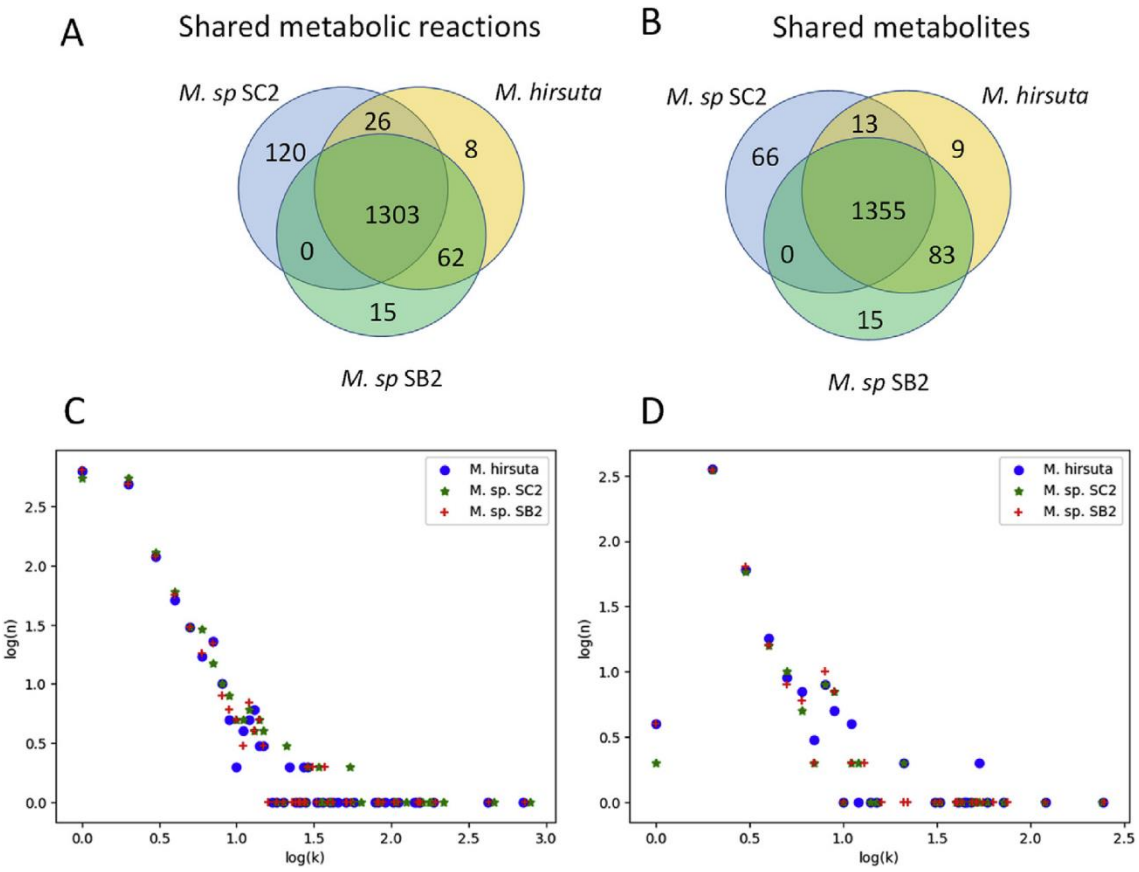



**Figure 1** | Topological characteristics of the reconstructed metabolic networks. The Venn diagrams show the overlap in reactions **(A)** and metabolites **(B)** among the three *Methylocystis* GSMMs. The scatter plots show in logarithmic scales the number (n) of metabolites with connectivity k (participating in k metabolic reactions), for the whole network **(C)** and for the sub-network of flux-carrying reactions during biomass optimization **(D)**. The plots show a scale free topological structure, with a parameter γ equal to 2.5. Highly connected metabolites, such as energy and redox cofactors, constitute outliers.

*Methylocystis* sp. SB2 contained the chelating agent Mg protoporphyrin IX, and its synthesis intermediates deoxylimonate, chlorophyllide, and protochlorophyllide, within its specific metabolites. In this case there were 12 specific genes responsible for this pathway, which were only present in *Methylocystis* sp. SB2 and included: Mg

protoporphyrin IX monomethyl ester oxidative cyclase (EC 1.14.13.81), Mg protoporphyrin IX O-methyltransferase (EC 2.1.1.11), protoporphyrin IX Mg-chelatase subunits H, I and D (EC 6.6.1.1), lysophospholipase (EC 3.1.1.5), etc. The presence of these genes suggested a higher ability of *Methylocystis sp.* SB2 to uptake Mg from the extracellular medium.

The species *Methylocystis* sp. SC2 exhibited 66 specific metabolites and 120 specific reactions, including the enzymes acyl-phosphate glycerol-3-phosphate O-acyltransferase PlsY or methylmalonyl-CoA mutase (EC 5.4.99.2). An interesting specific metabolite found only in *Methylocystis* sp. SC2, together with its biosynthesis intermediates, was cobalt-precorrin, which could indicate a higher capability of Co assimilation compared to the other two species. Aniline, which is the product of an aromatic-L-amino-acid decarboxylase (EC 4.1.1.28), was identified among the metabolites with potential industrial interest found in *Methylocystis* sp. SC2 genome.

From a topological point of view all the networks showed a scale free structure in which the number of metabolites (n) participating in k reactions was proportional to $k^{-\gamma}$, with the parameter γ equal to 2.5. The most highly connected metabolites such as water, protons, phosphate and energy or redox cofactors, were outliers with higher abundance than what could be expected from a scale free statistical distribution. This is a universal feature of metabolic networks, which has been observed in organisms as different as *E. coli*, *S. cerevisiae*, *H. pilori*, *S. aureus,* etc. (Becker, et. al., 2006). The same pattern was observed if only the metabolic subnetwork, whose reactions carry a non-zero flux under optimal growth conditions, were analyzed (only the reactions that carry flux when the biomass production was optimized, and their metabolites were considered). In this case, the metabolites participating in only one reaction were underrepresented, as those are the metabolites secreted or up-taken from the medium.

### *5.4.2. Model validation*

The theoretical yields on methane predicted by the metabolic model were compared to experimental biomass and PHB production yields of *M. hirsuta*.

**Growth yields and oxygen utilization**

Different methane oxidation mechanisms result in different growth yields and oxygen

demands. For example, the type I methanotroph *Methylomicrobium alcaliphilum* was recently found to rely on "direct coupling" by comparing its experimental oxygen-methane consumption ratio with the simulations of different mechanisms using a GSMM (Akberdin et. al., 2018).

**Table 2** |Experimental growth rates, yields and oxygen-methane consumption ratios.

| $O_2$:$CH_4$ | μ ($day^{-1}$) | Y (g-DW/mol$CH_4$) | Consumption ratio (mol$O_2$/ mol$CH_4$) |
|---|---|---|---|
| 2:1 | 1.2 ± 0.01 | 7.99 ± 0.005 | 1.46 ± 0.05 |
| 1.5:1 | 1.11 ± 0.01 | 7.47 ± 0.085 | 1.44 ± 0.06 |
| 1:1 | 1.11 ± 0.02 | 7.62 ± 0.145 | 1.43 ± 0.02 |

*M. hirsuta* growth experiments were carried out with headspace $O_2$:$CH_4$ molar ratios of 2:1, 1.5:1 and 1:1 in order to test if the headspace composition had an impact on the oxygen-methane consumption ratios. Nitrate was used as a nitrogen source (since ammonium results in competitive inhibition of methane monooxygenase as discussed later on). The observed growth rates, yields and oxygen-methane consumption rations are reported in **Table 2** (error intervals represent standard deviations among 3 biological replicates).

No significant differences among the three experiments were obtained (ANOVA test). In particular, oxygen and methane were consumed with the same stoichiometric ratio until the limiting substrate in the headspace was fully depleted.

In order to model the theoretical yields and consumption ratios in *M. hirsuta*, the maintenance ATP consumption rate was initially estimated by following the oxygen consumption rate of a culture in the absence of carbon source. The observed specific oxygen consumption rate was 6±1 mmol $O_2$ g-$DW^{-1}$ $day^{-1}$. Assuming a typical P/O ratio of 2.5, the maintenance ATP consumption rate was estimated at 30 mmol ATP g-$DW^{-1}$ $day^{-1}$. This value was imposed as a constraint in the model. Growth rates were predicted using the reconstructed model for *M. hirsuta* by setting the specific methane uptake rate to its average experimental value of 150 mmol $CH_4$ g-$DW^{-1}$ $day^{-1}$. Simulations were carried out under the "direct coupling" assumption (allowing flux in the pMMO1 reaction), the "uphill electron transfer" assumption (allowing the reaction catalyzed by complex III of the respiratory chain to proceed in its reverse direction) and finally the "redox arm" assumption (with an irreversible complex III and methane oxidation taking

place in the reaction pMMO2). **Figure 2** shows the predictions in each scenario compared to the experimental observations.

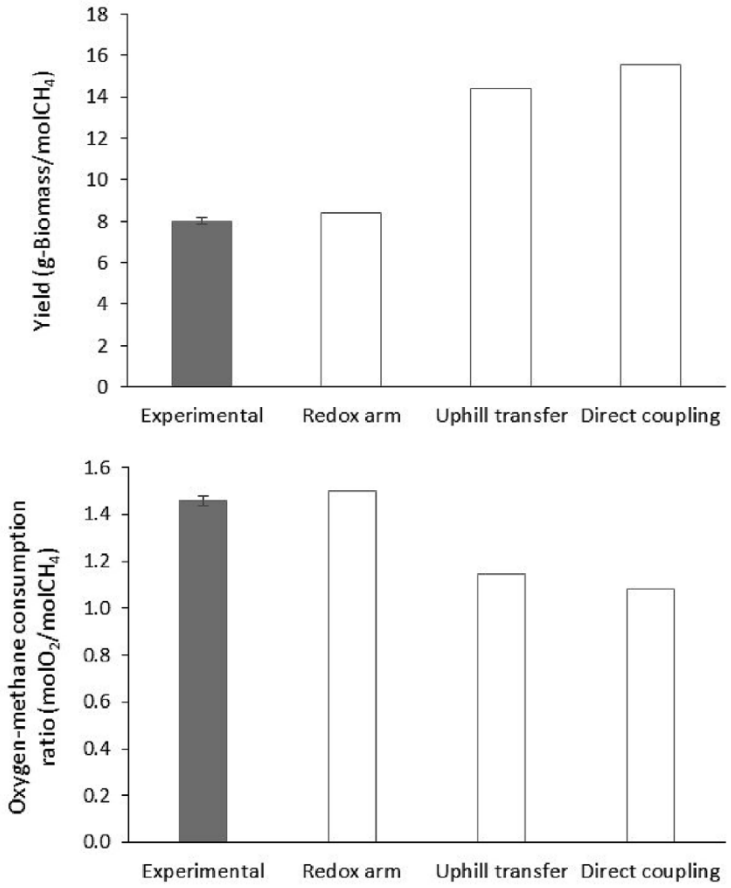



**Figure 2** | Experimental and theoretical biomass yields and oxygen-methane consumption ratios predicted using the "redox arm", "uphill electron transfer" and "direct coupling" mechanisms. Grey bars represent experimental measurements with error bars corresponded to the observed standard deviation of the measurements (for 3 biological replicates). White bars correspond to model predictions for each of the three considered methane oxidation mechanisms, using the GSMM of *Methylocystis hirsuta*.

The experimental results fit remarkably well with the predictions using the "redox arm" mechanism. The oxygen-methane consumption rates predicted by the redox arm mechanism was 1.5. This contrasts with the results obtained with Type I methanotrophs (de la Torre et. al., 2015; Akberdin et. al., 2018), where the experimental results were consistent with the assumption of the "direct coupling" mechanism. **Figure 3** illustrates the three potential methane oxidation mechanisms.

**Complex I inhibition affects methane oxidation**

Under the redox arm mechanism, methane consumption is coupled to the reduction of ubiquinone by the complex I of the respiratory chain. Therefore, complex I inhibition

would result in the inhibition of methane oxidation. Under the direct coupling hypothesis, even in the case of an inactive complex I, methane oxidation to formaldehyde could still be carried out autonomously, without the necessity of NADH inputs. On the other hand, complexes II, III and IV alone, are expected to be able to sustain respiration. Even if the inhibition of complex I will decrease the global energetic efficiency of the oxidative phosphorylation.

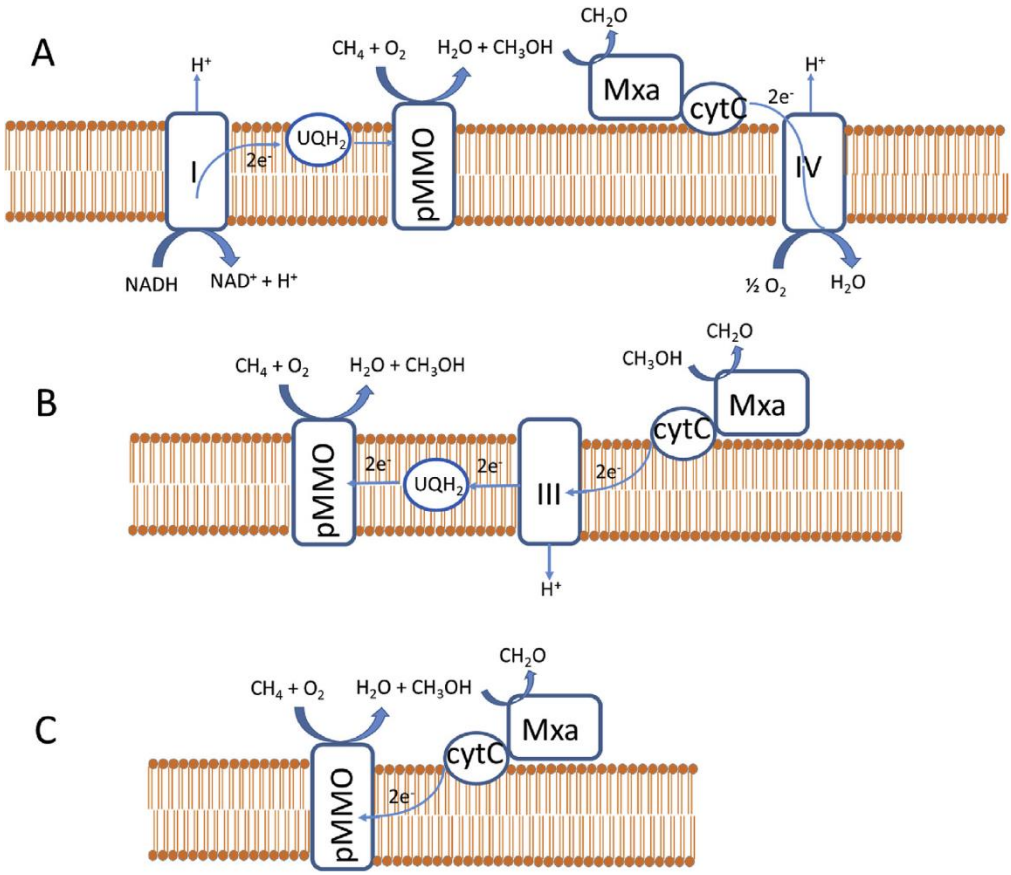



**Figure 3** | Alternative mechanisms of me-thane oxidation. **(A)** "Redox arm" mechanism in which ubiquinone acts as electron supplier to pMMO, these electrons originating ultimately from NADH via the complex I of oxidative phosphorylation. **(B)** "Uphill-electron transfer" mechanism where the electrons supplied to pMMO by ubiquinone originate from cytochrome-c via the reverse operation of complex III. **(C)** "Direct coupling" mechanism where cytochrome-c supplies directly electrons to pMMO.

Catechol has been reported to inhibit complex I at 0.4 mM concentrations (Degli Esposti, 1998). In order to validate the existence of a redox arm mechanism in *M. hirsuta* (and its close relatives, given the fact that they have identical pMMO complexes), catechol was used to inhibit complex I and its effects on methane oxidation were monitored.

The same experiment was performed with *Methylomicrobium alcaliphilum*, a type I methanotroph for which the direct coupling mechanism has been previously

established (Akberdin et. al., 2018). In this case, catechol is not expected to inhibit the oxidation of methane to glyceraldehyde and slowing down the rate of respiration could result in accumulation of acetaldehyde in the growth medium.

Aliquots of 20 mL of a *M. hirsuta* culture in the exponential growth phase (with a dry weight concentration of 205 mg $L^{-1}$) were placed in 125 mL serum flasks. A first test series was carried out (in triplicate) by adding catechol to the culture broth at 0.4 mM, while a second series (in triplicate) without catechol was used as a control. Methane and oxygen headspace concentrations were monitored over time and their specific biodegradation rates were calculated from their initial slopes. Liquid samples were taken after 24 hours in order to measure possible glyceraldehyde accumulation. The same experimental procedure was performed with 20 mL aliquots of *M. alcaliphilum* in the exponential growth phase with dry weight concentrations of 170 mg $L^{-1}$.

**Table 3** | Effects of catechol at 0.4 mM on the specific $CH_4$ consumption rates, $O_2$:$CH_4$ consumption ratios and formaldehyde accumulation rates of *Methylocystis hirsuta* and *Methkylomicrobium alcaliphilum*. Error intervals are standard deviations (3 biological replicates).

| | ***M. hirsuta*** **(control)** | ***M. hirsuta*** **(catechol 0.4 mM)** | ***M. alcaliphilum*** **(control)** | ***M. alcaliphilum*** **(catechol 0.4 mM)** |
|---|---|---|---|---|
| Specific $CH_4$ consumption rate (mmol g-$DW^{-1}day^{-1}$) | 181±1 | 17±5 | 4.95±0.18 | 2.65±0.4 |
| Specific $CH_2O$ accumulation rate (mmol g-$DW^{-1}day^{-1}$) | 0.12±0.31 | 0.02±0.35 | 0 | 0.12±0.05 |
| $O_2$:$CH_4$ consumption ratio | 1.58±0.02 | 1.8±0.1 | 1.16±0.01 | 1.37±0.02 |

The specific methane elimination rates, oxygen-methane consumption ratios and specific glyceraldehyde accumulation rates are summarized in **Table 3**. The oxygen-methane consumption ratios without catechol inhibition were consistent with the redox-arm (for *M. hirsuta*) and the direct coupling mechanism (for *M. alcaliphilum*). A specific methane oxidation rate of 4.95 mmol g-$DW^{-1}day^{-1}$ was observed for *M. alcaliphilum* (compared to 181 for *M. hirsuta*). This was also consistent with the low specific growth rate observed during the pre-growth of the inoculum (0.1 $day^{-1}$ compared to 1.1 for *M. hirsuta*). Catechol addition resulted in a methane oxidation rate

of 9% its initial value for the type II methanotroph *M. hirsuta*, while the type I methanotroph *M. alcaliphilum* was still able to oxydize methane at 54% of its initial rate.

The concentration of glyceraldehyde was also measured. Assuming that in the case of a direct coupling mechanism, catechol addition will have a stronger effect on the downstream consumption of glyceraldehyde than on its production, a higher accumulation rate of this compound is expected due catechol addition to *M. alcaliphilum*. After 24 hours of catechol addition, the concentration of formaldehyde in the culture broth of *M. alcaliphilum* increased from zero (at the beginning of the experiment) to values between 0.36 and 0.65 mg $L^{-1}$, while no formaldehyde accumulation was observed for the control. In the case of *M. hirsuta*, glyceraldehyde was already present at the beginning of the experiment at 2 mg $L^{-1}$ and in some of the biological replicates it decreased while in others increased. No statistically significant difference was observed due to catechol addition (average specific accumulation rates and standard deviations are shown in **Table 3**).

Both organisms showed higher oxygen-methane consumption rates under catechol inhibition. This is likely due to lower efficiencies in proton extrusion caused by complex I inhibition and to proton leaks caused by membrane disruption. In any case, the oxygen-methane consumption ratio of *M. alcaliphilum* inhibited with catechol is still lower than the value for the non-inhibited *M. hirsuta.* All these results come to support the hypothesis of *M. hirsuta* relying on a redox-branch mechanism for methane oxidation, in contrast to the type I methanotroph *M. alcalyphilum*, which uses a direct coupling mechanism.

**Influence of the nitrogen source**

The experimental and predicted biomass yields of *M. hirsuta* when ammonium was used as nitrogen source instead of nitrate were herein compared. Nitrogen is typically incorporated into amino acids in the same oxidation state than ammonium, and therefore, no redox equivalents for its reduction are necessary, as it is the case for nitrate, which needs to be reduced to ammonium before being incorporated into amino acids, thus higher biomass yields are expected using ammonium as nitrogen source. However, ammonium is known to inhibit pMMO via competitive inhibition (Campbell et. al., 2011), as pMMO can also catalyze the oxidation of ammonium to hydroxylamine.

Indeed, when *M. hirsuta* was grown in an ammonium mineral medium with a molar $O_2$:$CH_4$ ratio of 1:1 in the headspace, the experimentally observed growth rate was 0.76±0.02 $day^{-1}$ compared to 1.11±0.02 $day^{-1}$ when nitrate was used as nitrogen source. Interestingly, nitrate was accumulated in the cultivation medium at a rate of 1.5±0.1 mmol g-$DW^{-1}$ $day^{-1}$ during the experiment carried out with ammonium as nitrogen source, which suggested that the hydroxylamine produced by the pMMO was further oxidized to nitrite and finally to nitrate. This nitrate accumulation rate revealed that the rate of ammonium oxidation by pMMO was only 2.3% of the methane oxidation rate (it was assumed that all the accumulated nitrate is being originated from $NH_4^+$ oxidized by pMMO to hydroxylamine and subsequently to nitrite and nitrate). However, the competitive inhibition mediated by the presence of ammonium caused the specific methane oxidation rate to drop from 137±0.7 mmol g-$DW^{-1}$ $day^{-1}$ to only 67±5 mmol g-$DW^{-1}$ $day^{-1}$. Finally, when setting the experimental methane uptake rate and the nitrate production rate to their experimental values, the predicted and experimental biomass productions in *M. hirsuta* cultures showed a good agreement (**Figure 4**).

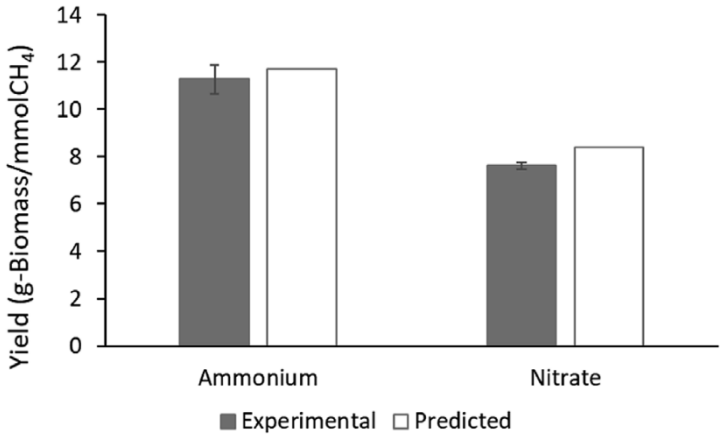



**Figure 4** | Predicted and experimental biomass yields using nitrate and ammonium as nitrogen sources in *M. hirsuta*. Ammonium resulted in higher yields on methane but at the expense of lower growth rates due to competitive inhibition of the pMMO enzyme. Grey bars correspond to the observed experimental values, with error bars equal to the standard deviations (for 3 biological replicates). White bars correspond to predictions using the GSMM of *Methylocystis hirsuta.*

Yields on methane using ammonium and nitrate as nitrogen sources were also measured for *Methylocystis* sp. SC2 (see **Figure S-1**), the obtained yields were practically identical to the theoretical ones.

**Production of PHB under nitrogen starvation conditions**

*Methylocystis hirsuta* can accumulate PHB over 40% of its total biomass dry weight under conditions of complete nitrogen depletion (López et. al., 2018). By monitoring methane consumption and PHB accumulation over time, it was found that *M. hirsuta* consumed methane at a specific rate of 46±1 mmol g-$DW^{-1}$ $day^{-1}$ under nitrogen deprivation. The experimental yield of PHB on methane was found to be 0.088±0.002 mol$PHB/mol CH_4$. Using the experimental specific methane consumption and ATP maintenance consumption rates as constraints, the optimization of PHB resulted in an estimated value of 0.11 mol$PHB/mol CH_4$ with an oxygenmethane consumption ratio of 1.5 mol$O_2$/mol$CH_4$ (the same as in the case of biomass production), compared to an experimental value of 1.52±0.012 mol$O_2$/mol$CH_4$. **Figure 5** shows the predictions and experimental results for PHB yield on methane and oxygen-methane consumption ratio.

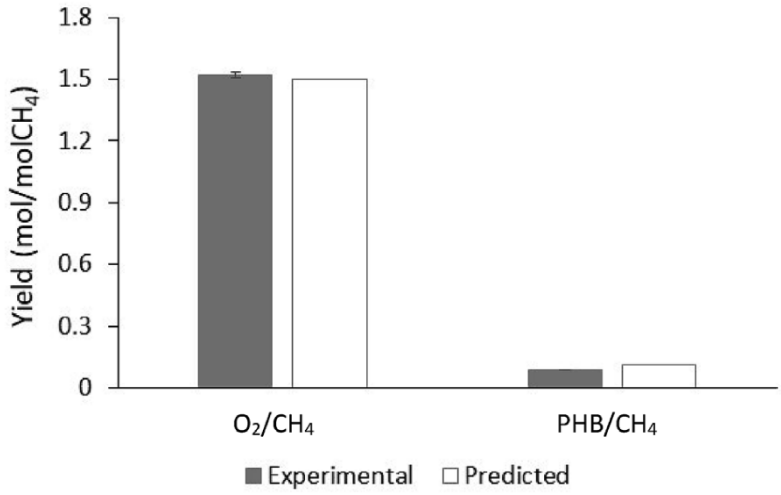



**Figure 5** | Theoretical versus experimental PHB yield on methane and oxygen-methane consumption molar ratios. Grey bars correspond to the observed experimental values, with error bars equal to the standard deviations (for 3 biological replicates). White bars correspond to predictions using the GSMM of *Methylocystis hirsuta*.

**Metabolic capability of *M. hirsuta* to supply building blocks for biosynthesis**

A large variety of industrially relevant products can be synthesized by *M. hirsuta* from a small number of metabolic precursors (Jouhten et. al., 2016), some key precursors and final products are summarized in **Table 4**.

For each of these final products, there are metabolic pathways that could be expressed in the host organism and which would drain the corresponding precursor in

the central carbon metabolism (their production would also be associated with the consumption of ATP and redox cofactors, therefore the estimations presented here are just meant to be indicative). Thus, biomass production (or PHB production in the case of aceto-acetyl-CoA) was first optimized using the reconstructed GSMMs. A parsimonious (minimizing the sum of metabolic fluxes) optimal solution was obtained using COBRApy and only the reactions active in this solution were allowed to carry flux in the subsequent optimizations. This was done in order to identify the metabolic sub-network that was active under optimal growth conditions (in the presence of nitrogen) and under optimal PHB synthesis (in the absence of nitrogen). Therefore, the results depicted in **Figure 6** were meant to represent the maximal amounts of each precursor that the introduction of a new pathway could drain from the active metabolic network of the cell (without extra metabolic enzymes having to be overexpressed). All the three metabolic networks resulted in identical results, suggesting that the architecture of the central carbon metabolism is very similar among the three species of *Methylocystis*.

**Table 4** |Metabolic precursors and derived industrial products.

| Metabolic precursor | Industrial products |
|---|---|
| 3-Dehyidroshikimate | Vaniline |
| L-Tyrosine | Pinocembrin, eriodictyol, naringenin, resveratrol, homoeriodicytol, chrisyn, luteolin, apigentin. |
| Pyruvate | Lactate |
| Oxaloacetate | Nicotinamine, 4-methylthiobutyl-desulfoglycosinolate. |
| Glyceraldehyde-3P | Propane-1,2-diol, propane-1,3-diol, |
| Acetyl-CoA | 6-methylsalicylate |
| Succinate | Succinate |
| Aceto-Acetyl-CoA | PHB, butanol, geraniol, beta-carotene, valencene, 8-epi-cedrol, cubebol, patchoulol, amorpha-4,11-diene |

As it could be expected, the precursor acetoacetyl-CoA is produced under nitrogen limitation conditions. The fact that *Methylocystis* strains are able to deviate high metabolic fluxes towards PHB (and therefore towards its precursor acetoacetyl-CoA) under nitrogen starvation, makes these operation conditions particularly suitable for the production of any of the derivatives of acetoacetyl-CoA (**Table 3**). In terms of total carbon yields, oxaloacetate was the precursor most efficiently produced, which makes these species potential cell factories for the production of chemicals such as nicotinamine or 4-methylthiobutyl-desulfoglycosinolate. Succinate, which besides being

a metabolite of the central carbon metabolism is also a very important building block for the chemical industry, can also be produced at high carbon yields. A simple overexpression of a suitable membrane transporter or the knockout of succinate consuming reactions, could make these strains suitable succinate producers.

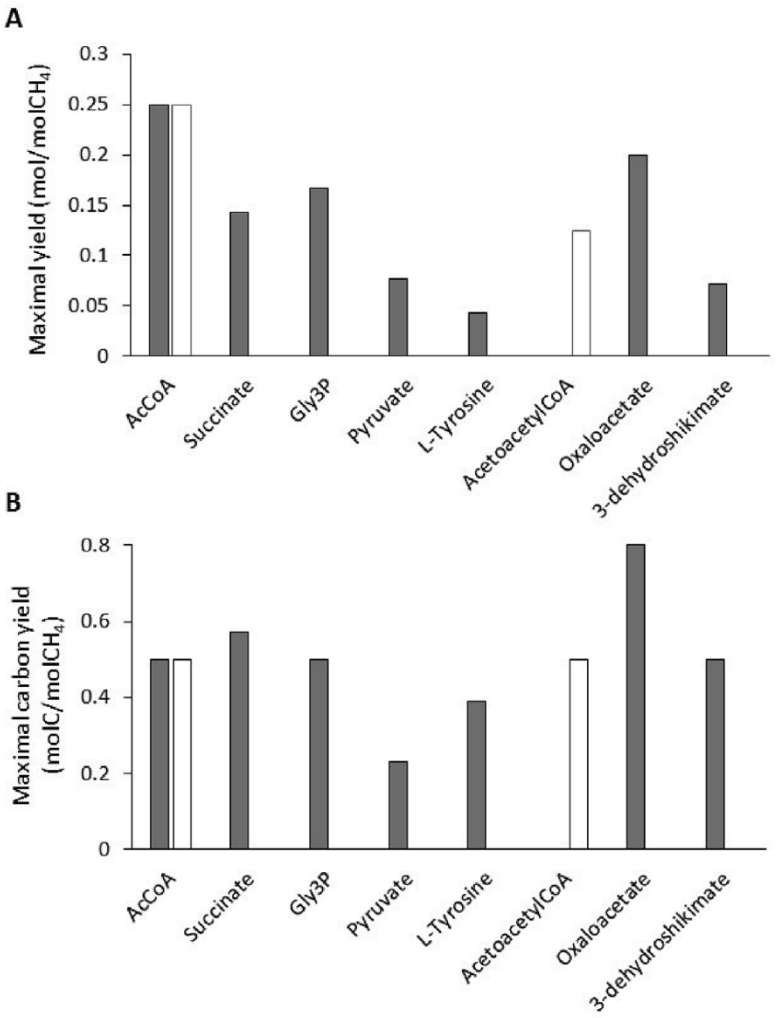



**Figure 6** | Maximal yields of metabolic precursors in the presence (grey bars) and absence of nitrogen (white bars). Molar yields on methane **(A)** and carbon yields on methane **(B)**. Each yield has been calculated by fixing the methane uptake rate to its experimental value for *Methylocystis hirsuta*, setting to zero the rates of the reactions non-used for growth optimization (grey bars), or PHB optimization (white bars), and maximizing the production rate of each metabolic precursors. In the cases of Acetyl-CoA and Acetoacetyl-CoA, the cofactor is previously removed from acetate and acetoacetate respectively (so that there is no net loss of CoA).

### Utilization of C2 compounds as carbon source

The strain *Methylocystis* sp. SB2 is known to be able to grow on C2 carbon sources such as acetate or ethanol (Im et. al., 2011). Acetyl-CoA can support the production of NADH through the TCA cycle, however it is not able to sustain the anaplerotic function of the

TCA cycle as supplier of precursors for biosynthesis. Therefore, in order to use C2 carbon sources, the cell must transform Acetyl-CoA into C3 compounds such as glycerate, which can be transformed into precursors for biosynthesis (via the lower glycolysis). *Methylocystis* sp. SB2 relies on a malonyl-CoA lyase to transform β-methylmalonyl-CoA into glyoxylate and propionyl-CoA. Glyoxylate is then transformed into glycerate via the so-called glyoxylate assimilation cycle. This mechanism has been confirmed in *Methylocystis* sp. SB2 by RNA-seq results, which showed the overexpression of all the metabolic genes involved in the glyoxylate assimilation cycle during growth on C2 substrates (Vorobev et. al., 2014). In this context, the models of *Methylocystis hirsuta* and *Methylocsytis* sp. SC2 included all the necessary reactions for glyoxylate assimilation, in particular the enzyme malonyl-CoA lyase. Therefore, it is reasonable to expect that all the three strains exhibit the potential to grow on C2 substrates.

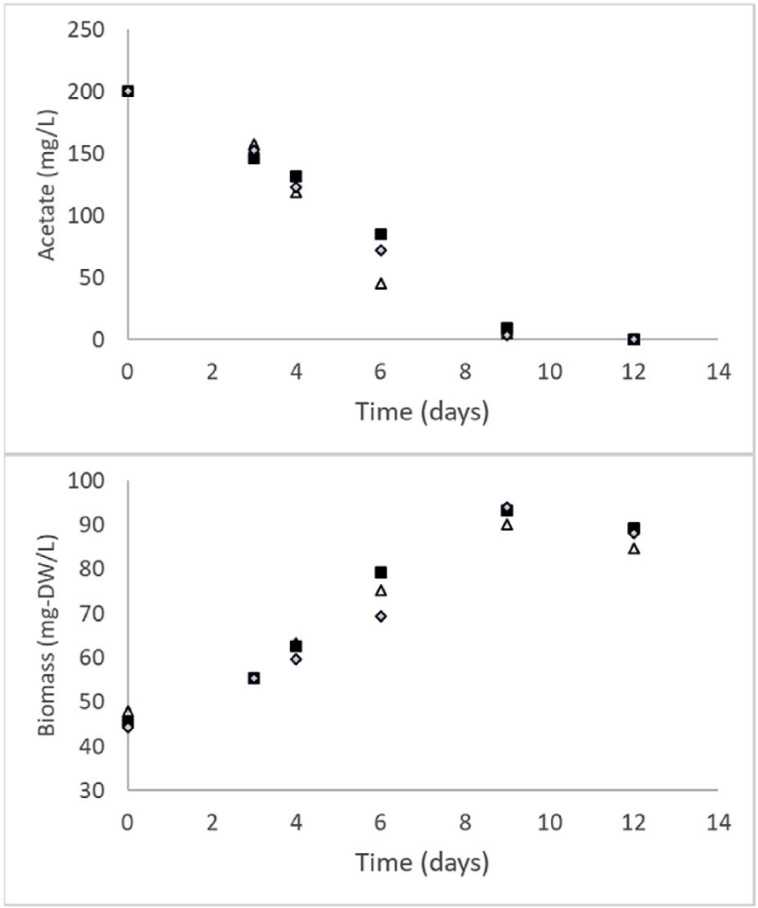



**Figure 7** | Growth and acetate consumption curves of *Methylocystis hirsuta*. Three biological replicates are shown. The observed biomass yield on acetate 14 ± 2 g-DW mol-1 was similar to the value reported for *Methylocistis* sp. SB2.

The ability of *Methylocystis hirsuta* to grow on acetate was assessed with very low

specific growth rates (0.08±0.006 $day^{-1}$) and a biomass yield on acetate of 14±2 g-DW $mol^{-1}$. This was very similar to the yield of 12±1 g-DW $mol^{-1}$ reported for *Methylocystis* sp. SB2 (Im et. al., 2011), but lower than the maximal theoretical yield of 17 g-DW $mol^{-1}$ predicted by the GSMM. It has been hypothesized that acetate has an uncoupling role in the respiratory chain, leading to lower respiration efficiencies (Im et. al., 2011). Despite the poor growth rates and biomass yields observed, the GSMM here developed suggests that *Methylocystis hirsuta* is indeed able to use C2 carbon compounds as carbon and energy sources.

Growth on acetate was also tested for *Methylocystis* sp. SC2 (see **Figure S-2**). After 400 hours the OD of the culture increased from 0.074 to 0.081, even if the difference was statistically significant (3 replicates), it cannot be concluded from the results that this strain is able to grow on C2 substrates (see **Figure 7**).

## 5.5. Discussion

The reconstruction of GSMMs for *M. hirsuta, M.* sp. SC2 and *M*. sp. SB2 revealed a very similar metabolism, with over 90% shared metabolic reactions and metabolites. The experimental biomass yields and oxygen-methane consumption ratios observed in *Methylocystis hirsute* matched the predictions obtained with the GSMM under the assumption of maximal biomass production and the "redox branch" mechanism for methane oxidation. The elucidation of the methane oxidation mechanism has been one of the main topics of controversy in the field of methanotrophy. Previous results with Type I methanotrophs pointed out to the "direct coupling" as the most likely mechanism (de la Torre et. al., 2015; Akberdin et. al., 2018). However, our results suggested that the nature of the redox co-factor involved in methane oxidation could be another of the key differences between Type I and II methanotrophs. The results for *Methylocystis hirsuta* regarding the methane oxidation mechanisms can very likely be extrapolated to the other two *Methylocystis* strains, as the pMMO enzymes coded in their genomes appeared to be very similar or identical (Bordel, et. al., 2018).

The experimental biomass yields of *Methylocystis hirsuta* were close to its maximal theoretical yields (both in the case of nitrate and ammonium used as carbon sources). This assumption of optimal yields, even if it is broadly used as a modeling approach for microorganisms, is often not accurate under substrate excess conditions

(Bordel, 2013). The growth rate of methanotrophs of the genus *Methylocystis* seems to be limited by the rate of methane oxidation. This was confirmed in the case of *Methylocystis* sp. SC2 by the fact that knocking out one of the two identical copies of a pMMO contained in the genome resulted in a decrease of its growth rate (Baani and Liesack, 2008). Organisms limited by substrate availability (or utilization) tend to have evolved to use substrates as efficiently as possible, leading to yields close to their theoretical maximum. Interestingly, the observed PHB yields under nitrogen deprivation conditions were also close to the theoretical maximum predicted by the model.

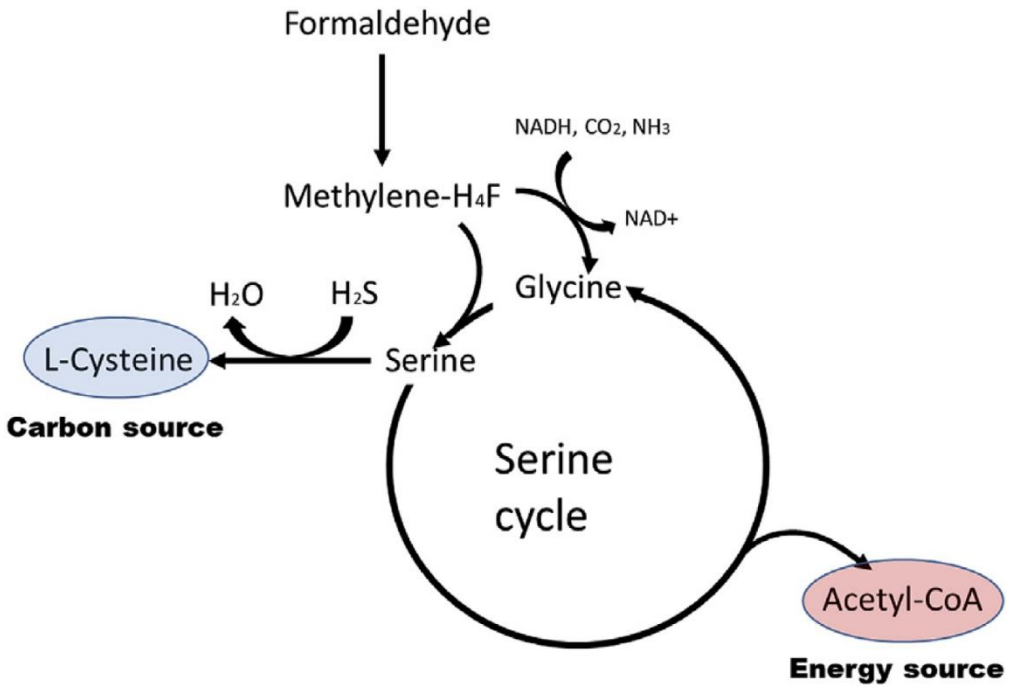



**Figure 8** | Illustration of the serine cycle producing acetyl-CoA for the TCA cycle and L-cysteine for biosynthesis of biomass building blocks. For the serine cycle to operate in steady state, the serine removed for the synthesis of L-cysteine has to be compensated by supplying glycine into the cycle by glycine synthase. In absence of this reaction, the serine cycle could operate supplying Acetyl-CoA (fueling the TCA cycle and respiration) but there would not be net production of precursors for biosynthesis.

It was also confirmed that *M. hirsuta*, as well as its relative *M.* sp. SB2, is able to grow on C2 carbon sources (acetate in this case) by using the glyoxylate assimilation pathway to transform acetyl-CoA into glycerate. However, the observed growth rate on acetate was dramatically lower than the growth rate on methane (17 times lower), which suggests that this pathway, even if it exists, is rather inefficient. This inefficiency at transforming acetyl-CoA into building blocks for biosynthesis reveals that the methane assimilation pathway in Type II methanotrophs, the serine cycle, is more complex than

its usual textbook descriptions. Indeed, the serine cycle yields acetyl-CoA as a product, which itself can be used to fuel the TCA cycle and obtain energy but not as a biomass building block (or at least not efficiently, as we have previously discussed). The flux distributions predicted by the reconstructed models revealed that part of the serine produced in the serine cycle is transformed into L-cysteine by a cystathionine β-synthase and this L-cysteine is the substrate for the biosynthesis of all the necessary biomass building blocks. In order to keep the concentrations of the intermediates of the serine cycle under steady state, glycine is supplied into the cycle by a glycine synthase that transforms methylene-tetrahydrofolate into glycine, playing an anaplerotic role in the serine cycle similar to the role played by pyruvate carboxylase in the TCA cycle (**Figure 7**). This high flux through L-cysteine should be considered for futuremetabolic engineering strategies using *Methylocystis* as cell factories (**Figure 8**).

Finally, Flux Balance Analysis simulations were performed in order to assess the potential of the *Methylocystis* strains here analyzed to supply precursors for the biosynthesis of chemicals of industrial and commercial interest. These simulations revealed that the *Methylocystis* strains have a considerable potential for the production of chemicals such as succinate, acetoacetyl-CoA derivatives (under nitrogen starvation conditions) and oxaloacetate derivatives.

## 5.6. Acknowledgements

This work was performed with the support of the Marie Curie grant H2020-MSCA-IF-2016 CH4BioVal (GA nº 750126). The financial support of the Regional Government of Castilla y León and the FEDER program and the Ministry of Science, Innovation and Universities are also gratefully acknowledged (CLU 2017-09, CTM2015-70442-R and VA281P18).

## 5.7. Supporting information

Supporting figures are provided below in section 5.7.1. In addition, the supporting information is included in the USB drive provided with the thesis and is also available on the Elsevier website at DOI: 10.1016/j.ymben.2019.04.001

### *5.7.1. Supporting figures*

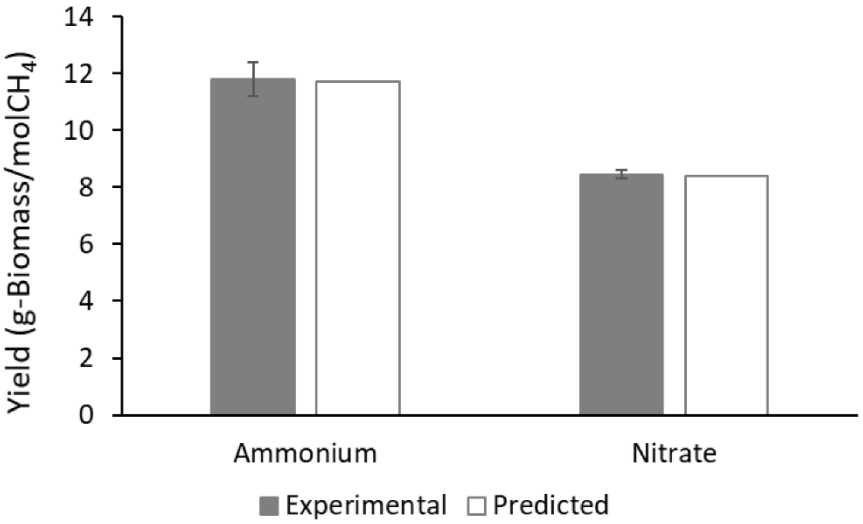



**Figure S-1** | Experimental and theoretical biomass yields on methane of the strain *Methylocysits* sp. strain SC2 using ammonium and nitrate as nitrogen sources.

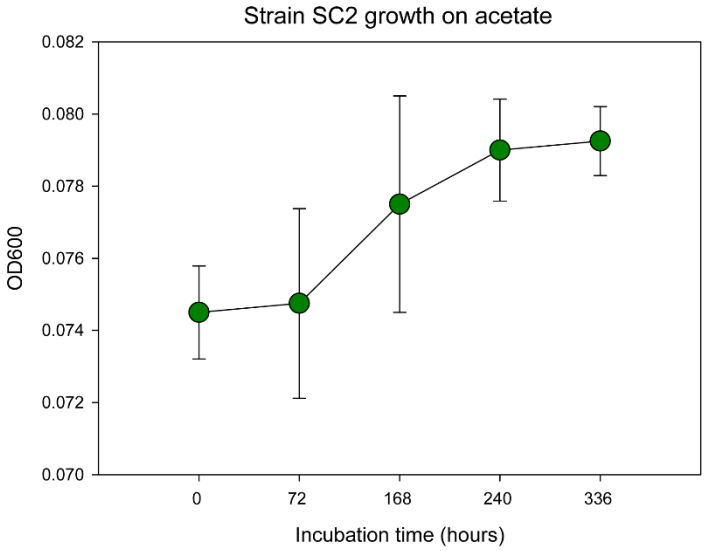



**Figure S-2** | Growth on acetate of the strain *Methylocystis* sp strain SC2.

## 5.7. References


Abbasi, T., Tauseef, S.M, Abbasi, S.A. (2012). Anaerobic digestion for global warming control and energy generation, an overview. *Renew Sustain Energy Rev* 16, 3228-3242.

Akberdin, I.R., Thompson, M., Hamilton, R., Desai, N., Alexander, D., Henard, C.A., Guarnieri, M.T., Kalyuzhnaya, M.G. (2018). Methane utilization in *Methylomicrobium alcaliphilum* 20ZR: a systems approach. *Sci Rep* 8, 2512.

Baani, M., Liesack, W. (2008). Two isozymes of particulate methane monooxygenase with different methane oxidation kinetics are found in *Methylocystis* sp. strain SC2. *Proc Natl Acad Sci USA* 105, 10203-10208.

Becker, S.A., Price, N.D., Palsson, B.O. (2006). Metabolite coupling in genome-scale metabolic networks. BMC *Bioinformatics* 7, 111.

Bordel, S. (2013). Experimental evidence suggests the existence of evolutionary conserved global operation principles governing global metabolism. *Sci Rep* 3, 3017.

Bordel, S., Rodríguez, E., Muñoz, R. (2018). Genome sequence of *Methylocystis hirsuta* CSC1, a polyhydroxyalkanoate producing methanotroph. *MicrobiologyOpen* 7, e771.

Campbell, M.A., Nyerges, G., Kozlowski, J.A., Poret-Peterson, A.T., Stein L.Y., Klotz, M.G. (2011). Model of the molecular basis for hydroxylamine oxidation and nitrous oxide production in methanotrophic bacteria. *FEMS Microbiol Lett* 322, 82-89.

Chelliah, V., Juty, N., Ajmera, I., Ali, R., Dumousseau, M., Glont, M., et al. (2015). BioModels: ten-year anniversary. *Nucl Acids Res* 43, D542-D548.

Comer A.D, Long M.R, Reed J.L, Brian F.P. (2017). Flux balance analysis indicates that methane is the lowest cost feedstock for microbial cell factories. *Metab Eng Commun* 5, 26-33.

Dam, B., Dam, S., Kube, M., Reinhardt, R., Liesack, W. (2012). Complete genome sequence of *Methylocystis* sp. strain SC2, an aerobic methanotroph with high affinity methane oxidation potential. *J Bacteriol* 194, 6008-6009.

Dam, B., Dam, S., Blom, J., Liesack, W. (2013). Genome analysis coupled with physiological studies reveals a diverse nitrogen metabolism in *Methylocystis* sp. strain SC2. *PLoS ONE* 8, e74767.

Degli Esposti, M. (1998). Inhibitors of NADH-ubiquinone reductase: an overview. *Biochim Biophys Acta* 1364, 222-235.

Ebrahim, A., Lerman, J.A., Palsson, B.O., Hyduke, D.R. (2013). COBRApy: Constraints-based reconstruction and analysis for python. *BMC Syst Biol* 7, 74.

García-Pérez, T., López, J.C., Passos, F., Lebrero, R., Revah, S., Muñoz, R. (2018). Simultaneous methane abatement and PHB production by Methylocystis hirsuta in a novel gas-recycling bubble column bioreactor. *Chem Eng J* 334, 691-697.

Henard, C.A., Smith, H., Dowe, N., Kalyuzhnaya, M.G., Pienkos, P.T., Guarnieri, M.T. (2016). Bioconversion of methane to lactate by an obligate methanotrophic bacterium. *Sci Rep* 6, 21585.

Im J., Lee, S.W., Yoon, S., DiSprito, A.A., Semrau, J.D. (2011). Characterization of a novel facultative *Methylocystis* species capable of growth on methane, acetate and ethanol. *Environ Microbiol Rep* 3, 174-181.

Jouhten, P., Boruta, T., Andrejev, S., Pereira, F., Rocha, I., Patil, K.R. (2016). Yeast metabolic chassis designs for diverse biotechnological products. *Sci Rep* 6, 29694.

Kalyuzhnaya, M.G., Puri, A.W., Lidstrom, M.E. (2015). Metabolic engineering in methanotrophic bacteria. *Metab Eng* 29, 142-152.

López, J.C., Quijano, G., Pérez, R., Miñoz, R. (2014). Assessing the influence of $CH_4$ concentration during culture enrichment on the biodegradation kinetics and population structure. *J Environ Manage* 146, 116-123.

Kalyuzhnaya, M.G., Khmelenina, V., Eshinimaev, B., Sorokin, D., Fuse, H., Lindsgrom, M., Trotsenko, Y. (2008). Classification of halo(alkali)philic and halo(alkali)tolerant methanotrophs provisionally assigned to the genera *Methylomicrobium* and *Methylobacter* and emended description of the

genus *Methylomicrobium*. *Int J Syst Evol Microbiol* 58, 591-596.

López, J.C., Arnáiz, E., Merchán, L., Lebrero, R., Muñoz, R. (2018). Biogas-based polyhydroxyalkanoates produced by *Methylocystis hirsuta*: A step further in anaerobic digestion biorefineries. *Chem Eng J* 333, 529-536.

Nguyen, A.D., Hwang, I.Y., Lee, O.K., Kalyuzhnaya, M.G., Mariyana, R., Hadiyati, S., et al. (2018). Systematic metabolic engineering of *Methylomicrobium alcaliphilum* 20Z for 2, 3-butanediol production from methane. *Metab Eng* 47, 323-333.

Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., et al. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* 33, 5691- 5702.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., et al. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res* 42, D206-D214

Patil. K.R., Rocha, I., Förster, J., Nielsen, J. (2005). Evolutionary programming as a platform for in silico metabolic engineering. *BMC Bioinform* 6, 308

Pieja, A.J., Morse, M.C., Cal, A.J. (2017). Methane to bioproducts: the future of the bioeconomy? *Curr Opin Chem Biol* 41, 123–131.

Puri, A.W., Owen, S., Chu, F., Chavkin, T., Beck, D.A.C., Kalyuzhnaya, M.G., et al. (2015). Genetic tools for the industrially promising methanotroph *Methylomicrobium buriatense*. *Appl Environ Microbiol* 81, 1775-1781.

Strong, P.J., Kalyuzhnaya, M., Silverman, J., Clarke, W.P. (2016). A methanotroph-based biorefinery: Potential scenarios for generating multiple products from a single fermentation. *Biores Technol* 215, 314-323.

de la Torre, A., Metivier, A., Chu, F., Laurens, L.M.L., Beck, D.A.C., Pienkos, P.T., et al. (2015). Genome-scale metabolic reconstruction and theoretical investigation of methane conversión in *Methylomicrobium buryatense* strain 5G (B1). *Microb Cell Fact* 14, 188.

Vorobev, A., Jagadevan, S., Jain, S., Anantharaman, K., Dick, G.J., Vuilleumier, S., et al. (2014). Genomic and transcriptomic analyses of the facultative methanotroph *Methylocystis* sp. strain SB2 grown on methane or ethanol. *Appl Environ Microbiol* 80, 3044-3052.

Whittenbury, R., Phillips, K.C., Wilkinson., J.F. (1970). Enrichment, isolation and some properties of methane-utilizing bacteria. *J Gen Microbiol* 61, 205-218.

Zuniga, C., Morales, M., Le Borgne, S., Revah, S. (2011). Production of poly-β-hydroxybutyrate by methylobacterium organophilum isolated form a methanotrophic consortium in a two phase-partition bioreactor. *J Hazard Mater* 190, 876-882.

# Chapter 6:

## Discussion and Outlook

# 6 Discussion and Outlook

## 6.1. Discussion

Methanotrophic bacteria (methanotrophs) are a unique group of microorganisms structurally and functionally specialized in the use of $CH_4$ as the sole source of carbon and energy. In this thesis, several research questions were addressed essential for the methanotrophs` research field. First and foremost, we were interested in developing a molecular technique that will allow us to reliably elucidate the global response of type IIa methanotrophs to environmental changes with high throughput (**Chapter 2** and **Chapter 3**). Subsequently, we aimed to use our newly developed technique to investigate facultative methanotrophy in type IIa methanotrophs, with a particular focus on potential applications in metabolic engineering and biotechnology (**Chapter 4** and **Chapter 5**).

We chose *Methylocystis* sp. strain SC2 as our model organism. This strain is one of the well-studied type IIa methanotrophs. The complete genomic information of strain SC2 was obtained by whole-genome shotgun sequencing (Dam et al., 2012). Based on its complete genome sequence, two comprehensive transcriptomics studies were conducted (Dam et al., 2014; Han et al., 2017). Undoubtedly, the transcriptome research increased our understanding of gene regulation in strain SC2. Nonetheless, transcriptomics is time-consuming and does not allow a high throughput. In addition, it is costly and does not elucidate the response of the main functional units of the cells to the environmental factors. In a nutshell, it does not uncover the structure, function, and expression of all translated proteins, which define the cellular phenotype of any given organism.

Hence, in **Chapter 2** we focused on the development of an optimized global proteome analysis workflow for strain SC2. Due to the high amount of membrane proteins in strain SC2, we foremost focused on optimizing cell lysis and the protein solubilization conditions. The advantageous effect of our new workflow was primarily based on the choice of the solubilization conditions. Importantly, the workflow was

based on downstream analysis of the crude lysate, since the overall protein amount strongly decreased after the removal of cell debris from the lysates. This strong decrease is explained by the high amount of hydrophobic membrane-associated proteins in SC2 cells. Application of MS-compatible detergent SDC combined with an extended heating of the crude lysates (up to 60 min) resulted in a two-fold higher concentration of the protein extract than in all other tested conditions (**Figure 1**) (Hakobyan et al., 2018).

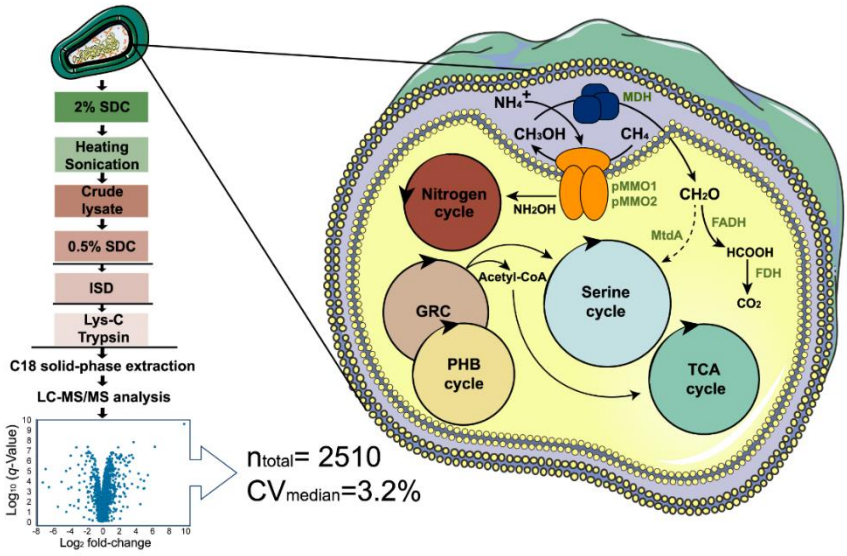



**Figure 1** | Flowchart illustrating the (i) different solubilization and digestion conditions, (ii) procedural steps used to obtain an optimized proteomics workflow for *Methylocystis* sp. strain SC2 and (iii) results obtained using the newly developed workflow. *Adopted from (Hakobyan et al., 2018).*

The next crucial step in our workflow development was the proteolytic digestion of extracted and solubilized proteins. The endopeptidase trypsin is the most commonly used protease in the field of shotgun proteomics. Nonetheless, we decided to apply a LysC/trypsin tandem digestion scheme for strain SC2, in order to increase the cleavage of hydrophobic membrane-associated proteins, such as pMMO subunits. LysC enables the specific pre-digestion of hydrophobic proteins prior to trypsin digestion. The tandem digestion was efficiently applied in several membrane proteomics studies and proved to increase the coverage of membrane-associated proteins (Vit and Petrak, 2017). In addition, LysC/trypsin tandem digestion was found most efficient in yielding fully cleaved peptides while reducing the abundance of miscleaved peptides (Glatter et al., 2012). Though, in initial study, it was applied in differential denaturation environment of urea, we used the tandem digestion with our crude lysate in SDC environment.

Our data in **Chapter 2** demonstrate that tandem digestion performs efficiently in constant detergent concentrations. Nonetheless, **Chapter 2** does not cover the comprehensive assessment of tandem digests in detergents using modern quantitative proteomics approaches (Hakobyan et al., 2018). Therefore, in **Chapter 3** we decided to compare the efficiency of LysC/trypsin digestion between urea and detergent environments (e.g., SDC and SLS). This comparison showed that the efficiency of tandem digestion is higher than in trypsin-only digestion, regardless of the digestion environment. Therefore, based on our findings in **Chapter 2** and **Chapter 3,** we suggest that the positive effect of tandem strategies is primarily due to the different target specificities of LysC and trypsin, rather than the effect of sample denaturation environment and state (Hakobyan et al., 2019).

Upon LC-MS analysis and LFQ, our new crude-MS workflow presented in **Chapter 2** proved to increase the overall proteome coverage of strain SC2, capturing 62% of the predicted SC2 proteome, with the least quantification bias for the different protein fractions. Most importantly, crude-MS led to a more than 10-fold higher coverage of membrane-associated proteins, in particular pMMO isozymes, relative to less effective conditions (**Figure 1**). Moreover, the positive effects of higher coverage combined with less quantification bias were also observed for some cytoplasmic and periplasmic proteins involved in PHB/GR and TCA cycles (Hakobyan et al., 2018).

Expanding our research to other members of the *Methylocystis/Methylosinus* group and to *E. coli* allowed us to show that the preference for particular protein solubilization conditions is a specific property of methanotrophic bacteria. In particular, the strong disparity in the contribution of the membrane sub-proteome to the total cellular proteome between strain SC2 and *E. coli* is due to the presence of pMMO in SC2 cells, which is directly linked to its methanotrophic lifestyle (details in **Chapter 1**). These findings suggest that for efficient and reproducible proteomics analysis, the proteomics workflow should be optimized when applied the first time to a microorganism with specialized cell architecture. This will help to avoid major quantification biases (Hakobyan et al., 2018).

As next, we aimed to address some of the unsolved problems in methanotrophy, using our newly developed proteomics workflow in combination with genomics and metabolic modeling. Metabolic engineering of methane-utilizing methanotrophs has

recently attracted attention, due to the cheap price of methane as a feedstock and the identification of previously unknown methanotroph traits such as facultative growth and novel biosynthetic capabilities (Kalyuzhnaya et al., 2015). For successful methanotroph applications in metabolic engineering, it is pivotal to understand what is not known about methanotrophy and how to fill knowledge gaps. In general, metabolic pathways other than the methane oxidation pathway are poorly resolved, and very little is known about the mechanisms how methanotrophic bacteria adapt to changes in environmental parameters or cultivation conditions (Kalyuzhnaya et al., 2015).

The impact of nitrogen as either ammonium- or nitrate-based fertilizer on the methane fluxes in soils is a widely studied topic of global concern. Various methanotroph studies have shown that aerobic methane oxidation is very sensitive to the addition of N-based fertilizers in upland soils (Mohanty et al., 2006). In particular, ammonium-based fertilizers are known to strongly inhibit the sink capacity of these soils for atmospheric methane. We know that this inhibition effect primarily occurs at low-methane and high-ammonium concentrations. Nonetheless, there is a lack of information about the molecular mechanisms on how ammonium affects methanotrophic cells. First insights into the effect of high-ammonium exposure on strain SC2 was obtained by Dam *et al*. (2014) using whole-genome transcriptomics. In **Chapter 2**, we extended this study and applied our newly developed SC2 proteomics workflow to uncover the functional and metabolic responses of strain SC2 to ammonium and nitrate treatments (Dam et al., 2014; Hakobyan et al., 2018). The results agreed well with those obtained by Dam et al. (2014) and thus validated our proteomics workflow as a fast and sensitive technique suitable for further research on strain SC2. Upon exposure to high ammonium, a good correspondence was observed between differential protein expression and the previously reported changes in mRNA transcription (Dam et al., 2014). The results of both transcriptomics and proteomics suggest that high concentrations of ammonium have an inhibitory effect on the methane oxidation capacities of strain SC2. This is due to a competitive inhibition of pMMO (Campbell et al., 2011) and the accumulation of toxic intermediates (hydroxylamine and nitrite) produced via the pMMO-based oxidation of ammonium (Nyerges et al., 2010; Stein, 2012). As further discussed in **Chapter 5**, the competitive inhibition mediated by the presence of ammonium indeed caused a drop in the methane oxidation rate. Nonetheless, in agreement with metabolic

predictions, type IIa methanotrophs showed a higher experimental biomass yield with ammonium than with nitrate (Bordel et al., 2019). This can be explained by the fact that ammonium will be incorporated into amino acids without a reduction step. By contrast, nitrate needs to be reduced to ammonium before being incorporated into amino acids. Thus, higher biomass yields are expected with ammonium as a nitrogen source, even when methane oxidation is partially inhibited. As a complementary information to already discussed findings, proteomics (**Chapter 2**) also revealed that the negative effect of ammonium is in part also due to its ionic-osmotic pressure on strain SC2 (Hakobyan et al., 2018).

In the following project, we aimed to elucidate the molecular mechanisms involved in the adaptation of type IIa methanotrophs to limiting methane and oxygen concentrations and, with that, in the switch to carbon and energy sources other than methane such as, for example, methanol, multi-carbon compounds, and hydrogen. In the early 2000`s, it was believed that methanotrophs gain metabolic energy (ATP) only via the respiratory chain by oxidizing NADH and/or reduced cytochromes that are formed during the oxidation of methane to carbon dioxide. However, methanotrophs often inhabit environments where methane and oxygen availability significantly varies over time (Roslev and King, 1995). Therefore, facultative methanotrophy was suggested to be a survival strategy under methane limitation. In fact, a facultative lifestyle was proven for members of the genus *Methylocella* (Dedysh et al., 2005; Dedysh and Dunfield, 2011). Later, facultative methanotrophy was also shown to occur in *Methylocystis* spp. and *Methylocapsa* spp (Dunfield et al., 2010; Belova et al., 2011; Im et al., 2011). However, it remains to be elucidated how frequently and which kind of alternative energy sources methanotrophic bacteria utilize under unfavorable conditions. By the means of genome-scale metabolic modeling (GSSM) and growth experiments, in **Chapter 5**, we showed that three *Methylocystis* spp., including strain SC2, have the metabolic capacity to grow on acetate (Bordel et al., 2019). The constructed metabolic models predicted that these strains use the glyoxylate assimilation pathway to transform acetyl-CoA to glycerate, thereby yielding high biomass as predicted. Nonetheless, the experimental findings in **Chapter 5** showed that the biomass yield from acetate was much lower than with methane, suggesting that the predicted pathway of acetate utilization is quite inefficient. Nonetheless, when flux

distributions were incorporated into the reconstructed metabolic models, it was shown that with acetate utitlization, some part of the serine produced during serine cycle feeds the TCA cycle resulting in NADH genertation. Furthermore, some part of serine transforms to L-cysteine, which serves as precursor for the cell biosynthesis from acetate (Bordel et al., 2019). Therefore, we suggest that the L-cysteine flux in type IIa methanotrophs should be considered for efficient predictions of their metabolic activities.

More recently, the alphaproteobacterial methanotrophs have been shown to contain genes encoding different types of [NiFe]-hydrogenases (Carere et al., 2017; Greening et al., 2016). Using $H_2$ as an easy accessible energy source, many microorganisms can prevail fluctuations in the availability of their preferred carbon and energy sources that occur due to changes in the environmental conditions (Hoehler and Jorgensen, 2013). Hydrogen for use as an electron donor to support methanotroph growth can be produced by non-expensive green energy sources such as wind and solar energy powered water electrolysis (Matassa et al., 2015). The utilization of $H_2$ as an alternative energy source may be an effective strategy of type IIa methanotrophs to increase their biomass yield and to survive under unfavorable conditions.

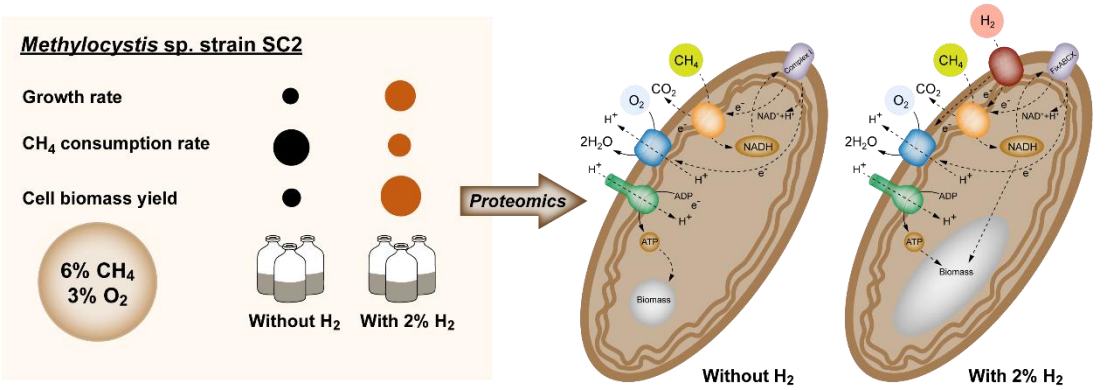



**Figure 2** | The illustration shows the physiological and proteome changes of strain SC2 upon $H_2$ supplementation under limited $CH_4$ and $O_2$ conditions. The continuous supply of $H_2$ as an alternative energy source leads to an increased biomass yield and decreased methane consumption rate in strain SC2. In long-term incubation (37d) with 2% $H_2$, strain SC2 reconfigures the electron transport chain, which leads to the increase in biomass yield. Strain SC2 uses the membrane-integrated Group 1d uptake hydrogenase to provide electrons to the electron transport chain through the quinone pool. Reduced quinone then transfers the electrons either to pMMO for $CH_4$ oxidation or to cytochrome c and terminal oxidase for ATP generation. Moreover, NADH oxidation may provide additional electrons through activity of the FixABCX complex.

In **Chapter 4**, we disclosed the ability of strain SC2 to utilize $H_2$ as an additional energy source under $CH_4$ and $O_2$ limited conditions (**Figure 2**). The increase in biomass yield agreed well with both thermodynamic model predictions of max biomass yield and the observed changes in SC2 proteome. As discussed in **Chapter 4**, strain SC2 responded to the addition of 2% $H_2$ with an immediate up-regulation of regulatory and uptake hydrogenases (Group 2b and Group 1d), and hydrogenase accessory proteins (**Figure 2**). Moreover, the electron transport chain of strain SC2 had been majorly reconstructed after long-term incubation. Given that hydrogenase-encoding genes are widely distributed among type IIa methanotrophs, it is most likely that the ability to utilize $H_2$ as an energy source is not restricted to strain SC2 but a more general phenomenon among *Methylocystis* spp.

The changes in the SC2 proteome upon the addition of $H_2$, together with the metabolic modeling predictions in **Chapter 5,** provided first experimental evidence for the transfer mechanism of those electrons that drive pMMO activity in type IIa methanotrophs. Despite major attempts, the identity of the electron donor(s) remained hitherto elusive (Kalyuzhnaya et al., 2015). Three mechanisms were proposed for methane oxidation: (i) "*redox arm*", in which the electrons driving methane oxidation are derived from NADH that is produced by formate and/or formaldehyde oxidation, while electrons produced from methanol oxidation are linked to the redox arm and used for ATP production; (ii) "*direct coupling"*, which means the direct transfer of reducing power from MDH (Mxa) to pMMO, and (iii) "*uphill electron transfer*" model, in which methanol oxidation partially supports methane oxidation (Kalyuzhnaya et al., 2015). In **Chapter 5,** we showed that in contrast to type I methanotrophs, type IIa methanotrophs use the "redox arm" mechanism to provide electrons to pMMO through the activity of NADH dehydrogenase (complex I). First experimental evidence for the "redox arm" mechanism is described in **Chapter 4.** Complex I was significantly down-regulated during the long-term incubation with $H_2$. Its electron transfer function was taken over by the activity of the flavin-based electron bifurcating complex (FixABCX) in combination with Group 1d uptake hydrogenase. In summary, $H_2$ can be utilized by strain SC2 and, most likely, other type IIa methanotrophs as an energy source for survival, but may also be considered for biotechnologically oriented methanotroph projects as an alternative energy source to significantly increase biomass yield from $CH_4$.

## 6.2. Outlook and concluding remarks

In this work, we developed an efficient crude-MS proteomics workflow for type IIa methanotrophs, which proved to be a rapid, efficient and accurate molecular technique to study the effect of environmental factors on metabolic changes in strain SC2. This is well evidenced by the proteome changes that we observed upon exposure of strain SC2 to high ammonium (**Chapter 2**). In addition, our crude-MS strategy may be applied to a variety of proteomic workflows that incorporate cell types with challenging solubilization properties (Hakobyan et al., 2018). Furthermore, we were able to address unsolved questions in the field of methanotroph research by means of our proteomics workflow and GSSMs.

In some upland and hydromorphic soils and in aquifer environments, nitrogen ($N_2$) fixed either as nitrate or ammonium may be limiting (Chu and Alvarez-Cohen, 1996). Therefore, most type II methanotrophs, including *Methylocystis* spp., are able to express the nitrogenase complex and to fix atmospheric $N_2$ (Auman et al., 2001). Thus, methanotrophs capable of $N_2$ fixation would have an advantage. However, it is yet unknown how methanotrophs adapt their metabolism to limitations in fixed nitrogen by switching to atmospheric $N_2$ fixation. Due to the high energy demand of $N_2$ fixation, a methanotrophic diazotroph requires high energy-yielding activity in order to sustain $N_2$ fixation (Kox et al., 2018). Therefore, environmental factors controlling methanotrophy supposedly contribute to the observed variability in the $CH_4$-dependent $N_2$ fixation (Ho and Bodelier 2015). However, the effect of environmental fluctuations in methane and oxygen on the ability of methanotrophs to fix atmospheric $N_2$ is not well explored (Kox et al., 2018). Therefore, proteomics in combination with transcriptomics and measurements of $N_2$ fixation activity could be applied to study the response of type IIa methanotrophs to nitrogen limitation and the environmental factors affecting this activity.

The genome sequence of strain SC2 revealed a full chromosome-encoded complement of $N_2$ fixation-related genes (Dam et al., 2013), including the genes encoding the active catalytic site of Mo-Fe nitrogenase, namely *nifH*, *nifD*, and *nifK*. Therefore, strain SC2 represents an ideal model organism to elucidate the effect of nitrogen limitation on the methane and nitrogen metabolisms of type IIa methanotrophs. Moroever, preliminary physiological studies showed that strain SC2 can

indeed grow under fixed nitrogen limitation ($N_2$ fixation), with the optimal oxygen concentration for growth between 5% and 10%. Though, it was shown that nitrogenase activity in strain SC2 was the greatest at around 1% oxygen in the headspace (Dam et al., 2013). Biological $N_2$ fixation activity in diazotrophic microorganisms is most commonly measured by the acetylene reduction assay (Hardy et al., 1968). However, with the use of acetylene to measure $N_2$ fixation rates, the role of diazotrophic methanotrophs may have been underestimated, because acetylene itself is an irreversible inhibitor of pMMO (Leppanen et al., 2015). Therefore, in strain SC2, acetylene reduction test was applied under the growth with methanol as a carbon and energy source. However, in future studies, $^{15}N_2$ stable isotope probing may be used as an alternative and more accurate method to measure $N_2$ fixation activity in $CH_4$ atmosphere (Vile et al., 2014; Kox et al., 2018).

In the native environment, methanotrophic bacteria and particularly *Methylocystis* spp. are often exposed to spatial heterogeneities in the available methane and oxygen (Roslev and King, 1995). Methane-oxygen gradients have been shown in different soils and sediments to change daily and seasonally (Roslev and King, 1994; Reim et al., 2012; Ho et al., 2015). In paddy soils, long-lasting droughts reduce methanotrophic activity due to the decrease in methane concentration (Ho et al., 2015). Moreover, the oxygen concentration can fluctuate in methanotrophic habitats between supersaturation and hypoxic condition in very short time (Roslev and King, 1994; Reim et al., 2012). Aerobic methanotrophs inhabiting these habitats should have specialized strategies to encounter the shortage in methane, oxygen, and/or alternative energy sources. Some methanotrophs are able to form exospores and/or cysts, thereby able to persist for extended periods during unfavorable conditions (Whittenbury et al., 1970). *Methylocystis* spp., including strain SC2, are also shown to form stress-resistant cysts to survive long-term starvation and desiccation. To date, a few physiological studies have focused on the survival mechanisms of methanotrophs under complete depletion of methane, in both oxic and anoxic conditions (Roslev and King, 1994; 1995; Tavormina et al., 2017). Nonetheless, their cellular changes have not yet been explored in detail. Therefore, in perspective, it would be interesting to investigate the strategies of *Methylocystis* spp. to withstand short-term and long-term fluctuations or even absence of methane, with and without the availability of oxygen.

**In Chapter 4,** we used a combination of proteomics and physiological growth experiments to show that, under low mixing ratios of methane and oxygen, strain SC2 is able to use $H_2$ as an alternative energy source to meet the demands of energy metabolism. As a step further, we could apply global proteome analyses to assess the physiological and functional changes in type IIa methanotrophs during active growth, methane starvation, cyst formation, and methane replenishment with and without oxygen availability. Finally, it would be interesting to specifically track the changes in pMMO turnover rates during the different phases of cellular and metabolic adaptation.

## 6.3. References


Auman, A.J., Speake, C.C., and Lidstrom, M.E. (2001). nifH sequences and nitrogen fixation in type I and type II methanotrophs. *Appl Environ Microbiol* 67(9)**,** 4009-4016.

Belova, S.E., Baani, M., Suzina, N.E., Bodelier, P.L.E., Liesack, W., and Dedysh, S.N. (2011). Acetate utilization as a survival strategy of peat-inhabiting *Methylocystis* spp. *Environ Microbiol Rep* 3(1)**,** 36-46.

Bordel, S., Rodriguez, Y., Hakobyan, A., Rodriguez, E., Lebrero, R., and Munoz, R. (2019). Genome scale metabolic modeling reveals the metabolic potential of three type II methanotrophs of the genus *Methylocystis*. *Met Eng* 54**,** 191-199.

Campbell, M.A., Nyerges, G., Kozlowski, J.A., Poret-Peterson, A.T., Stein, L.Y., and Klotz, M.G. (2011). Model of the molecular basis for hydroxylamine oxidation and nitrous oxide production in methanotrophic bacteria. *FEMS Microbiol Lett* 322(1)**,** 82-89.

Chu, K.H., and Alvarez-Cohen, L. (1996). Trichloroethylene degradation by methane-oxidizing cultures grown with various nitrogen sources. *Water Environ Res* 68(1)**,** 76-82.

Dam, B., Dam, S., Blom, J., and Liesack, W. (2013). Genome analysis coupled with physiological studies reveals a diverse nitrogen metabolism in *Methylocystis* sp. strain SC2. *PLoS One* 8(10)**,** e74767.

Dam, B., Dam, S., Kim, Y., and Liesack, W. (2014). Ammonium induces differential expression of methane and nitrogen metabolism-related genes in *Methylocystis* sp. strain SC2. *Environ Microbiol* 16(10)**,** 3115-3127.

Dam, B., Dam, S., Kube, M., Reinhardt, R., and Liesack, W. (2012). Complete genome sequence of *Methylocystis* sp. strain SC2, an aerobic methanotroph with high-affinity methane oxidation potential. *J Bacteriol* 194(21)**,** 6008-6009.

Dedysh, S.N., and Dunfield, P.F. (2011). Facultative and obligate methanotrophs how to identify and differentiate them. *Meth Enzym* 495**,** 31-44.

Dedysh, S.N., Knief, C., and Dunfield, P.F. (2005). *Methylocella* species are facultatively methanotrophic. *J Bacteriol* 187(13)**,** 4665-4670.

Dunfield, P.F., Belova, S.E., Vorobev, A.V., Cornish, S.L., and Dedysh, S.N. (2010). *Methylocapsa aurea* sp. nov., a facultative methanotroph possessing a particulate methane monooxygenase, and emended description of the genus *Methylocapsa*. *Int J Syst Evol Microbiol* 60**,** 2659-2664.

Glatter, T., Ludwig, C., Ahrne, E., Aebersold, R., Heck, A.J., and Schmidt, A. (2012). Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion. *J Proteome Res* 11(11)**,** 5145-5156.

Hakobyan, A., Liesack, W., and Glatter, T. (2018). Crude-MS Strategy for in-depth proteome analysis of the methane-oxidizing *Methylocystis* sp. strain SC2. *J Proteome Res* 17(9)**,** 3086-3103.

Hakobyan, A., Schneider, M.B., Liesack, W., and Glatter, T. (2019). Efficient tandem LysC/trypsin digestion in detergent conditions. *Proteomics* 19(20)**,** e1900136.

Han, D., Link, H., and Liesack, W. (2017). Response of *Methylocystis* sp. strain SC2 to salt stress: physiology, global transcriptome, and amino acid profiles. *Appl Environ Microbiol* 83(20), e00866-17.

Hardy, R.W., Holsten, R.D., Jackson, E.K., and Burns, R.C. (1968). The acetylene-ethylene assay for $N_2$ fixation: laboratory and field evaluation. *Plant Physiol* 43(8)**,** 1185-1207.

Ho, A., van den Brink, E., Reim, A., Krause, S.M., and Bodelier, P.L. (2015). Recurrence and frequency of disturbance have cumulative effect on methanotrophic activity, abundance, and community structure. *Front Microbiol* 6**,** 1493.

Hoehler, T.M., and Jorgensen, B.B. (2013). Microbial life under extreme energy limitation. *Nat Rev Microbiol* 11(2)**,** 83-94.

Im, J., Lee, S.W., Yoon, S., Dispirito, A.A., and Semrau, J.D. (2011). Characterization of a novel facultative *Methylocystis* species capable of growth on methane, acetate and ethanol. *Environ Microbiol Rep* 3(2)**,** 174-181.

Kalyuzhnaya, M.G., Puri, A.W., and Lidstrom, M.E. (2015). Metabolic engineering in methanotrophic bacteria. *Met Eng* 29**,** 142-152.

Kox, M.A.R., Aalto, S.L., Penttila, T., Ettwig, K.F., Jetten, M.S.M., and van Kessel, M. (2018). The influence of oxygen and methane on nitrogen fixation in subarctic Sphagnum mosses. *AMB Express* 8(1)**,** 76.

Leppanen, S., Rissanen, A., and Tiirola, M. (2015). Nitrogen fixation in Sphagnum mosses is affected by moss species and water table level. *Plant Soil* 389(1-2)**,** 185-196.

Mohanty, S.R., Bodelier, P.L., Floris, V., and Conrad, R. (2006). Differential effects of nitrogenous fertilizers on methane-consuming microbes in rice field and forest soils. *Appl Environ Microbiol* 72(2)**,** 1346-1354.

Nyerges, G., Han, S.K., and Stein, L.Y. (2010). Effects of ammonium and nitrite on growth and competitive fitness of cultivated methanotrophic bacteria. *Appl Environ Microbiol* 76(16)**,** 5648-5651.

Reim, A., Luke, C., Krause, S., Pratscher, J., and Frenzel, P. (2012). One millimetre makes the difference: high-resolution analysis of methane-oxidizing bacteria and their specific activity at the oxic-anoxic interface in a flooded paddy soil. *ISME J* 6(11)**,** 2128-2139.

Roslev, P., and King, G.M. (1994). Survival and recovery of methanotrophic bacteria starved under oxic and anoxic conditions. *Appl Environ Microbiol* 60(7)**,** 2602-2608.

Roslev, P., and King, G.M. (1995). Aerobic and anaerobic starvation metabolism in methanotrophic bacteria. *Appl Environ Microbiol* 61(4)**,** 1563-1570.

Stein, L.Y., Real, R., and Dunfield, P. F. (2012). Aerobic Methanotrophy and Nitrification: Processes and Connections. *In eLS, John Wiley & Sons,* Ltd: Chichester.

Tavormina, P.L., Kellermann, M.Y., Antony, C.P., Tocheva, E.I., Dalleska, N.F., Jensen, A.J., et al. (2017). Starvation and recovery in the deep-sea methanotroph *Methyloprofundus sedimenti*. *Mol Microbiol* 103(2)**,** 242-252.

Vile, M.A., Wieder, R.K., Zivkovic, T., Scott, K.D., Vitt, D.H., Hartsock, J.A., et al. (2014). $N_2$-fixation by methanotrophs sustains carbon and nitrogen accumulation in pristine peatlands. *Biogeochemistry* 121(2)**,** 317-328.

Vit, O., and Petrak, J. (2017). Integral membrane proteins in proteomics. How to break open the black box? *J Proteomics* 153**,** 8-20.

Whittenbury, R., Phillips, K.C., and Wilkinso.Jf (1970). Enrichment, isolation and some properties of methane-utilizing bacteria. *J Gen Microbiol* 61**,** 205-18.

## Acknowledgements

Working as a doctoral student in Max Planck Institute for Terrestrial Microbiology in Marburg was a magnificent as well as challenging experience for me. During these 4-5 years, many people directly or indirectly influenced my professional and personal development. It would not have been possible to succeed in my doctoral work without the precious support of all kind and knowledgeable people surrounding me, to only few of whom it is possible to give a particular mention here.

First and foremost, I would like to thank my supervisor, **PD Dr. Werner Liesack**. My doctoral studies and this thesis would never be possible to be accomplished without your enormous support, patience, guidance, encouragement and understanding in all kind of professional and personal situations. I would like to express my gratitude for your trust and encouragement for freedom to embark my scientific journey and academic development. Thank you very much for giving me a chance to be the member of your research group during these years and to complete my doctoral studies in respectable manner.

Furthermore, I would like to convey my deepest regards to my thesis committee members, **Prof. Dr. Lennart Randau, Prof. Dr. Andreas Brune, and Prof. Dr. Uwe Maier**, for their precious time, helpful guidance and valuable suggestions. In particular, I am very grateful to **Prof. Dr. Andreas Brune** for his insightful comments and for sharing his tremendous experience in environmental microbiology, during our joint seminars and beyond.

I would like to thank **Dr. Timo Glatter**, who became my proteomics "*guru*" and opened a completely new research field for me, which became eventually the basis of my doctoral studies. Without your great supervision, cheerful enthusiasm and ever-friendly nature I would not be able to lead my research work to its current high-quality state. During my PhD curriculum, I was able to interact with you in routine basis, which helped me to learn many valuable insights in proteomics field.

Above all, this thesis is dedicated to the two most important people in my life, **my son** and **my husband**. A very special thank you my dear husband, for your practical and emotional support in balancing our family duties with my work-study-professional crazy tasks. Thank you my love for your boundless patience, your encouragement and your inspiration throughout my PhD journey. You and our small son are my only motivation in personal and professional development. Without you, I would never finish this long way full of obstacles. Thank you for being a loving and understanding spouse and friend for me throughout my PdD journey. Moreover, I would like to express my endless gratitude to **my wonderful parents**, who brought me up, provided excellent education, and supported me in my determination to find and realize my potential, and to make my part of contribution to the surrounding world. I would like to thank **my lovely sister**, for providing her assistance and help whenever I or my family needed her. Thank you for being always close to me besides the distance between us. Thank you for being not only an awesome sister, but also a best friend for me.

Furthermore, I would like to thank all former and current members of our group. In particular, I thank you my dear friend **Dr. Rehab Abdallah** for your unbiased friendship and devotion. It was delightful to have a colleague and friend like you. I am also very grateful to my former colleague **Dr. Hyokwan Bae**, who provided the essential and useful support in the beginning of my methanotrophs research in the lab. In addition, I thank **Vera Jansen**, who was not only an excellent bachelor student, but also a very kind friend to me. It was very joyful and fulfilling to work side by side with you. Thank you for your effective work and input. I would like to thank all my collaborators, and most importantly **Dr. Svetlana N. Dedysh** and her team, for insightful and fruitful discussions, and resulted joined publications. In addition, I would like to

thank **Igor Oshkin** for his friendship and cheerful discussions during his short, but very productive stay in our lab.

I would like to thank the International Max Planck Research School for Environmental, Cellular and Molecular Microbiology (**IMPRS-Mic**) for provided finical support during my PhD studies and for excellent international PhD program with interdisciplinary training in modern microbiology, state-of-the-art research facilities, and mentoring by outstanding scientists. Moreover, I would like to thank **Dr. Marcela Hernández García** and **Dr. Zrinka Gattin** the former IMPRS-Mic coordinators and **Dr. Dusica Rados**, the current IMPRS-Mic coordinator, for all their help and assistance in solving not only academic, but also family-related questions. Thank you for your time, your advice and your beneficial suggestions.

In addition, I would like to thank **Christiane Nüsslein-Volhard-Stiftung** for granting me twice with CNV-foundation grant for excellent women scientists with children in the field of experimental natural sciences. I would like to express my cordial gratitude for the excellent opportunity to be one of the grantees of CNV-foundation, which provided me with additional capacity to organize my work/family balance.

Furthermore, I would like to thank **the administration** and **IT department** of MPI-TM. Special thanks to **Jens Hemer** for guiding me through the challenging bureaucratic details of German system, always providing excellent advice and valuable information. Moreover, I appreciate **Dr. Steffi Reissmann** for priceless support in all kind of family- and childcare-related questions and issues. I would like to express my sincere gratitude to **the late Mr. Christian Bengelsdorff**, who always did everything possible to support the doctoral students at the institute. I will always remember his kindness, humaneness and candor to me and my family.

Last but not least, I would like to thank **Prof. Dr. Nils-Kåre Birkeland**, for giving me opportunity to be trained in the state-of-the-art techniques of microbial ecology field available in his group after my master studies. Thank you and also your wonderful wife Mildred for great hospitality, humanity and warm atmosphere, which made me feel home in cold Norway. Moreover, I would like to thank my first supervisor **Dr. Hovik Panosyan**, for teaching me the fundamental basis of microbiology, for the chance to study abroad after my master studies. Besides professional training in his group, I also learned to be strong enough to resist the spectacular challenges of academic life and relationships.