Insights into anaerobic degradation of benzene and naphthalene

Xiyang Dong
天道酬勤

God rewards those who work hard
Abstract

Aromatic hydrocarbons, e.g. benzene and naphthalene, have toxic, mutagenic and/or carcinogenic properties. Fortunately, these compounds can be degraded in environmental systems by indigenous microorganisms. Especially under anaerobic conditions, the physiology and ecology of the microbes involved are still poorly understood. In this thesis, important knowledge gaps are addressed in this field using microbiological approaches combined with “omics tools” (metagenomics and metaproteomics). A novel “reverse stable isotope labelling” approach is also introduced to investigate biodegradation activities.

Firstly, the enrichment culture BP L was studied, which can degrade benzene coupled with sulfate reduction. It is dominated by an organism of the genus Pelotomaculum. Members of this genus are usually known to be fermenters, undergoing syntrophy with anaerobic respiring microorganisms or methanogens. It remains unclear if Pelotomaculum identified here (namely, Pelotomaculum candidate BPL) could perform both benzene degradation and sulfate reduction. By using a metagenomic approach, a high-quality genome was reconstructed for it. Interestingly, the genome has all the genes for a complete sulfate reduction, similar as in other Gram-positive sulfate-reducing bacteria. The proteome analysis revealed that the essential enzymes for sulfate reduction were all expressed during growth with benzene. The proteogenomic data also suggested: (1) anaerobic benzene degradation was activated by a yet unknown mechanism for conversion of benzene to benzoyl-CoA; (2) the central benzoyl-CoA degradation pathway involved reductive dearomatization by a class II benzoyl-CoA reductase followed by hydrolytic ring cleavage and modified β-oxidation; (3) the oxidative acetyl-CoA pathway was utilized for complete oxidation to CO₂. All these data indicate that, besides its potential to anaerobically degrade benzene, Pelotomaculum candidate BPL is the first member of the genus that can perform sulfate reduction.

Secondly, the functional role of environmental Spirochaetes was investigated. Such microbes are frequently detected in anaerobic hydrocarbon- and
organohalide-contaminated environments, but their function remains unclear. For insights into their roles in such systems, a sulfate-reducing, naphthalene-degrading enrichment culture was taken as an example. It comprises sulfate-reducing Deltaproteobacteria and the spirochaete Rectinema cohabitans. Genome sequencing and proteome analysis suggested that the Spirochaete is an obligate fermenter that catalyzes proteins and carbohydrates resulting in acetate, ethanol, and hydrogen production. Physiological experiments showed that H₂ derived from fermentation by R. cohabitans was used as reductant for sulfate respiration by the Deltaproteobacteria. Differential proteomics and physiological experiments revealed that R. cohabitans in turn utilizes biomass (proteins and carbohydrates) released from cell lysis. Further comparative genome analyses indicated that Rectinema species are widespread in contaminated environments and adapted to a hydrogenogenic fermentative lifestyle similar to R. cohabitans. Together, this indicates that environmental Spirochaetes thrive in highly toxic contaminated habitats by scavenging detrital biomass concomitant with production of hydrogen. The reductants produced can in turn fuel biomass formation by other hydrogenotrophic microorganisms in the ecosystem. Thus, environmental Spirochaetes might constitute a critical component of a microbial loop central to nutrient cycling in subsurface environments.

Finally, a sensitive approach to monitor biodegradation was introduced using reverse stable isotope labelling (RIL) via dissolved inorganic carbon (DIC). Microbial mineralization of organic compounds in laboratory microcosms is preferably assessed by adding ¹⁴C or ¹³C-labelled substrates and following the release of labelled CO₂. Labelled compounds, however, are usually expensive or not commercially available. The biodegradation assay used here contains unlabeled organic contaminants and ¹³C-labelled DIC with ¹³C atom fractions (x(¹³C)DIC) clearly above natural abundance (typically 2-50%). The CO₂ produced during incubation (x(¹³C) ≈ 1.11%) gradually dilutes the initial x(¹³C)DIC allowing to quantify microbial mineralization using mass-balance calculations. For ¹³C-enriched CO₂ samples, a newly-developed Isotope Ratio Mid-Infrared Spectrometer was shown to have a precision of x(¹³C) < 0.006%. As proof-of-principle, measured CO₂ production was close to the theoretical stoichiometry for anaerobic naphthalene degradation in a sulfate-reducing enrichment culture.
Furthermore, the method was used to measure aerobic degradation of dissolved organic carbon (DOC) adsorbed to granular activated carbon in a drinking water production plant. This process cannot be investigated using $^{13}$C-labelled substrates. Thus, the RIL approach can be applied to sensitively monitor biodegradation of various organic compounds (e.g. benzene and naphthalene) as well as complex substrate mixtures, which are not available with $^{13}$C-labelling.
Zusammenfassung


das Potenzial für anaeroben Benzolabbau besitzt, sondern auch das erste Mitglied dieser Gattung ist, das Sulfat reduzieren kann.


nicht erhältlich. Die Methode, die hier verwendet wurde, enthält nicht-markierten, organischen Kontaminationsstoff und \(^{13}\)C-markierten anorganischen Kohlenstoff mit einem höheren \(^{13}\)C-Anteil, als er natürlicherweise vorkommt. Das produzierte \(\text{CO}_2\) verdünnt allmählich den ursprünglichen \(^{13}\)C-Anteil, was eine Quantifizierung der mikrobiellen Mineralisierungsrate mittels Berechnung der Massenbilanz erlaubt. Ein neu entwickeltes Isotopenverhältnis-Mittelinfrarotspektrometer ermöglichte das Messen von \(^{13}\)C-Anteilen von weniger als 0.006%. Die gemessene \(\text{CO}_2\)-Produktion war nahe der theoretisch ermittelten Stöchiometrie für anaeroben Naphthalinabbau in einer sulfatreduzierenden Anreicherungskultur. Außerdem wurde die Methode verwendet, um aeroben Abbau von gelöstem organischen Kohlenstoff zu messen, der an granulare Aktivkohle in einer Anlage zur Trinkwassergewinnung gebunden war. Dieser Prozess kann nicht mit Hilfe von \(^{13}\)C-markierten Substraten gemessen werden. Das Verfahren umgekehrter stabiler Isotopenmarkierung kann also verwendet werden den biologischen Abbau verschiedenster organischer Substanzen (z.B. Benzol und Naphthalin) sowie komplexer Substratmischungen, die nicht mit \(^{13}\)C markiert werden können, empfindlich zu messen.
# Table of Contents

Abstract.................................................................................................................. 1

Zusammenfassung................................................................................................ 4

List of Figures ...................................................................................................... 11

List of Tables ....................................................................................................... 12

Abbreviations ...................................................................................................... 13

1. Introduction .................................................................................................. 15

1.1. Anaerobic benzene and naphthalene degradation with various electron acceptors ................................................................. 15

1.2. Multiple possible activation mechanisms for anaerobic benzene degradation .................................................................................. 17

1.3. Syntrophic biodegradation of benzene and naphthalene ......................... 20

1.3.1 Methanogenic degradation of benzene and naphthalene .................. 20

1.3.2 Syntrophic benzene degradation ...................................................... 21

1.4. Besides key players, who else is there and what are they doing? ........... 23

1.5. The importance of microbial H₂ metabolism in the subsurface system.... 24

1.6. Using ¹³C isotopes to quantify anaerobic hydrocarbon degradation ...... 25

1.6.1 Stable isotope fractionation .............................................................. 26

1.6.2 Stable isotope labelling..................................................................... 28

1.6.3 Stable isotope techniques................................................................. 29

1.7. Objectives of this thesis ........................................................................... 30

2. Materials and Methods ................................................................................. 33

2.1. Cultivation of anaerobic cultures .............................................................. 33

2.1.1 Cultivation of the enrichment culture BPL ........................................ 33
2.1.2 Cultivation of the enrichment culture N47 and its pure isolate .... 33
2.1.3 Cultivation of bacterial cultures for RIL experiments ............. 34
2.2. Metagenomic and genomic analyses ........................................ 36
  2.2.1 Binning draft genomes from the BPL metagenome ............... 36
  2.2.2 Sequencing and reconstruction of Spirochaetes genomes ...... 37
2.3. Phylogenetic analysis .......................................................... 39
  2.3.1 16S rRNA gene analysis of *Pelotomaculum* ..................... 39
  2.3.2 16S rRNA gene analysis of Spirochaetes ......................... 39
  2.3.3 Hydrogenase classification and analysis .......................... 40
2.4. Proteomic analysis .............................................................. 40
  2.4.1 Shotgun proteomic analysis of the enrichment culture BPL 40
  2.4.2 Mass spectrometry-based proteome analyses of N47 .......... 42
2.5. Spatial distribution profiles of microbial activities in GAC biofilms 45
2.6. Analytical methods ............................................................. 46
2.7. Analysis of carbon stable isotope ratios ................................ 47
  2.7.1 Sample preparation for carbon isotope ratio measurements ... 47
  2.7.2 Expression of carbon isotope ratios .................................. 48
  2.7.3 Isotope Ratio Mass Spectrometry ..................................... 48
  2.7.4 Isotope Ratio Mid-Infrared Spectroscopy .......................... 49
2.8. Calculations of CO₂ production using carbon isotope ratios ....... 49

3. Results .................................................................................... 52

3.1. Reconstructing metabolic pathways of a member of the genus *Pelotomaculum* suggesting its potential to oxidize benzene to carbon dioxide with direct reduction of sulfate ........................................... 52
  3.1.1 Phylogenetic binning and proteome ................................... 52
  3.1.2 A novel member of the genus *Pelotomaculum* .................. 55
  3.1.3 Initial activation of benzene ............................................. 57
  3.1.4 Central benzoyl-CoA degradation pathway ....................... 58
  3.1.5 Complete oxidation via the Wood-Ljungdahl pathway ........ 59
  3.1.6 Sulfate reduction in the genus *Pelotomaculum* ................. 61
3.1.7 Hydrogenases and formate dehydrogenases ........................................... 62

3.2. Fermentative Spirochaetes drive nutrient cycling by a subsurface microbial loop in hydrocarbon-contaminated habitats ........................................... 63
    3.2.1 The genome and proteome of *R. cohabitans* suggest it is an obligate fermenter ................................................................. 63
    3.2.2 Interspecies hydrogen transfer supports a naphthalene-degrading enrichment culture ............................................................ 66
    3.2.3 *R. cohabitans* feeds on organic carbon derived from necromass .... 70
    3.2.4 Spirochaetes may generally drive microbial loops in hydrocarbon- and organohalide-contaminated environments ..................... 73

3.3. Monitoring microbial mineralization using reverse stable isotope labelling analysis by mid-infrared laser spectroscopy ........................................ 75
    3.3.1 Verification of the method .............................................................................................................................. 76
    3.3.2 Evaluation of highly enriched $\delta^{13}$C values measured by IRMS ...... 79
    3.3.3 Using IRIS for analyzing highly enriched $^{13}$C samples ................ 80
    3.3.4 Sensitivity of CO₂ quantification under anoxic conditions .......... 82
    3.3.5 Determining aerobic mineralization rate of GAC biomass ............ 84

4. Discussion .............................................................................................................................. 87
    4.1. A member of the genus *Pelotomaculum* oxidizes benzene to carbon dioxide with direct reduction of sulfate ........................................... 88
    4.2. A hydrogen-driven subsurface microbial loop ........................................ 91
    4.3. Reverse stable isotope labelling analysis as a novel tool to monitor biodegradation ................................................................. 93

5. Conclusions and Outlook ...................................................................................................... 96
    5.1. General conclusions .............................................................................................................................. 96
    5.2. Future perspective .............................................................................................................................. 97

6. References ............................................................................................................................ 99

Authorship Clarification ........................................................................................................... 114
List of Figures

Figure 1. Possible activation mechanisms for initial steps of anaerobic benzene degradation ........................................................................................................................................ 18
Figure 2. Two working scenarios for Peptococcaceae as primary degraders ...... 22
Figure 3. Application of $^{13}$C stable isotope tools .......................................................... 26
Figure 4. Near-infrared spectrum of 5% carbon dioxide ........................................... 29
Figure 5. Epifluorescence microscopy images of naphthalene-degrading, iron-reducing enrichment cultures .................................................................................. 35
Figure 6. Evolution of influent and effluent DOC concentrations .................... 46
Figure 7. Identification of marker genes within recovered bins with CheckM ...... 53
Figure 8. Maximum likelihood tree of near-complete 16S rRNA genes ............. 54
Figure 9. Metabolic pathway reconstruction for Pelotomaculum candidate BPL .. 56
Figure 10. Genomic context of the metF gene ............................................................ 60
Figure 11. Metabolic pathway reconstruction for Spirochaetes......................... 64
Figure 12. Determinants of H$_2$ metabolism in Spirochaetes ......................... 68
Figure 13. H$_2$ oxidation and evolution under different growth conditions ......... 69
Figure 14. Quantitative proteomic analysis of highly enriched Deltaproteobacteria N47 .................................................................................................................. 70
Figure 15. Condensed maximum likelihood tree ................................................. 72
Figure 16. Anaerobic degradation profiles............................................................. 76
Figure 17. The calculated CO$_2$ production for anaerobic degradation .......... 77
Figure 18. Performance comparison of IRIS and IRMS ..................................... 79
Figure 19. Anaerobic degradation of naphthalene ............................................. 83
Figure 20. Aerobic mineralization of DOC adsorbed on the GAC filters .......... 85
Figure 21. Aerobic mineralization potential of DOC........................................ 86
Figure 22. A simplified scheme of subsurface microbial loop ......................... 91
Figure 23. Reverse stable isotope labelling analysis ........................................ 93
List of Tables

Table 1. Standard free energy changes and calculated possible synthesis of ATP ................................................................. 16
Table 2. Stoichiometric calculations of the measured and the theoretical ratios .. 78
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abc</td>
<td>Anaerobic benzene carboxylase</td>
</tr>
<tr>
<td>APS</td>
<td>Adenosine-5'-phosphosulfate</td>
</tr>
<tr>
<td>BAC</td>
<td>Biologically active carbon</td>
</tr>
<tr>
<td>BCRs</td>
<td>Benzoyl-CoA reductases</td>
</tr>
<tr>
<td>Bss</td>
<td>Benzylsuccinate synthase</td>
</tr>
<tr>
<td>CRDS</td>
<td>Cavity ring-down spectroscopy</td>
</tr>
<tr>
<td>CSIA</td>
<td>Compound-specific isotope analysis</td>
</tr>
<tr>
<td>CTAB</td>
<td>Hexadecyltrimethylammonium bromide</td>
</tr>
<tr>
<td>DAPI</td>
<td>4', 6-diamidino-2-phenylindole</td>
</tr>
<tr>
<td>DIC</td>
<td>Dissolved inorganic carbon</td>
</tr>
<tr>
<td>DOC</td>
<td>Dissolved organic carbon</td>
</tr>
<tr>
<td>DsrABC</td>
<td>Dissimilatory sulfite reductase cytoplasmic subunits</td>
</tr>
<tr>
<td>DsrMk</td>
<td>Transmembrane dissimilatory sulfite reductase complex</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescent in situ hybridization</td>
</tr>
<tr>
<td>GAC</td>
<td>Granular activated carbon</td>
</tr>
<tr>
<td>GC-MS</td>
<td>Gas chromatography–mass spectrometry</td>
</tr>
<tr>
<td>IRMS</td>
<td>Isotope ratio mass spectrometry</td>
</tr>
<tr>
<td>IRIS</td>
<td>Isotope ratio infrared spectroscopy</td>
</tr>
<tr>
<td>MCR</td>
<td>Methyl coenzyme M reductase</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass-to-charge ratios</td>
</tr>
<tr>
<td>Nod</td>
<td>Nitric oxide dismutase</td>
</tr>
<tr>
<td>OTU</td>
<td>Operational taxonomic unit</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PAHs</td>
<td>Polycyclic aromatic hydrocarbons</td>
</tr>
<tr>
<td>QmoAB</td>
<td>Quinone-interacting membrane-bound oxidoreductase</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>RIL</td>
<td>Reverse stable isotope labelling analysis</td>
</tr>
<tr>
<td>SIP</td>
<td>Stable isotope probing</td>
</tr>
</tbody>
</table>
1. Introduction

1.1. Anaerobic benzene and naphthalene degradation with various electron acceptors

Aromatic hydrocarbons, e.g. benzene and naphthalene, have toxic, mutagenic, and/or carcinogenic properties (Samanta et al. 2002). They can be released into groundwater due to leakages from pipelines, oil tanks or gas stations and other improper treatments, thus posing a hazardous risk to drinking water resources. But fortunately, these compounds can be removed by microorganisms which exist in nature.

Due to their limited chemical reactivity, these compounds were considered to be degraded only in the presence of free oxygen for many decades (Widdel and Rabus 2001). However, because of the low oxygen solubility in water (8 mg/L = 250 μM at 25 °C) and high carbon loads of aromatic hydrocarbons, the saturated systems rapidly turn anoxic even when they are exposed to small carbon loads (e.g. 33 μM benzene) (Meckenstock and Mouttaki 2011). Thus, anaerobic degradation of these compounds is of outstanding interest. Up to date, it has been shown that Fe(III), nitrate, sulfate, perchlorate and electrodes can be utilized by anaerobic microorganisms as terminal electron acceptors during the degradation of benzene and naphthalene (Meckenstock et al. 2016).

From a thermodynamic point of view, the redox potential of electron acceptors determines not only the biochemical strategies of anaerobic degradation pathways, but also the amount of potential energy the microorganisms can gain (Table 1). Normally, in case of acetate as the substrate, more energy can be theoretically provided for cell growth per mole of acetate for microorganisms when nitrate or ferric iron are used as electron acceptors, compared to the acceptors sulfate and carbon dioxide. Enigmatically, previous studies showed that anaerobic benzene degradation under various electron acceptors is against the energetic rule as shown in Table 1. One possible clue is that lower biomass yield under nitrate-reducing condition was produced than that under methanogenic conditions (Vogt et al. 2011). Additionally, anaerobic naphthalene degradation is
demonstrated so far only reliably to occur under sulfate- and ferric-reducing conditions (Kleemann and Meckenstock 2011, Musat et al. 2009). Nitrate-dependent naphthalene degradation is never reproduced or can only be stimulated by a non-fermentable substrate (Mittal and Rockne 2008), even though the oxidation of naphthalene with nitrate should be more feasible compared to with sulfate and ferric (Philipp and Schink 2012).

Table 1. Standard free energy changes and calculated possible synthesis of ATP associated with complete mineralization of 1 mol benzene and naphthalene with different electron acceptors.

<table>
<thead>
<tr>
<th>Electron acceptors (oxidized/reduced)</th>
<th>Stoichiometric equation</th>
<th>$\Delta G^0$ (KJ)</th>
<th>Estimated ATP (mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Benzene</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$/CH$_4$</td>
<td>C$_6$H$_6$ + 10.5H$_2$O + 3.75CO$_2$ → 6HCO$_3^-$ +3.75CH$_4$ + 6H$^+$</td>
<td>-106</td>
<td>1.5</td>
</tr>
<tr>
<td>SO$_4^{2-}$/H$_2$S</td>
<td>C$_6$H$_6$ + 3H$_2$O + 3.75SO$_2^{3-}$ → 6HCO$_3^-$ + 3.75HS$^-$ + 2.25H$^+$</td>
<td>-186</td>
<td>2.7</td>
</tr>
<tr>
<td>Ferrihydrite/Fe$^{2+}$</td>
<td>C$_6$H$_6$ + 30Fe(OH)$_3$ → 6HCO$_3^-$ + 30Fe$^{2+}$ + 18H$_2$O + 54OH$^-$</td>
<td>-525--1104</td>
<td>7.5~15.8</td>
</tr>
<tr>
<td>MnO$_2$/Mn$^{2+}$</td>
<td>C$_6$H$_6$ + 15MnO$_2$ + 24H$^+$ → 6HCO$_3^-$ + 15Mn$^{2+}$ + 12H$_2$O</td>
<td>-2104</td>
<td>30.1</td>
</tr>
<tr>
<td>NO$_3^-$/N$_2$</td>
<td>C$_6$H$_6$ + 6NO$_3^-$ → 6HCO$_3^-$ + 3N$_2$</td>
<td>-2978</td>
<td>42.5</td>
</tr>
<tr>
<td><strong>Naphthalene</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$/CH$_4$</td>
<td>C$_{10}$H$_8$ + 18H$_2$O + 6CO$_2$ → 10HCO$_3^-$ + 6CH$_4$ + 10H$^+$</td>
<td>-157</td>
<td>1.2</td>
</tr>
<tr>
<td>S$_0$/HS$^-$</td>
<td>C$_{10}$H$_8$ + 30H$_2$O + 24S → 24 HS$^-$ + 10HCO$_3^-$ + 34H$^+$</td>
<td>-40</td>
<td>0.6</td>
</tr>
<tr>
<td>SO$_4^{2-}$/H$_2$S</td>
<td>C$_{10}$H$_8$ + 6SO$_2^{3-}$ + 6H$_2$O → 10HCO$_3^-$ + 6HS$^-$ + 4H$^+$</td>
<td>-286</td>
<td>4.1</td>
</tr>
<tr>
<td>Ferrihydrite/Fe$^{2+}$</td>
<td>C$_{10}$H$_8$ + 48Fe(OH)$_3$ → 10HCO$_3^-$ + 48Fe$^{2+}$ + 28H$_2$O + 86OH$^-$</td>
<td>-828--1754</td>
<td>11.9~25.1</td>
</tr>
<tr>
<td>MnO$_2$/Mn$^{2+}$</td>
<td>C$_{10}$H$_8$ + 24MnO$_2$ + 38H$^+$ → 10HCO$_3^-$ + 24Mn$^{2+}$ + 18H$_2$O</td>
<td>-3355</td>
<td>47.9</td>
</tr>
<tr>
<td>NO$_3^-$/N$_2$</td>
<td>C$_{10}$H$_8$ + 1.2H$_2$O + 9.6NO$_3^-$ → 10HCO$_3^-$ + 0.4H$^+$ + 4.8N$_2$</td>
<td>-4752</td>
<td>67.9</td>
</tr>
</tbody>
</table>

Table 1 was modified from Meckenstock et al. (2016). ^a The Gibbs free energy of formation for naphthalene, 2-methylnaphthalene, phenanthrene are obtained from Mavrovouniotis (1990), Dean (1987), and McFarland and Sims (1991), respectively. The Gibbs free energy of formation for other substances are obtained from Madigan et al. (2008). Note that the calculation of energy yield is based on reduction potential differences when ferrihydrite is used as the electron acceptor (Madigan et al. 2008). ^b The standard redox potential for ferrihydrite/Fe$^{2+}$ is -0.1 V~+0.1 V (Straub et al. 2001). ^c The energy required to synthesize ATP is taken as 70 KJ/mol, assuming the intracellular pH in anaerobic bacteria as 7.0 (Jackson and McInerney 2002).
1.2. Multiple possible activation mechanisms for anaerobic benzene degradation

Unlike oxygenase-dependent reactions involved in aerobic degradation, anaerobic microorganisms have to employ alternative biochemical pathways to mineralize these compounds in the absence of molecular oxygen. For anaerobic degradation of benzene, three types of initial activation reactions have been proposed: methylation to toluene, hydroxylation to phenol and direct carboxylation to benzoate (Figure 1).

Methylation of benzene to toluene was proposed from the recovery of labeled toluene using compound-specific isotope analyses (Aburto-Medina and Ball 2014, Meckenstock and Mouttaki 2011, Vogt et al. 2011). But hitherto benzylsuccinate, the key product of further anaerobic toluene activation, has never been detected in benzene-degrading cultures. Although Taubert et al. (2012) found two proteins that might be involved in putative β-oxidation pathway of benzylsuccinate using protein-based stable isotope probing (protein-SIP) in a benzene-degrading, sulfate-reducing enrichment culture, they only showed a low homology (~30% identity) to BbsA and BbsB of Thauera aromatica.

Evidence favoring hydroxylation of benzene to phenol as initial activation reaction was found for Geobacter metallireducens which metabolizes benzene with Fe(III) citrate as the electron acceptor (Zhang et al. 2012, Zhang et al. 2013). This scenario was supported by the production of 18O-labelled phenol during growth with H218O, and the loss of capacity for benzene degradation due to deletion of genes encoding subunits of two phenol catabolic enzymes (Zhang et al. 2013). Subsequently, Gmet 0231 and 0232 were identified as genes specific for benzene rather than phenol oxidation (Zhang et al. 2014). Nevertheless, unless further investigation is performed on the gene products, these findings should be taken with great caution.
Figure 1. Possible activation mechanisms for initial steps of anaerobic benzene degradation.
A) Hydroxylation into phenol; B) carboxylation into benzoate; C) Wood-Ljungdahl-type carbonylation; D) methylation to toluene; E) forming of CoM-mimic molecules via coenzyme M-like reductases. Modified from Abu Laban et al. (2010).

Compared to methylation and hydroxylation, more literatures are in favor of direct carboxylation of benzene into benzoate as the first step. The best investigated examples comprise a Fe(III)-reducing enrichment culture BF (Abu Laban et al. 2010), a hyperthermophilic archaeon Ferroglobus placidus (Holmes et al. 2011), and a nitrate-reducing enrichment culture (Luo et al. 2014). Combined genomic and proteomic studies of culture BF revealed that gene products of ORFs 137 and 138 were probably two subunits of putative anaerobic benzene carboxylase (Abc), i.e., AbcD and AbcA, respectively (Abu Laban et al. 2010). The metatranscriptome of a nitrate-reducing enrichment culture during the metabolism of benzene versus benzoate suggested that members of the family Peptococcaceae were responsible for initial benzene activation. It showed 97% similarity in 16S rRNA gene sequences with the culture BF (Luo et al. 2014). In that case, two benzene-specific
genes similar to \textit{abcD} and \textit{abcA} were also highly transcribed. Consequently, these two benzene-induced genes might be used as a biomarker for anaerobic benzene degradation in contaminated sites. However, for field application, benzene carboxylation activity still needs to be confirmed in vitro. In a pure archaeal culture, the transcript abundance of Ferp\textunderscore1630 with homology to \textit{abcA} was higher in benzene-grown cells compared to acetate- and benzoate-grown cells. However, other genes encoding for carboxylase proteins could not be found in the proximity of Ferp\textunderscore1630 (Holmes et al. 2011). Recently, van der Waals et al (2017) detected a high abundance of the benzene carboxylase gene (\textit{abcA}) in a denitrifying biofilm reactor for benzene degradation dominated by members of the \textit{Peptococcaceae} and \textit{Ananerolineaceae}.

However, there is no direct biochemical evidence for any of these three reactions, so far. For some benzene-degrading cultures, it has been shown that their ability to degrade benzoate, toluene or phenol is not as high as expected, especially for sulfate-reducing and methanogenic cultures (Abu Laban et al. 2009, Devine 2013). On the one side, this fact limits the potential to perform a differential experiment using proteomics or transcriptomics. On the other side, other possible mechanisms for initial benzene activation might be possible. Laso-Pérez et al. (2016) proposed that an enzyme similar to methyl coenzyme M reductase (MCR) is responsible for initial butane cleavage in \textit{Candidatus Syntrophoarchaeum}. It can activate butane to butyl–CoM, followed by an unknown step for conversion of the intermediate butyl–CoM to butyl–CoA (Figure 1). If this is the case, the benzene-degrading microbes should harbor genes for both reverse methanogenesis and e.g. sulfate reduction. Only in this way it can make the hydrocarbon activation process thermodynamically favorable in one single cell (Laso-Pérez et al. 2016, Ragsdale 2016). Another possible mechanism was proposed by Devine (2013): a so-called Wood-Ljungdahl-type carbonylation pathway for conversion of benzene to benzoyl-CoA. In this process, a metal-centered cofactor might allow for a simultaneous addition of CO and CoA to form benzoyl-CoA (Figure 1).

In contrast to the above-mentioned mechanisms, \textit{Alicycliphilus denitrificans} strain BC and \textit{Dechloromonas aromatica} strain RCB, which grow on benzene with chlorate and nitrate as the electron acceptors, respectively, are suggested to
adopt cryptic aerobic pathways. They possibly use molecular oxygen produced intracellularly during reduction of chlorate or nitrate (Oosterkamp et al. 2013, Salinero et al. 2009, Weelink et al. 2008). Under nitrate-reducing conditions, oxygen can be produced by nitric oxide dismutase (Nod) via NO dismutation into $N_2$ and $O_2$. Recently, Zhu et al. (2017) found that nod genes exist in contaminated aquifers in high diversity and abundance.

1.3. Syntrophic biodegradation of benzene and naphthalene

Syntrophy is a term used to describe microbial cross-feeding, more specifically the tight interdependency between producer and consumer (Dolfing 2014). It is of importance for degradation of organic compounds at methanogenic conditions where electron acceptors are limited or absent. Fermenting bacteria convert complex substrates to small molecules, such as acetate, formate and $H_2$. But the accumulation of end products (e.g. hydrogen) makes the fermentation process energetically unfavorable. Methanogenic archaea can maintain those end products at low concentrations making the whole reaction exergonic (Morris et al. 2013). This example emphasizes the need of thermodynamic rational for microbial interdependency in the concept of syntrophy. But syntrophic interactions can also occur in the presence of electron acceptors (e.g. sulfate, nitrate, or ferric iron) without participation of methanogens (Gieg et al. 2014). At such a situation, the relationship might be not as tight as methanogenic conditions.

1.3.1 Methanogenic degradation of benzene and naphthalene

Methanogenic hydrocarbon degradation is important in hydrocarbon-contaminated environments, as electron acceptors are often not available due to the high carbon load (Jiménez et al. 2016, Sieber et al. 2012). However, due to slow-growth and the syntrophic nature of this process, laboratory-enriched methanogenic cultures have been rarely reported for benzene and naphthalene degradation. And most observations come from experiments based on stable isotope probing, molecular analysis and omics techniques including metagenomics, metatranscriptomics and metaproteomics (Jiménez et al. 2016).
Luo et al. (2016) monitored a long-term benzene-degradation experiment using different types of techniques, including 16S rRNA gene sequencing, metagenome sequencing and quantitative polymerase chain reaction (qPCR). They showed that members belonging to the new genus *Deltaproteobacterium* ORM2 initiated anaerobic benzene attack. Hydrogenotrophic methanogen (*Methanoregula, Methanospirillum, Methanocella, Methanolinea* or *Methanobacterium*) and acetoclastic *Methanoseta* were supposed to maintain low concentrations of hydrogen and acetate generated from benzene fermentation. Berdugo-Clavijo et al. (2012) tested if a crude oil-degrading methanogenic enrichment culture can be used for naphthalene degradation but they couldn’t find substantial methane production. Siegert et al. (2011) reported a similar observation when ¹³C labeled naphthalene was the only substrate added into microcosms containing Zeebrugge sediment.

### 1.3.2 Syntrophic benzene degradation

For syntrophy in the presence of electron acceptors such as nitrate, Fe(III), or sulfate, members of Gram-positive *Peptococcaceae* from the order *Clostridiales* are frequently reported as primary degraders of hydrocarbons like benzene (Lueders 2017) (Figure 2b). The family *Peptococcaceae* encompasses the genera *Peptococcus, Cryptanaerobacter, Dehalobacter, Desulfitibacter, Desulfispora, Desulfitobacterium, Desulfonispora, Desulfosporosinus, Desulfotomaculum, Desulfurispora, Pelotomaculum, Sporotomaculum, Syntrophobotulus*, and *Thermincola* (Stackebrandt 2014). Members of this family are physiologically heterogeneous with several types of metabolic features, including chemoorganotroph, chemolithoheterotroph, chemolithoautotroph, or syntrophy with hydrogenotrophs (Ezaki 2015, Stackebrandt 2014).
Figure 2. Two working scenarios for Peptococcaceae as primary degraders in the presence of electron acceptors such as nitrate, Fe(III), or sulfate. (a) Complete degradation coupled with reduction of electron acceptors is performed by a single cell process; (b) Peptococcaceae are responsible for fermentation of hydrocarbons and share electrons in the form of e.g. acetate and hydrogen with other respiring bacteria. Modified from Gieg et al. (2014).

In a denitrifying benzene-degrading enrichment culture (van der Zaan et al. 2012), addition of hydrogen inhibited anaerobic degradation of benzene presumably performed by Peptococcaceae based on results from DNA-stable isotope probing with $^{13}$C-labelled benzene. But they did not give clear evidence for who was the other denitrifying partner responsible for H$_2$ scavenge. In a study performed on another denitrifying enrichment culture (Luo et al. 2014), a benzene-degrading Peptococcaceae strain worked together with a benzoate-degrading denitrifying Azoarcus strain for the complete mineralization of benzene. But in a stricter interpretation of syntrophy, such an interaction is only to share work together but not because of the thermodynamic reason as for methanogenic conditions. Similar interactions were also reported for iron-reducing and sulfidogenic benzene-degrading enrichment cultures dominated by Peptococcaceae with evidence derived from DNA-SIP experiments (Herrmann et al. 2010, Kunapuli et al. 2007). They all proposed that electrons from benzene degradation could be transferred to either co-existing sulfate or iron reducers. But none of them could exclude the possibility that Peptococcaceae themselves could be directly involved with sulfate.
or iron reduction. Therefore, it is still likely that benzene can be mineralized by a single organism without syntrophic interactions (Figure 2a).

1.4. Besides key players, who else is there and what are they doing?

While it is important to have key players for activation of hydrocarbons and/or reduction of electron acceptors, it has become clear that more subtle and complex microbial interactions are also critical for complete removal of hydrocarbons in situ. Thus, besides a focus on the key players, it is relevant to understand who else is there and why they are there. Molecular surveys of anaerobic hydrocarbon-degrading enrichment cultures have revealed a majority of microbial species in the subsurface do not show a clear correlation to the geochemical processes observed.

**Spirochaetes.** *Sphaerochaeta pleomorpha* strain Grapes and *Sphaerochaeta globosa* strain Buddy were obtained during enriching trichloroethene (TCE)-dehalo-respiring anaerobes (Caro-Quintero et al. 2012, Ritalahti et al. 2012). They are members of the phylum Spirochaetes based on 16S rRNA gene sequences and whole-genome level phylogenies, but they never display a helical morphology or motility during laboratory cultivation. This was further confirmed by genome sequencing showing the lack of genes encoding a flagellar apparatus (Caro-Quintero et al. 2012, Ritalahti et al. 2012). To date, the roles of these organisms in this enrichment have remained enigmatic. In addition, Spirochaetes are also abundant in microbial communities degrading other hydrocarbon compounds such as terephthalate (Nobu et al. 2015), trichloroethene (Caro-Quintero et al. 2012, Ritalahti et al. 2012), and other organohalides (Duhamel and Edwards 2006). But the function of the Spirochaetes in these cultures has remained unresolved.

**Epsilon-proteobacteria.** Epsilon*-proteobacteria* are frequently found in hydrocarbon-degrading, sulfate-reducing enrichment cultures (Lueders 2017). For example, they were detected in a benzene-degrading, sulfate-reducing microbial community but not supposed to be essential for benzene degradation (Keller et al. 2015, Starke et al. 2016). Metagenomics, metaproteomics and pulsed $^{13}$C$_2$-acetate protein-stable isotope labelling suggested that they were responsible for acetate
capture and polysulfide cycling (Keller et al. 2015, Starke et al. 2016). The oxidation of sulfide could avoid the toxicity of polysulfide to other microorganisms.

**Parcubacteria (OD1).** Members belong to uncultivated Parcubacteria (OD1) were found in a methanogenic benzene-degrading enrichment culture but their abundance was not directly linked to benzene consumption (Luo et al. 2016). A similar role in sulfur cycling as for Epsilon-proteobacteria was proposed for OD1, in agreement with usage of ferrous sulfide in the medium (Luo et al. 2016). In contrast, a nearly complete genome for an organism from Parcubacteria (OD1) derived from an oil reservoir suggested that it is an obligate fermenter (Hu et al. 2016). Therefore, the function of OD1 in hydrocarbon degradation can be diverse.

**Chloroflexi and Chlorobi.** Stable isotope probing of two benzene-degrading cultures suggested that specific members of the phyla Chloroflexi (mainly Anaerolineae) and Chlorobi (mainly Ignavibacteria) were unlikely involved in direct benzene degradation, even though they were at relatively high abundances (Taubert et al. 2012, van der Zaan et al. 2012). However, they might thrive on dead biomass (Taubert et al. 2012, van der Zaan et al. 2012).

**Caldiserica (formerly candidate division OP5).** In a methanogenic reactor treating terephthalate-containing wastewater, metagenomics revealed that members of the candidate division OP5 could not degrade terephthalate or its metabolites (e.g. acetate and butyrate) (Lykidis et al. 2011). Instead, they were proposed to use CO$_2$ and H$_2$ from terephthalate fermentation for anaerobic autotrophic butyrate production. The produced butyrate would fuel other syntrophs like Syntrophus, Thermotogae and WWE1 in the system.

**1.5. The importance of microbial H$_2$ metabolism in the subsurface system**

Hydrogen is an important energy source in lithoautotrophic microbial ecosystems when organic compounds are limited or absent (Greening et al. 2016). Bagnoud et al. (2016) introduced high hydrogen concentrations to deep Opalinus Clay rock formation. Based on metaproteomics they found that it can greatly fuel sulfate reducers belonging to the family Desulfobulbaceae, which thereby fixes inorganic carbon. These generated organic carbons then become necromass, which can be
used by other microorganisms, sustaining the microbial metabolic web in this system. Hernsdorf et al. (2017) took samples from groundwater at depths of 140-250 m below the surface and they found that a majority of microorganisms in this system rely on hydrogen consumption based on the detection of genes encoding oxidative [NiFe] hydrogenases and electron-bifurcating [FeFe] hydrogenases. Like Bagnoud et al., they also proposed that the oxidation of hydrogen will contribute to carbon fixation, which is important for closing the carbon loop in this oligotrophic ecosystem.

For ecosystems rich in organic carbons, genes encoding hydrogenases also exist in large quantities and high diversities (Greening et al. 2016). Wrighton et al. (2014) used metagenomics and proteomics to understand microbial interdependencies in an aquifer, which was fed with acetate aiming to elevate carbon loads. The results show that members from Candidate Phyla Radiation (CPR) (Hug et al. 2016) were responsible for fermentation and their products e.g. hydrogen could be further consumed by other respiratory organisms in this system, e.g. sulfate reducers. This emphasizes the importance of hydrogen exchange for linking fermentation and respiratory metabolisms. In hydrocarbon–contaminated system, sulfate reducers can not only feed on a variety of organic pollutants, but also use the fermentation products such as hydrogen, acetate and fatty acids (Wasmund et al. 2017). But so far little is known about how they can affect the bioremediation process.

1.6. Using $^{13}$C isotopes to quantify anaerobic hydrocarbon degradation

One of the most important steps in understanding microbial degradation processes is to quantify the extent of biodegradation. The direct measurement of concentrations of contaminants, in most cases, is not practical. On one hand, the biodegradation rate is usually very slow and it can only lead to very small change in the concentration, most probably below the reproducibility of the measurements; on the other hand, most compounds are volatile and can strongly adsorb to sediments in situ or stoppers in laboratory microcosm studies. In this context, people turn to using the powerful stable isotope tools to reliably assess pollutant biodegradation. There are two main stable isotope tools in this perspective: (i)
analysis of natural abundance changes of stable isotopes for one specific compound based on the isotopic fractionation effects during the biodegradation process; (ii) determination of incorporation of stable isotopes to the degradation intermediates and/or mineralization products derived from a labelled compound (Figure 3). These two tools rely on (iii) the development of very precise and accurate stable isotope analytical techniques (Figure 3). Here the three aspects are summarized as below with focus on the $^{13}$C stable isotopes.

**Figure 3.** Application of $^{13}$C stable isotope tools to quantify the extent of anaerobic hydrocarbon degradation.

1.6.1 Stable isotope fractionation

In biochemical reactions, lighter isotopes (e.g. $^{12}$C and $^1$H) react faster, which results in a relative enrichment of heavier isotopes (e.g. $^{13}$C and $^2$H) in the nonbiodegraded, residual fraction of the substrates (e.g. benzene and naphthalene). The shifts in stable isotope ratios (e.g. $^{13}$C/$^{12}$C, $^2$H/$^1$H) of organic contaminants can be measured by means of compound-specific isotope analysis using mainly gas chromatography-isotope ratio mass spectrometry, and therefore can be employed for qualitative and even quantitative determination of anaerobic degradation of benzene and PAHs (Elsner 2010, Meckenstock and Richnow 2010, Schmidt and Jochmann 2012, Thauer et al. 1977). The stable isotope ratios are expressed in δ notations as δ$^{13}$C values (‰) for carbon with Vienna Pee Dee
Belemnite as the standard, and δ²⁷H values (‰) for hydrogen with Vienna Standard Mean Ocean Water as the standard (Coplen 2011).

For quantification of in situ biodegradation, the stable isotope fractionation factor α needs to be determined in the laboratory first. For field applications, the Rayleigh equation is then used to estimate the extent of contaminant biodegradation based on stable isotope ratio shifts and corresponding isotope fractionation factors (Hunkeler 2008, Meckenstock and Richnow 2010). Based on ¹³C/¹²C fractionation data, Griebler et al. (2004) calculated that decreasing benzene concentrations in a contaminated aquifer were almost exclusively due to biodegradation. Fischer et al. (2008) could quantitatively connect the stable isotope fractionation of a mixture of ring-deuterated (d₅) and completely (d₈) deuterium-labeled toluene injected into a contaminated aquifer, with the release of deuterium from toluene mineralization. However, variations in fractionation factors for different microorganisms and biochemical reactions have to be taken into account (Bergmann et al. 2011b, Fischer et al. 2008, Griebler et al. 2004, Mancini et al. 2008). Griebler et al. (2004) also obtained ¹³C/¹²C isotope fractionation factors for anaerobic degradation of naphthalene and 2-methylnaphthalene using the sulfate-reducing enrichment culture N47. However, Morasch et al. (2011) could not find significant stable isotope shifts to evidence the occurrence of methylnaphthalenes. The study indicated the limitation of carbon isotope fractionation for molecules larger than 11 carbon atoms such as 2-methylnaphthalene. The extent of observable isotope fractionation is diluted with an increasing number of carbon atoms (Elsner 2010, Kummel et al. 2015). This makes an accurate detection of biodegradation impossible using compound-specific isotope analysis for higher-molecular-weight PAHs, e.g. molecules with more than 11 carbon atoms.

Analysis of hydrogen stable isotope ratios, however, might be useful for higher-molecular-mass PAHs as the number of hydrogen atoms per molecule only slightly increases for larger PAHs and hydrogen stable isotope fractionation effects can be very strong. Bergmann et al. (2011b) demonstrated that hydrogen isotope enrichment factors were remarkably large during anaerobic degradation of naphthalene by the sulfate-reducing cultures N47 and NaphS2. This indicated that
hydrogen stable isotope fractionation might be employed to qualitatively and quantitatively assess microbial degradation of PAHs in the field.

1.6.2 Stable isotope labelling

Complete oxidation of hydrocarbons to CO$_2$ and CH$_4$ is the final goal of bioremediation leading to complete removal of contaminants without toxic intermediates. In terms of monitoring complete mineralization, one of the best ways is to use stable isotope-based methods by either fully or partially labelling the target contaminant with $^{13}$C. If the compound is biodegraded, it will be converted to $^{13}$CO$_2$ or $^{13}$CH$_4$. Changes in $^{13}$C/$^{12}$C carbon isotope ratios therefore indicate the mineralization rate of the target compound. This method is quite useful to assess the biodegradation potential in laboratory microcosms of enrichment cultures or pure isolates. Besides the data from evolution profiles of $^{13}$CO$_2$ or $^{13}$CH$_4$, further evidence should be provided to demonstrate a reliable mineralization of pollutants such as the consumption of electron acceptors (e.g. sulfate or nitrate) and the accumulation of biomass (Bolliger et al. 1999).

Stable isotope labelling has been successfully used to assess biodegradation of recalcitrant organic compounds such as benzene and polycyclic aromatic hydrocarbon (Fischer et al. 2016). For example, Nijenhuis et al. (2007) sufficiently proofed anaerobic mineralization of $^{13}$C$_6$-monochlorobenzene in laboratory microcosms indicated by the production of $^{13}$CO$_2$ even at very small portions of degraded contaminant. But with stable isotope fractionation analysis, they failed to provide clear evidence demonstrating the high sensitivity of stable isotope labelling tools. Morasch et al. (2007) and Fischer et al. (2016) followed the production of $^{13}$CO$_2$ from $^{13}$C-labelled polycyclic aromatic hydrocarbons (e.g. naphthalene) by isotope-ratio mass spectrometry (IRMS) providing in situ biodegradation evidence via the high sensitivity and reproducibility of $^{13}$C/$^{12}$C stable isotope analysis.

However, this method has at least two disadvantages: (1) usually in field experiments people have to deal with a mixture of compounds which is impossible to be labelled; (2) this method is less applicable for higher-molecular compounds
because the $^{13}$C-labelled contaminants usually need special syntheses; and thus, they are very expensive or even not commercially available.

### 1.6.3 Stable isotope techniques

Stable isotope methods rely on highly accurate and precise techniques for determination of $^{13}$C/$^{12}$C carbon isotope ratios. Isotope ratio mass spectrometry (IRMS) is the method of choice for highly accurate determinations (Zanasi et al. 2006, Zare et al. 2009). But the IRMS instrument requires a relatively large space in a temperature-controlled laboratory and professional experts for operation and maintenance. Besides, they are costly and not portable. Recently, the development of optical-based technologies has provided more options for the determination of carbon isotope ratios in e.g. CH$_4$ and CO$_2$. The most attractive one is referred to as isotope ratio infrared spectroscopy (IRIS) relying on infrared laser absorption. These instruments are smaller in size and less expensive. Moreover, they can be installed in the field allowing for high-resolution and real-time in situ measurements which IRMS cannot perform (van Geldern et al. 2014).

![Near-infrared spectrum of 5% carbon dioxide at 6.7 kPa operating pressure, showing three different isotopes. Taken from Wahl et al. (2006).](image)

**Figure 4.** Near-infrared spectrum of 5% carbon dioxide at 6.7 kPa operating pressure, showing three different isotopes. Taken from Wahl et al. (2006).

Different molecules have different absorption lines at specific wavelengths due to the quantum mechanical rotational and vibrational states. This principle is also
applied to the same molecule but with different isotopic species. For example, if $^{12}$C in CO$_2$ is replaced by $^{13}$C, the absorption wavelengths change (Crosson et al. 2002, Wahl et al. 2006) (Figure 4). Therefore, shifts in different absorption peaks allow to quantify the respective abundances of different isotopologues following Beer’s law. As we can see from Figure 4, the spectral features are quite narrow and therefore special techniques are needed to distinguish between closely spaced wavelengths of light. The method used to achieve this goal is called tunable diode laser absorption using the continuous wave for such a high selectivity (Bowling et al. 2003, Fischer et al. 2016). Meanwhile this technique uses multi-pass cells to have an enhanced path length for high resolution (Bowling et al. 2003, Fischer et al. 2016).

So far, the most frequently used commercial IRIS instruments are Cavity ring-down spectroscopy (CRDS) (Crosson et al. 2002) and Delta Ray (van Geldern et al. 2014) based on tunable laser direct absorption in the near-infrared and mid-infrared wavelength regions, respectively. These instruments have a comparable accuracy for determination of $^{13}$C/$^{12}$C in CO$_2$ at natural abundance to conventional IRMS (Crosson et al. 2002, Flores et al. 2017). Moreover, they can be used to roughly estimate CO$_2$ concentrations (Bowling et al. 2003). IRMS has been shown that it is not suitable for microbial labeling experiments as highly enriched samples cannot be measured (Reinsch and Ambus 2013). In contrast, IRIS instruments can be applied for labelled samples as long as the concentration of e.g. CO$_2$ is not high enough. But compared to IRMS, so far, the application of IRIS in compound-specific isotope analysis (CSIA) is still at its infancy. One pioneering work done by Zare et al. (2009) integrated a CRDS instrument, a chromatographic separation technique, and a combustor into one system (GC-C-CRDS). With this setting, they successfully measured the $^{13}$C/$^{12}$C isotope ratios in organic compounds, such as ethane and propane.

1.7. Objectives of this thesis

This thesis has two objectives. Firstly, I aimed to employ omics techniques (metagenomics and metaproteomics) and physiological experiments to understand the ecology of anaerobic hydrocarbon degradation performed by slow-growing
bacteria. Secondly, I planned to develop a novel reverse stable isotope labelling method for monitoring microbial mineralization rates of organic compounds. In detail, they are divided into three parts in this thesis as below.

**Part I** Reconstructing metabolic pathways of a member of the genus *Pelotomaculum* suggesting its potential to oxidize benzene to carbon dioxide with direct reduction of sulfate (Chapters 2.1.1, 2.2.1, 2.3.1, 2.4.1, 2.6, 3.1 and 4.1)

Genomic information, usually combined with transcriptomics or proteomics, has the potential to directly and comprehensively unravel the metabolic potential and physiology of cultures that have the capability for anaerobic mineralization of benzene. The enrichment culture BPL used in this part is able to degrade benzene under sulfate-reducing conditions (Abu Laban et al. 2009). Previous community analysis based on 16S rRNA gene sequencing and Fluorescent in situ hybridization (FISH) revealed the dominant member (~95%) in this community was the Gram positive bacterial genus *Pelotomaculum*. However, members of the genus *Pelotomaculum* are mostly obligate syntrophs that need methanogens as syntrophic partners to oxidize propionate, alcohols, or aromatic compounds (Imachi et al. 2002, Stams and Plugge 2009). The aim of Part I is to understand who is responsible for benzene degradation and sulfate reduction using a combination of genomic and proteomics.

**Part II** Fermentative Spirochaetes drive nutrient cycling by a subsurface microbial loop in hydrocarbon-contaminated habitats (Chapters 2.1.2, 2.2.2, 2.3.2, 2.3.3, 2.4.2, 2.6, 3.2 and 4.2)

Under sulfate-reducing conditions, naphthalene can be anaerobically mineralized to CO₂ by the stable enrichment culture N47 (Meckenstock et al. 2000). Based on 16S rRNA gene sequence analysis, ~93% of the culture comprises a sulfate-reducing deltaproteobacterium that has been shown genetically and biochemically to be responsible for naphthalene degradation. The remainder of the enrichment can be assigned to representatives of the phylum Spirochaetes (Bergmann et al. 2011a, Selesi et al. 2010). To date, the roles of these organisms in this enrichment have remained enigmatic. The aim of Part II is to gain insights into its function in
this enrichment using combination of physiological, genomic, and proteomic analyses.

**Part III Monitoring microbial mineralization using reverse stable isotope labelling analysis by mid-infrared laser spectroscopy (Chapters 2.1.3, 2.5, 2.6, 2.7, 2.8, 3.3 and 4.3)**

A crucial step in studying microbial degradation of organic compounds is using microcosm experiments to prove the degradation ability and to obtain degradation rates. Of special interest here is to assess if the compounds are totally mineralized to CO₂ or if they are just transformed to metabolites which might be even more toxic in some cases (Fischer et al. 2016). A preferred method is to prove mineralization of the compound of interest by the development of labelled CO₂ in airtight containers. To this end, target compounds are either labelled with stable (¹³C) or radioactive (¹⁴C) carbon isotopes and the evolution of ¹³CO₂ or ¹⁴CO₂ is measured over time, providing a very sensitive way for detection of low substrate turnover compared to residual substrate concentration measurements (Bahr et al. 2015, Berry et al. 2017, Fischer et al. 2016, Johnsen et al. 2013). Compared to radioactive labelling, stable isotope methods are a preferred alternative in recent years, because the compounds are easier to handle and less expensive. However, this method is less applicable for higher-molecular weight compounds, because the ¹³C-labelled substrates usually require special synthesis, and are thus very expensive or even not commercially available. Furthermore, complex substrates such as dissolved organic carbon (DOC) or crude oil cannot be synthesized or labelled. To address the above-mentioned issue, in Part III I aimed to develop a new method for quantitatively assessing biodegradation of organic compounds to CO₂, namely reverse stable isotope labelling (RIL).
2. Materials and Methods

2.1. Cultivation of anaerobic cultures

2.1.1 Cultivation of the enrichment culture BPL

The anaerobic sulfate-reducing culture BPL was enriched from soil at a former coal gasification site in Gliwice, Poland, and was cultivated in bicarbonate-buffered (30 mM) freshwater medium as described earlier (Abu Laban et al. 2009). Benzene (0.5 mM) (Sigma-Aldrich, Steinheim, Germany) was added as sole electron donor and sulfate (10 mM) as electron acceptor. The enrichment culture was transferred in 1:10 (vol/vol) dilutions and the metabolic activity was monitored by consumption of sulfate (Abu Laban et al. 2009).

2.1.2 Cultivation of the enrichment culture N47 and its pure isolate

The naphthalene-degrading, sulfate-reducing enrichment culture N47 was cultivated in bicarbonate-buffered freshwater medium as described previously (Selesi et al. 2010). Three mL of 1.5% naphthalene dissolved in 2,2,4,4,6,8,8-heptamethylnonane (Sigma-Aldrich, Steinheim, Germany) were added to each cultivation bottle after autoclaving. The highly enriched Deltaproteobacteria N47 culture (without Spirochaetes) (Bergmann et al. 2011a, Selesi et al. 2010) performs sulfate reduction and naphthalene oxidation. It was obtained from the enrichment culture N47 by serial dilution in agar roll tubes and cultivated in the same way as the enrichment culture N47. *Rectinema cohabitans* strain HM isolated from the enrichment culture N47 (Koelschbach et al. 2017) was cultivated in bicarbonate-buffered freshwater medium (pH 7.0), which was reduced with a final concentration of 0.5 mM sodium sulfide and amended with 10 mM glucose and 0.1% yeast extract. All culture bottles were incubated at 30 °C in the dark.

To test if *R. cohabitans* can utilize dead biomass of highly enriched Deltaproteobacteria as carbon source, we first harvested cells of highly enriched Deltaproteobacteria from 150 ml cultures by centrifugation at 3214 × g (30 min), suspended them in 2.5 mL anoxic fresh medium (without substrate and yeast extract), and autoclaved them in a closed serum bottle three times at 121 °C for
20 min. We then transferred 0.5 mL dead biomass of highly enriched Deltaproteobacteria to serum bottles filled with 30 ml fresh medium and inoculated 3 mL starter cultures of *R. cohabitans*. For control experiments, we incubated fresh media with the same amount of starter cultures of *R. cohabitans* without biomass of highly enriched Deltaproteobacteria or dead biomass without the inoculation of *R. cohabitans*. To investigate if *R. cohabitans* can gain carbon sources from the supernatant of highly enriched Deltaproteobacteria cultures, we transferred 3 mL starter cultures of *R. cohabitans* to serum bottles filled with 30 ml centrifuged supernatant of highly enriched Deltaproteobacteria cultures, which was in addition filtered through 0.22 μm pore-size filter membrane (Sarstedt, Germany). For control experiments, we incubated 30 ml filtered supernatant of highly enriched Deltaproteobacteria without *R. cohabitans*.

### 2.1.3 Cultivation of bacterial cultures for RIL experiments

*Thauera aromatica* strain K172 (DSM 6984) and *Geobacter metallireducens* strain GS-15 (ATCC 53774/DSM 7210) were purchased from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Braunschweig, Germany). Both strains were cultivated as described before under anoxic conditions in a carbonate-buffered freshwater mineral medium at pH of 7.2 to 7.4 spiked with ~0.8 mM pure toluene (Sigma-Aldrich, Steinheim, Germany) (Widdel and Bak 1992) and amended with 5 mM nitrate or 50 mM ferric citrate as electron acceptors, respectively. The sulfate-reducing enrichment culture N47 was cultivated as described in Chapter 2.1.2 but with addition of 1 mL of 1.5% 2-methyl-naphthalene or 3 mL of 1.5% naphthalene dissolved in 2,2,4,4,6,8,8-heptamethylnonane (Sigma-Aldrich, Steinheim, Germany) to each cultivation bottle after autoclaving.

The uncharacterized iron-reducing, naphthalene-degrading enrichment culture SN was enriched from Pitch Lake in Trinidad and Tobago (Meckenstock et al. 2014) with ferrihydrite as electron acceptor according to previously described methods (Kleemann and Meckenstock 2011). It was transferred to freshwater mineral medium every 3-4 months when significant production of ferrous iron was indicated by the presence of black iron oxides. The culture was maintained in the lab for more than three years before our experiment, leading to a stable and
sediment-free, iron-reducing culture (Figure 5). For the RIL experiment, 3 mL of 1.5% naphthalene dissolved in 2,2,4,4,6,8,8-heptamethylnonane were added to each cultivation bottle after autoclaving.

For every culture, RIL experiments were performed using five separate 250 mL serum bottles first filled with 135 mL of fresh medium without bicarbonate buffer and flushed with CO2/N2 (20:80, v/v). The NaH13CO3 (Sigma-Aldrich Co.) and NaH12CO3 (Sigma-Aldrich Co.) were injected separately through the stopper from filter-sterilized stock solutions (1 M each, headspace briefly flushed with N2) to produce a mixture of 30 mM bicarbonate in the medium. Considering the linearity range of IRMS/IRIS, we shifted the added labelled bicarbonate ratios from 50 to 10%. Afterwards, substrates stated before were injected through the stoppers. Then, three bottles were inoculated with 15 mL parent cultures whereas the other two were inoculated with triple-autoclaved cultures serving as controls. All culture bottles were incubated at 30 °C in the dark.

Figure 5. Epifluorescence microscopy images of naphthalene-degrading, iron-reducing enrichment cultures after six months' cultivation stained with 4', 6-diamidino-2-phenylindole (DAPI).
2.2. Metagenomic and genomic analyses

2.2.1 Binning draft genomes from the BPL metagenome

Genomic DNA was extracted from a 1200-mL BPL culture when ~2 mM sulfate was reduced using the modified CTAB (hexadecyltrimethylammonium bromide) method (Zhou et al. 1996). Whole metagenome sequencing was performed with a 454 GS FLX Titanium system (Roche, Penzberg, Germany) in two different runs. The resulting reads were co-assembled into contigs with the Newbler v2.7 assembly software. The automated assembly of the sequences was checked manually by mapping the raw 454 reads against the contigs with Mosaik v2.2.3 (Lee et al. 2014) followed by visual inspection with Tablet v1.14 (Milne et al. 2013). The assembled contigs were submitted to IMG-M 4.530 (Project ID: Gp0111374) (Markowitz et al. 2014) for automated gene calling and annotation. Automatic annotations were manually curated for genes described in this study by bioinformatics tools in the IMG platform (e.g. KEGG and BLAST) (Markowitz et al. 2014).

All contigs from the BPL metagenome were annotated taxonomically without length restriction with taxator-tk v1.3.0e using default settings and the refpack (microbial_20150430) provided by the authors (Droge et al. 2015). Additionally, we screened all contigs for (partial) marker genes (5S, 16S, 18S, and thirty-one conserved genes) (Wu and Scott 2012) and classified them taxonomically, both using PhyloPythiaS+ v1.4 (Gregor et al. 2016). The two assignment methods provided similar results. Then, we validated the results by checking i. their consistency with the taxonomic profiles obtained from both independent 16S rRNA gene amplicon sequencing data and the contigs-extracted SSU (16S) sequences data for the same community, ii. the read coverage and iii. the GC content of contigs assigned to the same clades.

Finally, we identified the major taxa based on the taxonomic annotations and used them to extract sample-specific training data for the training stage of the composition-based taxonomic classifier PhyloPythiaS+ using default settings. By considering the taxonomic annotation, GC content, and average read coverage, we divided the metagenome contigs into four genome bins.
2.2.2 Sequencing and reconstruction of Spirochaetes genomes

For comparative genomic analyses, three draft genomes were used: *R. cohabitans*, uncultured Spirochaetes bacterium bdmA 4 and uncultured Spirochaetes bacterium SA-8. The draft genome of uncultured Spirochaetes bacterium SA-8 was derived from the metagenome of a terephthalate-degrading community TA, which is available on IMG/M as taxon object ID 3300001095 (Nobu et al. 2015). The draft genome of uncultured Spirochaetes bacterium bdmA 4 was reconstructed from the metagenome of another sulfate-reducing, naphthalene-degrading enrichment culture (designated Sob). Culture Sob was enriched from a creosote-contaminated wood preservation facility located in Soběslav, south of Bohemia (Czech Republic), as described by Kummel et al (2015). The shotgun sequencing information and followed-up assembly and binning process for the (meta)genomes for *R. cohabitans* and enrichment culture Sob are described as below.

Construction and sequencing of paired-end DNA libraries. DNA (30 to 250 ng) was sonicated to a 100- to 800-bp size range on the E210 Covaris instrument (Covaris, Inc., USA). Fragments were end-repaired, then 3'-adenylated, and Illumina adapters were added by using a NEBNext Sample Reagent Set (New England Biolabs). Ligation products were purified by Ampure XP (Beckmann Coulter) and DNA fragments (>200 bp) were PCR-amplified using Illumina adapter-specific primers and Platinum Pfx DNA polymerase (Invitrogen). The amplified library fragments were size selected on 3% agarose gel at around 300 bp. After library profile analysis by Agilent 2100 Bioanalyzer (Agilent Technologies, USA) and qPCR quantification (MxPro, Agilent Technologies, USA), the library was sequenced using 101 base-length read chemistry in a paired-end flow cell V3 on the Illumina Hiseq2000 sequencer (RTA version 1.13.48) (Illumina, USA) in order to obtain overlapping reads and generate meta-reads of 180 bp.

Nextera mate-paired library preparation and sequencing. Three mate pair libraries were prepared for each sample following the Nextera protocol (Nextera Mate Pair sample preparation kit, Illumina). Briefly, genomic DNA was simultaneously enzymatically fragmented and tagged with a biotinylated adaptor. Fragments were
size selected (3-5 Kb, 5-8 Kb and 8-11Kb) through regular gel electrophoresis, and circularized overnight with a ligase. Linear, non-circularized fragments were digested, and circularized DNA was fragmented to 300-1000-bp size range using Covaris E210. Biotinylated DNA was immobilized on streptavidin beads, end-repaired, then 3’-adenylated, and Illumina adapters were added. DNA fragments were PCR-amplified using Illumina adapter-specific primers and then purified. Finally, libraries were quantified by qPCR and libraries profiles were evaluated using an Agilent 2100 bioanalyzer (Agilent Technologies, USA). For enrichment culture Sob, 3-5kb, 5-8kb and 8-11Kb libraries were sequenced using 101 base length read chemistry (v1) in a paired-end flow cell on the Illumina HISEQ2500 sequencer (rapid mode, RTA version 1.17.21.3) (Illumina, USA). For *R. cohabitans*, the 8-11kb library was sequenced using 151 base-length read chemistry (v1) on a paired-end flow cell on the Illumina Miseq platform (RTA version 1.18.54) (Illumina, USA).

**Assembly of the *R. cohabitans* genome.** Reads from the paired-end library were down-sampled before being combined with sequences from the largest (8-11 kbp) mate-paired library and assembled with the ALLPATHS engine (Ribeiro et al. 2012).

**Reconstruction of the genome for uncultured Spirochaetes bacterium bdmA 4.** Although the enrichment culture Sob was dominated by sulfate-reducing Deltaproteobacteria N47 (Bergmann et al. 2011a, Selesi et al. 2010), with about 30% of the raw reads (from either the paired-end or mate-paired libraries) mapping to its genome, the metagenome was endowed with significant phylogenetic diversity and was estimated to harbor about 30-50 distinct genomes. A combination of assembly-free (binning) and targeted assembly techniques were used in order to direct assembly efforts towards lower complexity read partitions. This was achieved by using sequence clustering methods extracting coverage and/or compositional signals from the reads to partition the sequences in k-mer space in an overlap and alignment independent way (Gkanogiannis et al. 2016). The resulting partitions were then assembled independently, leading to the reconstruction of several dozens of megabase sized genomic fragments, including the nearly complete genome for uncultured Spirochaetes bacterium bdmA 4.
(estimated 97% complete based on the distribution of 141 lineage specific marker genes compiled by the checkM software) (Parks et al. 2015).

The draft genomes of *R. cohabitans*, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8 were uploaded into the MicroScope genome annotation platform and automatically annotated (Vallenet et al. 2013). The annotations were manually curated and edited using the Magnifying Genome (MaGe) environment (Vallenet et al. 2006), and metabolic pathways were predicted using the integrated pathway tools of MaGe that are based on the KEGG and MicroCyc databases. Presence of glycoside hydrolases and extracellular peptidases was evaluated using dbCAN (E<1e⁻¹⁰ and cover fraction >0.4) (Yin et al. 2012) and MEROPS (E<1e⁻¹⁰) (Rawlings et al. 2016), respectively.

2.3. Phylogenetic analysis

2.3.1 16S rRNA gene analysis of *Pelotomaculum*

We picked the near-complete *Pelotomaculum* candidate BPL 16S rRNA gene, 16S RefSeq sequences for the closely related genera *Pelotomaculum* and *Desulfotomaculum*, every *Desulfotomaculum* sequence that was deposited in SILVA SSU r123 as well as an outgroup sequence *Thermincola potens* JR, which we used later to root the tree. We aligned those sequences to the SILVA reference alignment using SINA v1.2.12 and trimmed the resulting multiple sequence alignment to SILVA positions 1007 to 43293. Finally, a maximum-likelihood tree was calculated with RAxML v.8.2.8 using the combined rapid bootstrapping (1000 replicates) and maximum likelihood search algorithm with GTRCAT approximation model and with random seed 29582.

2.3.2 16S rRNA gene analysis of *Spirochaetes*

The phylogeny of the *Spirochaetes* was inferred using 16S rRNA gene sequences retrieved from the NCBI Genome Database. Sequences were aligned using the MUSCLE algorithm included in MEGA7 and trimmed to uniform length (Kumar et al. 2016). A maximum likelihood phylogenetic tree was computed by MEGA7 using a general time reversible substitution model and uniform rates among sites. The tree was bootstrapped with 500 replicates.
2.3.3 Hydrogenase classification and analysis

Potential genes encoding hydrogenase catalytic subunits were identified in the genomes of *R. cohabitans* and sulfate-reducing Deltaproteobacteria N47 (Bergmann et al. 2011a, Selesi et al. 2010) by screening translated sequences through the web database HydDB (Søndergaard et al. 2016). Sequences were validated as hydrogenases if they encoded cysteine-containing motifs characteristic of hydrogenases (Greening et al. 2016) and branched with hydrogenase reference sequences on phylogenetic trees. To infer evolutionary relationships, the source and reference sequences were aligned using the MUSCLE algorithm (Edgar 2004) and visualized on neighbor-joining phylogenetic trees (Saitou and Nei 1987) constructed with MEGA7 (Kumar et al. 2016) and bootstrapped with 500 replicates. For genetic organization maps, protein/domain function was predicted using the Conserved Domain Database (CDD) (Marchler-Bauer et al. 2011) and iron-sulfur clusters were predicted by identifying conserved cysteine residues as previously described (Greening et al. 2016).

2.4. Proteomic analysis

2.4.1 Shotgun proteomic analysis of the enrichment culture BPL

For proteomics, two independent experiments were performed in a similar way as follows. Extraction of proteins from bacterial cultures was performed as previously described (Abu Laban et al. 2010, Selesi et al. 2010). Comparative proteomic analysis is impossible here, because the culture cannot grow with any of the possible down-stream metabolites like toluene, benzoate, or phenol. Furthermore, it is an enrichment culture, and easily degradable substrates like butyrate and acetate will change the compositions of the microbial community. Thus we performed a qualitative proteomics experiment with benzene as substrate. Cells were harvested from a total of 1200 ml culture by centrifugation (30 min at 3,739 × g and 4°C). The cell pellet was washed three times with 1× phosphate buffered saline (PBS buffer), and resuspended in a mixture of 400 μl lysis buffer, including 9 M urea, 2 M thiourea, 4% 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS), and 1% dithiothreitol (DTT), and 66.8 μl of a 7× stock solution containing a complete EDTA-free mini-protease inhibitor cocktail tablet.
(Roche Diagnostics GmbH, Penzberg, Germany). After 30 min of incubation at room temperature, the cell-buffer mixture was subjected to sonication in an ice-water bath for two minutes (amplitude: 0.3, cycle: 60%; UP50H, Hielscher GmbH, Germany). The homogenized solution was centrifuged for 10 min at 20,000 × g and 4°C. The supernatant was treated for 30 min at room temperature with 4 μl nuclease mix (GE Healthcare) and centrifuged for 30 min at 15,000 × g and 4°C. Estimation of protein concentrations in the supernatant was performed with the two-dimensional Quant kit according to the protocol of the manufacturer (GE Healthcare). However, quantification of protein was affected by interference with ferrous sulfide, prohibiting a proper estimation of protein concentrations.

For sample preparation, a modified filter-aided sample preparation (FASP) approach was performed (Wisniewski et al. 2009). After tryptic digestion, samples were stored at -20°C until further use.

LC-MS/MS analysis was performed as described previously on a LTQ-Orbitrap XL (Thermo Fisher) (von Toerne et al. 2013). Briefly, the pre-fractionated samples were automatically injected and loaded onto the trap column (Acclaim PepMap100, C18, 5 μm, 100 Å pore size, 300 μm ID × 5 mm µ-Precolumn - No 160454; Thermo Scientific). After 5 min, the peptides were eluted and passed to the analytical column (Acclaim PepMap100, C18, 3 μm, 100 Å pore size, 75 μm ID × 15 cm, nanoViper-No 164568; Thermo Scientific) by reversed phase chromatography which was operated on a nano-HPLC (Ultimate 3000, Dionex). A nonlinear 170 min gradient was used for elution with a mobile phase of 35% acetonitrile in 0.1% formic acid in water (A) and 0.1% formic acid in 98% acetonitrile (B) at a flow rate of 300 nL/min. The gradient settings were: 5-140 min: 14.5-90% A, 140-145 min: 90% A-95% B, 145-150 min: 95% B followed by equilibration for 15 min to starting conditions. The 10 most abundant peptide ions were selected from the MS pre-scan for fragmentation in the linear ion trap if they exceeded an intensity of at least 200 counts and were at least doubly charged. A high-resolution (60,000 full-width half maximum) MS spectrum was acquired in the Orbitrap with a mass range from 200 to 1500 Da during fragment analysis.
MS-MS spectra were searched against the BPL metagenome database (20979 entries) via the MASCOT search engine (version 2.5.1; Matrix Science). A mass tolerance of 10 ppm for peptide precursors and 0.6 Da for MS-MS peptide fragments was applied allowing no more than one missed cleavage. Fixed modifications were set to carbamidomethylation of cysteine and variable modifications to oxidation of methionine and deamidation of asparagine and glutamine.

The Scaffold software (version 4.4.1.1., Proteome Software Inc.) was used to validate peptide identifications and visualize results (Keller et al. 2002). Peptide threshold was set to a false discovery rate (FDR) of 1 percent. A decoy protein FDR of 2.2% was observed. Protein identifications were accepted if two or more unique peptides were identified.

2.4.2 Mass spectrometry-based proteome analyses of N47

For proteomics, the enrichment culture N47 and the culture of high enriched Deltaproteobacteria N47 were grown with naphthalene (four parallel cultures each). In addition, four cultures of R. cohabitans strain HM were grown with 10 mM glucose and 0.1% yeast extract. The cultures were harvested in the exponential growth phase and immediately frozen at -20 °C. The frozen cell pellets were warmed to room temperature and taken up in 1 mL phosphate buffered saline (PBS). The samples were lysed by ultrasonification in a Bioruptor® (Diagenode) for 5 min to break the cells, release the proteome and shear the genomic DNA. The turbid solution was cleared by centrifugation (12,000 × g, 4 °C, 5 min). The supernatant was transferred to a fresh 1.5 mL Eppendorf tube. The protein concentration was determined by using a modified Bradford assay (RotiNanoquant; Roth) following the manufacturer's instructions.

Reduction/alkylation and tryptic digestion. An aliquot corresponding to 25 µg total protein content was removed from each sample. Sodium dodecyl sulfate (SDS) and dithiothreitol were added to a final concentration of 2% and 5 mM respectively. The protein solution was incubated at 90 °C for 5 min. After the samples had cooled down to room temperature, 20 mM iodoacetamide was added and the
samples were incubated for 30 min at ambient temperature. A chloroform/methanol precipitation was performed in order to remove the reducing and alkylating agents and to remove the SDS (Wessel and Flugge 1984). The obtained protein pellet was then dissolved in 25 µL of 8M urea and 100 mM ammonium bicarbonate and extensively vortexed. The solutions were cleared by centrifugation and the supernatants transferred to 96 well plates (Eppendorf). To start the protein digestion we added 833 ng Lys-C (1:30; Wako Laboratory Chemicals) and incubated the samples for 3 h at 37 °C. Next the samples were diluted to 25 mM ammonium bicarbonate, supplemented with 1.0 M urea, and 1.25 µg sequencing grade Trypsin (1/20; Promega), and incubated overnight at 37 °C with shaking. On the next morning, the samples were acidified by adding formic acid (final 0.5% v/v).

**Sample clean-up for LC-MS.** Acidified tryptic digests were desalted on custom-made C18 StageTips as described (Rappsilber et al. 2007). Approximately 15 µg of peptides (based on the initial protein concentration) were loaded to each to two disc StageTip. After elution from the StageTips, samples were dried using a vacuum concentrator (Eppendorf) and the peptides were taken up in 10 µL 0.1 % formic acid solution.

**LC-MS/MS.** Experiments were performed on an Orbitrap Elite instrument (Thermo) (Michalski et al. 2012) that was coupled to an EASY-nLC 1000 liquid chromatography (LC) system (Thermo). The LC was operated in the one-column mode. The analytical column was a fused silica capillary (75 µm × 30 cm) with an integrated PicoFrit emitter (New Objective) packed in-house with Reprosil-Pur 120 C18-AQ 1.9 µm resin (Dr. Maisch). The analytical column was encased by a column oven (Sonation) and attached to a nanospray flex ion source (Thermo). The column oven temperature was adjusted to 45 °C during data acquisition and in all other modi at 30 °C. The LC was equipped with two mobile phases: solvent A (0.1% formic acid, FA, in water) and solvent B (0.1% FA in acetonitrile, ACN). All solvents were of UHPLC (ultra-high performance liquid chromatography) grade (Sigma). Peptides were directly loaded onto the analytical column with a maximum flow rate that would not exceed the set pressure limit of 980 bar (usually around 0.5 – 0.8 µL/min). Peptides were subsequently separated on the analytical column.
by running a 300-min gradient of solvent A and solvent B (start with 7% B; gradient 7% to 35% B for 280 min; gradient 35% to 100% B for 10 min and 100% B for 10 min) at a flow rate of 300 nl/min. The mass spectrometer was operated using Xcalibur software (version 2.2 SP1.48). The mass spectrometer was set in the positive ion mode. Precursor ion scanning was performed in the Orbitrap analyzer (FTMS; Fourier Transform Mass Spectrometry) in the scan range of m/z 300-1800 and at a resolution of 60,000 with the internal lock mass option turned on (lock mass was 445.120025 m/z, polysiloxane) (Olsen et al. 2005). Product ion spectra were recorded in a data dependent fashion in the ion trap (ITMS; Ion Trap Mass Spectrometry) in a variable scan range and at a rapid scan rate. The ionization potential (spray voltage) was set to 1.8 kV. Peptides were analyzed using a repeating cycle consisting of a full precursor ion scan (1.0 × 10^6 ions or 30 ms) followed by 15 product ion scans (1.0 × 10^4 ions or 50 ms) where peptides are isolated based on their intensity in the full survey scan (threshold of 500 counts) for tandem mass spectrum (MS2) generation that permits peptide sequencing and identification. CID (collision-induced dissociation) collision energy was set to 35% for the generation of MS2 spectra. During MS2 data acquisition dynamic ion exclusion was set to 60 seconds with a maximum list of excluded ions consisting of 500 members and a repeat count of one. Ion injection time prediction, preview mode for the FTMS, monoisotopic precursor selection and charge state screening were enabled. Only charge states higher than 1 were considered for fragmentation.

Peptide and Protein Identification using MaxQuant. RAW spectra were submitted to an Andromeda (Cox et al. 2011) search in MaxQuant (version 1.5.0.25) using the default settings (Cox and Mann 2008). Label-free quantification and match-between-runs was activated (Cox et al. 2014). MS/MS spectra data were searched against the in-house generated protein sequences of Deltaproteobacteria N47 (5297 entries) (Bergmann et al. 2011a, Selesi et al. 2010) and Rectinema cohabilitans HM (2580 entries). All searches included a contaminants database (as implemented in MaxQuant, 267 sequences). The contaminants database contains known MS contaminants and was included to estimate the level of contamination. Andromeda searches allowed oxidation of methionine residues (16
Da), acetylation of protein N-terminus (42 Da), Carbamyla
tion (K) and Deamidation (NQ) as dynamic modification and the static modification of
cysteine (57 Da, alkylation with iodoacetamide). Enzyme specificity was set to “Trypsin/P”.
The instrument type in Andromeda searches was set to Orbitrap and the precursor
mass tolerance was set to ±20 ppm (first search) and ±4.5 ppm (main search). The
MS/MS match tolerance was set to ±0.5 Da. The peptide spectrum match FDR
and the protein FDR were set to 0.01 (based on target-decoy approach). Minimum
peptide length was seven amino acids. For protein quantification, unique and razor
peptides were allowed. Modified peptides were allowed for quantification. The
minimum score for modified peptides was 40.

Data Analysis: Initial data analysis was performed by using the PERSEUS
computational platform (version 1.5.5.3) (Tyanova et al. 2016). The mass
spectrometry proteomics data have been deposited to the ProteomeXchange
Consortium via the PRIDE (Vizcaino et al. 2016) partner repository
(https://www.ebi.ac.uk/pride/archive/) with the dataset identifier PXD005624.

2.5. Spatial distribution profiles of microbial activities in GAC biofilms

Three parallel down-flow granular activated carbon (GAC) filters (GAC 1 – 3, 30
cm diameter, 180 cm length each) placed after an ozonation unit were used in this
study. They are parts of a pilot plant (Muelheim, Germany) treating raw water from
Ruhr River for drinking water production using a similar procedure as reported
previously (Hammes et al. 2008). The water for the pilot plant originates from the
river Ruhr. Before it enters the pilot plant it is filtered by slow sand filtration
followed by ozonation. After ozonation the water is filtered by the multilayer filter
which removes particulate impurities. Afterwards the water enters the GAC filters 1
and 2 filled with NORIT® GAC 830 as well as GAC filter 3 filled with NORIT® ROW
0.8 Supra. In the GAC filters mainly adsorption occurs. However, over time
biofilms develop on the GAC turning it biologically active leading to the bio-
degradation of organic compounds. When effluent DOC concentrations remained
relatively stable, DOC was predominately removed by biological oxidation and
GAC became biologically active carbon. At this stage, samples were taken to
investigate the spatial distribution of microbial activities (Figure 6).
Figure 6. Evolution of influent and effluent DOC concentrations of the granular activated carbon filters as a function of time.

The GAC samples were taken from various depths of GAC filters (0 cm, 20 cm, 40 cm, 80 cm and 140 cm). For RIL experiments, the GAC particles from different depths were divided into triplicate portions (3 g each), and then transferred to three separate 250 mL serum bottles filled with 50 mL influent water (after ozonation) under a laminar flow box. Large air headspace was reserved to ensure oxic conditions in the closed bottles. For each microcosm, NaH\textsubscript{13}CO\textsubscript{3} and NaH\textsubscript{12}CO\textsubscript{3} at a ratio of 10:90 were injected separately through the stopper to produce a mixture of 10 mM bicarbonate in the medium. All bottles were incubated at 17 °C in the dark to mimic the temperature conditions of GAC filters. The pH values were regularly checked and remained stable at around 7 during the whole experiments.

2.6. Analytical methods

Toluene concentrations were determined using headspace analyses by GC/MS (GC, Trace-DSQ; MS, Thermo Finnigan, San Jose, CA) in selective ion monitoring mode with a fused-silica capillary column DB-5 as previously described (Abu Laban et al. 2009). Total concentrations of toluene in the cultures were calculated from the concentrations in the liquid phase and in the headspace using the dimensionless air water partitioning constant $K_{aw}$ at 25 °C of 0.235 (Staudinger and Roberts 1996).
As we added naphthalene and 2-methynapthalene in 2,2,4,4,6,8,8-heptamethylnonane as carried phase prohibiting a reliable direct quantification of substrate concentrations (Kleemann and Meckenstock 2011), sulfate consumption was monitored using the barium-gelatin method to indirectly quantify naphthalene and 2-methynapthalene degradation (Tabatabai 2009). Iron(II) production was measured with the ferrozine assay (Stookey 2002). Cell growth was monitored by microscopically counting cell numbers.

Hydrogen was measured with a Shimadzu GC-8A Gas Chromatography according to Schuler and Conrad (1990). Briefly, aliquots of headspace air (250 µL) were firstly sampled from the headspace of the cultures after shaking by using a gas-tight syringe (VICI Precision Sampling). Then high concentrations of hydrogen (>100 ppmv) were measured with a thermal conductivity detector whereas low concentrations of hydrogen (<100 ppmv) with a reduction gas detector (RGD2 detector, Trace Analytical, Stanford, CA, USA). The measurements were calibrated with standard H₂ gases (Messer Schweiz) (2 and 50 ppmv for samples lower than 100 ppmv, 1,000 ppmv for samples higher than 100 ppmv).

2.7. Analysis of carbon stable isotope ratios

2.7.1 Sample preparation for carbon isotope ratio measurements

For sample preparation, we adopted a procedure which is used to analyze carbon isotope ratios in natural water samples by taking liquid samples and liberate CO₂ using phosphoric acid as reported previously (Assayag et al. 2006). (1) For GC-IRMS measurements, 0.5 mL liquid samples were directly taken from the culture bottles and immediately injected into 15 mL serum bottles sealed with butyl rubber stoppers. The vial had been pre-filled with 4.5 mL phosphoric acid (1M) and flushed with nitrogen or helium gas. (2) For IRIS measurements, 50 µL of 85% phosphoric acid were introduced into an uncapped 12 mL Labco Exetainer vial (Labco Limited, United Kingdom). Then, the vial was closed with screw-capped pierceable butyl rubber septa and flushed with CO₂-free synthetic air via an ASX-7100 autosampler (Teledyne CETAC Technologies, Omaha, USA) for two minutes. Afterwards, an aliquot (0.5 mL) from the culture bottle was taken and
injected into the vial through the septum. Then, we left the sample to equilibrate at stable room temperature overnight to liberate all CO₂ into the headspace.

2.7.2 Expression of carbon isotope ratios

All carbon isotope ratios (\(^{13}\text{C}/^{12}\text{C}\), referred to as \(R\)) are reported in conventional delta notation (\(\delta^{13}\text{C}\)) as per mil (‰) values, relative to the Vienna Pee Dee Belemnite (VPDB) standard (Equation 1).

\[
\delta^{13}\text{C} = \frac{R_{sample}}{R_{VPDB}} - 1
\]  

where 0.0111802 is the ratio of \(^{13}\text{C}/^{12}\text{C}\) in VPDB standard. To facilitate the comparison of results which are obtained from samples highly enriched in \(^{13}\text{C}\) compared to natural samples, all carbon isotope ratios in this study were converted into \(^{13}\text{C}\) atom fractions (\(x(^{13}\text{C})\)) (%) from \(\delta^{13}\text{C}\) values according to previous study (Coplen 2011).

2.7.3 Isotope Ratio Mass Spectrometry

CO₂ samples (100 μL) were taken from the headspace of 15 mL serum bottles with a gas-tight syringe and manually injected into a GC-IRMS system consisting of a TRACE GC Ultra gas chromatograph with split/splitless injector (GC) (both Thermo Fisher Scientific Corporation, Milan, Italy) coupled to a Finnigan MAT 253 isotope ratio mass spectrometer (IRMS) connected via a Finnigan GC combustion III interface (Thermo Fisher Scientific Corporation, Bremen, Germany). The temperature of the injector was held at 150 °C for 360 seconds. Helium 5.0 (AirLiquide, Oberhausen, Germany) was used as carrier gas with a constant flow rate of 1.6 mL min\(^{-1}\). The split flow was fixed at 160 mL min\(^{-1}\). CO₂ was separated from other gases using a 30 m Rtx-Qbond column with an inner diameter of 0.32 mm and a film thickness of 10 μm (Restek GmbH, Bad Homburg, Germany). Certified pure CO₂ gas (\(\delta^{13}\text{C} = -35.28 \text{‰}\)) (AirLiquide, Oberhausen, Germany) was used as working reference gas. All samples were measured in triplicate. When compared to isotope ratio mid-infrared spectroscopy (IRIS), a secondary standard
of $^{13}$CO$_2$ gas ($\chi(^{13}C) = 10\%$, Sigma-Aldrich, Taufkirchen, Germany) was analyzed to correct data using a two-point calibration approach.

2.7.4 Isotope Ratio Mid-Infrared Spectroscopy

Samples were analyzed using a Thermo Fisher Delta Ray CO$_2$ Isotope Ratio Infrared Spectrometer with Universal Reference Interface (URI) Connect (Di Martino et al. 2016, Fischer and Lopez 2016, Rizzo et al. 2015, van Geldern et al. 2014). IRIS is capable of measuring both carbon and oxygen isotope compositions of CO$_2$ in air. The instrument’s analyzer uses a tunable diode laser absorption (TDLA) technique operated at a mid-infrared wavelength of 4.3 μm. During analysis, the laser scans over the absorption lines for the various CO$_2$ isotopologues of the sample (i.e. $^{12}$C$^{16}$O$^{16}$O, $^{13}$C$^{16}$O$^{16}$O, and $^{12}$C$^{16}$O$^{18}$O) and fits the average signal spectrum to a reference to simultaneously quantify CO$_2$ isotope ratios. CO$_2$-free synthetic air (Air Liquide, Düsseldorf, Germany) was used as the carrier gas. CO$_2$ in synthetic air at 414.2 ppm (Air Liquide, Düsseldorf, Germany) was used for CO$_2$ concentration calibration. Before measurements, two pure CO$_2$ reference gases were used for calibration of carbon isotope ratios: one with $\delta^{13}$C values of -9.7‰ ($\chi(^{13}C) = 1.10\%$, Thermo Fisher, Bremen, Germany) and the other with $\chi(^{13}C) = 10\%$ (Sigma-Aldrich, Taufkirchen, Germany). For sample measurements, pure CO$_2$ gas with $\chi(^{13}C) = 10\%$ was used as working reference gas and samples were bracketed with working reference gas. The CO$_2$ concentration for the working reference gas and the sample gas coming into the analyzer was set to 380 ppm to achieve optimal precision. The $\delta^{13}$C values were reported as average of five minutes’ measurements. The long measurement time under constant gas flows implies that IRIS needs a relatively larger gas sample size compared to IRMS.

2.8. Calculations of CO$_2$ production using carbon isotope ratios

The conversions of $\chi(^{13}C)$ changes to CO$_2$ production were based on isotope mass balance equations as follows.
(1). CO₂ production under anoxic conditions was calculated following two mass balance equations:

\[ CO_{2\text{total}} = CO_{2\text{background}} + CO_{2\text{produced}} \]  
\[ CO_{2\text{total}} \times x^{(13)C}_{\text{total}} = CO_{2\text{background}} \times x^{(13)C}_{\text{background}} + CO_{2\text{produced}} \times x^{(13)C}_{\text{produced}} \]

where \( CO_{2\text{total}} \) is the final amount of total inorganic carbon in the system (mmole), \( CO_{2\text{background}} \) is the initial total inorganic carbon in the system (mmole), \( CO_{2\text{produced}} \) is the total released CO₂ from microbial mineralization of aromatic compounds (mmole), \( x^{(13)C}_{\text{total}} \) is the final carbon isotope ratio of total inorganic carbon in the system, \( x^{(13)C}_{\text{background}} \) is the initial carbon isotope ratio of total inorganic carbon in the system, \( x^{(13)C}_{\text{produced}} \) is the carbon isotope ratio of CO₂ released from microbial production, assumed to be 1.1%. \( CO_{2\text{background}} \) can be calculated from the initial total DIC content in the liquid (\( DIC_{\text{measured}} \)) and the volume of 20% CO₂ gas in the headspace (confirmed by measurements, 0.2 \( n_{\text{gas,headspace}} \)):

\[ CO_{2\text{background}} = DIC_{\text{measured}} + 0.2 \times n_{\text{gas,headspace}} \]  

Thus, the total CO₂ production can be obtained via Equation 5 which is based on the combination of Equations 2 - 4:

\[ CO_{2\text{produced}} = \frac{(DIC_{\text{measured}}+0.2 \times n_{\text{gas,headspace}}) \times (x^{(13)C}_{\text{background}}-x^{(13)C}_{\text{total}})}{x^{(13)C}_{\text{total}}-x^{(13)C}_{\text{produced}}} \]

For data interpretation, we evaluated isotope ratios of CO₂ from the liquid phase without further correction, since equilibrium carbon isotope fractionation between the carbonate species acts on all samples in the same way and cancels out in the comparison between standard and samples. Therefore, the carbon isotope ratio of DIC can be taken as representative of the whole closed system:

\[ x^{(13)C}_{\text{total}} = x^{(13)C}_{\text{total DIC}} \]
\[ x(^{13}C)_{background} = x(^{13}C)_{background \ DIC} \] (7)

(2). \( \text{CO}_2 \) production under oxic conditions was calculated based on the same mass balance equations as under anoxic conditions (Equations 2 and 3). As negligible amounts of \( \text{CO}_2 \) exist in the headspace, \( \text{CO}_2^\text{background} \) can be calculated using different equation:

\[ \text{CO}_2^\text{background} = \text{CO}_2^\text{bicarbonate solution} + \text{CO}_2^\text{natural} \] (8)

Where \( \text{CO}_2^\text{bicarbonate solution} \) is the amount of inorganic carbon in the added bicarbonate solution (mmole) and \( \text{CO}_2^\text{natural} \) is the amount of inorganic carbon in the natural water (mmole). Thus, the total \( \text{CO}_2 \) production can be obtained via Equation 5 which is based on the combination of Equations 2, 3 and 8:

\[ \text{CO}_2^\text{produced} = \frac{(\text{CO}_2^\text{natural} + \text{CO}_2^\text{bicarbonate solution}) \times (x(^{13}C)_{background} - x(^{13}C)_\text{total})}{x(^{13}C)_\text{total} - x(^{13}C)_\text{produced}} \] (9)

\( \text{CO}_2^\text{natural} \) can be obtained via the following equation:

\[ \text{CO}_2^\text{background} \times x(^{13}C)_{background} = \text{CO}_2^\text{bicarbonate solution} \times x(^{13}C)_{bicarbonate solution} + \text{CO}_2^\text{natural} \times x(^{13}C)_\text{natural} \] (10)

\( x(^{13}C)_{bicarbonate solution} \) is the carbon isotope ratio of the added bicarbonate solution and \( x(^{13}C)_\text{natural} \) is the carbon isotope ratio of the natural water (assuming that the carbon isotope ratio of the natural water is 1.1%).
3. Results

3.1. Reconstructing metabolic pathways of a member of the genus *Pelotomaculum* suggesting its potential to oxidize benzene to carbon dioxide with direct reduction of sulfate

Benzene is a widely distributed pollutant in groundwater posing a hazardous risk to drinking water resources (Manoli and Samara 1999). Several recent microcosm studies and lab cultures have shown that benzene and other aromatic compounds can be mineralized to CO₂ under nitrate-reducing, iron-reducing, sulfate-reducing, and methanogenic conditions (Kunapuli et al. 2007, Luo et al. 2014, Musat and Widdel 2008, Ulrich and Edwards 2003, van der Zaan et al. 2012). The enrichment culture BPL used in this study is able to degrade benzene with sulfate as electron acceptor (Abu Laban et al. 2009). Previous community analysis based on 16S rRNA gene sequencing and FISH revealed that the dominant bacterial member (~95%) belongs to the phylum Firmicutes genus *Pelotomaculum* (Abu Laban et al. 2009). Members of the genus *Pelotomaculum* are mostly obligate syntrophs that need methanogens as syntrophic partners to oxidize propionate, alcohols, or aromatic compounds (Imachi et al. 2002, Stams and Plugge 2009). However, it is hypothesized here that the anaerobic benzene degradation in the enrichment culture BPL is performed as a single cell process by a novel member of the genus *Pelotomaculum* which is able to reduce sulfate as electron acceptor.

3.1.1 Phylogenetic binning and proteome

Metagenomic sequencing of enrichment culture BPL yielded an assembly of 16.9 million base pairs into 5,842 contigs with a N50 of 12.3 thousand base pairs. Gene prediction with the IMG/M pipeline produced 20,979 open reading frames with 12,736 (60.03%) being assigned to a putative functional name.

The metagenome contigs were manually divided into four genome bins: (1) *Pelotomaculum* operational taxonomic unit (OTU) (described as *Pelotomaculum* candidate BPL) with a relative abundance of 54.27%; (2) *Gracilibacter* OTU with 32.40%; (3) *Desulfomonile* OTU with 7.41%; and (4) *Ignavibacterium* OTU with 5.92%. The discrepancy for population abundance estimates for *Pelotomaculum*
candidate BPL between 454 read coverage here and PCR-based 16S rRNA gene frequency reported previously (Abu Laban et al. 2009) can be explained by experimental bias and varying 16S rRNA gene operon copy numbers. For instance, rrnDB v.4.4.4 (Nov 2015) (Stoddard et al. 2015) reports two operons for *Pelotomaculum*, only one for *Desulfomonile* and *Ignavibacterium*, and an unknown number for *Gracilibacter*. The program CheckM v1.0.3 (Parks et al. 2015) was used to determine the completeness and contamination of the four genome bins. Based on the lineage-specific single-copy marker gene analysis by CheckM, the constructed genomes for *Pelotomaculum* and *Gracilibacter* OTUs are 99% and 97% complete, respectively (Figure 7). The genomes for *Desulfomonile* and *Ignavibacterium* OTUs could not be recovered in full length (Figure 7).

![Figure 7](image-url)

**Figure 7.** Identification of marker genes within recovered bins with CheckM. Each bar represents a marker gene. Bars in green represent markers identified exactly once, while bars in grey represent missing markers. Markers identified multiple times in the genome are represented by shades of blue or red depending on the AAI between pairs of multi-copy genes and the total number of copies present (2-5+). Pairs of multi-copy genes with an AAI ≥90% are indicated with shades of blue, while genes with less amino acid similarity are shown in red. A gene present 3 or more times may have pairs with an AAI ≥90% and pairs with an AAI < 90%.

The shotgun proteomic analysis of the enrichment culture BPL community from two independent experiments produced a total of 545 proteins encoded by the genes located on the contigs of the metagenome. Almost 97% of the proteins were assigned to both the *Pelotomaculum* candidate BPL (354 proteins, 64%, Table A1) and *Gracilibacter* OTU bin (166 proteins, 30%). We did not analyze the *Desulfomonile* and *Ignavibacterium* OTUs because of their minor abundances in
the metagenome and their low coverage in the proteome (only 3% of the total proteins).

In the genome of *Gracilibacter* OTU (IMG ID: clostridium Ga0073690), no genes for sulfate reduction could be detected. If *Gracilibacter* activated benzene, *Gracilibacter* should have the capability to perform the corresponding downstream degradation. However, among the proteins classified to the *Gracilibacter* OTU, only a few matched enzymes that could be related to benzoate degradation. Possibly the *Gracilibacter* OTU grew by fermentation of metabolites produced by *Pelotomaculum* candidate BPL. For example, an indication could be the identification of pyruvate ferredoxin oxidoreductase in the proteome.

**Figure 8.** Maximum likelihood tree of near-complete 16S rRNA genes with bootstrap support values (%) showing the phylogenetic affiliation of *Pelotomaculum* candidate BPL. Strain names are followed by the corresponding RefSeq or SILVA sequence identifiers. Thick branch lines indicate a closer phylogenetic neighborhood of *Pelotomaculum* candidate BPL. The sequence of the 16S rRNA gene of *Pelotomaculum* candidate BPL can be found in Database A1.
3.1.2 A novel member of the genus *Pelotomaculum*

The reconstructed *Pelotomaculum* candidate BPL genome (IMG ID: pelotomaculum Ga0073689) consisted of 99 high-quality contigs with a total length of 2.97 Mbp and an average GC content of 53.71%. They were assembled from 454 reads (mean > 500 bp) according to standard procedures in isolate genome sequencing, using a 72-fold average positional coverage. Their length ranges from 1 kb to 234 kb with a mean of 30 kb. *Pelotomaculum thermopropionicum* SI is the only strain in the genus *Pelotomaculum* for which a full genome is available in the public resources. The genome has a total length of 3.03 Mbp with an average G+C content of 53.0%. It is reported to be a fermenting syntrophic organism whose metabolism depends on coupling with other microbes. Thus, we performed genomic comparisons between the *Pelotomaculum* candidate BPL and the *Pelotomaculum thermopropionicum* SI genomes (NCBI Reference Sequence: NC_009454.1). Comparative analyses of the average nucleotide identity (ANI) based on the IMG/M platform indicated that the two genomes represent different *Pelotomaculum* species, with an ANI value of 75.98%, far below an ANI threshold range for species delineation (95–96%) (Kim et al. 2014). This was in agreement with their low 16S rRNA gene sequence identity (< 95%). The corresponding identity scatter plot showed that the contigs from *Pelotomaculum* candidate BPL covered the entire genome of *Pelotomaculum thermopropionicum* SI (Figure A1). However, the nucleotide sequence similarity was limited leaving many gaps (Figure A1), which emphasized the differences between the reconstructed *Pelotomaculum* candidate BPL genome and the RefSeq isolate genome sequence. The gene distributions of the genomes were compared for Clusters of Orthologous Groups (COG). We found significant differences in the COG categories distributions at the level of \( p < 0.05 \) by the Chi-squared test, with several differences in the reference genome of *Pelotomaculum thermopropionicum* SI (Figure A2), indicating different metabolic capacities. Employing a phylogenetic analysis of the near-complete 16S rRNA gene (position 113-1661, Database A1), we found that among known members of the *Peptococcaceae* family, the BPL candidate is clearly separate from *Pelotomaculum thermopropionicum* SI (Figure 1). Furthermore, it has the closest
phylogenetic relationship (with 95% 16S rRNA gene sequence similarities) to *Pelotomaculum isophthalicicum* and *Pelotomaculum terephthalicicum* that both can utilize benzoate and other low-molecular weight aromatic compounds when grown in co-culture with hydrogenotrophic methanogens (Qiu et al. 2006) (Figure 8).

**Figure 9.** Metabolic pathway reconstruction for *Pelotomaculum* candidate BPL. Benzene degradation and sulfate reduction pathways are evidenced by proteome analysis. Numbers correspond to the following enzymes: B1, class II benzoyl-CoA reductase (Bam type); B2, cyclohex-1,5-dienecarbonyl-CoA hydratase; B3, 6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenases; B4, 6-oxo-cyclohex-1-ene-carbonyl-CoA hydrolase; W1, carbon monoxide dehydrogenase/acetyl-CoA synthetase; W2, 5-methyltetrahydrofolate methyltransferase; W3, methylene-tetrahydrofolate reductase; W4, methylene-tetrahydrofolate dehydrogenase; W5, formate-tetrahydrofolate synthetase; W6, formate dehydrogenase; S1, ATP sulfurylase, S2, adenosine-5’-phosphosulfate reductase; S3, dissimilatory sulfite reductase. Abbreviations are: APS, adenosine-5’-phosphosulfate; QmoAB, quinone-interacting membrane-bound oxidoreductase; DsrABC, dissimilatory sulfite reductase cytoplasmic subunits; DsrMK, transmembrane dissimilatory sulfite reductase complex; e^−, electron.
3.1.3 Initial activation of benzene

In order to find evidence for methylation as the initial activation reaction, we tried to identify genes related to anaerobic degradation of toluene in the genome. Benzylsuccinate synthase (Bss) is a key enzyme in the anaerobic toluene degradation pathway. However, we could neither identified bss alpha-subunit (bssA)-like genes in the genome nor genes for beta-oxidation of benzylsuccinate to benzoyl-CoA. Thus, we exclude a degradation of benzene via methylation to toluene. Anaerobic phenol degradation to benzoyl-CoA was known to proceed with the participation of phenylphosphate synthase, phenylphosphate carboxylase, 4-hydroxybenzoate-CoA ligase and 4-hydroxybenzoyl-CoA reductase (Abu Laban et al. 2010). Only a few hits could be identified for genes showing homology to individual subunits of these enzymes. However, when looking at the genetic context, genes encoding for the other subunits of these proteins which would be needed to build catalytically active enzymes are lacking. The absence of relevant key genes for anaerobic phenol degradation excludes activation of benzene via hydroxylation in this culture. The earlier reported inability of culture BPL to grow on toluene or phenol confirmed this genomic analysis (Abu Laban et al. 2009).

Previous metabolite analyses (e.g. benzoate as an intermediate) and substrate tests indicated a direct carboxylation of benzene as the initial activation mechanism for *Pelotomaculum* candidate BPL (Abu Laban et al. 2009). Thus, we searched for genes encoding for the proposed anaerobic benzene carboxylase (AbcA and AbcD) responsible for direct carboxylation to benzoate in the iron-reducing culture BF (Abu Laban et al. 2010). In the genome of *Pelotomaculum* candidate BPL, we identified one gene copy (Ga0073689_2617920776) assigned as 4-hydroxy-3-polyprenylbenzoate decarboxylase which shared 33% amino acid sequence identity to the anaerobic benzene carboxylase gene abcA. However, no abcD homologue could been detected in the proximity of gene Ga0073689_2617920776 and no gene product of gene Ga0073689_2617920776 could be identified in the proteome. We thus conclude that *Pelotomaculum* candidate BPL does not utilize an anaerobic benzene carboxylase similar to culture BF.
Furthermore in culture BF, a gene for a putative benzoate-CoA ligase was located next to the carboxylase genes. This ligase is intended to convert the carboxylation product benzoate to benzoyl-CoA which can be further degraded through the benzoyl-CoA degradation pathway. If benzene carboxylation occurred, we would expect genes encoding benzoate-CoA ligase which further activates benzoate to benzoyl-CoA in the genome. In *Geobacter metallireducens* GS-15, there are two different enzymes for benzoyl-CoA production: the ATP-dependent benzoate-CoA ligase (BamY) and succinyl-CoA: benzoate CoA transferase (Bct) (Oberender et al. 2012). Three hits (Ga0073689_2617922070, 2617921892, and 2617920988) were obtained in our genome showing only 27-29% amino acid sequence identities to the *bamY* gene of *Geobacter metallireducens* GS-15. Additionally, we could not identify genes homologous to *bct* in *Geobacter metallireducens* GS-15 or the putative benzoate-CoA ligase *bzlA* in the iron-reducing enrichment culture BF (Abu Laban et al. 2010).

### 3.1.4 Central benzoyl-CoA degradation pathway

According to present knowledge, anaerobic degradation of benzene has to proceed via the central intermediate, benzoyl-CoA, regardless of the initial activation mechanisms. Two distinct classes of benzoyl-CoA reductases (BCRs) are presently known for the reduction of benzoyl-CoA to cyclohexa-1,5-diene-1-carboxyl-CoA (dienoyl-CoA). The ATP-dependent Class I BCR (BcrABCD) is predominantly found in facultative anaerobes; the ATP-independent and tungsten cofactor-containing class II BCR (BamBCDEFGHI) occurs in obligate anaerobic bacteria (Fuchs et al. 2011, Kung et al. 2009, Löffler et al. 2011). As expected for an obligate anaerobic sulfate reducer degrading aromatic compounds, a protein similar to the active site subunit of the class II BCR in *Geobacter metallireducens* GS-15 (BamB, Ga0073689_2617921884) was detected in the proteome (Figure 9 and Table A1). The gene is located adjacent to a *bamC* paralogue (Ga0073689_2617921883) with 66% amino acid sequence identity to the *bamC* gene in *Geobacter metallireducens*. Candidate genes for *bamDEFGHI* putatively encoding benzoyl-CoA reductase were identified in at least one copy in the genome of *Pelotomaculum* candidate BPL. They shared amino acid sequence identities ranging from 33% to 71% to the respective proteins described in
Geobacter metallireducens GS-15 (Loeffler et al. 2011, Wischgoll et al. 2005), indicating that Pelotomaculum candidate BPL degraded benzoate-CoA via the class II benzoyl-CoA reductase (Bam type).

Products of a gene cluster (Ga0073689_2617919856-8) were detected in the proteome putatively encoding 6-oxo-cyclohex-1-ene-carbonyl-CoA hydrolase (Oah), cyclohex-1,5-dienecarbonyl-CoA hydratase (Dch), and 6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenases (Had), respectively. These genes and gene products suggest further transformation of dienoyl-CoA to 3-hydroxypimelyl-CoA (Figure 9 and Table A1). The hydratase (Dch) could catalyze water addition to a carbon-carbon double bond, followed by dehydrogenation (Had), and hydrolytic ring cleavage (Oah) (Holmes et al. 2012, Wischgoll et al. 2005). Proteins for further β-oxidation of 3-hydroxypimelyl-CoA to acetyl-CoA and CO₂ via glutaryl-CoA were also identified in the proteome, i.e. glutaryl-CoA dehydrogenase (Ga0073689_2617919859), enoyl-CoA hydratase (Ga0073689_2617921893), and 3-hydroxyacyl-CoA dehydrogenase (Ga0073689_2617921894) (Table A1).

3.1.5 Complete oxidation via the Wood-Ljungdahl pathway

For complete oxidation of acetyl-CoA to CO₂, many sulfate reducers run the Wood-Ljungdahl pathway in reverse, also termed as oxidative carbon monoxide dehydrogenase/acetyl-CoA pathway. These are e.g. Desulfatibacillum alkenivorans AK-01, Desulfobacula toluolica Tol2, Desulfobacterium autotrophicum HRM2, or Desulfotomaculum acetoxidans DSM 771 (Callaghan et al. 2012, Spormann and Thauer 1988, Strittmatter et al. 2009, Wohlbrand et al. 2013). Likewise, all necessary enzymes involved in the Wood-Ljungdahl pathway were detected in the proteome and encoded in the Pelotomaculum candidate BPL genome (Figure 9 and Table A1). Namely, they contained the key enzyme carbon monoxide dehydrogenase/acetyl coenzyme A (CoA) synthetase CODH/ACS (Ga0073689_2617920888 and 2617921376-80), and the enzymes for the methyl branch of the pathway including 5-methyltetrahydrofolate methyltransferase AcsE (Ga0073689_2617921126), methylenetetrahydrofolate reductase MetF (Ga0073689_2617921113), methylene-tetrahydrofolate dehydrogenase FolD (Ga0073689_2617920380), formate-tetrahydrofolate synthetase Fhs
An interesting feature is the detection of genes encoding heterodisulfide reductase (Hdr) related proteins which co-localize with methylenetetrahydrofolate reductase MetF (Ga0073689_2617921113) (Figure 10). The methylenetetrahydrofolate reductase gene metF is tightly associated with a gene (Ga0073689_2617921112) designated as the N-terminus of MetF (metV) and they are located next to mvhD encoding the delta subunit of methyl-viologen-reducing hydrogenase and hdrA encoding heterodisulfide reductase iron-sulfur subunit A (Figure 10). The function of the delta subunit MvhD and HdrA is described as transferring reducing equivalents to methylenetetrahydrofolate reductase MetF in an electron-bifurcation process (Buckel and Thauer 2013, Mock et al. 2014, Wohlbrand et al. 2013). Similar gene clusters of metVF genes together with the hdrA-mvhD genes are also present in other acetate-oxidizing sulfate reducers, i.e. Desulfatibacillum alkenivorans AK-01, Desulfobaculum toluolica Tol2, Desulfobacterium autotrophicum HRM2 (Wohlbrand et al. 2013). We conclude that Pelotomaculum candidate BPL has a similar way of energy conservation from acetate oxidation and sulfate reduction as described previously for sulfate-reducing bacteria.

Among possible genes of the tricarboxylic acid (TCA) cycle, we have found malate dehydrogenase (NAD) (Ga0073689_2617921286) from the oxidative branch and isocitrate dehydrogenase (NADP) (Ga0073689_2617922024) from the reductive
branch indicating that this part of the genome might be incomplete. Absence of the key enzyme α-ketoglutarate dehydrogenase at least indicates that Pelotomaculum candidate BPL cannot operate a TCA cycle for oxidation of acetyl-CoA to CO₂ which agrees with the presence of a complete Wood-Ljungdahl pathway for oxidation of acetate.

### 3.1.6 Sulfate reduction in the genus Pelotomaculum

So far, members of the genus Pelotomaculum are supposed to be unable to grow by sulfate reduction (Imachi et al. 2006). However, the closely related genus Desulfotomaculum is well-known for sulfate reduction. The key enzymes of dissimilatory sulfate reduction ATP sulfurylase (Sat) (Ga0073689_2617919560), adenosine-5'-phosphosulfate reductase (Apr) (Ga0073689_2617919558-9), and dissimilatory sulfite reductase (Dsr) (Ga0073689_2617920264-5), have been found in the proteome of our Pelotomaculum candidate BPL (Figure 9 and Table A1). ATP sulfurylase catalyzes the initial activation of sulfate to adenosine-5'-phosphosulfate (APS) releasing pyrophosphatase. Pyrophosphate can be used to create a proton motive force by an annotated, putative proton-translocating membrane pyrophosphatase (Ga0073689_2617919937). Genome analysis revealed the genes aprA and aprB encoding the sulfite-forming APS reductase are located adjacent to the sat gene. In close proximity to the sat and apr genes are qmoAB genes (Ga0073689_2617919556-7) probably encoding for an adenylylsulfate reductase-associated electron transfer protein which is homologous to heterodisulfide reductases (Ramos et al. 2015). The QmoAB redox complex is proposed to be the electron donor of APS reductase in Gram-positive sulfate-reducing bacteria (Junier et al. 2010). Interestingly, gene Ga0073689_2617919555 was automatically annotated by the IMG platform as the membrane subunit of the QmoABC complex (qmoC), which has been observed only in Gram-negative sulfate-reducing bacteria so far (Pereira et al. 2011). Our further search found that the product of the gene lacked transmembrane helices and should therefore encode a different function (Junier et al. 2010). Further reduction of sulfite to sulfide could be catalyzed by dissimilatory sulfite reductase encoded by dsrAB genes (Ga0073689_2617920264-5). The electrons for DsrAB reductase are probably transferred from the quinone pool by the DsrMK complex.
and DsrC (Junier et al. 2010). According to the Gene Ortholog Neighborhoods tool in the IMG platform, *Pelotomaculum* candidate BPL exhibited the same gene organization for the *dsr* operon as other Gram-positive sulfate-reducing bacteria, e.g. *Desulfotomaculum reducens* Mi-1 or *Desulfotomaculum acetoxidans* DSM 771 (Junier et al. 2010). Similar to *Desulfotomaculum reducens* Mi-1, the identified *dsrM* (Ga0073689_2617920256) encoding a membrane protein functioning as a conduit to transfer electrons from quinones in *Desulfotomaculum reducens* Mi-1 was also annotated as the nitrate reductase gamma subunit in the IMG platform. The electrons for the quinone pool might come from NADH oxidation coupled with proton translocation catalyzed by the energy-conserving NADH-quinone oxidoreductase (Complex I, Ga0073689_2617921751-61). Another essential protein, dissimilatory sulfite reductase subunit C (DsrC, Ga0073689_2617920254) was also identified in the proteome. The proteogenomic data presented here clearly indicate that *Pelotomaculum* candidate BPL has a complete sulfate reduction pathway which is active during anaerobic benzene degradation.

**3.1.7 Hydrogenases and formate dehydrogenases**

No genes encoding [FeFe] and [NiFe] hydrogenases could be identified in the genome. Their absence suggests that *Pelotomaculum* candidate BPL neither has the capability for H₂ oxidation nor for fermentation with H₂ production. This was unexpected as members of the genus *Pelotomaculum* are typical fermenting organisms and usually possess hydrogenases. Three genes putatively encoding formate dehydrogenases (Ga0073689_2617920664, 2617920666, and 2617922225) were identified in the genome but neither of them could be identified in the proteome. This suggests that formate transfer rather than hydrogen transfer may be important if *Pelotomaculum* candidate BPL would possibly grow in syntrophic fermentation when sulfate is not available.
3.2. Fermentative Spirochaetes drive nutrient cycling by a subsurface microbial loop in hydrocarbon-contaminated habitats

Our studies have shown that Spirochaetes are prominent members of the naphthalene-degrading, sulfate-reducing enrichment culture N47 derived from sediments of a contaminated aquifer (Meckenstock et al. 2000, Selesi et al. 2010). Based on 16S rRNA gene sequence analysis, ~93% of the culture comprises a sulfate-reducing deltaproteobacterium that catalyzes naphthalene degradation and the remainder comprises Spirochaetes (Bergmann et al. 2011a, Selesi et al. 2010). We separated the enrichment culture into a highly enriched Deltaproteobacteria N47 culture (without Spirochaetes) (Bergmann et al. 2011a, Selesi et al. 2010) and a pure isolate of the spirochaete *Rectinema cohabitans* strain HM (Koelschbach et al. 2017). In this work, the aim was to elucidate the functional role of the Spirochaetes in this enrichment culture. Results from physiological, genomic, and proteomic analyses indicated that *R. cohabitans* grows by fermentation of organic compounds derived from dead biomass (necromass), recycling electrons as H₂ and short chain fatty acids or alcohols to the sulfate-reducing Deltaproteobacteria N47. The co-culture of the two organisms thus provides an example of a simple microbial loop. Further comparative and community genome analyses indicated that other Spirochaetes, both uncultured and isolated, are hydrogen-producing fermentative microbes fulfilling similar roles as *R. cohabitans* in supporting biodegradation.

3.2.1 The genome and proteome of *R. cohabitans* suggest it is an obligate fermenter

A draft genome of *R. cohabitans* HM was sequenced and assembled to develop hypotheses about its ecophysiological role in the naphthalene-degrading culture N47. The 2.82 Mb genome comprises 41 contigs and, based on CheckM analysis (Parks et al. 2015), is estimated to be 98% complete with no measurable strain-level heterogeneity. The organism encodes two 16S-23S-5S ribosomal RNA (rRNA) operons, 46 transfer RNA (tRNA) genes for all 20 amino acids, and 2583 protein-coding sequences. As in other nonspiral Spirochaetes, genes associated
Figure 11. Metabolic pathway reconstruction for Spirochaetes in hydrocarbon- and organohalide-contaminated environments. Genes for the illustrated pathways were detected in the genomes of *R. cohabitans*, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8. Predicted proteins indicated by numbers in the figure can be found in Table A2.
with motility and chemotaxis are absent (Caro-Quintero et al. 2012). Analysis of the metabolic capacity from the genome suggests *R. cohabitans* adopts a fermentative lifestyle, with a large fraction of enzymes associated with carbohydrate, protein, and molecular hydrogen (H₂) metabolism (Figure 11). In contrast, key genetic determinants of aerobic and anaerobic respiration, the tricarboxylic acid cycle, and acetogenic carbon fixation are absent from the genome. This is consistent with previous work showing that various Spirochaetes isolated from the genera *Sphaerochaeta*, *Spirochaeta*, *Borrelia*, and *Treponema* also adopt an obligately fermentative lifestyle (Caro-Quintero et al. 2012).

The genome suggests that *R. cohabitans* degrades proteins, peptides, and carbohydrates. For example, genes encoding several extracellular peptidases (Merops family M23B, S8A, S26A, and S33) may mediate decomposition of bacterial proteins and cell wall constituents (Figure 11 and Table A2). These enzymes may be secreted by the predicted Type II secretory system encoded in the genome (Table A2). The genome also contains a diverse suite of genes encoding transporters for branched amino-acids, dipeptides, and oligopeptides, suggesting that *R. cohabitans* has the capacity to import hydrolysis products into the cell (Figure 11). Also identified were numerous genes encoding intracellular peptidases, aminotransferases, and four ketoacid-ferredoxin oxidoreductases (Figure 11 and Table A2). The genome also encodes 92 predicted carbohydrate-active enzymes, including hydrolases, extracellular binding proteins and ABC transporters involved in carbohydrate acquisition and breakdown (Figure 11), consistent with its ability to utilize saccharides as carbon sources (Koelschbach et al. 2017).

For central metabolism, the genome encodes a complete Embden-Meyerhof-Parnas glycolysis pathway for oxidizing glucose to pyruvate concomitant with ATP and NADH production (Figure 11 and Table A2). The determinants of the classical pentose phosphate pathway involved in NADPH and pentose synthesis were also found (Table A2). The resulting pyruvate from glycolysis can be converted to acetyl-CoA through ferredoxin-dependent (pyruvate-ferredoxin oxidoreductase) or NAD-dependent (pyruvate dehydrogenase complex) enzymes. Like many other fermenting bacteria, *R. cohabitans* is predicted to convert acetyl-CoA to the
fermentation product acetate with concomitant ATP production by phosphate acetyltransferase and acetate kinase (James et al. 2016). The presence of genes for aldehyde and alcohol dehydrogenases suggests that acetyl-CoA can also be reduced to ethanol in two steps using. Proteomic analysis of R. cohabitans grown on glucose revealed that genes related to mixed-acid fermentation were expressed at high peptide abundances (Table A3).

Furthermore, a set of genes are also present in the genome encoding a membrane-bound RnfABCDGE-type redox complex (Figure 11 and Table A2). This complex is frequently found in the genomes of anaerobic organisms, including other obligately fermentative bacteria (Dodsworth et al. 2013, Sorokin et al. 2014), and was recently shown to link the oxidation of reduced ferredoxin and reduction of NAD\(^+\) to electrogenic pumping of protons or sodium ions across the cell membrane (Biegel et al. 2011, Tremblay et al. 2012). This complex is likely to have dual roles in maintaining the redox balance and generating membrane potential to fuel ATP synthesis and active transport. A V-type ATPase and a membrane-associated pyrophosphatase are also likely to support the membrane potential maintenance via ATP and pyrophosphate hydrolysis. Proteomic analysis revealed that genes related to the V-type ATPase and pyrophosphatase were expressed at high peptide abundances (Table A3).

3.2.2 Interspecies hydrogen transfer supports a naphthalene-degrading enrichment culture

Hydrogen gas, a diffusible end product of bacterial fermentation and a primary electron source for sulfate-reducing deltaproteobacteria, might be transferred between the members of the N47 enrichment culture (Schwartz et al. 2013). Supporting this hypothesis, we identified putative hydrogenase genes in the genomes of both R. cohabitans and Deltaproteobacteria N47 (Bergmann et al. 2011a, Selesi et al. 2010). To gain insight into their potential ecophysiological roles, we classified them using their primary sequence-based phylogeny and genetic organization according to our recently developed hydrogenase classification scheme (Greening et al. 2016, Søndergaard et al. 2016). Seven putative [FeFe]-hydrogenase genes were identified in the genome of the
spirochaete. Three are monomeric [FeFe] Group A1 or B hydrogenases (Figures 12A and 12C) which couple oxidation of reduced ferredoxin to evolution of H₂ during carbohydrate and protein fermentation in diverse bacteria (Wolf et al. 2016). Two others encode the trimeric electron-bifurcating [FeFe] Group A3 hydrogenases (Figures 12A and 12C), which reversibly couple the oxidation of ferredoxin and NADH to the evolution of H₂, thereby providing a novel mechanism to balance redox state and ATP demand. The genome also encodes [FeFe] Group C1 and C3 hydrogenases, both candidate subgroups implicated in H₂ sensing (Søndergaard et al. 2016).

Re-analysis of the genome for Deltaproteobacteria N47 (Bergmann et al. 2011a, Kummel et al. 2015) revealed the presence of two Group 1a and Group 1b [NiFe]-hydrogenases (Figure 12B) that couple H₂ oxidation to cytochrome c₃ reduction in the respiratory chains of sulfate-reducing deltaproteobacteria. Whereas the former sequence is incomplete due to sequence coverage gaps, a complete operon encoding the 1b enzyme and its cytochrome c₃ partner is detectable (Figure 12D) (Greening et al. 2016).

To confirm that *R. cohabitans* is a fermentative H₂ producer, we performed pure culture growth experiments with glucose as carbon substrates. Gas chromatography demonstrated that H₂ rapidly accumulated in the headspace during cultivation with glucose resulting in H₂ concentrations up to 400,000 ppmv after two weeks (Figure 13A). Proteomic analysis of cells grown with glucose confirmed the expression of all hydrogenase-related genes with high peptide abundances (Table A3). Also consistent with genome-based predictions, we observed a gradual decrease in H₂ concentration when highly enriched Deltaproteobacteria N47 was cultivated under naphthalene-degrading, sulfate-reducing conditions and a headspace supplemented with a limited amount of H₂ gas (~ 8285 ppmv) (Figure 13A). Consumption of H₂ by Deltaproteobacteria N47 was inhibited when sodium molybdate (NaMoO₄) was added at Day 31 (Figure 13A), a known potent inhibitor of sulfate reduction (Lovley and Klug 1983). This suggests that, in common with classical sulfate-reducing isolates (Schwartz et al. 2013), Deltaproteobacteria N47 directly couples H₂ oxidation to sulfate reduction in an anaerobic respiratory process.
Figure 12. Determinants of H₂ metabolism in Spirochaetes in hydrocarbon- and organohalide-contaminated environments. (A) Phylogenetic tree of the [FeFe]-hydrogenase catalytic subunit sequences detected in *R. cohabitans*, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8; (B) Phylogenetic tree of the [NiFe]-hydrogenase catalytic subunit sequences detected in Deltaproteobacteria N47; (C) Genetic organization of hydrogenases from *R. cohabitans*; (D) Genetic organization of the hydrogenase operon detected in Deltaproteobacteria N47. The illustrated tree for Figure 12A is condensed, with the full tree in Figures A3. Genetic organization diagrams are shown to-scale and genes/domains are color-coded as follows: green = catalytic site; blue = secondary subunit; yellow = electron acceptor or donor; light orange = maturation factor. Redox-active centers are shown in circles, where: yellow = [2Fe2S] cluster, green=[3Fe4S] cluster, and red = [4Fe4S] cluster.
Finally, in order to reveal a potential mutualistic relationship between the two strains, we monitored H₂ concentrations in the enrichment culture N47, comprising *R. cohabitans* and Deltaproteobacteria N47, under naphthalene-degrading, sulfate-reducing conditions. Under these conditions, the concentration of H₂ remained constant (~ 2 ppmv) for one month (Figure 13A). However, a sharp increase in H₂ concentration was observed when sulfate reduction by Deltaproteobacteria N47 was inhibited due to the addition of sodium molybdate. This suggests that there is a tight coupling between spirochaete-mediated fermentative H₂ evolution and deltaproteobacterium-mediated respiratory H₂ consumption under naphthalene-degrading, sulfate-reducing conditions.

![Figure 13](image)

**Figure 13.** H₂ oxidation and evolution under different growth conditions. (A) Interspecies hydrogen exchange between *R. cohabitans* and Deltaproteobacteria N47. Blue triangles: *R. cohabitans* + glucose; red circles: Deltaproteobacteria N47 + naphthalene + H₂; black squares: co-culture of the two strains + naphthalene. (B) Hydrogen evolution with dead biomass and metabolites as carbon sources for *R. cohabitans*. Blue up-triangles: *R. cohabitans* + killed Deltaproteobacteria N47 cells; black squares: *R. cohabitans* + filtered Deltaproteobacteria N47 supernatant; green diamonds: carry-over control of *R. cohabitans*; red circles: control of filtered Deltaproteobacteria N47 supernatant; pink down-triangles: control of killed Deltaproteobacteria N47 cells. Values are means of two or three individual incubations measured twice each. Error bars indicate SD of biological replicates.
3.2.3 *R. cohabitans* feeds on organic carbon derived from necromass

![Volcano plot diagram](image)

**Figure 14.** Quantitative proteomic analysis of highly enriched Deltaproteobacteria N47 and enrichment culture N47 grown in the presence of naphthalene. The volcano plot was generated using the PERSEUS software. The calculated Log2 (fold change) (x-axis) was then plotted against the corresponding p-values (y-axis). Proteins to the left and above the significance line are significantly depletes in the co-culture and proteins to the right and above the significance line are significantly enriched in the co-culture. The filled squares indicate proteins from *R. cohabitans*; the filled circles represent proteins from Deltaproteobacteria N47.

While the above results demonstrate that H$_2$ is an important electron link between *R. cohabitans* and Deltaproteobacteria N47, the question remained about the origin of reducing equivalents (i.e. carbon sources) used for H$_2$ production by *R. cohabitans*. Considering that *R. cohabitans* cannot utilize naphthalene, its carbon source must be supplied by Deltaproteobacteria N47 given no other exogenous sources were added. Thus, we analyzed the proteomes of *R. cohabitans* and Deltaproteobacteria N47 grown as highly enriched Deltaproteobacteria N47 and enrichment culture N47 respectively. Comparing the proteomes of the two cultures...
under naphthalene-degrading conditions revealed that only five proteins of Deltaproteobacteria N47 were differentially expressed, none of which could be assigned a distinct ecophysiological role (Figure 14). However, 53 proteins belonging to *R. cohabitans* were identified under the co-culture conditions (Table A4). Most of them are related to transporters for oligopeptides, dipeptides, amino acids, sugars, and dicarboxylic acids (Figure 14 and Table A4), suggesting that these compounds are actively taken up from the medium by *R. cohabitans*. Other enriched enzymes include protease M41, glutamate dehydrogenase (amino acid catabolism), pyruvate-ferredoxin oxidoreductase (ferredoxin reduction), and components of the Sec translocon (secretion of hydrolases, insertion of transporters) (Table A4).

We subsequently confirmed that *R. cohabitans* can use organic carbon sources derived from Deltaproteobacteria N47 to support fermentative H₂ production by cultivating *R. cohabitans* with autoclaved biomass of Deltaproteobacteria N47. After three days, the H₂ production by *R. cohabitans* was twofold higher in fresh media incubations supplemented either with heat-killed cells or filtered supernatant from Deltaproteobacteria N47 (Figure 13B). This provides further evidence that *R. cohabitans* can use necromass and potentially exudates from Deltaproteobacteria N47 to support its obligately fermentative lifestyle. This finding is also in line with growth experiments revealing that *R. cohabitans* obligately requires the presence of 0.1% yeast extract to grow in pure culture but not in co-culture with Deltaproteobacteria N47 (Koelschbach et al. 2017). Pimelate and other dicarboxylic acids are important downstream metabolites of anaerobic naphthalene degradation and have been detected in the supernatant of the enrichment culture N47 (Meckenstock et al. 2016). It is also possible that dicarboxylic acids released during naphthalene degradation are taken up by some of the identified transporters and further catabolized. But the genomic search in *R. cohabitans* did not support this hypothesis, as we could not find genes for the β-oxidation of dicarboxylic acids (Harrison and Harwood 2005), in agreement with the proteome analysis. Together, these observations suggest that *R. cohabitans* primarily feeds on proteins, sugars, and possibly other compounds derived from detrital Deltaproteobacteria N47 biomass.
Figure 15. Condensed maximum likelihood tree of partial 16S rRNA gene sequences of Spirochaetes showing the phylogenetic affiliation of *R. cohabitans*, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8. Phylogenetic trees were constructed by the maximum-likelihood method and are bootstrapped with 500 replicates. The expanded tree is shown in Figure A4.
**3.2.4 Spirochaetes may generally drive microbial loops in hydrocarbon- and organohalide-contaminated environments**

To further illuminate the ecological niche of Spirochaetes in polluted habitats, we compared the genome of *R. cohabitans* with the draft genomes of two uncultured environmental Spirochaetes reconstructed from metagenomics sequences: uncultured Spirochaetes bacterium SA-8, originating from a terephthalate-degrading, methane-producing bioreactor (Nobu et al. 2015), and uncultured Spirochaetes bacterium bdmA 4, derived from a naphthalene-degrading, sulfate-reducing culture enriched from a creosote spill (Kummel et al. 2015). Genome-based metabolic reconstructions suggest that, akin to *R. cohabitans*, both organisms are nonspiral Spirochaetes and obligate fermenters (Figure 11 and Table A2). In addition, both organisms encode [FeFe] Groups A and B hydrogenases that are phylogenetically related to those of *R. cohabitans* (Figures 12A and A3), suggesting that they also fermentatively produce H2, sustaining associated hydrogenotrophic methanogens and sulfate reducers. Revisiting the published genomes of the nonspiral Spirochaetes *S. pleomorpha* and *S. globosa* from trichloroethene-degrading cultures also indicates a potential for H2 production via saccharolytic mixed-acid fermentation (Figures 12A and A3) (Caro-Quintero et al. 2012).

The 16S rRNA gene sequences of *R. cohabitans*, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8 are 95% identical to each other, and form a common clade distinct from *Treponema* species in the 16S rRNA based phylogenetic tree (Figures 15 and A4). In the same clade, we could also identify numerous 16S rRNA gene sequences associated with contaminated aquifers or bioreactors containing hydrocarbons or organohalides and oil reservoirs (a small selection is shown in Figure 15). This indicates that *Rectinema* species are ubiquitous in hydrocarbon- and organohalide-polluted anoxic sites. *S. globosa* and *S. pleomorpha* strains are distantly related to the three members of the genus *Rectinema* (Koelschbach et al. 2017), suggesting that nonspiral morphology evolved multiple times and that a similar metabolism is present in distinct groups of environmental Spirochaetes (Figure A4). Altogether, comparative genomic analyses suggest that these environmental nonspiral
Spirochaetes fulfill ecophysiological functions similar to *R. cohabitans* in their communities: recycling necromass and providing H₂ to respiring bacteria (e.g. sulfate reducers) or hydrogenotrophic methanogens in these communities.
3.3. Monitoring microbial mineralization using reverse stable isotope labelling analysis by mid-infrared laser spectroscopy

In the classical stable isotope labelling experiments, fully or partially $^{13}$C-labelled compounds are added to the cultivation medium which is buffered with bicarbonate at natural $^{13}$C/$^{12}$C isotopic abundance (i.e., with a $^{13}$C fraction $x(^{13}$C) ≈ 1.11%). The change of the $^{13}$C/$^{12}$C isotope ratio of CO$_2$ in the headspace is then followed over time and used to calculate the quantity of evolved $^{13}$CO$_2$. In our reverse stable isotope labelling approach, we simply set out to reverse this set up by adding $^{13}$C-labelled bicarbonate with $x(^{13}$C) values ranging from 2-50%, depending on the expected production of CO$_2$. Then, a regular substrate or mixture for microbial degradation is supplied which has a normal stable isotope ratio at natural abundance ($x(^{13}$C) ≈ 1.11%). Consequently, CO$_2$ with a stable isotope ratio at natural abundance is released during mineralization. As the stable isotope ratio of the released CO$_2$ differs significantly from the isotope ratio of the labelled background dissolved inorganic carbon (DIC), one can measure the changes in the $^{13}$C/$^{12}$C isotope ratio of the total DIC (i.e. the sum of inorganic carbon species in the medium) over time which can be used to provide evidence for CO$_2$ production and to calculate its absolute amount.

Here, to demonstrate the feasibility and usefulness of the RIL method, we studied anaerobic degradation of aromatic hydrocarbons, which belong to the most recalcitrant priority pollutants in contaminated aquifers. The broader environmental relevance of this method was further shown by aerobic mineralization of DOC adsorbed to granular activated carbon (GAC) during drinking water production.
3.3.1 Verification of the method

![Graphs showing degradation profiles](image)

Figure 16. Anaerobic degradation profiles of (a) toluene by *G. metallireducens*, (b) 2-methylnaphthalene by the enrichment culture N47, (c) naphthalene by an uncharacterized iron-reducing enrichment culture SN, and (d) toluene by *T. aromatica*. The added bicarbonate buffer (30 mM) consisted of non-labelled and $^{13}$C-labelled sodium bicarbonate at a ratio of 50:50 for *G. metallireducens* (a) and the enrichment culture N47 (b), 80:20 for the enrichment culture SN (c), and 90:10 for *T. aromatica* (d). Sulfate or iron reduction as surrogate measurements indicated methylnaphthalene or naphthalene degradation. Carbon isotope ratio changes were expressed as a percentage compared to the initial $x(^{13}$C) value: $\Delta x(^{13}$C) = ($x(^{13}$C) final - $x(^{13}$C) initial)/$x(^{13}$C) initial × 100. The $x(^{13}$C) values of DIC samples for (a)-(c) were only measured by IRMS. Both IRMS and IRIS were performed for (d). Data are means of two or three parallel incubations measured three times each. Error bars represent standard deviation of the biological replicates.

The proof of principle for the reverse stable isotope labelling (RIL) method was first verified with two reference cultures: (1) anaerobic toluene degradation by *G. metallireducens* and (2) anaerobic 2-methylnaphthalene degradation by the enrichment culture N47 (Figure 16). The medium was supplemented with unlabelled and $^{13}$C-labelled sodium bicarbonate at a ratio of 50:50 and carbon
stable isotope ratios of DIC samples were measured with IRMS. During anaerobic toluene degradation by *G. metallireducens*, the initial $^{13}$C fractions of DIC samples ($x^{(13}C)_{\text{initial}} = 33.16\%$) decreased by $\Delta x^{(13}C) = 13.03 \%$, which corresponded to the consumption of total toluene from 0.75 to 0.18 mM (Figure 16a). No significant changes could be observed for toluene concentrations or $^{13}$C fractions in control bottles. During anaerobic degradation of 2-methylnaphthalene by the enrichment culture N47, the initial $^{13}$C fractions of DIC samples ($x^{(13}C)_{\text{initial}} = 34.01\%$) decreased by 13.62% concomitant with a measured reduction of sulfate from 9.97 to 8.01 mM (equal to consumption of 0.29 mM of 2-methylnaphthalene) (Figure 16b). No significant changes could be observed for sulfate concentrations or $^{13}$C fractions in control bottles. These two experiments provided proof-of-principle that measuring stable carbon isotope ratios of DIC by the RIL method can be used for demonstrating microbial degradation.

![Figure 17](image.jpg)

**Figure 17.** The calculated CO$_2$ production for anaerobic degradation of (a) toluene by *G. metallireducens* (Figure 16a), (b) 2-methylnaphthalene by the enrichment culture N47 (Figure 16b), (c) naphthalene by an uncharacterized iron-reducing enrichment culture SN (Figure 16c). Data points depict means of two or three parallel incubations measured three times each. Error bars represent standard deviation of the biological replicates.

We then used the RIL method to test if an uncharacterized enrichment culture can mineralize naphthalene under iron-reducing conditions (Figure 16c). In contrast to previous experiments, unlabelled and $^{13}$C-labelled sodium bicarbonate were added
at a ratio of 80:20, making the initial $x^{(13)C}$ values lower than in the proof-of-principle study. Carbon stable isotope ratios of DIC samples were measured with IRMS. Naphthalene was the sole organic carbon source added. In control bottles inoculated with autoclaved cultures, the $^{13}$C fractions of DIC samples remained stable at $\sim$15.46%. In the bottles inoculated with active cultures, a considerable lag phase of about 90 days took place before a decrease in $x^{(13)C}$ values and an increase of ferrous iron could be observed. Then, the initial $^{13}$C fraction of DIC samples ($x^{(13)C}_{\text{initial}} = 15.26\%$) dropped by 10.86% coupled to a concomitant ferrous iron production of 31 mM. This data provided clear evidence that naphthalene mineralization to CO$_2$ occurred coupled with iron reduction. But due to the non-linearity of IRMS measurement at high isotope ratios (see below), the calculation of CO$_2$ production and the electron balance were not accurate (Figure 17 and Table 2).

**Table 2.** Stoichiometric calculations of the measured and the theoretical ratios between the reduction of electron acceptors or donors and CO$_2$ production at the last sampling point for Figures 16a-c.

<table>
<thead>
<tr>
<th>Figure 16a</th>
<th>Figure 16b</th>
<th>Figure 16c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toluene oxidation: CO$_2$ production</td>
<td>Sulfate reduction: CO$_2$ production</td>
<td>Iron reduction: CO$_2$ production</td>
</tr>
<tr>
<td>Theoretical ratios</td>
<td>0.17</td>
<td>0.61</td>
</tr>
<tr>
<td>Calculated ratios</td>
<td>0.59</td>
<td>0.31</td>
</tr>
</tbody>
</table>
3.3.2 Evaluation of highly enriched $x^{(13C)}$ values measured by IRMS

![Figure 18](image)

**Figure 18.** Performance comparison of IRIS and IRMS. (a) Effect of CO$_2$ concentrations (proportional to Peak Area All, i.e. sum of peak areas of $m/z$ 44, 45, and 46 of CO$_2$) on the IRMS-measured $^{13}$C atom fractions for $^{13}$C-enriched gas ($x^{(13C)} = 11.073\%$; black squares) and natural abundance gas ($x^{(13C)} = 1.0671\%$; black triangles), $n = 3$. (b) Effect of CO$_2$ concentrations on the IRIS-measured $^{13}$C atom fractions ($x^{(13C)} = 10.3\%$; blue circles) and the linear regression line showing IRIS-measured enriched $^{13}$C samples plotted against theoretical $x^{(13C)}$ values ($n = 5$; black triangles). No error bars are shown as they lie within the dimensions of the symbol.

For analysis of carbon stable isotope ratios by IRMS, CO$_2$ samples are firstly ionized to different ions mass-to-charge ratios ($m/z$) at 44, 45, and 46. The resulting streams of ions are subsequently separated in a magnetic field. Then, the Faraday cup collectors are used to specifically collect signals of each ion beam. In the case of natural abundance samples, the $^{13}$C atom fractions are at $x^{(13C)} = \sim 1.11\%$, making signals (voltage outputs) of $m/z$ 45 and 46 much weaker compared to $m/z$ 44 (Figure A5a). In order to get comparative intensities for the highest quality results, each Faraday cup is designed to connect its own amplification ohmic resistor, at a ratio of 3:300:1000 for $m/z$ 44, 45, and 46, respectively (Figure A5a). In contrast, when measuring samples highly enriched in $^{13}$C (e.g. $x^{(13C)} = 15\sim35\%$) by using the same standard setting, the amplified signals will not be in the same range, but with $m/z$ 45 dominating the peak (Figure A5b), even though the absolute ratio of $^{13}$C/$^{12}$C is smaller than 1. Consequently, the IRMS response could be imprecise. Aware of this, we preformed two linearity
tests on an IRMS using two different reference gases to investigate if reported $x(^{13}\text{C})$ values could significantly alter with changes in CO$_2$ concentrations. Results showed a good linearity when we measured the natural abundance reference gas ($x(^{13}\text{C}) = 1.067\%$) (Figure 18a). For highly enriched samples ($x(^{13}\text{C}) = 10.3\%$), however, the measured $x(^{13}\text{C})$ values increased with elevated gas partial pressures showing a larger positive slope of 0.02 (significantly linearly dependent on the CO$_2$ concentration) (Figure 18a). This indicated that the results shown in Figures 16a-c could be inaccurate, as we could not apply working reference gases at the same concentration. This was further confirmed by the stoichiometric calculations which showed great differences between the measured and the theoretical ratios between the reduction of electron acceptors or donors and CO$_2$ production (Figure 17 and Table 2). Nevertheless, we could obtain a semi-quantitative assessment of the degradation activities as shown from the results.

One possible solution to alleviate this problem is to switch the feedback resistor of the preamplifier to adjust m/z 44 signal to an appropriate range (Goodman and Brenna 1992). Unfortunately, this cannot be achieved due to the limitations of the IRMS we used. Reinsch and Ambus adjusted resistor settings of the mass spectrometer to obtain higher measurement precision, but still they found that $x(^{13}\text{C})$ values measured by IRMS for $^{13}$C-enriched samples ($x(^{13}\text{C}) = 50\%$) were also significantly dependent on CO$_2$ concentrations (Reinsch and Ambus 2013). They subsequently used the empirical relationship to correct data but the measured $x(^{13}\text{C})$ values never showed the expected values. Using IRMS is therefore not straightforward for determination of $x(^{13}\text{C})$ values in the highly-enriched samples, which motivated us to search for a better alternative.

### 3.3.3 Using IRIS for analyzing highly enriched $^{13}$C samples

In contrast to IRMS, IRIS is capable of measuring enrichments up to $x(^{13}\text{C}) = 25\%$. Some other early-designed $^{13}$CO$_2$/$^{12}$CO$_2$ infrared gas analyzer using TDLA techniques have been shown to measure values of $x(^{13}\text{C})$ as high as 5% or 10%, as long as the $^{13}$CO$_2$ concentration is within the $^{13}$CO$_2$ range of the calibration gas (Dannoura et al. 2011, Plain et al. 2009). Therefore, infrared gas analyzers can be a good alternative for measuring highly enriched samples in $^{13}$CO$_2$-labelling
experiments. The TDLA technique employed by the Delta Ray IRIS follows the Beer-Lambert law (Bowling et al. 2003). While a laser beam scans through absorption lines of $^{13}$CO$_2$ and $^{12}$CO$_2$ of samples, the decrease of intensity of the laser can be directly linked to the concentration of the respective molecular species and hence carbon isotope ratios. For quantifications, absorptions within the scan range can be fitted by an algorithm adapted to certain environmental conditions within the laser cell (100 mbar pressure). For natural abundance samples, the peak areas for $^{13}$CO$_2$ are relatively small compared to that for $^{12}$CO$_2$ (Figure A6a) (van Geldern et al. 2014). For $^{13}$C-enriched samples (e.g. $\delta^{13}$C = 10.0 %), the peak areas for $^{13}$CO$_2$ increase strongly (Figure A6b). However, as long as the $^{13}$CO$_2$ band does not reach the saturation level, it can be used to determine carbon isotope ratios. In all our tests, we did not observe any saturation during measurements. Besides, IRIS can automatically determine CO$_2$ concentrations and adjust the CO$_2$ concentrations of samples to an optimal and constant concentration (380 ppm) going into the optical cell, which could address the issue of isotope ratio drifts caused by CO$_2$ concentrations (van Geldern et al. 2014).

To further confirm that IRIS can measure highly enriched $^{13}$C samples, we measured bicarbonate mixtures with different $\delta^{13}$C values. The measured $\delta^{13}$C versus expected $\delta^{13}$C values showed an excellent linear correspondence (correlation coefficient $R^2 = 0.9996$) over a wide range from $\delta^{13}$C = 6% to $\delta^{13}$C = 12%, giving a trendline of $y = 1.0315x - 0.1177$ (Figure 18b), demonstrating its capability to measure high $^{13}$C enrichments. The standard deviations of $\delta^{13}$C values varied with sample enrichments but were all below 0.006%. The precision is more than adequate for biological studies, where sample variation is considerably large than this analytical error (for details see next section). We subsequently sampled different aliquots (0.2 to 0.6 mL) from a 10% $^{13}$C-labelled bicarbonate solution and added these into the vials to produce samples with different CO$_2$ concentrations. Results illustrated that even though the CO$_2$ concentrations of samples changed significantly the resultant variation was very low (11.073% ± 0.001%) and remained within the analytical precision (Figure 18b),
therefore demonstrating the efficiency of the automatic concentration adjustment by the analyzer.

For a direct comparison of using the RIL method to provide evidence for biodegradation via IRMS or IRIS, non-labelled and $^{13}$C-labelled sodium bicarbonate were added at a ratio of 90:10 to an anaerobic toluene degradation experiment with *T. aromatica*. IRIS and IRMS measurements produced a similar decreasing trend of $x(13C)$ values, both indicating toluene mineralization (Figure 16d). Despite identical drops of $x(13C)$ values, IRIS data showed a systematic offset from IRMS data at $x(13C) = 0.145$ % on average. We are currently investigating the reasons for this unexpected difference. However, the variations could be related to the failure to produce the same concentration of injected CO$_2$ samples as working reference gases for IRMS.

3.3.4 Sensitivity of CO$_2$ quantification under anoxic conditions

The enrichment culture was cultivated in the presence of 10% $^{13}$C-labelled sodium bicarbonate as an example to investigate if the RIL method can be used to sensitively quantify CO$_2$ production based on IRIS measurement. Naphthalene was added as the sole organic carbon source while sulfate was supplied as the electron acceptor. In bottles inoculated with active cultures, we observed a reduction of 6.98 mM sulfate, concomitant with a production of 11.87 mM CO$_2$ (Figure 18). The ratio for sulfate consumption to CO$_2$ production (5.9:10) is very close to the theoretical stoichiometry (a ratio of 6:10) (Meckenstock et al. 2016). The high precision of measuring the $x(13C)$ values with IRIS together with the closed electron balance indicates that the RIL method is reliable to quantify the extent of microbial mineralization.
Figure 19. Anaerobic degradation of naphthalene by the enrichment culture N47. The added bicarbonate buffer (30 mM) consisted of non-labelled and \( ^{13} \text{C} \)-labelled sodium bicarbonate at a ratio of 90:10. The \( x^{(13\text{C})} \) values of DIC samples were measured by IRIS. Data are means of two or three parallel incubations. Error bars represent standard deviation of the biological replicates.

In control bottles with autoclaved cultures, a maximum variability of \( x^{(13\text{C})} = 0.05\% \) was determined during the whole incubation time, which depicts the experimental error due to the sampling process and analytical uncertainty. To conclude that there is distinct evidence of biodegradation in bottles with active cultures, depletion in \( ^{13} \text{C} \)-enrichment of CO\(_2\) \( (x^{(13\text{C})}_{\text{background}} - x^{(13\text{C})}_{\text{total}}) \) should be at least two times larger than the experimental error, meaning a production of 0.38 mM CO\(_2\) would be sufficient to provide evidence of biodegradation. This corresponds to 38 \( \mu \)M of naphthalene or 1.6\% of the 2.34 mM of added naphthalene and clearly demonstrates the sensitivity of the method. This small extent of biodegradation would be impossible to assess using other methods, except for methods using direct stable or radio isotope labelled substrates. However, further theoretical consideration of the precision of the analysis predicts that the sensitivity of the method could be improved by a factor of 5 to 10 concerning the developed CO\(_2\). For example, optimization of measurements by IRIS showed that contamination of ambient air in the autosampler system could greatly contribute to analytical uncertainties. In fact, adding a pre-flushing step to the autosampler system with CO\(_2\)-free synthetic air reduced the standard deviations of measured \( x^{(13\text{C})} \) from
0.028\% down to 0.008\% (n=12, samples were taken one by one from the same culture bottle).

### 3.3.5 Determining aerobic mineralization rate of GAC biomass

Biologically activated carbon filters, usually placed after the step of ozonation, are widely used in the drinking water production system for removal of undesired biodegradable DOC (Simpson 2008, Velten et al. 2011). During the steady-state of GAC filters when biomass concentration of the biofilm and effluent DOC concentration remain constant, the removal process is predominately performed by the indigenous microbial communities attached on the surface of activated carbon (i.e., GAC biofilms). One particular interest in this stage is to characterize the spatial distribution of microbial activities along the GAC filters, which can guide the operation and design of GAC filters. Most current studies focus on investigations of active biomass densities to indirectly describe it, e.g. adenosine tri-phosphate (ATP) analyses (Gibert et al. 2013, Velten et al. 2007, Wen et al. 2015). Here we show that our RIL method can be used as an alternative approach but to directly assess the aerobic mineralization potential of DOC by biomass attached on the GAC filters and resolve their spatial distribution profiles.

We simply cultivated BAC particles in lab microcosms filled with influent water (after ozonation) and 10\% $^{13}$C-labelled sodium bicarbonate. DIC samples were taken regularly and shifts in isotopic composition of DIC samples were measured by IRIS for calculation of CO$_2$ production. In 22 days, no observable CO$_2$ production could be found for incubations of only ozonized river water or multilayer filtered water (Figure 20). Insignificant amounts of CO$_2$ production were observed in incubations of fresh GAC, which was most likely caused by carbonate adsorbed to the surface, in agreement with the decrease of pH values. For active microcosms with incubations of GAC particles + influent water, however, we could observe significant CO$_2$ production. During the first two days, the degradation rate decreased and stayed almost linear afterwards (Figure 20). Thus, these data indicated that the RIL method can be used to monitor microbial DOC mineralization rates performed by the biofilm matrix attached to the solid surface of GAC.
Figure 20. Aerobic mineralization of DOC adsorbed on the GAC filters by biomass attached on GAC surfaces. The added bicarbonate buffer (10 mM) consisted of non-labelled and $^{13}$C-labelled sodium bicarbonate at a ratio of 90:10. The $\delta^{13}$C values of DIC samples were measured by IRIS. Data are means of three parallel incubations. Error bars represent standard deviation of the biological replicates.

As no clear DOC mineralization can be found in control microcosms so that the measured mineralization for active microcosms all came from biofilms attached on the granular activated carbon. The regression polynomials (2nd degree) for the first 24 hours of incubation were used to calculate the mineralization rates (i.e. regression slope) in each microcosm can be calculated. As shown in Figure 21, for GAC filters 1 and 2, the highest mineralization rate was observed at the second sampling point (20 cm depth from the top) and decreased by factors of 1.2 and 1.3 to the bottom of filters, respectively. From a spatial distribution point of view, the results are in line with previous studies where densities of biomass attached on the filters decreased along with the increase of the depth from the top to the bottom (Gibert et al. 2013, Velten et al. 2011). But for GAC filter 3 we observed the opposite trend. This can be due to the distinct types of GAC used for these three filters, which suggested the importance of choosing the right type of granular activated carbon.
Figure 21. Aerobic mineralization potential of DOC by biomass attached on GAC surfaces at different depths of the GAC filters. The added bicarbonate buffer (10 mM) consisted of non-labelled and $^{13}$C-labelled sodium bicarbonate at a ratio of 90:10. The $x(^{13}$C) values of DIC samples were measured by IRIS. Data are means of three parallel incubations. Error bars represent standard deviation of the biological replicates.
4. Discussion

A major question in biodegradation of benzene and naphthalene under anoxic conditions is (1) to answer who are the key players and how they complete the degradation process. Combined genomic and proteomic techniques were used to analyze an enrichment culture anaerobically utilizing benzene under sulfate-reducing conditions. The results provide evidence that benzene degradation is coupled to sulfate reduction in a single cell process, which is unusual for bacteria affiliated to *Pelotomaculum* genus (Chapter 4.1.).

Another important question is (2) to understand who else is there and what its function is in the community or ecosystem. A naphthalene-degrading, sulfate-reducing enrichment culture was taken as an example, and results from physiological, genomic, and proteomic analyses are presented indicating that environmental Spirochaetes recycle detrital biomass and convert it into end products or electrons that support growth of associated degraders. This finding emphasizes the importance of microbes that are not involved in the key processes (Chapter 4.2.).

A crucial step in studying microbial degradation of benzene and naphthalene is using microcosm experiments to prove the degradability and to obtain degradation rates. Of special interest is (3) to assess if the compounds are fully mineralized to CO₂ or if they are just transformed to metabolites which might be even more toxic in some cases. A novel reverse stable isotope labelling analysis is developed as a sensitive means to monitor biodegradation, focusing on hydrocarbon substrates (Chapter 4.3.).
4.1. A member of the genus *Pelotomaculum* oxidizes benzene to carbon dioxide with direct reduction of sulfate

Three types of initial activation reactions for anaerobic benzene degradation have been proposed: methylation to toluene, hydroxylation to phenol, and direct carboxylation to benzoate. Methylation of benzene to toluene was mostly based on the recovery of labeled toluene when labeled benzene was added to cultures (Aburto-Medina and Ball 2014, Meckenstock and Mouttaki 2011). However, the key product of anaerobic toluene activation, benzylsuccinate, has never been detected in benzene-degrading cultures, and so far no other evidence has been reported to support ring methylation as the initial activation reaction. There is also indication for hydroxylation of benzene to phenol as initial activation reaction by *Geobacter metallireducens* which metabolized benzene with Fe(III) citrate as the electron acceptor (Caldwell and Suflita 2000, Zhang et al. 2012, Zhang et al. 2013). Several groups suggested a direct carboxylation of benzene to benzoate as the first step for anaerobic benzene metabolism (Abu Laban et al. 2010, Holmes et al. 2011, Luo et al. 2016). In the Fe(III)-reducing enrichment culture BF, combined genomic and proteomic studies revealed genes encoding for two subunits of a putative anaerobic benzene carboxylase (AbcD and AbcA) (Abu Laban et al. 2010).

Based on proteogenomic analysis of the enrichment culture BPL, we propose that *Pelotomaculum* candidate BPL might not employ any of the three above-mentioned initial activation reactions proposed for benzene degradation. *Pelotomaculum* candidate BPL is not the only case of lacking genomic evidence for any of the activation reactions proposed so far. In a benzene-degrading methanogenic culture, benzene degradation was affiliated to a *Deltaproteobacterium* bacterium ORM2 (Devine 2013, Luo et al. 2016). There, the authors could not identify genes involved in toluene or phenol activation, either. Additionally, they failed to identify genes encoding for UbiD-like carboxylases and benzoate-CoA ligases. They therefore proposed an alternative carbonylation pathway for transformation of benzene to benzoyl-CoA via either a Wood-Ljungdhal-type reaction or a Gatterman-Koch-type mechanism (Devine 2013). Similarly, in a sulfate-reducing, benzene-degrading consortium enriched by Phelps
et al. (2001), benzoate was detected in culture supernatants but it could not support cell growth. Thus, the authors proposed that benzene was directly transformed to benzoyl-CoA (Phelps et al. 2001). In another study performed by Taubert et al. (2012), a similar benzene-degrading, sulfate-reducing enrichment culture was analyzed using protein-based stable isotope probing. Although this might be due to the poor detection limit of the method, they could not find proteins of known or proposed pathways for benzene activation.

Taking together, the genomic information on anaerobic benzene degradation from literature and the data presented here indicate a common observation for benzene activation. The iron- and nitrate-reducing bacteria known so far seem to activate benzene via carboxylation (Abu Laban et al. 2010, Luo et al. 2014). Fermenting- and sulfate-reducing bacteria including candidate Pelotomaculum BPL studied here might have developed so far unknown, oxygen-independent strategies for anaerobic benzene activation.

Members of the genus Pelotomaculum are mostly obligate syntrophs that need methanogens as syntrophic partners to oxidize propionate, alcohols, or aromatic compounds (Imachi et al. 2002, Stams and Plugge 2009). Recent studies propose that syntrophic metabolism during biodegradation of benzene and other hydrocarbons might occur in the presence of electron acceptors, such as sulfate or nitrate (Gieg et al. 2014). For example, stable isotope probing (SIP) with the Fe(III)-reducing enrichment culture BF revealed that members of the Peptococcaceae were the primary benzene degraders and shared electrons with members of the family Desulfobulbaceae (Kunapuli et al. 2007). Similar results were obtained for two denitrifying cultures based on either batch or chemostat studies (Luo et al. 2014, van der Zaan et al. 2012).

P. thermopropionicum is the only member whose genome has been sequenced in the genus Pelotomaculum. Similar to other known representatives in the genus, it lacks the capability to reduce sulfate, even though the related genes for sulfate reduction have been identified in the genome (Plugge et al. 2011). Previously, a draft genome for an uncultured Pelotomaculum spp. was retrieved in a terephthalate-degrading culture (Lykidis et al. 2011). Combined with the
subsequent proteomic studies, *Pelotomaculum* spp. was confirmed to degrade terephthalate via syntrophic interactions with methanogens (Wu et al. 2013). The given examples provide more and more evidence that members of the genus *Pelotomaculum* carry out syntrophic metabolism with anaerobic respiring microorganisms.

In another study, *Pelotomaculum* members were also proposed to degrade benzene in a similar sulfate-reducing enrichment culture (Herrmann et al. 2010). The authors hypothesized that benzene degradation by *Pelotomaculum* occurred by fermentation coupled to sulfate reduction performed by another organism in the enrichment culture (Herrmann et al. 2010, Rakoczy et al. 2011, Taubert et al. 2012). Interspecies hydrogen transfer was suggested to play an important role for benzene mineralization (Rakoczy et al. 2011).

However, by using the integrated genomic and proteomic data, we conclude in our study that the anaerobic benzene degradation in the enrichment culture BPL is performed as a single cell process by a novel member of the genus *Pelotomaculum* which is able to reduce sulfate as electron acceptor. This indicates that detection of *Pelotomaculum* species in benzene-contaminated areas therefore does not literally indicate the occurrence of benzene fermentation but can also be linked to direct sulfate-dependent benzene mineralization.
4.2. A hydrogen-driven subsurface microbial loop

Molecular analyses of microbial communities showed that environmental Spirochaetes are common and abundant in contaminated sites and in enrichment cultures degrading hydrocarbons and organohalides. Our study of several enrichment cultures now provides a functional trait to these frequently detected organisms. Their primary role is to degrade necromass via fermentation, resulting in production of hydrogen, short-chain fatty acids, and alcohols. These products are then utilized by sulfate-reducing bacteria and likely hydrogenotrophic methanogens, thereby recycling the necromass. Hydrocarbon-containing and organohalide-contaminated sites are extreme habitats where high concentrations of organic solvents challenge the integrity of cell membranes, inducing cell leakage or lysis (Rodriguez Martinez et al. 2008, Sherry et al. 2014). Thus, there might be high biomass turnover in such ecosystems (Griebler and Lueders 2009) and mechanisms to utilize electrons released as necromass might be important for ecosystem function. Such processes are likely to occur both in natural
hydrocarbon-containing environments and in ecosystems that have been contaminated with hydrocarbons and organohalides due to anthropogenic influences.

We thus propose that environmental Spirochaetes and similar ecotypes fill an important ecological niche by forming a subsurface microbial loop that recycles dead biomass and exudate (Figure 22). Through this loop, the electrons derived from necromass are continuously recycled by chemolithotrophs and chemoorganotrophs. This process alone will only support limited biomass production due to the high energy loss between trophic levels. Hydrocarbons and organohalides are therefore important as exogenous energy, electron, and carbon inputs in such systems. This differs from the traditional microbial loops described in marine and freshwater ecosystems, which are driven by photo-litho-auto-trophy rather than chemo-organo-hetero-trophy (Fenchel 2008). As electron donors are present in excess in hydrocarbon-contaminated aquifers, hydrocarbon degradation might be stimulated by the nutrients released from the microbial loop, as long as electron acceptors such as sulfate do not become rate-limiting. Thus, environmental Spirochaetes and other necromass-degrading microorganisms might contribute to enhancing biodegradation rates at contaminated plume fringes in groundwater or in oil reservoirs where the electron donors (e.g. hydrocarbons) and acceptors (e.g. sulfate or nitrate) meet in opposing gradients and where the highest degradation activity and biomass is found (Meckenstock et al. 2015). The subsurface microbial loop may also be important for recycling of nutrients such as phosphorus and nitrogen, similar to the well-described microbial loops of aquatic ecosystems (Fenchel 2008, Kerner et al. 2003).

The presented example of a simple microbial loop puts the importance of necromass and H₂-cycling in the subsurface into a new perspective. Moreover, the data reported here align with recent metagenomic studies on hydrogen-driven chemo-litho-auto-trophic deep subsurface processes and gaseous CO₂-saturated deeply sourced groundwater (Bagnoud et al. 2016, Probst et al. 2017). Altogether, this suggests the microbial loop as a generic theoretical concept for subsurface.
4.3. Reverse stable isotope labelling analysis as a novel tool to monitor biodegradation

Figure 23. Reverse stable isotope labelling analysis as a novel tool to (1) analyze if microorganisms in environmental samples or enrichment cultures can degrade specific organic compounds to CO₂, like hydrocarbon degradation experiments and (2) sensitively detect biodegradation of complex organic substrates such as DOC.

The novel RIL method can assess the biodegradation of organic compounds in a suit of different applications. A typical question for research or practice would be to analyze if microorganisms in environmental samples or enrichment cultures can degrade specific organic compounds to CO₂, similar to the hydrocarbon degradation experiments presented here (Figure 23). Compared to other methods such as following substrate disappearance, the RIL can proof the mineralization to CO₂. It does also not come along with problems of substrate adsorption to matrix or stoppers, evaporation or other unspecific losses. However, the method runs into limitations if other easily degradable organic carbon is present in the sample (sediment or water) leading to additional release of non-labelled CO₂. This is a common problem in microcosm experiments also appearing if e.g. the disappearance of electron acceptors such as molecular oxygen (respiration test) or production of sulfide or ferrous iron are employed to follow biodegradation activity. In many cases, this problem is solved by appropriate controls without substrate if the organic carbon background is not too high.
However, this also turns into a big potential of RIL: the very sensitive detection of biodegradation of complex organic substrates such as DOC, as demonstrated in our experiments with granular activated carbon (Figure 23). Measuring DOC degradation is notoriously difficult and needs severe experimental efforts. Due to its high sensitivity, RIL can detect even tiny degradation activities that are difficult to assess, otherwise. Only using labelled organic compounds could compete with this sensitivity, but organics such as DOC cannot be labelled. RIL has the advantage that it can analyze degradation of natural organic matter in such environmental samples even under pristine conditions by only adding of a little $^{13}$C-labelled DIC. Another big advantage is that degradation of compounds can be analyzed that are adsorbed to solid matrices, which is very difficult or almost impossible to measure with other methods at this precision and sensitivity. In the case of e.g. wood or plastics degradation, this could also include the degradation of the solid itself.

As demonstrated here, RIL performs under different biodegradation scenarios. Extremely slow biodegradation experiments with very difficult, anaerobic degradation of polycyclic aromatic hydrocarbons can be analyzed as well as faster, aerobic degradation of DOC. The precise analysis of carbon stable isotope ratios by IRMS or IRIS allows for the detection of very small CO$_2$ production. This makes RIL much more sensitive compared to directly measuring DIC concentrations or respiration tests (e.g. ISO 9408 and OECD 301B).

If the production of DIC/CO$_2$ during microbial degradation is large, e.g. by fast-growing aerobic microorganisms and higher amounts of organic substrates, conventional GC-MS can be used for analyzing degradation of stable isotope-labelled compounds (Liou et al. 2008). For tiny or slow biodegradation, however, GC-MS cannot compete with tools for analyzing stable isotope ratios of CO$_2$ such as IRMS and IRIS. IRMS does not produce accurate results if the carbon isotope ratios are too high and out of the linearity range. Thus, the Delta Ray™ IRIS provides the easiest and cheapest as well as the most precise analysis for $^{13}$C/$^{12}$C ratios much higher than natural abundance.
A limitation for RIL is the need of closed systems. For example, aerobic incubations with continuous oxygen supply or open microcosms cannot be assessed by this method. In most cases, however, one can perform the same experiment in closed, stoppered vials when providing sufficient headspace (Bahr et al. 2015). As the major advantage of RIL is assessing small amounts of biodegradation, aerobic degradation of big amounts of substrates will anyway not be a typical field of application.

It is important that CO$_2$ fixation by photosynthetic organisms, other autotrophs, or heterotrophic CO$_2$ fixation does not affect RIL experiments. The release of CO$_2$ during biodegradation changes the stable isotope ratio of the DIC because of the additional amount of $^{12}$CO$_2$. CO$_2$ fixation, however, does not change the isotope ratio of the DIC because both isotopomers are consumed with the same rate. Stable isotope fractionation effects by discrimination of $^{13}$CO$_2$ during fixation are orders of magnitude smaller than the changes in the isotope ratios observed here and do not have to be accounted for. This even applies to systems such as methanogenesis where CO$_2$ is both released and consumed. While the release of CO$_2$ during substrate degradation by fermenting organisms changes the stable isotope ratio of the DIC indicating biodegradation the consumption by methanogen does not significantly.
5. Conclusions and Outlook

5.1. General conclusions

In this thesis, new insights into anaerobic benzene and naphthalene degradation were obtained through a series of experiments, with the aid of omics techniques, physiological experiments, and stable isotopes.

(1) Proteogenomic analyses of *Pelotomaculum* candidate BPL shed new light on anaerobic benzene degradation by sulfate-reducing microorganisms and in particular by the genus *Pelotomaculum*. The member of the genus *Pelotomaculum* identified here, candidate strain BPL, did not only degrade benzene in the absence of molecular oxygen but also performed sulfate reduction, making it the first representative of the genus *Pelotomaculum* which has the capacity to grow by dissimilatory sulfate reduction. Comparison with other benzene-degrading cultures indicates that such co-cultures might be wide spread. Detection of *Pelotomaculum* species in cultures or natural samples therefore does not literally indicate fermentation, especially in the case of the anaerobic hydrocarbon degradation.

(2) The current view upon anaerobic degradation processes is mostly focused on key player perspective. However, plenty of molecular studies, and recently metagenomics studies indicate some microbial species exist in high abundances in the community but they are apparently not involved in the key degradation process. By the example of environmental Spirochaetes who recycle detrital biomass in a simple subsurface microbial loop, an explanation was provided for these microbes. They are most likely involved in necromass recycling providing electrons and nutrients to associated degraders. Thus, they are similarly important as the key organisms. Such loops are likely to be especially important in hydrocarbon-containing sites and other extreme habitats where physicochemical conditions challenge bacterial integrity as well as in oligotrophic subsurface habitats like the deep terrestrial subsurface.

(3) Assessing the mineralization of many substrates like PAHs and crude oil mixtures can be very difficult because of issues such as low water solubility and
complexity. While CO2 production can be readily measured by starting with radiolabeled substrates (14C), special facilities are typically required. In contrast, stable isotopes are safer and easier to handle but aside from a few 13C-labeled substrates (especially hydrocarbons) are not commercially available. Compared to them, the advantage of the RIL method lies in the very precise analysis of the 13C/12C ratio which allows for the detection of very small CO2 production and in the relatively cheap addition of 13CO2 instead of 13C-labelled organic compounds. This could also be measured with classical IRMS but the application of the new Delta Ray promises a much easier and cheaper analysis. Mass balance can be used to denote equations to calculate microbial CO2 production. The method can be used in a wide range of applications, such as monitoring crude oil degradation under anoxic conditions and discovering organic carbon turnover under aerobic conditions.

5.2. Future perspective

Cultivation-independent genomics has greatly improved our understanding of microbial physiology and ecology. Especially, relying on the fast development of bioinformatics methods, enormous amounts of draft genomes with high qualities have been produced these days. More strikingly, there are even a lot of candidate phyla found with no cultivated representatives (Hug et al. 2016). With these genomic data, one can easily reconstruct metabolic pathways as shown in this thesis for Pelotomaculum candidate BPL and R. cohabitans. But it should be noted that they can only provide potential metabolic functions. Further solid biochemistry should be warranted to provide unambiguous evidence, as shown for experiments of R. cohabitans in this thesis.

Besides, a lot of predicted genes in the genome cannot be annotated properly because we just do not know their functions. For example, one goal of the thesis at the beginning is to understand the initial activation mechanism for benzene degradation. However, in this thesis I failed to provide in-depth information about it. The most important reason is due to the limited availability of closely related known proteins. This makes a definitive annotation of the sequences currently unfeasible. Therefore, it is suggested that the further study on the enrichment
culture BPL or other benzene-degrading cultures is to isolate a pure culture for benzene degradation. The cultivation efforts can be not only guided with the genomic data but also beneficial from novel isolation methods (Sizova et al. 2012, Solden et al. 2016). However, most likely the isolation process is extremely difficult and it is even impossible as microbes can only live via interactions with their partners. In this case, we have to play with mixed cultures combining gold standard methods for pure cultures and novel techniques, e.g. comparative omics (transcriptomics and proteomics), stable isotope probing, and high-resolution microscopy.
6. References


Authorship Clarification

I) Reconstructing metabolic pathways of a member of the genus *Pelotomaculum* suggesting its potential to oxidize benzene to carbon dioxide with direct reduction of sulfate

In Part I, metaproteogenomic analysis was performed on an enrichment culture BPL which carried out benzene oxidation coupled with sulfate reduction. The PhD candidate together with Rainer U. Meckenstock developed the concept of this study and planned laboratory experiments. Kerstin Nicolaisen extracted DNA for the metagenome and protein for one of the metaproteomic analyses (the other one done by the PhD candidate). Thomas Weinmaier conducted initial analyses of the metagenome and Johannes Dröge finished follow-up bioinformatics analysis, including binning process and phylogenetic analysis. Proteins identification was done by Christine von Toerne. The PhD candidate performed data processing with both metagenome and metaproteome. Sviatlana Marozava provided help with metaproteomic analysis. The work was published in FEMS Microbiology Ecology (Dong et al. 2017). The original manuscript draft for Part I was developed by the PhD candidate. All co-authors in this article contributed to writing, reviewing and editing of the manuscript.


II) Fermentative Spirochaetes drive nutrient cycling by a subsurface microbial loop in hydrocarbon-contaminated habitats

In Part II, the functional role of Spirochaetes was explored in contaminated habitats. The PhD candidate together with Rainer U. Meckenstock developed the concept of this study and planned laboratory experiments. The PhD candidate conducted the experiments and data analyses. Chris Greening performed phylogenetic analysis of hydrogenases and revised 16S rRNA gene sequences analysis of Spirochaetes. Ralf
Conrad contributed to gas chromatography analysis of hydrogen concentrations. Thomas Brüls contributed to the part of sequencing and reconstruction of Spirochaetes genomes. Svenja Blaskowski, Farnusch Kaschani, and Markus Kaiser contributed to mass spectrometry-based proteome analyses. The original manuscript draft for Part II was developed by the PhD candidate. All co-authors contributed to writing, reviewing and editing of the manuscript.


**III) Monitoring microbial mineralization using reverse stable isotope labelling analysis by mid-infrared laser spectroscopy**

In Part III, a novel method was developed with the name of reverse stable isotope labelling analysis. The PhD candidate together with Rainer U. Meckenstock developed the concept of this study and planned laboratory experiments. Maik A. Jochmann, Martin Elsner, and Armin H. Meyer contributed to experimental parts of Isotope Ratio Mass Spectrometry analyses. The experiments for applying RIL method to determine aerobic mineralization rates were conducted by the bachelor students Leonard E. Bäcker and Mona Rahmatullah. They were supervised by Rainer U. Meckenstock and the PhD candidate. Daniel Schunk and Guido Lens provided the pilot plant used for drinking water production and helped design the experiments for DOC removal. All other experiments were conducted and analyzed by the PhD candidate. The original manuscript draft for Part III was developed by the PhD candidate. All co-authors contributed to writing, reviewing and editing of the manuscript.


**IV) Anaerobic degradation of benzene and polycyclic aromatic hydrocarbons**
This is a review article (Meckenstock et al. 2017) published in a special issue of Journal of Molecular Microbiology and Biotechnology. The outline of it was developed by Rainer U. Meckenstock. The PhD candidate wrote “Table 1” and the Section “Stable Isotope Fractionation to Assess Biodegradation of Benzene and PAHs”, which are parts of the introduction of the current PhD thesis (Table 1 in Chapter 1.1 and Chapter 1.6.1). They were revised and edited by Rainer U. Meckenstock.

CURRICULUM VITAE

Name: Xiyang Dong
Address: Bäuminghausstr. 150, 45326 Essen, Germany
Telephone no.: +49 17661433583
E-Mail: xiyang.dong@uni-due.de, thegreatlake@126.com
Date of birth: Jan 3rd, 1988
Place of birth: Shandong, China
Nationality: Chinese

Education

Mar 2015 – Sep 2017 Guest PhD student at Fakultät für Chemie - Biofilm Centre, University of Duisburg-Essen, Germany
Supervisor: Prof. Rainer Meckenstock

Oct 2013 – Sep 2017 PhD student at Institute of Groundwater Ecology, Helmholtz Zentrum München, Germany
Supervisor: Prof. Rainer Meckenstock

Sep 2010 – Jun 2013 Master of engineering in Environmental Engineering, Dalian University of Technology, China
Supervisor: Prof. Jiti Zhou and Associate Prof. Yu Zhang

Sep 2006 – Jul 2010 Bachelor of engineering in Environmental Engineering, Nanjing Normal University, China
List of Scientific Communications

Scientific Communications Related to the Dissertation

Peer-reviewed publications


Submitted manuscripts for peer-review


Selected conference contributions

1. **Dong X** (2017) The functional role of environmental spirochetes at contaminated sites. *Oral presentation* at Microbiology and Infection 2017 - 5th Joint Conference of the German Society for Hygiene and Microbiology (DGHM) and the Association for General and Applied Microbiology (VAAM), Würzburg, Germany.

3. **Dong X** (2016) Metaproteogenomic analysis revealed a member of the genus *Pelotomaculum* completely oxidizing benzene to carbon dioxide with direct reduction of sulfate. **Poster presentation** at the Annual Conference 2016 of the Association for General and Applied Microbiology (VAAM), Jena, Germany.

**Other Scientific Communications Not Related to the Dissertation**

**Peer-reviewed publications**


Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor Prof. Rainer Meckenstock. Being a great mentor, he was always willing to assist and provide support and guidance whenever it was needed. He also gave me opportunities to grow skills. His passion and motivation in science greatly inspired me. More importantly, he provided career advice and facilitated opportunities on my career development.

I would like to thank Dr. Tillmann Lüders for being my official supervisor at TUM after Rainer moved to UDE and providing very helpful suggestions on my thesis. I am very grateful to Prof. Wolfgang Liebl for taking over the chair of my doctoral examination committee. Thanks also go to Prof. Siegfried Scherer for being the second examiner. And special thanks to Dr. Maik Jochmann for being a member of my personal Thesis Committee. Much gratitude to Dr. Alfredo Perez de Mora for taking care of me during my first several months in Germany.

Much gratitude to Dr. Johannes Dröge for binning process, Dr. Chris Greening for hydrogenase analysis, and Dr. Thomas Brüls for bioinformatics analysis. Special thanks to the following people for proteome analyses: Dr. Christine von Toerne, Dr. Sviatlana Marozava, Svenja Blaskowski and Dr. Farnusch Kaschani. Many thanks to Prof. Ralf Conrad and Melanie Klose for hydrogen measurement. Thanks Dr. Martin Elsner and Dr. Armin Meyer for IRMS analysis. Thanks also goes to Leonard Bäcker and Mona Rahmatullah for their work on DOC removal project.

I would like to thank Dr. Verena Brauer for translating the abstract into German, and Dr. Hubert Müller for proof-reading my thesis. My gratitude also goes to all group members in the lab of Meckenstock for creating a positive working environment.

I dedicate this thesis to my parents for their unconditional love and support, and to my beloved girlfriend Qing for her love, friendship, help, understanding and patience. Finally, I would acknowledge the Chinese Scholarship Council for financial support.
Appendix

**Supporting Tables**

Table A1 Expressed genes belonging to *Pelotomaculum* candidate BPL during anaerobic degradation of benzene.

Table A2 List of genes identified in the genomes and used for metabolic pathway reconstruction for Spirochaetes in hydrocarbon- and organohalide-contaminated environments shown in Figure 11. Genes for the illustrated pathways were detected in the genomes of *Rectinrema cohabitans* HM, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8.

Table A3 Proteomic analysis of *R. cohabitans* grown on glucose. Four cultures of *R. cohabitans* strain HM were grown with 10 mM glucose and 0.1% yeast extract. Considering the size is too large to include here, the mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (https://www.ebi.ac.uk/pride/archive/) with the dataset identifier PXD005624.

Table A4 List of identified proteins for *R. cohabitans* HM as shown in Figure 14. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (https://www.ebi.ac.uk/pride/archive/) with the dataset identifier PXD005624.
Table A1: Expressed genes belonging to *Pelotomaculum* candidate BPL during anaerobic degradation of benzene.

<table>
<thead>
<tr>
<th>Gene ID in Pelotomaculum spp. Genome</th>
<th>IMG product name</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sulfate reduction</strong></td>
<td></td>
</tr>
<tr>
<td>2617919556</td>
<td>putative adenylylsulfate reductase-associated electron transfer protein QmoB</td>
</tr>
<tr>
<td>2617919557</td>
<td>putative adenylylsulfate reductase-associated electron transfer protein QmoA</td>
</tr>
<tr>
<td>2617919558</td>
<td>adenylylsulfate reductase subunit A</td>
</tr>
<tr>
<td>2617919559</td>
<td>dissimilatory adenylylsulfate reductase beta subunit (EC 1.8.99.2)</td>
</tr>
<tr>
<td>2617919560</td>
<td>ATP sulfurylase (sulfate adenylyltransferase)</td>
</tr>
<tr>
<td>2617919937</td>
<td>K(+)‐stimulated pyrophosphate‐energized sodium pump</td>
</tr>
<tr>
<td>2617920254</td>
<td>Dissimilatory sulfite reductase (desulfoviridin), gamma subunit</td>
</tr>
<tr>
<td>2617920256</td>
<td>Nitrate reductase gamma subunit</td>
</tr>
<tr>
<td>2617920264</td>
<td>dissimilatory sulfite reductase beta subunit</td>
</tr>
<tr>
<td>2617920265</td>
<td>dissimilatory sulfite reductase alpha subunit (EC 1.8.99.1)</td>
</tr>
<tr>
<td>2617921751</td>
<td>NADH:ubiquinone oxidoreductase subunit 2 (chain N)</td>
</tr>
<tr>
<td>2617921752</td>
<td>NADH:ubiquinone oxidoreductase subunit 4 (chain M)</td>
</tr>
<tr>
<td>2617921753</td>
<td>NADH‐quinone oxidoreductase subunit L</td>
</tr>
<tr>
<td>2617921755</td>
<td>NADH:ubiquinone oxidoreductase subunit 6 (chain J)</td>
</tr>
<tr>
<td>2617921756</td>
<td>NADH‐quinone oxidoreductase subunit I</td>
</tr>
<tr>
<td>2617921757</td>
<td>NADH:ubiquinone oxidoreductase subunit 1 (chain H)</td>
</tr>
<tr>
<td>2617921758</td>
<td>NADH‐quinone oxidoreductase subunit D</td>
</tr>
<tr>
<td>2617921761</td>
<td>NADH‐quinone oxidoreductase subunit A</td>
</tr>
<tr>
<td><strong>Wood-Ljungdahl pathway</strong></td>
<td></td>
</tr>
<tr>
<td>2617920380</td>
<td>5,10-methylene-tetrahydrofolate dehydrogenase/Methenyl tetrahydrofolate cyclohydrolase</td>
</tr>
<tr>
<td>2617920888</td>
<td>Ni-dependent carbon monoxide dehydrogenase precursor</td>
</tr>
<tr>
<td>2617921113</td>
<td>methyltetrahydrofolate–corrinoid iron-sulfur protein Co-methyltransferase</td>
</tr>
<tr>
<td>2617921126</td>
<td>5,10-methyltetrahydrofolate reductase</td>
</tr>
<tr>
<td>2617921323</td>
<td>Formyltetrahydrofolate synthetase</td>
</tr>
<tr>
<td>2617921376</td>
<td>CO dehydrogenase/acetyl-CoA synthase delta subunit (corrinoid Fe-S protein)</td>
</tr>
<tr>
<td>2617921377</td>
<td>CO dehydrogenase maturation factor</td>
</tr>
<tr>
<td>2617921378</td>
<td>CO dehydrogenase/acetyl-CoA synthase gamma subunit (corrinoid Fe-S protein)</td>
</tr>
<tr>
<td>2617922225</td>
<td>formate dehydrogenase, alpha subunit</td>
</tr>
<tr>
<td><strong>benzoyl-CoA degradation pathway</strong></td>
<td></td>
</tr>
<tr>
<td>2617919856</td>
<td>6-ketocyclohex-1-ene-1-carbonyl-CoA hydrolase</td>
</tr>
<tr>
<td>2617919857</td>
<td>cyclohexa-1,5-dienecarbonyl-CoA hydratase</td>
</tr>
<tr>
<td>2617919858</td>
<td>6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase</td>
</tr>
<tr>
<td>2617919859</td>
<td>Acyl-CoA dehydrogenases</td>
</tr>
<tr>
<td>2617921884</td>
<td>tungsten-dependent benzoyl-CoA reductase subunit bamB</td>
</tr>
<tr>
<td>2617921892</td>
<td>acetyl-CoA synthetase</td>
</tr>
<tr>
<td>2617921893</td>
<td>Enoyl-CoA hydratase/carnithine racemase</td>
</tr>
<tr>
<td>2617921894</td>
<td>3-hydroxyacyl-CoA dehydrogenase</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Function</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2617921895</td>
<td>Acyl dehydratase</td>
</tr>
<tr>
<td>2617921991</td>
<td>tungsten-dependent benzoyl-CoA reductase-related protein bamD</td>
</tr>
<tr>
<td><strong>unclassified enzymes</strong></td>
<td></td>
</tr>
<tr>
<td>2617919427</td>
<td>Hemerythrin HHE cation binding domain-containing protein</td>
</tr>
<tr>
<td>617921157</td>
<td>Putative zinc-finger</td>
</tr>
<tr>
<td>2617919439</td>
<td>TusA-related sulfurytransferase</td>
</tr>
<tr>
<td>2617919500</td>
<td>Oxygen-sensitive ribonucleoside-triphosphate reductase</td>
</tr>
<tr>
<td>2617919503</td>
<td>S-adenosylhomocysteine hydrolase</td>
</tr>
<tr>
<td>2617919510</td>
<td>Acyl carrier protein</td>
</tr>
<tr>
<td>2617919522</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919530</td>
<td>large subunit ribosomal protein L28</td>
</tr>
<tr>
<td>2617919532</td>
<td>6,7-dimethyl-8-ribityllumazine synthase</td>
</tr>
<tr>
<td>2617919573</td>
<td>Aspartokinases</td>
</tr>
<tr>
<td>2617919577</td>
<td>Stage IV sporulation protein A (spore_IV_A)</td>
</tr>
<tr>
<td>2617919589</td>
<td>4-hydroxy-3-methylbut-2-enyl diphosphate reductase</td>
</tr>
<tr>
<td>2617919592</td>
<td>3-phosphoshikimate 1-carboxyvinyltransferase</td>
</tr>
<tr>
<td>2617919594</td>
<td>-deoxygen-D-arabinopyrophosphate-7-phosphate synthase</td>
</tr>
<tr>
<td>2617919604</td>
<td>Protein of unknown function (DUF1657)</td>
</tr>
<tr>
<td>2617919613</td>
<td>Pyruvate kinase</td>
</tr>
<tr>
<td>2617919632</td>
<td>homocitrate synthase NifV</td>
</tr>
<tr>
<td>2617919635</td>
<td>aspartylglutamyl-tRNA(Asn/Gln) amidotransferase subunit C</td>
</tr>
<tr>
<td>2617919637</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919644</td>
<td>Copper amine oxidase N-terminal domain</td>
</tr>
<tr>
<td>2617919645</td>
<td>Copper amine oxidase N-terminal domain</td>
</tr>
<tr>
<td>2617919656</td>
<td>YeeE/YedE family (DUF395)</td>
</tr>
<tr>
<td>2617919662</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919670</td>
<td>RND family efflux transporter, MFP subunit</td>
</tr>
<tr>
<td>2617919674</td>
<td>Predicted Fe-Mo cluster-binding protein, NIX family</td>
</tr>
<tr>
<td>2617919678</td>
<td>Subtilisin-like serine proteases</td>
</tr>
<tr>
<td>2617919680</td>
<td>Cell fate regulator YlbF, YheA/YmcA/DUF963 family (controls sporulation, competence, biofilm development)</td>
</tr>
<tr>
<td>2617919682</td>
<td>Cell wall-associated hydrolases (invasion-associated proteins)</td>
</tr>
<tr>
<td>2617919702</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919703</td>
<td>Molybdopterin biosynthesis enzyme</td>
</tr>
<tr>
<td>2617919704</td>
<td>Molybdopterin biosynthesis enzyme</td>
</tr>
<tr>
<td>2617919705</td>
<td>S-layer homology domain</td>
</tr>
<tr>
<td>2617919708</td>
<td>molybdopterin adenylyltransferase</td>
</tr>
<tr>
<td>2617919709</td>
<td>Co-chaperonin GroES (HSP10)</td>
</tr>
<tr>
<td>2617919710</td>
<td>Chaperonin GroEL (HSP60 family)</td>
</tr>
<tr>
<td>2617919717</td>
<td>Benzoyl-CoA reductase/2-hydroxyglutaryl-CoA dehydratase subunit, BcrC/BadD/HgdB</td>
</tr>
<tr>
<td>2617919759</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919780</td>
<td>dTDP-4-dehydrosharrnose 3,5-epimerase and related enzymes</td>
</tr>
<tr>
<td>2617919818</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2617919824</td>
<td>ABC-type phosphate transport system, periplasmic component</td>
</tr>
<tr>
<td>2617919828</td>
<td>phosphate uptake regulator, PhoU</td>
</tr>
<tr>
<td>2617919842</td>
<td>Predicted DNA-binding protein with PD1-like DNA-binding motif - COG1661</td>
</tr>
<tr>
<td>2617919843</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919851</td>
<td>Sporulation and spore germination</td>
</tr>
<tr>
<td>2617919853</td>
<td>3-oxoacyl-[acyl-carrier protein] reductase</td>
</tr>
<tr>
<td>2617919854</td>
<td>Acetyl-CoA acetyltransferase</td>
</tr>
<tr>
<td>2617919855</td>
<td>Predicted nucleic-acid-binding protein containing a Zn-ribbon - COG1545</td>
</tr>
<tr>
<td>2617919865</td>
<td>membrane fusion protein, cobalt-zinc-cadmium efflux system</td>
</tr>
<tr>
<td>2617919870</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919871</td>
<td>Transcriptional regulator</td>
</tr>
<tr>
<td>2617919884</td>
<td>Nucleoside-diphosphate-sugar epimerases</td>
</tr>
<tr>
<td>2617919891</td>
<td>S-adenosylmethionine synthetase</td>
</tr>
<tr>
<td>2617919895</td>
<td>Cold shock proteins</td>
</tr>
<tr>
<td>2617919896</td>
<td>putative sigma-54 modulation protein</td>
</tr>
<tr>
<td>2617919902</td>
<td>Transketolase, C-terminal subunit</td>
</tr>
<tr>
<td>2617919903</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919910</td>
<td>Predicted RNA-binding protein - COG1532</td>
</tr>
<tr>
<td>2617919923</td>
<td>bifunctional phosphoglucose/phosphomannose isomerase</td>
</tr>
<tr>
<td>2617919931</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase</td>
</tr>
<tr>
<td>2617919932</td>
<td>3-phosphoglycerate kinase</td>
</tr>
<tr>
<td>2617919933</td>
<td>Triosephosphate isomerase</td>
</tr>
<tr>
<td>2617919935</td>
<td>Enolase</td>
</tr>
<tr>
<td>2617919952</td>
<td>Small-conductance mechanosensitive channel</td>
</tr>
<tr>
<td>2617919957</td>
<td>Phosphoribosylformimino-5-aminimidazole carboxamide ribonucleotide (ProFAR) isomerase</td>
</tr>
<tr>
<td>2617919966</td>
<td>AICAR transformylase/IMP cyclohydrolase PurH (only IMP cyclohydrolase domain in Aful)</td>
</tr>
<tr>
<td>2617919972</td>
<td>Phosphoribosylcarboxymimidazole (NCAIR) mutase</td>
</tr>
<tr>
<td>2617919980</td>
<td>S-layer homology domain</td>
</tr>
<tr>
<td>2617919989</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919992</td>
<td>Rhodanese-related sulfurtransferase</td>
</tr>
<tr>
<td>2617919993</td>
<td>Chitobiase/beta-hexosaminidase C-terminal domain</td>
</tr>
<tr>
<td>2617919996</td>
<td>Ig-like domain-containing protein</td>
</tr>
<tr>
<td>2617919997</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617919998</td>
<td>Cohesin domain</td>
</tr>
<tr>
<td>2617919999</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920000</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920001</td>
<td>Chitobiase/beta-hexosaminidase C-terminal domain</td>
</tr>
<tr>
<td>2617920002</td>
<td>S-layer homology domain</td>
</tr>
<tr>
<td>2617920003</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920007</td>
<td>Ribosomal protein S9</td>
</tr>
<tr>
<td>2617920017</td>
<td>Uncharacterized protein conserved in bacteria</td>
</tr>
<tr>
<td>2617920021</td>
<td>Acyl-CoA dehydrogenases</td>
</tr>
<tr>
<td>ID</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2617920032</td>
<td>Protein of unknown function (DUF2680)</td>
</tr>
<tr>
<td>2617920037</td>
<td>Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase</td>
</tr>
<tr>
<td>2617920039</td>
<td>Pyridoxine biosynthesis enzyme</td>
</tr>
<tr>
<td>2617920046</td>
<td>DNA polymerase sliding clamp subunit (PCNA homolog)</td>
</tr>
<tr>
<td>2617920084</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920087</td>
<td>Aspartyl aminopeptidase</td>
</tr>
<tr>
<td>2617920092</td>
<td>Uncharacterized conserved protein - COG3874</td>
</tr>
<tr>
<td>2617920122</td>
<td>Uncharacterized conserved protein - COG0432</td>
</tr>
<tr>
<td>2617920126</td>
<td>DNA-binding transcriptional regulator, Lrp family</td>
</tr>
<tr>
<td>2617920143</td>
<td>LysM domain</td>
</tr>
<tr>
<td>2617920195</td>
<td>NADPH-dependent 2,4-dienoyl-CoA reductase, sulfur reductase</td>
</tr>
<tr>
<td>2617920210</td>
<td>F0F1-type ATP synthase, epsilon subunit (mitochondrial delta subunit)</td>
</tr>
<tr>
<td>2617920211</td>
<td>F0F1-type ATP synthase, beta subunit</td>
</tr>
<tr>
<td>2617920212</td>
<td>F0F1-type ATP synthase, gamma subunit</td>
</tr>
<tr>
<td>2617920213</td>
<td>F0F1-type ATP synthase, alpha subunit</td>
</tr>
<tr>
<td>2617920214</td>
<td>F0F1-type ATP synthase, delta subunit (mitochondrial oligomycin sensitivity protein)</td>
</tr>
<tr>
<td>2617920215</td>
<td>F0F1-type ATP synthase, subunit b</td>
</tr>
<tr>
<td>2617920216</td>
<td>F0F1-type ATP synthase, subunit a</td>
</tr>
<tr>
<td>2617920236</td>
<td>Transaldolase</td>
</tr>
<tr>
<td>2617920237</td>
<td>Fructose/tagatose bisphosphate aldolase</td>
</tr>
<tr>
<td>2617920239</td>
<td>Response regulator containing CheY-like receiver, AAA-type ATPase, and DNA-binding domains</td>
</tr>
<tr>
<td>2617920245</td>
<td>prevent-host-death family protein</td>
</tr>
<tr>
<td>2617920247</td>
<td>Transcriptional regulators</td>
</tr>
<tr>
<td>2617920255</td>
<td>Fe-S oxidoreductase</td>
</tr>
<tr>
<td>2617920257</td>
<td>4Fe-4S dicluster domain-containing protein</td>
</tr>
<tr>
<td>2617920270</td>
<td>UDP-N-acetylglucosamine enolpyruvyl transferase</td>
</tr>
<tr>
<td>2617920279</td>
<td>Adenylosuccinate synthase</td>
</tr>
<tr>
<td>2617920283</td>
<td>Ribosomal protein L9</td>
</tr>
<tr>
<td>2617920287</td>
<td>Single-stranded DNA-binding protein</td>
</tr>
<tr>
<td>2617920289</td>
<td>Regulators of stationary/sporulation gene expression</td>
</tr>
<tr>
<td>2617920301</td>
<td>Ribosomal protein L17</td>
</tr>
<tr>
<td>2617920302</td>
<td>DNA-directed RNA polymerase, alpha subunit/40 kD subunit</td>
</tr>
<tr>
<td>2617920304</td>
<td>Ribosomal protein S11</td>
</tr>
<tr>
<td>2617920312</td>
<td>Ribosomal protein S5</td>
</tr>
<tr>
<td>2617920314</td>
<td>large subunit ribosomal protein L6</td>
</tr>
<tr>
<td>2617920315</td>
<td>Ribosomal protein S8</td>
</tr>
<tr>
<td>2617920324</td>
<td>Ribosomal protein L22</td>
</tr>
<tr>
<td>2617920330</td>
<td>Ribosomal protein S10</td>
</tr>
<tr>
<td>2617920332</td>
<td>type IV pilus assembly protein PilA</td>
</tr>
<tr>
<td>2617920353</td>
<td>Predicted Zn-dependent proteases and their inactivated homologs - COG0312</td>
</tr>
<tr>
<td>2617920358</td>
<td>elongation factor P</td>
</tr>
<tr>
<td>2617920361</td>
<td>NADPH-dependent 2,4-dienoyl-CoA reductase, sulfur reductase</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------</td>
</tr>
<tr>
<td>2617920371</td>
<td>oxaloacetate decarboxylase, alpha subunit</td>
</tr>
<tr>
<td>2617920373</td>
<td>Uncharacterized conserved protein YloU, alkaline shock protein (Asp23) family</td>
</tr>
<tr>
<td>2617920377</td>
<td>ferredoxin–NADP+ reductase</td>
</tr>
<tr>
<td>2617920378</td>
<td>glutamate synthase (NADPH/NADH) small chain</td>
</tr>
<tr>
<td>2617920381</td>
<td>Formiminotetrahydrofolate cyclodeaminase</td>
</tr>
<tr>
<td>2617920382</td>
<td>Exonuclease VII small subunit</td>
</tr>
<tr>
<td>2617920414</td>
<td>Uncharacterized conserved protein - COG3360</td>
</tr>
<tr>
<td>2617920427</td>
<td>Putative sterol carrier protein</td>
</tr>
<tr>
<td>2617920434</td>
<td>Aspartate-semialdehyde dehydrogenase</td>
</tr>
<tr>
<td>2617920436</td>
<td>Dihydridopicolinate synthase/N-acetylneuraminate lyase</td>
</tr>
<tr>
<td>2617920444</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920456</td>
<td>stage V sporulation protein S</td>
</tr>
<tr>
<td>2617920469</td>
<td>IMP dehydrogenase/GMP reductase</td>
</tr>
<tr>
<td>2617920539</td>
<td>S-layer homology domain</td>
</tr>
<tr>
<td>2617920542</td>
<td>Negative regulator of flagellin synthesis (anti-sigma28 factor)</td>
</tr>
<tr>
<td>2617920549</td>
<td>Flagellar capping protein</td>
</tr>
<tr>
<td>2617920552</td>
<td>Flagellin and related hook-associated proteins</td>
</tr>
<tr>
<td>2617920568</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920570</td>
<td>Stress response protein SCP2</td>
</tr>
<tr>
<td>2617920571</td>
<td>tellurium resistance protein TerD</td>
</tr>
<tr>
<td>2617920573</td>
<td>tellurium resistance protein TerD</td>
</tr>
<tr>
<td>2617920589</td>
<td>Superoxide dismutase</td>
</tr>
<tr>
<td>2617920615</td>
<td>Electron transfer flavoprotein, alpha subunit</td>
</tr>
<tr>
<td>2617920616</td>
<td>Electron transfer flavoprotein, beta subunit</td>
</tr>
<tr>
<td>2617920631</td>
<td>Cell division GTPase</td>
</tr>
<tr>
<td>2617920668</td>
<td>NADH-quinone oxidoreductase subunit F</td>
</tr>
<tr>
<td>2617920708</td>
<td>Outer membrane efflux protein</td>
</tr>
<tr>
<td>2617920709</td>
<td>Copper amine oxidase N-terminal domain</td>
</tr>
<tr>
<td>2617920710</td>
<td>Copper amine oxidase N-terminal domain</td>
</tr>
<tr>
<td>2617920712</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920728</td>
<td>NifU homolog involved in Fe-S cluster formation</td>
</tr>
<tr>
<td>2617920742</td>
<td>precorrin-2/cobalt-factor-2 C20-methyltransferase</td>
</tr>
<tr>
<td>2617920747</td>
<td>Uncharacterized conserved protein</td>
</tr>
<tr>
<td>2617920779</td>
<td>Predicted periplasmic solute-binding protein - COG1427</td>
</tr>
<tr>
<td>2617920796</td>
<td>Uncharacterized protein conserved in bacteria</td>
</tr>
<tr>
<td>2617920806</td>
<td>FOG: LysM repeat</td>
</tr>
<tr>
<td>2617920807</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920823</td>
<td>stage V sporulation protein G</td>
</tr>
<tr>
<td>2617920835</td>
<td>Parvulin-like peptidyl-prolyl isomerase</td>
</tr>
<tr>
<td>2617920838</td>
<td>Bacterial nucleoid DNA-binding protein</td>
</tr>
<tr>
<td>2617920844</td>
<td>Methyl-accepting chemotaxis protein</td>
</tr>
<tr>
<td>2617920854</td>
<td>prevent-host-death family protein</td>
</tr>
<tr>
<td>ID</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2617920867</td>
<td>2,4-dienoyl-CoA reductase, NADH:flavin oxidoreductases, Old Yellow Enzyme family</td>
</tr>
<tr>
<td>2617920868</td>
<td>S-layer homology domain</td>
</tr>
<tr>
<td>2617920891</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920910</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920940</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617920966</td>
<td>Methyl-accepting chemotaxis protein</td>
</tr>
<tr>
<td>2617920969</td>
<td>Methyl-accepting chemotaxis protein</td>
</tr>
<tr>
<td>2617920990</td>
<td>NADPH-dependent FMN reductase</td>
</tr>
<tr>
<td>2617920998</td>
<td>rod shape-determining protein MreB</td>
</tr>
<tr>
<td>2617921007</td>
<td>arginine decarboxylase</td>
</tr>
<tr>
<td>2617921008</td>
<td>Spermidine synthase</td>
</tr>
<tr>
<td>2617921016</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617921017</td>
<td>Predicted membrane protein - COG2855</td>
</tr>
<tr>
<td>2617921019</td>
<td>N-acetylglutamate synthase (N-acetylaspartate amidotransferase)</td>
</tr>
<tr>
<td>2617921020</td>
<td>Acetylglutamate kinase</td>
</tr>
<tr>
<td>2617921021</td>
<td>Omitheine/acetylaspartate amidotransferase</td>
</tr>
<tr>
<td>2617921024</td>
<td>Omitheine carbamoyltransferase</td>
</tr>
<tr>
<td>2617921025</td>
<td>argininosuccinate synthase</td>
</tr>
<tr>
<td>2617921045</td>
<td>Branched-chain amino acid amidotransferase/4-amino-4-deoxychorismate lyase</td>
</tr>
<tr>
<td>2617921047</td>
<td>acetylactate synthase, large subunit</td>
</tr>
<tr>
<td>2617921048</td>
<td>Acetylactate synthase, small (regulatory) subunit</td>
</tr>
<tr>
<td>2617921049</td>
<td>Ketol-acid reductoisomerase</td>
</tr>
<tr>
<td>2617921051</td>
<td>2-isopropylmalate synthase</td>
</tr>
<tr>
<td>2617921052</td>
<td>3-isopropylmalate dehydratase large subunit</td>
</tr>
<tr>
<td>2617921054</td>
<td>Isocitrate/isopropylmalate dehydrogenase</td>
</tr>
<tr>
<td>2617921056</td>
<td>NifU-like domain-containing protein</td>
</tr>
<tr>
<td>2617921084</td>
<td>Ribosomal protein S2</td>
</tr>
<tr>
<td>2617921085</td>
<td>Translation elongation factor Ts</td>
</tr>
<tr>
<td>2617921087</td>
<td>Ribosome recycling factor</td>
</tr>
<tr>
<td>2617921114</td>
<td>Heterodisulfide reductase, subunit A and related polyferredoxins</td>
</tr>
<tr>
<td>2617921115</td>
<td>CoB-CoM heterodisulfide reductase subunit C</td>
</tr>
<tr>
<td>2617921116</td>
<td>Heterodisulfide reductase, subunit B</td>
</tr>
<tr>
<td>2617921119</td>
<td>4Fe-4S dicluster domain-containing protein</td>
</tr>
<tr>
<td>2617921120</td>
<td>4Fe-4S dicluster domain-containing protein</td>
</tr>
<tr>
<td>2617921121</td>
<td>NAD(P)H-flavin reductase</td>
</tr>
<tr>
<td>2617921123</td>
<td>heterodisulfide reductase subunit A</td>
</tr>
<tr>
<td>2617921133</td>
<td>FKBP-type peptidyl-prolyl cis-trans isomerase (trigger factor)</td>
</tr>
<tr>
<td>2617921134</td>
<td>Protease subunit of ATP-dependent Clp proteases</td>
</tr>
<tr>
<td>2617921139</td>
<td>Ruberythrin</td>
</tr>
<tr>
<td>2617921145</td>
<td>Zn-dependent alcohol dehydrogenases</td>
</tr>
<tr>
<td>2617921178</td>
<td>Glutamate synthase domain 2</td>
</tr>
<tr>
<td>2617921183</td>
<td>Clostridium P-47 protein</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>2617921188</td>
<td>Clostridium P-47 protein</td>
</tr>
<tr>
<td>2617921189</td>
<td>Clostridium P-47 protein</td>
</tr>
<tr>
<td>2617921206</td>
<td>Universal stress protein UspA and related nucleotide-binding proteins</td>
</tr>
<tr>
<td>2617921225</td>
<td>ABC-type molybdate transport system, periplasmic component</td>
</tr>
<tr>
<td>2617921237</td>
<td>ABC-type Fe3+-hydroxamate transport system, periplasmic component</td>
</tr>
<tr>
<td>2617921239</td>
<td>Cobalamin biosynthesis protein CobN and related Mg-chelatases</td>
</tr>
<tr>
<td>2617921243</td>
<td>Protoporphyrin IX magnesium-chelatase (EC 6.6.1.1)</td>
</tr>
<tr>
<td>2617921245</td>
<td>Copper amine oxidase N-terminal domain</td>
</tr>
<tr>
<td>2617921249</td>
<td>Protein of unknown function (DUF4198)</td>
</tr>
<tr>
<td>2617921251</td>
<td>Methylase involved in ubiquinone/menaquinone biosynthesis, UbiE</td>
</tr>
<tr>
<td>2617921252</td>
<td>Peptide/nickel transport system substrate-binding protein</td>
</tr>
<tr>
<td>2617921263</td>
<td>Prenyltransferase and squalene oxidase repeat-containing protein</td>
</tr>
<tr>
<td>2617921272</td>
<td>Uncharacterized protein probably involved in high-affinity Fe2+ transport</td>
</tr>
<tr>
<td>2617921288</td>
<td>Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, alpha subunit</td>
</tr>
<tr>
<td>2617921314</td>
<td>S-layer homology domain-containing protein</td>
</tr>
<tr>
<td>2617921321</td>
<td>Cell division protease FtsH</td>
</tr>
<tr>
<td>2617921327</td>
<td>Dihydroneopterin aldolase</td>
</tr>
<tr>
<td>2617921335</td>
<td>Predicted oxidoreductase, contains short-chain dehydrogenase (SDR) and DUF2520 domains</td>
</tr>
<tr>
<td>2617921338</td>
<td>Aspartate 1-decarboxylase</td>
</tr>
<tr>
<td>2617921339</td>
<td>Quinolinate synthetase</td>
</tr>
<tr>
<td>2617921350</td>
<td>Phosphoenolpyruvate synthase/pyruvate phosphate dikinase</td>
</tr>
<tr>
<td>2617921361</td>
<td>Probable rRNA maturation factor</td>
</tr>
<tr>
<td>2617921370</td>
<td>Predicted Fe-Mo cluster-binding protein, NifX family</td>
</tr>
<tr>
<td>2617921373</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>2617921374</td>
<td>Uncharacterized conserved protein - COG1433</td>
</tr>
<tr>
<td>2617921379</td>
<td>CO dehydrogenase/acyetyl-CoA synthase beta subunit</td>
</tr>
<tr>
<td>2617921380</td>
<td>Ni-dependent carbon monoxide dehydrogenase precursor</td>
</tr>
<tr>
<td>2617921406</td>
<td>Uncharacterized conserved protein</td>
</tr>
<tr>
<td>2617921411</td>
<td>Peroxamine synthetase</td>
</tr>
<tr>
<td>2617921413</td>
<td>GDPmannose 4,6-dehydratase</td>
</tr>
<tr>
<td>2617921451</td>
<td>Uncharacterized conserved protein YloU, alkaline shock protein (Asp23) family</td>
</tr>
<tr>
<td>2617921459</td>
<td>Cell fate regulator YlbF, YheA/YmcA/DUF963 family (controls sporulation, competence, biofilm development)</td>
</tr>
<tr>
<td>2617921466</td>
<td>RNA-binding protein Hfq</td>
</tr>
<tr>
<td>2617921504</td>
<td>Protein of unknown function (DUF1292)</td>
</tr>
<tr>
<td>2617921510</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>2617921525</td>
<td>Predicted membrane GTPase involved in stress response - COG1217</td>
</tr>
<tr>
<td>2617921570</td>
<td>Chaperonin GroEL (HSP60 family)</td>
</tr>
<tr>
<td>2617921571</td>
<td>Molecular chaperone GrpE (heat shock protein)</td>
</tr>
<tr>
<td>2617921572</td>
<td>Molecular chaperone</td>
</tr>
<tr>
<td>2617921578</td>
<td>Ribosomal protein S21</td>
</tr>
<tr>
<td>2617921592</td>
<td>Cysteinyl-tRNA synthetase</td>
</tr>
<tr>
<td>2617921655</td>
<td>N-formylmethionyl-tRNA deformylase</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>2617921657</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617921662</td>
<td>Uncharacterized protein conserved in bacteria</td>
</tr>
<tr>
<td>2617921665</td>
<td>Uncharacterized conserved protein - COG1416</td>
</tr>
<tr>
<td>2617921670</td>
<td>ATPase, P-type (transporting), HAD superfamily, subfamily IC</td>
</tr>
<tr>
<td>2617921717</td>
<td>Copper chaperone</td>
</tr>
<tr>
<td>2617921725</td>
<td>Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family</td>
</tr>
<tr>
<td>2617921733</td>
<td>Benzoyl-CoA reductase/2-hydroxyglutarlyl-CoA dehydratase subunit, BcrC/BadD/HgdB</td>
</tr>
<tr>
<td>2617921734</td>
<td>Benzoyl-CoA reductase/2-hydroxyglutarlyl-CoA dehydratase subunit, BcrC/BadD/HgdB</td>
</tr>
<tr>
<td>2617921806</td>
<td>Chorismate mutase</td>
</tr>
<tr>
<td>2617921820</td>
<td>Predicted periplasmic protein (DUF2233)</td>
</tr>
<tr>
<td>2617921825</td>
<td>Uncharacterized protein containing a ferredoxin domain</td>
</tr>
<tr>
<td>2617921828</td>
<td>Di- and tricarboxylate transporters</td>
</tr>
<tr>
<td>2617921834</td>
<td>Secreted protein containing C-terminal beta-propeller domain distantly related to WD-40 repeats</td>
</tr>
<tr>
<td>2617921875</td>
<td>Cold shock proteins</td>
</tr>
<tr>
<td>2617921876</td>
<td>Universal stress protein UspA and related nucleotide-binding proteins</td>
</tr>
<tr>
<td>2617921886</td>
<td>FAD/FMN-containing dehydrogenases</td>
</tr>
<tr>
<td>2617921887</td>
<td>heterodisulfide reductase subunit D</td>
</tr>
<tr>
<td>2617921888</td>
<td>FAD/FMN-containing dehydrogenases</td>
</tr>
<tr>
<td>2617921889</td>
<td>KamA family protein</td>
</tr>
<tr>
<td>2617921890</td>
<td>Enoyl-CoA hydratase/carnithine racemase</td>
</tr>
<tr>
<td>2617921891</td>
<td>acetyltransferase</td>
</tr>
<tr>
<td>2617921896</td>
<td>Crotonobetainyl-CoA:carnitine CoA-transferase CaiB</td>
</tr>
<tr>
<td>2617921897</td>
<td>CoA:oxalate CoA-transferase</td>
</tr>
<tr>
<td>2617921905</td>
<td>phosphoenolpyruvate--protein phosphotransferase</td>
</tr>
<tr>
<td>2617921910</td>
<td>N-acetylmuramoyl-L-alanine amidase</td>
</tr>
<tr>
<td>2617921915</td>
<td>RNAse PH</td>
</tr>
<tr>
<td>2617921959</td>
<td>Sporulation protein and related proteins</td>
</tr>
<tr>
<td>2617921962</td>
<td>Preprotein translocase subunit YajC</td>
</tr>
<tr>
<td>2617921964</td>
<td>Preprotein translocase subunit SecD</td>
</tr>
<tr>
<td>2617921965</td>
<td>Preprotein translocase subunit SecF</td>
</tr>
<tr>
<td>2617921968</td>
<td>Universal stress protein UspA and related nucleotide-binding proteins</td>
</tr>
<tr>
<td>2617921985</td>
<td>carboxyl-terminal processing protease</td>
</tr>
<tr>
<td>2617921996</td>
<td>Ribosomal protein L11</td>
</tr>
<tr>
<td>2617921997</td>
<td>Ribosomal protein L1</td>
</tr>
<tr>
<td>2617921998</td>
<td>Ribosomal protein L10</td>
</tr>
<tr>
<td>2617921999</td>
<td>Ribosomal protein L7/L12</td>
</tr>
<tr>
<td>2617922000</td>
<td>DNA-directed RNA polymerase, beta subunit/140 kD subunit</td>
</tr>
<tr>
<td>2617922002</td>
<td>DNA-directed RNA polymerase subunit beta</td>
</tr>
<tr>
<td>2617922005</td>
<td>Ribosomal protein S7</td>
</tr>
<tr>
<td>2617922006</td>
<td>translation elongation factor 2 (EF-2/EF-G)</td>
</tr>
<tr>
<td>2617922007</td>
<td>GTPases - translation elongation factors</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
</tr>
<tr>
<td>2617922012</td>
<td>RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)</td>
</tr>
<tr>
<td>2617922055</td>
<td>large subunit ribosomal protein L21</td>
</tr>
<tr>
<td>2617922070</td>
<td>long-chain acyl-CoA synthetase</td>
</tr>
<tr>
<td>2617922073</td>
<td>Rhodanese-related sulfurtransferase</td>
</tr>
<tr>
<td>2617922086</td>
<td>HlyD family secretion protein</td>
</tr>
<tr>
<td>2617922096</td>
<td>thiazole-phosphate synthase</td>
</tr>
<tr>
<td>2617922097</td>
<td>Sulfur transfer protein involved in thiamine biosynthesis</td>
</tr>
<tr>
<td>2617922102</td>
<td>Conserved protein/domain typically associated with flavoprotein oxygenases, DIM6/NTAB family</td>
</tr>
<tr>
<td>2617922105</td>
<td>Transcription elongation factor</td>
</tr>
<tr>
<td>2617922118</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2617922137</td>
<td>Acetyl-CoA acetyltransferase</td>
</tr>
<tr>
<td>2617922161</td>
<td>ABC-type molybdate transport system, periplasmic component</td>
</tr>
<tr>
<td>2617922188</td>
<td>Thioredoxin domain-containing protein</td>
</tr>
<tr>
<td>2617922218</td>
<td>Uncharacterized conserved protein - COG2442</td>
</tr>
<tr>
<td>2617922227</td>
<td>4Fe-4S binding domain-containing protein</td>
</tr>
<tr>
<td>2617922233</td>
<td>Formylmethanofuran dehydrogenase subunit E</td>
</tr>
<tr>
<td>2617922234</td>
<td>cobalt/nickel transport protein</td>
</tr>
<tr>
<td>2617922251</td>
<td>Uncharacterized protein conserved in bacteria</td>
</tr>
<tr>
<td>2617922257</td>
<td>adenine phosphoribosyltransferase</td>
</tr>
<tr>
<td>2617922283</td>
<td>D-alanine-D-alanine ligase and related ATP-grasp enzymes</td>
</tr>
<tr>
<td>2617922285</td>
<td>Translation elongation factors (GTPases)</td>
</tr>
<tr>
<td>2617922317</td>
<td>Clp amino terminal domain</td>
</tr>
<tr>
<td>undefined</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>undefined</td>
<td>Thioredoxin domain-containing protein</td>
</tr>
<tr>
<td>undefined</td>
<td>Predicted acid phosphatase - COG0496</td>
</tr>
</tbody>
</table>
Table A2 List of genes identified in the genomes and used for metabolic pathway reconstruction for Spirochaetes in hydrocarbon- and organohalide-contaminated environments shown in Figure 11. Genes for the illustrated pathways were detected in the genomes of *Rectinema cohabitans* HM, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8.

<table>
<thead>
<tr>
<th>Glycolysis/Gluconeogenesis</th>
<th>No.</th>
<th>Enzyme</th>
<th>EC Number(s)</th>
<th>Spirochaetes Bin 1 SA-8</th>
<th>uncultured Spirochaetes bacterium bdmA 4</th>
<th><em>R. cohabitans</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G1</td>
<td>glucokinase</td>
<td>2.7.1.2</td>
<td>SPSA8_v1_410004</td>
<td>SPBDM4_v1_70210</td>
<td>SPBIB_v1_120011</td>
</tr>
<tr>
<td></td>
<td>G2</td>
<td>glucose-6-phosphate isomerase</td>
<td>5.3.1.9</td>
<td>SPSA8_v1_690007_pgi</td>
<td>SPBDM4_v1_50976 - pgi, SPBDM4_v1_40538 - pgi</td>
<td>SPBIB_v1_130044 - pgi</td>
</tr>
<tr>
<td></td>
<td>G3</td>
<td>6-phosphofructokinase</td>
<td>2.7.1.11</td>
<td>not identified</td>
<td>SPBDM4_v1_40290</td>
<td>SPBIB_v1_350003</td>
</tr>
<tr>
<td></td>
<td>G4</td>
<td>fructose-bisphosphate aldolase</td>
<td>4.1.2.13</td>
<td>SPSA8_v1_230024_fba</td>
<td>SPBDM4_v1_50273, SPBDM4_v1_50942 - fba, SPBDM4_v1_4057 - fbaA</td>
<td>SPBIB_v1_140014 - fba</td>
</tr>
<tr>
<td></td>
<td>G5</td>
<td>triose-phosphate isomerase</td>
<td>5.3.1.1</td>
<td>SPSA8_v1_560006</td>
<td>SPBDM4_v1_51054 - tpiA</td>
<td>SPBIB_v1_100039 - tpiA</td>
</tr>
<tr>
<td></td>
<td>G6</td>
<td>glyceraldehyde-3-phosphate dehydrogenase (phosphorylating)</td>
<td>1.2.1.12</td>
<td>SPSA8_v1_560002</td>
<td>SPBDM4_v1_51050 - gapA</td>
<td>SPBIB_v1_100034 - gapA</td>
</tr>
<tr>
<td></td>
<td>G7</td>
<td>phosphoglycerate kinase</td>
<td>2.7.2.3</td>
<td>SPSA8_v1_560003</td>
<td>SPBDM4_v1_51051 - pgk, SPBDM4_v1_41029</td>
<td>SPBIB_v1_100035 - pgk, SPBIB_v1_210059, SPBIB_v1_290179</td>
</tr>
<tr>
<td></td>
<td>G8</td>
<td>phosphoglycerate mutase</td>
<td>5.4.2.1</td>
<td>SPSA8_v1_410010_gpmi</td>
<td>SPBDM4_v1_40326, SPBDM4_v1_70215 - gpmi, SPBDM4_v1_40696 - gpmA</td>
<td>SPBIB_v1_90032 - gpmi, SPBIB_v1_270010 - gpmA</td>
</tr>
<tr>
<td></td>
<td>G9</td>
<td>phosphopyruvate hydratase</td>
<td>4.2.1.11</td>
<td>not identified</td>
<td>SPBDM4_v1_40824 - eno</td>
<td>SPBIB_v1_280034 - eno</td>
</tr>
<tr>
<td></td>
<td>G10</td>
<td>pyruvate kinase</td>
<td>2.7.1.40</td>
<td>SPSA8_v1_280005</td>
<td>SPBDM4_v1_50076 - pyk</td>
<td>SPBIB_v1_290204 - pyk</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pentose phosphate pathway</th>
<th>No.</th>
<th>Enzyme</th>
<th>EC Number(s)</th>
<th>Spirochaetes Bin 1 SA-8</th>
<th>uncultured Spirochaetes bacterium bdmA 4</th>
<th><em>R. cohabitans</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P1</td>
<td>Glucose-6-phosphate dehydrogenase (NADP(+))</td>
<td>1.1.1.49</td>
<td>SPSA8_v1_290001</td>
<td>SPBDM4_v1_50483 - zwf</td>
<td>SPBIB_v1_210196 - zwf</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>6-phosphogluconolactonase</td>
<td>3.1.1.31</td>
<td>SPSA8_v1_290002</td>
<td>SPBDM4_v1_50484</td>
<td>SPBIB_v1_210195</td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>Phosphogluconate dehydrogenase (decarboxylating)</td>
<td>1.1.1.44</td>
<td>SPSA8_v1_640011</td>
<td>SPBDM4_v1_70232 - gnd</td>
<td>SPBIB_v1_900014 - gnd</td>
</tr>
<tr>
<td></td>
<td>P4</td>
<td>D-glyceraldehyde-3-phosphate glycolaldehyde transferase</td>
<td>2.2.1.1</td>
<td>SPSA8_v1_280002</td>
<td>SPBDM4_v1_40259, SPBDM4_v1_40260, SPBDM4_v1_41032 - tkt, SPBDM4_v1_50175 - tkt</td>
<td>SPBIB_v1_10035, SPBIB_v1_10036, SPBIB_v1_250033 - tktB</td>
</tr>
<tr>
<td>No.</td>
<td>Enzyme</td>
<td>EC Number(s)</td>
<td>Spirochaetes Bin 1 SA-8</td>
<td>uncultured Spirochaetes bacterium</td>
<td>R. cohabitanis</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-------------------------------------</td>
<td>--------------</td>
<td>-------------------------------</td>
<td>-----------------------------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>ribulose-phosphate 3-epimerase</td>
<td>5.1.3.1</td>
<td>SPSA8_v1_330009</td>
<td>SPBDM4_v1_70027 - rpe</td>
<td>SPBIB_v1_100090 - rpe</td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>ribose-5-phosphate isomerase</td>
<td>5.3.1.6</td>
<td>SPSA8_v1_780001</td>
<td>not identified</td>
<td>SPBIB_v1_100034 - rpiB, SPBIB_v1_410042 - rpiB</td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td>ribose-5-phosphate diphosphotransferase</td>
<td>2.7.6.1</td>
<td>not identified</td>
<td>SPBDM4_v1_40813</td>
<td>SPBIB_v1_280022</td>
<td></td>
</tr>
</tbody>
</table>

**Starch and sucrose metabolism pathway**

- **No.**
- **Enzyme**
- **EC Number(s)**
- **Spirochaetes Bin 1 SA-8**
- **uncultured Spirochaetes bacterium**
- **R. cohabitanis**

**Fermentation**

- **No.**
- **Enzyme**
- **EC Number(s)**
- **Spirochaetes Bin 1 SA-8**
- **uncultured Spirochaetes bacterium**
- **R. cohabitanis**

**Membrane energization**

- **No.**
- **Enzyme**
- **EC Number(s)**
- **Spirochaetes Bin 1 SA-8**
- **uncultured Spirochaetes bacterium**
- **R. cohabitanis**

**Protein degradation**

133
<table>
<thead>
<tr>
<th>No.</th>
<th>Enzyme</th>
<th>EC Number(s)</th>
<th>Spirochaetes Bin 1 SA-8</th>
<th>uncultured Spirochaetes bacterium</th>
<th>R. cohabitans</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>FeFe A1 hydrogenase</td>
<td></td>
<td>not identified</td>
<td>SPBDM4_v1_80097</td>
<td>SPBIB_v1_20025</td>
</tr>
<tr>
<td>H2</td>
<td>FeFe A3 hydrogenase</td>
<td>SPSA8_v1_300018</td>
<td></td>
<td>SPBDM4_v1_40494</td>
<td>SPBIB_v1_260013, SPBIB_v1_210130</td>
</tr>
<tr>
<td>H3</td>
<td>FeFe B hydrogenase</td>
<td>SPSA8_v1_810002</td>
<td></td>
<td>not identified</td>
<td>SPBIB_v1_290056, SPBIB_v1_360015</td>
</tr>
<tr>
<td>H4</td>
<td>FeFe C hydrogenase</td>
<td>SPSA8_v1_270001, SPSA8_v1_810005</td>
<td></td>
<td>SPBDM4_v1_40759</td>
<td>SPBIB_v1_360012, SPBIB_v1_280015</td>
</tr>
</tbody>
</table>

**Secretion systems**

<table>
<thead>
<tr>
<th>No.</th>
<th>Enzyme</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Preprotein translocase, SecG subunit</td>
<td>SPBIB_v1_100040</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Protein translocase subunit SecA</td>
<td>SPBIB_v1_110017</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Protein-export membrane protein SecF</td>
<td>SPBIB_v1_210207</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Protein translocase subunit SecD</td>
<td>SPBIB_v1_210208</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>putative SEC-C motif domain protein</td>
<td>SPBIB_v1_280056</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Protein translocase subunit SecE</td>
<td>SPBIB_v1_310045</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>preprotein translocase membrane subunit</td>
<td>SPBIB_v1_370006</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Carbohydrate-active enzymes (polysaccharide degradation)**

<table>
<thead>
<tr>
<th>No.</th>
<th>Enzyme</th>
<th></th>
<th>Database A5</th>
<th>Database A4</th>
<th>Database A3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>hydrolases and extracellular binding proteins</td>
<td></td>
<td>Database A5</td>
<td>Database A4</td>
<td>Database A3</td>
</tr>
<tr>
<td>S2</td>
<td>ABC transporters</td>
<td></td>
<td>Database A5</td>
<td>Database A4</td>
<td>Database A3</td>
</tr>
<tr>
<td>S3</td>
<td>Hydrolases</td>
<td>Database A5</td>
<td>Database A4</td>
<td>Database A3</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td></td>
</tr>
</tbody>
</table>

135
Table A4 List of identified proteins for *R. cohabitans* HM as shown in Figure 14.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPBIB_v1_250074</td>
<td>2,3-cyclic-nucleotide 2-phosphodiesterase</td>
<td>purine metabolism and pyrimidine metabolism.</td>
</tr>
<tr>
<td>SPBIB_v1_150165</td>
<td>31 kDa immunogenic protein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_350024</td>
<td>5-nucleotidase/2,3-cyclic phosphodiesterase-like hydrolase</td>
<td>purine metabolism and pyrimidine metabolism.</td>
</tr>
<tr>
<td>SPBIB_v1_150087</td>
<td>ABC sugar transporter, periplasmic ligand binding protein</td>
<td>sugar metabolism</td>
</tr>
<tr>
<td>SPBIB_v1_260060</td>
<td>ABC sugar transporter, periplasmic ligand binding protein</td>
<td>sugar metabolism</td>
</tr>
<tr>
<td>SPBIB_v1_250100</td>
<td>ABC transporter substrate binding protein</td>
<td>malto-oligosaccharides, multiple oligosaccharides, glycerol 3-phosphate, and iron</td>
</tr>
<tr>
<td>SPBIB_v1_330021</td>
<td>ABC-type dipeptide transport system, periplasmic component</td>
<td>oligo/dipeptide and amino acid transporters</td>
</tr>
<tr>
<td>SPBIB_v1_210094</td>
<td>ABC-type transporter, periplasmic subunit</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_210125</td>
<td>ABC-type transporter, periplasmic subunit</td>
<td>oligo/dipeptide and amino acid transporters</td>
</tr>
<tr>
<td>SPBIB_v1_290103</td>
<td>Aconitate hydratase</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_20007</td>
<td>Basic membrane lipoprotein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_240055</td>
<td>Basic membrane lipoprotein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_310088</td>
<td>Branched-chain amino acid ABC transporter substrate-binding protein</td>
<td>oligo/dipeptide and amino acid transporters</td>
</tr>
<tr>
<td>SPBIB_v1_150042</td>
<td>conserved exported protein of unknown function</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_210118</td>
<td>conserved exported protein of unknown function</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_210161</td>
<td>conserved exported protein of unknown function</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_270015</td>
<td>conserved exported protein of unknown function</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_350064</td>
<td>conserved exported protein of unknown function</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_310019</td>
<td>Desulfurodoxin</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_100150</td>
<td>DNA-binding protein HRL18</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_80014</td>
<td>D-ribose-binding protein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_250059</td>
<td>exported protein of unknown function</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_340054</td>
<td>exported protein of unknown function</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_380039</td>
<td>exported protein of unknown function</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_240026</td>
<td>Extracellular ligand-binding receptor</td>
<td>oligo/dipeptide and amino acid transporters</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Function</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>SPBIB_v1_150035</td>
<td>Extracellular solute-binding protein</td>
<td>malto-oligosaccharides, multiple oligosaccharides, glycerol 3-phosphate, and iron</td>
</tr>
<tr>
<td>SPBIB_v1_10042</td>
<td>Extracellular solute-binding protein family 1</td>
<td>malto-oligosaccharides, multiple oligosaccharides, glycerol 3-phosphate, and iron</td>
</tr>
<tr>
<td>SPBIB_v1_380003</td>
<td>Extracellular solute-binding protein family 1</td>
<td>malto-oligosaccharides, multiple oligosaccharides, glycerol 3-phosphate, and iron</td>
</tr>
<tr>
<td>SPBIB_v1_400009</td>
<td>Extracellular solute-binding protein family 1</td>
<td>malto-oligosaccharides, multiple oligosaccharides, glycerol 3-phosphate, and iron</td>
</tr>
<tr>
<td>SPBIB_v1_60010</td>
<td>Extracellular solute-binding protein family 1</td>
<td>malto-oligosaccharides, multiple oligosaccharides, glycerol 3-phosphate, and iron</td>
</tr>
<tr>
<td>SPBIB_v1_290108</td>
<td>Extracellular solute-binding protein family 5</td>
<td>oligo/dipeptide and amino acid transoprters</td>
</tr>
<tr>
<td>SPBIB_v1_250085</td>
<td>FAD-dependent pyridine nucleotide-disulfide oxidoreductase</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_270007</td>
<td>Gamma-glutamyltranspeptidase</td>
<td>intracellular peptidases</td>
</tr>
<tr>
<td>SPBIB_v1_290147</td>
<td>Gamma-glutamyltranspeptidase</td>
<td>intracellular peptidases</td>
</tr>
<tr>
<td>SPBIB_v1_80037</td>
<td>glutamate dehydrogenase, NADP-specific</td>
<td>Amino acid degradation</td>
</tr>
<tr>
<td>SPBIB_v1_100034</td>
<td>glyceraldehyde-3-phosphate dehydrogenase A</td>
<td>glycerol fermentation</td>
</tr>
<tr>
<td>SPBIB_v1_380029</td>
<td>glycerol-3-phosphate transporter subunit;periplasmic-binding component of ABC superfamily</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_340036</td>
<td>Iron ABC transporter, substrate binding protein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_210026</td>
<td>Leucine-, isoleucine-, valine-, threonine-, and alanine-binding protein</td>
<td>oligo/dipeptide and amino acid transoprters</td>
</tr>
<tr>
<td>SPBIB_v1_290158</td>
<td>oligopeptide transporter subunit;periplasmic-binding component of ABC superfamily</td>
<td>oligo/dipeptide and amino acid transoprters</td>
</tr>
<tr>
<td>SPBIB_v1_50038</td>
<td>Oligopeptide-binding protein AliB</td>
<td>oligo/dipeptide and amino acid transoprters</td>
</tr>
<tr>
<td>SPBIB_v1_150028</td>
<td>Periplasmic binding protein/LacI transcriptional regulator</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_10062</td>
<td>Periplasmic sugar-binding protein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_410087;SPBIB_v1_10012</td>
<td>Plasmid stabilization system protein;Plasmid stabilization system protein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_110024</td>
<td>protease, ATP-dependent zinc-metallo (M41)</td>
<td>Type II secretory system</td>
</tr>
<tr>
<td>SPBIB_v1_310056</td>
<td>protein chain elongation factor EF-Tu (duplicate of tufB)</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_110017</td>
<td>Protein translocase subunit SecA</td>
<td>Protease</td>
</tr>
<tr>
<td>SPBIB_v1_150059</td>
<td>Purine-binding protein BAB2_0673</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_340021</td>
<td>putative 2-oxoacid-flavodoxin fused oxidoreductase:conserved protein;4Fe-4S cluster binding protein</td>
<td>Amino acid degradation</td>
</tr>
<tr>
<td>SPBIB_v1_210218</td>
<td>putative ABC-type transport system, periplasmic component/surface lipoprotein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_100146</td>
<td>putative basic membrane lipoprotein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_260082</td>
<td>putative cation ABC transporter, periplasmic binding protein</td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
<td>Functions</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>SPBIB_v1_310055</td>
<td>putative Elongation factor G</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_350021</td>
<td>putative Extracellular solute-binding protein family 1</td>
<td>malto-oligosaccharides, multiple oligosaccharides, glycerol 3-phosphate, and iron</td>
</tr>
<tr>
<td>SPBIB_v1_210067</td>
<td>putative Periplasmic sugar binding protein-like protein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_340064</td>
<td>Split soret cytochrome c</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_260025</td>
<td>Sugar ABC transporter substrate-binding protein</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_100013</td>
<td>TRAP dicarboxylate transporter, DctP subunit</td>
<td></td>
</tr>
<tr>
<td>SPBIB_v1_380022</td>
<td>TRAP dicarboxylate transporter, DctP subunit</td>
<td></td>
</tr>
</tbody>
</table>
Supporting Figures

Figure A1 Identity dot plot between *Pelotomaculum* candidate BPL (contigs on the y axis) and *Pelotomaculum thermopropionicum* SI (full genome sequence on the x axis) generated using MUMmer (http://mummer.sourceforge.net/). Red dots indicate stretches of nucleotide identity and blue dots reverse complementary matches. Grey horizontal lines separate the individual contigs on the y axis.

Figure A2 Gene count of COG functional categories for *Pelotomaculum* candidate BPL and *Pelotomaculum thermopropionicum* SI.

Figure A3 Full phylogenetic tree showing the phylogeny of the [FeFe]-hydrogenase catalytic subunit sequences detected in *R. cohabitans* HM, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8.

Figure A4 Expanded maximum likelihood tree of partial 16S rRNA gene sequences of Spirochaetes. This shows the phylogenetic affiliation of *R. cohabitans*, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8.

Figure A5 Mass spectrum of CO$_2$ obtained via Isotope Ratio Mass Spectrometry: (a) natural sample with a low $^{13}$C abundance ($\delta^{13}$C = -9.7 ‰); (b) $^{13}$C-enriched sample ($x^{13}$C = 10.0 ‰). CO$_2$ samples are firstly ionized to different ions mass-to-charge ratios (m/z) at 44, 45 and 46. The first three peaks (square-shaped) belong to working gas while the last one to the sample.

Figure A6 Normalized transmission spectrum of air containing 380 ppm CO$_2$ obtained via Isotope Ratio Mid-Infrared Spectroscopy: (a) natural sample with a low $^{13}$C abundance ($\delta^{13}$C = -9.7 ‰); (b) $^{13}$C-enriched sample ($x^{13}$C = 10.0 ‰). The peak areas labeled as $(13)$CO$_2$(1) and CO$_2$(1) were used to determine the $^{13}$C/$^{12}$C isotopic ratios in this study.
Figure A1 Identity dot plot between *Pelotomaculum* candidate BPL (contigs on the y axis) and *Pelotomaculum thermopropionicum* SI (full genome sequence on the x axis) generated using MUMmer (http://mummer.sourceforge.net/). Red dots indicate stretches of nucleotide identity and blue dots reverse complementary matches. Grey horizontal lines separate the individual contigs on the y axis.
Figure A2 Gene count of COG functional categories for *Pelotomaculum* candidate BPL and *Pelotomaculum thermopropionicum* SI.
Figure A3 Full phylogenetic tree showing the phylogeny of the [FeFe]-hydrogenase catalytic subunit sequences detected in *R. cohabitans* HM, uncultured *Spirochaetes* bacterium *bdmA* 4, and uncultured *Spirochaetes* bacterium *SA*-8.
Figure A4 Expanded maximum likelihood tree of partial 16S rRNA gene sequences of Spirochaetes. This shows the phylogenetic affiliation of *R. cohabitans*, uncultured Spirochaetes bacterium bdmA 4, and uncultured Spirochaetes bacterium SA-8.
Figure A5 Mass spectrum of CO$_2$ obtained via Isotope Ratio Mass Spectrometry: (a) natural sample with a low $^{13}$C abundance ($\delta^{13}$C = -9.7 ‰); (b) $^{13}$C-enriched sample ($x(13C) = 10.0 \%$). CO$_2$ samples are firstly ionized to different ions mass-to-charge ratios ($m/z$) at 44, 45 and 46. The first three peaks (square-shaped) belong to working gas while the last one to the sample.
Figure A6 Normalized transmission spectrum of air containing 380 ppm CO₂ obtained via Isotope Ratio Mid-Infrared Spectroscopy: (a) natural sample with a low ¹³C abundance ($\delta^{13}$C = -9.7 ‰); (b) ¹³C-enriched sample ($x^{13}$C = 10.0 ‰). The peak areas labeled as (13)CO₂(1) and CO₂(1) were used to determine the $^{13}$C/$^{12}$C isotopic ratios in this study.
Supporting Databases

Database A1 FASTA-formatted sequences for 16S rRNA genes of *Pelotomaculum* candidate BPL.

Database A2 FASTA Amino Acid format for protein coding genes in the genome of *Pelotomaculum* candidate BPL. Considering the size is too large to include here, it can be found at Integrated Microbial Genomes (IMG) system. GOLD Study ID: Gs0114392, GOLD Project ID: Gp0111374, GOLD Analysis Project Id: Ga0073689.

Database A3 FASTA Amino Acid format for protein coding genes in the genome of *R. cohabitans* HM. Considering the size is too large to include here, it can be found at Microbial Genome Annotation & Analysis Platform (MicroScope) with name of *Uncultured spirochete bib* (WGS SPIROBIBN47.1). It was submitted to European Nucleotide Archive under study identifier PRJEB19364.

Database A4 FASTA Amino Acid format for protein coding genes in the genome of uncultured *Spirochaetes* bacterium SA-8. Considering the size is too large to include here, it can be found at Microbial Genome Annotation & Analysis Platform (MicroScope) with name of *Spirochaetes Bin 1 SA-8* (WGS SPSA8.1). It was submitted to European Nucleotide Archive under study identifier PRJEB19364.

Database A5 FASTA Amino Acid format for protein coding genes in the genome of uncultured *Spirochaetes* bacterium bdmA 4. Considering the size is too large to include here, it can be found at Microbial Genome Annotation & Analysis Platform (MicroScope) with name of *Uncultured spirochete bdmA 4* (WGS SPIRO4BDMA.1). It was submitted to European Nucleotide Archive under study identifier PRJEB19364.

Database A6 Amino Acid format for hydrogenase sequences used for Figures 12 and A3.
Database A1 FASTA-formatted sequences for 16S rRNA genes of *Pelotomaculum* candidate BPL.

> bpl_candidate_16S@contig01451_113_1661

TCATAGAAAGATTATAGGAGAGTTTGTACCTGCGTCAGGACAGCAACGCTGGCCGCTGCTTAAC
ACATGCAAGTCGACGCGGGAAGCTTAACGCTGGGGCTGGAAACACGGCCAGATCGCCAGACTGCG
ATAGGAGGACAAGACCCGGTACGAGCGTGGCCTGAGACGGGTCGACTGCTACGCTCTCAGGAGG
CAGCAGTGGGGAACTTCCGCAATGGGGCGCAGCCTGACGGAGCAACGCCGCGTGAGTGACGAAGGC
CTTCGGGTTGTAAAACTCTGTCTTCGGAGAAAGAAATGACGCTGCCCTTCCGAATGGGCGCAAGC
CGGCTAACTACGTGCGTAATGGGACAGCGAACGGGACTCTCGGGCCTGGGTTACTGACGCTGAGG
CGCGAAAGCGTGGGGAGCGAACGGGATTAGATACCCCGGTAGTCCTGCAAMATGCTAGGATGGGG
AACTGCTGGTGCCGCTTACATCGGTGCTAGTTGGAGGTCGTCAGCTCGTGTCGTGAGATGTTGGT
TAAGTCCCGCAACGAGCGCAACCCCTACGTTAGTTGCTACCGGACATGGGAGTTGCTTTACGCGC
ATGCACGGGCTCAGTTCGGATTGCAGGCTGCAACCCGCCTGCATGAAGTCGGAATCGCTAGTAATCG
CAGCATACTGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCAGAAAGCTGAC
ACAACACCCGAAGCCGGTGACTTAACCGCAAGGAGAGAGCCGTCGAAGGTGGGGTTGGTGATTGGG
GTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATC A CCTCCTTT
CTAA
Database A6 Amino Acid format for hydrogenase sequences used for Figures 12 and A3.

**SPBIB_v1_260013|ID:2716327|hndD** FeFe Group A3 hydrogenase catalytic subunit [Uncultured spirochete bib]

MNVKVKNGIPVQVAEGSTVLEAAAKANVKIPTLCYNPDLPWASCIGCICVVKIEEIGNMLRLSCCTPV SEGMIISNDPLDVLQRTKTVIILSTLHPDCLACRPQACELQLAEOFQIGREOPYYKMKVRDIPQD TSTGSLILNSPKICRCGCMEQVWAVGFLRGRGESIAPADDVKGDPSICKCQCGSASCACHE PVGAIYENQDKLWVADLMEGKPEACTCAVQAPARVLAESFGGLPGTDLTGKVYATLRLGFD AVFDTNFADLTIMEEGTFEHRLTGALKGMQATADKSMLTSCPAWVDYMEKYDPMIPN FSTAKSPQMOQMKITTYWAADPKVDPAKIVSYSVIMPCTAKKKFENRSDMSGSSYKDVDTLTT RELARMIKQAGIDFNLEPSEPDSPLGYSAGVIFGATAGGMVEAAALTAYLFTKEELKDNFTA VRGLSGLKEATVHINGELRVAAHQMNGTIDLQVARKAREGTPWHFIEMACRGGCCIGGQ GGYPYATDEVKRLRMRGIDYNDKEQYRCSHDNPYIKQIYAFLQPASHKHAHELLHTQYKERPL YLK*

**SPBIB_v1_20025|ID:27161852|hndD** FeFe Group A3 hydrogenase catalytic subunit [Uncultured spirochete bib]

MEEQLEITIDGQNVRDPSANIVEEACAHAGVKIPTLCYLGISQASCICGCICVEVEGAKSLVRCV QKPVPQMKGKRTSSPRVIRAKTAVELLNLNPDDCLSCIRCSDTCELHTIANILEVRADRFPPKYK MPDOTTSEVIRDDSKCLCGRCVAVEETEQGVHAIAFSGGRARTRSTVFDDRGLAQASCACVGCQSC SVVCPTGAITEKDESREVFKIDPFLVQVQTAPAIRASLGEALPLPSVGTGVAALRFLGFN RFVDQFTADLTIMEEGSELLERLSSCRGTLPMITSCSPGWINFIEGFYDPLLHSTCKSPQMGFS VAKTYYAKKHCPHMRVAAHMCPQGNMVTGWAQWKEQDPQVKVAPFDFVDW ALTRELARMIKCLAGIDISRLPEEFDDPLQGQSTGAGTFGTTGGVMEAALRTYVELVEKPLNEIF TKVGRFGEAIKTAEEVVVGGSPVRVAVAHGLANARRLDEIRAGKSPYHFEIIMSCPGGCVGGGQGQ VLADLEKCLRQALYEEERKLAIRKSHENPAVARLQYELQFLKPGLHELLHTSYKARVF*

**SPBIB_v1_210130|ID:27163022|hndD** FeFe Group A3 hydrogenase catalytic subunit [Uncultured spirochete bib]

MNKVTLLINGKKSVPEHTTFVSAEEKAGISIPALCRHDEPKAGCAGMCIVKIEGQPGYKRCSTVA VEEGMEELTSSEIDIRGILEVLAAPADCLQCICHKGCELTACLIERFEIDLYRDRYTRGLPV DRSSFGIVRDMKCCGRCVQCNEVQTVASIFHGRGNTVSPAYGGTMDSVCVNCQCCIV YCPVAGLYEKEAIEEWAHKADPDKVVAQAIPAVRNAEEGFLPGLGSEISKLYSALKLGIDVFV DTNSFADEVEEPFLERKLESGFFPLTPSICSPGWIFKGETYFPELLENVSSCSSQPMOQLALIKT HYAESRGLDRNNVLSLMPCTAKKFEAGREPOMSSGARDDVYLTTREIRMRQVQIDFRNLPP GTDPDLSSYSAATIGFAGGVCVEAAALTAYELGKALEKGEGVDFVRGMTGVKAEIDVGDTPI KVAVTNGLANARKILDRAEHDRGRSYYHEFIMACAGCVCVGGQGQPINTLRARRRIGEYDRE DRSLNLKSHENPEAIRLQYELQFLPGSKEAHELHTKSYPREQYQFSE*

**SPBIB_v1_290056|ID:27163581|FeFe Group B hydrogenase catalytic subunit [Uncultured spirochete bib]

VEPEFVEMRLGTYLPTCDDIRKIFVGVTRFVLAVKPKEDIDYLFDEIDIMGKTPYRCCSYRELIVKQR IRLAFGLPLIEERDPTVSAGIEAFFTRDKVISAPLNVIRAAACEKPEDKVIVTDMQCSMAHPCSI VCNPVNIASFNPPGKAFIDQKCKVCVKCCKMKCAYQPSITRMRVRPCCAAACGVDAHIPDPGYATIDQ DKCNVCNGLCTVSCPFAASIKDETELQIVLHLQMLPGEDPPRPRYALAPSVQGFGRASPAVAGLRA VGFRGVERADVLCADDRSLKALAEKSDPNQHPHTDSCPLSWKVTARRHPFEDNAAES FTPMVETAKI1KDQPSARVVFIPAIKAIAELEPEMQQYVDVHLTETFELAIAIVIFAKDIDLAEITPE FPDEASALGRYAVAGVGVTAAIETARQKYGENKIAVARADLNRACMLADIKSKGTSPELVEG MACPGGCVGGQPGTGLLSSLARNEVGGKAFALAPNPDATLTK*

148
MNROVIYETMRQCDYKVCVERCPVKAIQLDGHARVEDICILCGHCVMVCPQSAKVRSDVERVRRRLELKPVASLAPSAEFGCTAQLAISIRKGFAAVSETALGADLYSSAMRELQQLADNGRNFLIGACAPAVYRSFVPDLVPFLUSENGSPMAHARYLRSFGENALVFIPGCIACKKEADESGLVNVUAVLFAELROLFEEUNIFSPAEKTPSNFTOFFPGRSKGELYPFEIGGMAISIHKHGSTDVPCMSFSGMHIQIALRLDFGDPFETGIFLELACEECCINGPCDSASRTGTVSKRQILLEYKQGGESEPFNTIEAQCQVKSAQPQVQRQVPEEPLYAQQFTRYSQSDDIAAHLMGDFEVFDVAVAIATEITRSILAPGPAPRISISACPVPVRLIQRFPSLNPILMPPIEAAREARKRSEISRKIGIFLLSPCTAKVTARVMPGYSQSSIDAVIFSDILPFLKRALEERTKPLSKPKGKSSPVQAGHSPILPMMDNLEPGMGWARSDELADLPDSVSVQDGISNYIDLEFIEENIEQISYIEALECPGCCVPGLPAMVENPHIAIARNMRQGKDNLALESASSPMAKIPPSKPELEQPSIWEOSVPSKPVLVLDDLSKALEMAEHDITIHAQLPGIDCGACPCDCCLAEDIVRFGASIEDCRLLEHPYSGKTKHENIHEGLST*

SPBDM4_v1_80097|ID:27159873|FeFe Group A1 hydrogenase catalytic subunit (fragment) [Uncultured spirochete bdm4]

VCETEQGVNIAFTRGRARTVTFMDRGLAQSCACVCGCCSVCPTGAIETKEDSRDFVALRDSKLSVSVVQTAPIARSLGALPGASLVTMGMAAARLRGFAKVFTQTDATLTIMEESELERLSHGTMTIPTMSCPGWINIFEGYPDLGLHVSTCKSPQMGAVAKVYQAAGIAPDMKRVVSMIPCTAKXKYEARKMDGAWWKEQDDPGKVPRPFDVDWALTRELAMIKLAGEIHLPEDFDPGLRGSTAATFGTMTGGVMEALRTVYEIEKPLENFETOVRGEFISKAEVLLGGSVPVRAVAHGLSNARILLDEIRAQKSPQFYEIEMSCPGCIGGGGGQVPLADIEKLLARSKYTEDRILPRKSHENPVAINTYKDFKGLKPGHLHSELSHTYSARVF*

SPBD4_v1_40494|ID:27157799|FeFe Group A3 hydrogenase catalytic subunit [Uncultured spirochete bdm4]

MVNVKINGIPVQAEGSTMLEAAKAHKVIPTLCYNPDLSWPAACIGCVKVEGSKMLRSCCTPSEGMSITNADLGQRTVRITIELSTPDDCFLCPRNQSCSCELTQLTEAQQFERQEPYKKVMRDIPTDSTNSLILPSKICIRCRCVECVQEMQGWWAIEFLRGRSLISIRIAPAADVKLGDSPCKGCQCSAHCPVGAIIYENDTSLWVLGEMPAKAIKALCQVAPIAARPVALAESFGLPPGTDLKTYIALRLGDFDVDNFSADLTIMEEGLVHRLTVLAKGGMQATADHHKLPISTCCPAWVYYDEMYKSYDSMIPNSTAKSPQOOMGVMIKTYWAAKAVDPAKIYVSIMPCTAKXKFRNGRDESMYSGYKVDVTLTRELAMIKQAGIDFNLPDSCQPSLPGYPSAGVIFGATGGVMEALRTAYFLVTKELKDVNFT
AVRGLSGIKEATVHNGIELRVAHAOMGNIATVLEEVRKARAEGRETWHFIEVMACRGGCIGGQGQPYGATDEVRLRMRSLRYSYDHDEKSEYRCSHDNPYIKQIYAEEFKPGLSRSHRHELHHTHYEKPLYLK*

SPBDM4libraries_01_40759|ID:27158064|FeFe Group C1 hydrogenase catalytic subunit [Uncultured spirochete bdmA 4]

VSFEPAANQSNFQTFHSVLLDLLCDVCCTICIKFCTKAIRVRNGAKIFEDRCIDCGECCRCPKGAKKAVSDPLFMDAYDVKALPAPSYLAQFGTKYTSQDFNIAHSAGFDEVFDVAVGALVVTMTRSLLAQEALKPRISACAVPRVRLQIQFRPSLNLMPIPILPSSAEARERRRLGSLQKYGFLISPCTAKVTSSRTPLYEQSAIDAVFSFGDIYASLakersAPHASIADISESGHNASTHPLLIMNNLEPGGWARSDELDHALHENAVSDGINSNIENVLEFAENGIDSQIEYEAALACPCCGVCPPAVENPHIARSMRQRYQHPFSAYPGSKTADQDHHTAKERSKLSPESSATADKEYLAKWTRPLLPNPVLVLTDLRSKALMAEKIQIRRNLPGICACGCAPDCDAFAEDIVGRGLSIRDCLAASAPPARNEENPK*

JGI12104J13512_100111416FeFe Group A3 hydrogenase catalytic subunit [Uncultured spirochete SA-8]

MINCKVNGIPQVAEGATILEASAKKANVIKTPLCYNPDALPWAAGICVAKIEGNSKMLRACCTPVAEGMNITHPDEVTKVIEMILSTHPDLCAPRNNQCELTOILAEQFRGEOAQFPMKMLHLPDIDTTGSVLNPEKCSVRCRTCVPQCQMOMWAEFLGRGETIRAPAADAKLGSPICKGCQSACHPGAYAIONDKTIKVDALRTGDATKCVIQAIAPVRVLAGAEFLQPQTPDLTGYIKARKLGDVFVDFTNFAADLTIMEEGTFEVRKLTSALONGLTATREKSMPLSCTCAPWVDYMEKYPDMPINPSFASKPQMMGMAGIKTYWEAKANVRPKFISVSIMPTAKKFETRDMSGSGQYQVDVSISSTLARMIKOAIGDINLNPESPDPLGYTPGAYAGAFGATGGVMEAAALRTAYYLTGNEALKDNTFAVRGISEKEASVHNVGLVRLVAHAQOMINGNIEVNRKARDGREDTPHWIEMACRGGCICQGGQPYGATDEVKLRIRGIYDHDEQKEYRSCHQNPYIKKVNEYLEKPGSHEAHELHTHYTERPLFLK*

JGI12104J13512_10014452FeFe Group B hydrogenase catalytic subunit [Uncultured spirochete SA-8]

MAKENAVRLKRLILICEVTSVLEKLALEIDISPYAMTNENWETIQCCIHDRANLRLRLMSLLGYSTEGLMDVKDPLRTYAEQLSGKPDAPFLSLMSAANCNGCTKSRIVTNCQGCLARPCKVNCPCGAVSIINQGSHIDPELCVNCLGECSCPHAIKIIPVPEECVPGVAIQUKGEDIAKIDESCISKICLGCKGLSACPFQAPVEKSIHVLKILAMPAAMAVQFPFPFPQKIALEHFGSKVMEAAADRVAAEAREFAERLHENVKLATCCPSEVRAAGSGMEQSLSDTSPMLVAAKTATKQPQDFAVFIPCLARKREWASQNSPLVYALTSSEVGAIALMAGARIXVEENEDLATOMAGSNGYGRFATGQVTAAYKHALPSADEVVNCIVGSITKTDTASLKKKVQSEDKPLLEVMACDGGCINCQQLTNPKVAGFAGLYAKAEHTTALKSA*

JGI12104J13512_10009791FeFe Group C1 hydrogenase catalytic subunit [Uncultured spirochete SA-8]

AAREAALKVLADLPSLVRGMFFISPCTAKTAIRMPLYGERSLVALDAFVFSFGDIFPLAALKALKSRPSETPSRFLAPLQIERMRVHGDGEIEDGIEQAVSYDGQNVIALLEEIDNGKSSVPEIALACPGGOVGGPMAMANPHARAAMKHIAAEEKTAAMQSGQSEQPAAGFQWERLQPKPVFVLRDMLKALQAEMELITSQPLGLDGCAGSPDCRALAEDVKGAVEDLVMMRKKTAFNLTNE*

JGI12104J13512_10004525FeFe Group C3 hydrogenase catalytic subunit [Uncultured spirochete SA-8]

MDSAQVITETECQCYKLCRCPVKAIQVRQHVARLEDRCICGNCVEICPKQAKKVRSDLERAKILIIRLQTMLASPIAESFGIPMQKLACGKKLGFSHVSESVGADAVSESVSTKIAEQQPPLLMLSSAPAVQVYDYLPEMSFGISTACSMPMAHARIIKTQLTPGTAIFVAOGPICAKKREADSEGAVDVAFTQERLQWVFCEEGIPDEAAVREDSEFFNLAQGIDLYPEICEMGMAVISIKHLHSADMTYSGIHQIREAVRIGDVQYLENNPQNGQKTLFELMLACEGCGVNGPMVQSGTVAKRLVLEK
EKRTERGRVLSEPVSVDIAADLTMPQSRQKPVAAVAVSEEDIKASLASIGKYSRKDELNCSCGCG YDSCQFAAAMFLGKAETMCLSYTRKLQAKKANALLKAMPSAVAADDHIVENCNPFAELLG TDVMELYALKPSLEGADLRKLLPFWEAEFQVLMPESQDIVASDFCNGKIVHGSVFSIEKGLLAGG LFQDITAPWIQKDVRISQARKVMSQNLRTVQKIAAYLLGENAAAEAAALTSIESFQERSGDSYKSN KE*

DESUN47_v1_100151|ID:27211809| NiFe Group 1a hydrogenase catalytic subunit (fragment) [Uncultured Desulfobacterium sp N47]

MKVNLGPVTRIEGHNLIEIIETTVENKIVDARCMEGFRGFEVFLQGRSLDAQQITQRICGVCPYAH AVASSYAQESVYKLNBPNNPGRIMHNLIQGANHLYDYLLQFYQLAALDFVDTIALKYKG*

DESUN47_v1_100152|ID:27211810| NiFe Group 1b hydrogenase large subunit [Uncultured Desulfobacterium sp N47]

MAKRITIDPITRIEGHLRIEVEAVADGVNANWSSGQMFRGIEMLKGRDPRDAPLFTQRCGVCTYV HYLASVRAIEAAVGVIIPENARILRNLLHTGQYQHDHIHNYFHYHLHALDW/DILSAALKADPOQKTAGLAE NVCQARWGGTAYFKOQVRDRTKFVESQGQLQGNAYWGHPPAYLVSPEANMLAVSHYLEALRQA KAAQMHAVFGANKPHLQSLVGGTVCAMDLTDPDIAEFLYLYWKQTTQFVDVYLNPDVLAGSYFK DWGALGGTSSLAWFPEGEPERSLMPRGLMNNDISTVKQAEQDQITEHVASHWYVGNAD LHFPQGQTNQHDGYNPDDRYSWIKAPRYEGEPCEVGPLARMLVYAGSGKSTARKLVDQITQL SIPVTALFSTLGRRTAAARLETVDGAMEMW/IMKLVENKSGQDNTYQTVWTMPDKAIGCGLNDVP RGSLGHWIEIEDKKIKNYQYVYPSTWNLGPSCNGKGLGPVEQLGTVAPKDMLRTSVHSSDF PCIACAVHMIDPRESNEVYRIQVL*