npg

ORIGINAL ARTICLE

Protein-SIP enables time-resolved analysis of the carbon flux in a sulfate-reducing, benzene-degrading microbial consortium

Martin Taubert[1], Carsten Vogt[2], Tesfaye Wubet[3], Sabine Kleinsteuber[4], Mika T Tarkka[3], Hauke Harms[4], François Buscot[3,5], Hans-Hermann Richnow[2], Martin von Bergen[1,6] and Jana Seifert[1]

[1]Department of Proteomics, UFZ—Helmholtz Centre for Environmental Research, Leipzig, Germany; [2]Department of Isotope Biogeochemistry, UFZ—Helmholtz Centre for Environmental Research, Leipzig, Germany; [3]Department of Soil Ecology, UFZ—Helmholtz Centre for Environmental Research, Halle/Saale, Germany; [4]Department of Environmental Microbiology, UFZ—Helmholtz Centre for Environmental Research, Leipzig, Germany; [5]Institute of Biology, Chair of Soil Ecology, University of Leipzig, Leipzig, Germany and [6]Department of Metabolomics, UFZ—Helmholtz Centre for Environmental Research, Leipzig, Germany

Benzene is a major contaminant in various environments, but the mechanisms behind its biodegradation under strictly anoxic conditions are not yet entirely clear. Here we analyzed a benzene-degrading, sulfate-reducing enrichment culture originating from a benzene-contaminated aquifer by a metagenome-based functional metaproteomic approach, using protein-based stable isotope probing (protein-SIP). The time-resolved, quantitative analysis of carbon fluxes within the community supplied with either ^{13}C-labeled benzene or ^{13}C-labeled carbonate yielded different functional groups of organisms, with their peptides showing specific time dependencies of ^{13}C relative isotope abundance indicating different carbon utilization. Through a detailed analysis of the mass spectrometric (MS) data, it was possible to quantify the utilization of the initial carbon source and the metabolic intermediates. The functional groups were affiliated to Clostridiales, Deltaproteobacteria and Bacteroidetes/Chlorobi. The Clostridiales-related organisms were involved in benzene degradation, putatively by fermentation, and additionally used significant amounts of carbonate as a carbon source. The other groups of organisms were found to perform diverse functions, with Deltaproteobacteria degrading fermentation products and Bacteroidetes/Chlorobi being putative scavengers feeding on dead cells. A functional classification of identified proteins supported this allocation and gave further insights into the metabolic pathways and the interactions between the community members. This example shows how protein-SIP can be applied to obtain temporal and phylogenetic information about functional interdependencies within microbial communities.

Introduction

Besides their importance in all processes of environmental significance, such as carbon or nitrogen cycling, microbial communities also play essential roles in natural attenuation (NA) of contaminated environments. Microbial contaminant degradation is of specific interest in oxygen-limited or anoxic

habitats such as deep soils, sediments or aquifers, where contaminants display a long residence time and high persistence. Information on the carbon flux between community members is very useful for deciphering functioning of anaerobic subsurface microbial communities, and to understand the complex interactions within these communities. To perform an analysis that approximates reality, it is necessary to apply methods that allow conditions close to *in situ*. This excludes classical analyses involving the cultivation of individual community members.

One of the methods allowing the analysis of elemental fluxes between individual strains in

communities is stable isotope probing (SIP), in which substrates labeled with a heavy isotope (for example, ^{13}C or ^{15}N) are used in metabolic tracer experiments. The incorporation of the heavy isotopes in different classes of biomolecules is used as a marker for metabolic activity of the respective strain (Dumont and Murrell, 2005; Neufeld et al., 2007a). SIP has been applied to amino acids, phospholipid-derived fatty acids (Boschker et al., 1998), DNA and RNA (Manefield et al., 2002; Neufeld et al., 2007b; Chen and Murrell, 2010) and most recently to proteins (Jehmlich et al., 2008b, 2010).

Protein-SIP offers a more direct link to the function of an organism or a community than DNA/RNA-SIP. If a community sequence catalog/ reference is available, protein-SIP can provide detailed phylogenetic information together with the functional information. In protein-SIP, high-resolution mass spectrometry is used to detect and quantify the mass shift associated with heavy isotope labels in peptides. Simultaneously, the identity and origin of each peptide are revealed via comparison with database information, allowing a functional and phylogenetic classification.

During the past few years, the applicability of protein-SIP has been shown (Jehmlich et al., 2008a, b), and first applications on complex communities have demonstrated the virtue of this method (Bastida et al., 2010). In recent studies, it was shown that MS analysis allows detection of additional ^{13}C incorporation down to 0.1% relative isotope abundance (RIA), which exceeds the sensitivity of nucleotide-based methods by two orders of magnitude (Taubert et al., 2011a). Furthermore, its suitability for quantifying protein biosynthesis was demonstrated in a time-resolved induction study by assessing the intensity ratios between labeled and unlabeled peptide species (labeling ratios) in mass spectra (Taubert et al., 2011b).

This study focuses on a benzene-degrading, sulfate-reducing microbial community originating from a contaminated aquifer. The aquifer system and the inhabiting microbial community have been subject of a number of previous studies (Gödeke et al., 2006; Schirmer et al., 2006; Herrmann et al., 2008). Among the widespread BTEX (benzene, toluene, ethylbenzene, xylene) pollutants, benzene is the most hazardous compound and is especially problematic in the Earth's saturated critical zone. The exact mechanism of how benzene is degraded under anoxic conditions is still not entirely clear. Benzene is stabilized by a delocalized π electron system, thus probably an energy-dependent activation reaction is necessary in the degradation process. Methylation to toluene, carboxylation to benzoate or hydroxylation to phenol have been proposed as such activation mechanisms. Furthermore, taxonomically different organisms involved in anaerobic benzene degradation might use different mechanisms of ring activation (Vogt et al., 2011).

The laboratory microcosms used in this study were established from an aquifer community grown under conditions similar to those encountered in situ. In previous molecular genetic studies employing terminal restriction fragment length polymorphism (T-RFLP) analysis, 16S rRNA genes analysis and DNA-SIP, a phylotype belonging to the Cryptanaerobacter/Pelotomaculum group within the Peptococcaceae was found to be closely associated with the benzene-degrading activity (Kleinsteuber et al., 2008; Herrmann et al., 2010).

In this study, either $^{13}C_6$-labeled benzene or ^{13}C-labeled carbonate was used to trace the carbon flux within the microbial community. Microcosms were sampled at several time points to get a time-resolved picture of the metabolization of the labeled carbon. To facilitate the identification of the analyzed peptides, the metagenome of the community was sequenced by a shotgun approach. In combination with the SIP data, the functional metaproteomic analysis for the first time yielded a temporally and phylogenetically resolved picture of the carbon flux within a complex microbial community.

Materials and methods

Setup of microcosms

Biological material was derived from a contaminated industrial site near Zeitz, Saxony-Anhalt, Germany. The main organic contaminant at the site is benzene, which is present in concentrations up to 13 mM in two connected aquifers. The aquifers are anoxic, and sulfate is the main electron acceptor with concentrations up to 10 mM (for more information of the field site, see Schirmer et al. (2006)). NA processes have been extensively studied at the site (Gödeke et al., 2006; Schirmer et al., 2006). In an experimental plant on the site, a column system consisting of four columns wound in series and filled with gravel granules (grain size between 2 and 3.15 mm) has been continuously percolated with groundwater from the lower aquifer at a flow rate of $0.5 \, l \, h^{-1}$ since 2006. The columns are 190 cm in length and 27.3 cm in diameter, and are made of stainless steel. The percolating groundwater contains on an average 300 μM benzene and 4 mM sulfate (Vogt et al., 2007). Since 2007, benzene has been continuously degraded under sulfate-reducing conditions in all four columns. For cultivation of the benzene-degrading, sulfate-reducing community, gravel granules from the column system were transferred into sterile 1-l glass bottles, which were then immediately filled with anoxic, sulfidic groundwater and closed with butyl stoppers and screw caps.

Subsequently, groundwater was exchanged inside an anaerobic glove box (gas atmosphere 95% nitrogen and 5% hydrogen; Coy Laboratory Products Inc., Grass Lake, MI USA) by anoxic mineral salt medium, as described in Vogt et al. (2007), containing 20 mM sulfate as electron acceptor.

For the ^{13}C-benzene experiment, 10 bottles were prepared comprising five labeled incubations and five unlabeled controls. One of the latter was used for DNA extraction for a shotgun metagenome sequencing approach. For the ^{13}C-carbonate experiment, six bottles were prepared comprising three labeled incubations and three unlabeled controls. In the labeled incubations, $NaHCO_3$ in the mineral medium (final concentration 30 mM) was replaced by [^{13}C]$NaHCO_3$.

After exchange of the medium, 15 µl of benzene were added through the butyl stopper with a glass syringe (Hamilton, Reno, NV, USA). For the ^{13}C-benzene cultivation, [${}^{13}C_6$]benzene was used. Incubations were kept at room temperature in the dark during the cultivation. In addition, for both experiments sterile controls containing gravel granules that had been autoclaved three times were set up and treated identically to the incubations.

Benzene and sulfide concentrations, the carbon isotope composition of CO_2 and methane in the headspace of the microcosms, and the bacterial community composition based on T-RFLP patterns were monitored during cultivation as described in the Supplementary Information.

Protein extraction

At different times during the cultivation, sets of a labeled incubation and a control were sampled simultaneously. For the ^{13}C-benzene experiment, sampling times were at days 50, 63 (control incubation used for DNA extraction at day 220), 97, 180 and 300 of the cultivation. For the ^{13}C-carbonate experiment, incubations were sampled at days 75, 151 and 300 of the cultivation. An overview of the experimental setup and cultivation timeline can be found in Supplementary Figure S1. Protein extraction from the sediment was carried out using a modified phenol extraction protocol as described previously (Benndorf et al., 2009) with minor modifications. Details with regard to the extraction procedure are given in the Supplementary Information.

Sample preparation for one-dimensional gel electrophoresis

Because of low protein amounts, complete samples were used for one-dimensional (1D) gel electrophoresis without prior determination of protein amounts. On the basis of rough estimation from gel pictures, protein amounts varied between 5 and 100 µg. Air-dried protein pellets were suspended in 50 µl 1 × Laemmli buffer (Laemmli, 1970), dissolved in an ultrasonic bath and incubated for 10 min at 95 °C. Samples were centrifuged at 13 000 g to remove precipitates before loading on sodium dodecyl sulfate gels (4% stacking gel, 12% separating gel). Electrophoresis was performed at 10 mA per gel. Polypeptides were stained by colloidal Coomassie Brilliant Blue G-250 (Roth, Kassel, Germany). Entire lanes were cut into gel pieces for each sample and subsequently in-gel tryptic digestion was performed (Jehmlich et al., 2008a). A detailed protocol can be found in the Supplementary Information. Tryptic peptides were subjected to MS analysis. In samples of later time points, an optimized protocol was used: electrophoresis was performed only until the separation front had moved around 3 cm into the separating gel. Each entire lane was then cut into seven gel pieces before tryptic digestion, in the range of 200, 100, 60, 40, 30, 25 and 20 kDa (see Supplementary Table S1 for a detailed overview of all samples). A detailed protocol can be found in the Supplementary Information. Tryptic peptides were subjected to MS analysis.

MS analysis by UPLC-LTQ Orbitrap-MS/MS

Peptides were analyzed by UPLC-LTQ Orbitrap-MS/MS as described in Bastida et al. (2010). The peptides were eluted over gradients of different length (see Supplementary Table S1) with 2–60% solvent B (acetonitrile, 0.1% formic acid). Continuous scanning of eluted peptide ions was carried out between 150 and 2000 m/z, automatically switching to MS/MS CID mode on ions exceeding an intensity of 2000. Raw data were processed for database searches using the Thermo Proteome Discoverer software (v.1.2 build 208; Thermo Fisher Scientific, Waltham, MA, USA). Searches were performed by tandem mass spectrometry ion search algorithms from the Mascot house server (v.2.3) (Perkins et al., 1999) with the following parameters: tryptic cleavage with maximal two missed cleavage sites; a peptide tolerance threshold of ± 10 p.p.m. and an MS/MS tolerance threshold of ± 0.8 kDa; and carbamidomethylation at cysteines as static and oxidation of methionines as variable modifications. Searches were performed against bacterial and archaeal sequences of the NCBI non-redundant database (version 29 August 2011; National Center for Biotechnology Information, Rockville Pike, Bethesda, MD, USA) and a nucleic acid shotgun metagenome database from the benzene-degrading community (see Supplementary Information for more information).

In the NCBInr database search, proteins were considered as identified when at least two unique peptides were identified with at least medium confidence (false discovery rate below 0.05). In the metagenome databases, contigs were considered identified when at least two unique peptides with medium confidence or one unique peptide with high confidence (false discovery rate below 0.01) and Mascot ion score threshold >40 was identified. False discovery rates were checked by decoy database searches. For the identification of ^{13}C-labeled peptides, the respective ^{12}C peptides were measured as well to compare chromatographic

retention time and MS/MS fragmentation patterns, as described previously (Jehmlich *et al.*, 2008a).

Quantification of ^{13}C incorporation
Peak lists from mass spectra of identified peptides were manually selected and exported using Qual-Browser v.2.0.7 (Xcalibur, Thermo Fisher Scientific Inc., Waltham, MA, USA) to quantify ^{13}C incorporation. For each peptide, at least six single MS scans from the extracted ion chromatogram were averaged. For the calculation of ^{13}C RIAs (defined as the percentage of ^{13}C atoms in relation to the total number of carbon atoms in a peptide), a widely applied method based on comparison of theoretical and experimental spectral data was used (Snijders *et al.*, 2005). A Pearson correlation coefficient (*R*) was used to determine the best fit between calculated data and experimental data as described in MacCoss *et al.* (2005). To quantify protein biosynthesis, the labeling ratio (lr) between labeled and unlabeled peptide species (defined as the amount of labeled peptide species in total peptide species) was calculated as described in Taubert *et al.* (2011b). The doubling of the amount of a specific peptide, as a logarithmic measure of the increase of peptide amount, was calculated according to Equation (1):

$$d = -\log_2(1 - \text{lr}) \qquad (1)$$

Semiautomatic calculation of RIA and labeling ratio was carried out using an Excel spreadsheet as described previously (Taubert *et al.*, 2011a). For peak selection, a mass tolerance window of 5 p.p.m. was used. The spreadsheet is available for download from the website of the Helmholtz Centre for Environmental Research—UFZ (http://www.ufz.de/index.php?en=20647).

After quantification on peptide level, RIA, doublings and lr values were averaged for each identified protein in a sample. Proteins were grouped based on RIA and the shape of incorporation patterns, evaluated by comparison with theoretical pattern.

Annotation of metagenome sequences
In contigs containing identified peptides, corresponding open reading frames (ORFs) were manually annotated and analyzed regarding protein function and taxonomic affiliation by blast (National Library of Medicine, Bethesda, MD, USA) using the NCBI non-redundant database (version 29 August 2011). For functional annotation, blastp queries with predicted polypeptide sequences were used. A blast score of 50 was used as cutoff value. If a frameshift was detected among a single contig sequence, resulting in two adjacent ORFs in different frames matching two adjacent regions of the target protein sequence, the summed scores of both ORFs were compared with the cutoff value. On neighboring ORFs containing subunits of the same enzyme, only one of them had to reach the cutoff value. As short ORFs (<50 AS) in general can only produce low scores even for good alignments, for such ORFs a sequence identity cutoff value of 50% was used instead of a blast score cutoff value. For ORFs with blast scores <100, additionally blastx searches of the nucleic acid sequence were performed to exclude false positives by overlapping coding sequences in other reading frames. If best matches were unclassified (for example, hypothetical proteins, proteins of unknown function, and so on), the next match was chosen, as long as the other criteria were still fulfilled. For phylogenetic annotation, a blast score of 50 was used as cutoff value, and with regard to the phylogenic position of the hits, up to three best matches were analyzed if the score difference between them was below 100. The lowermost taxonomic category common to those hits was chosen. If the hits belonged to different phyla, the ORF was annotated as 'uncertain'. In addition, blastn searches of the whole contigs were performed and evaluated using the same blast thresholds. The phylogenetic affiliations estimated by blastn and blastp were compared to each other, and in case of differences, the results with higher scores were chosen. Nucleic acid sequences of the identified contigs are available at the website of the Helmholtz Centre for Environmental Research—UFZ (http://www.ufz.de/index.php?en=22376).

Results and discussion

Cultivation
To monitor the process of benzene degradation during cultivation, concentrations of benzene and sulfide as well as the carbon isotope composition of CO_2 were determined weekly. Benzene became depleted in all experimental setups except the abiotic controls, making repeated additions necessary (Supplementary Figures S2 and S3). In the controls, an initial drop in benzene concentration was observed that can be explained by adsorption of benzene to the butyl rubber stoppers used to seal the bottles gas-tight, as also observed by others (Holmes *et al.*, 2011). Benzene concentrations remained constant for more than 200 days after initial dispersion of the added benzene. The sulfide concentration was also found to increase in all cultures except the controls. High sulfide concentrations led to a cessation in benzene degradation, which could be restarted by precipitation of the sulfide with anoxic $FeCl_2$ solution. Benzene consumption and sulfide production rates showed a correlation of about 1.25 to 2.8 mol sulfide per mol benzene, which differed from the value of 3.75 expected from the stoichiometry of the reaction (Herrmann *et al.*, 2010). This discrepancy can be explained by the heterogeneous material of the samples and connected abiotic processes like precipitation of sulfide with iron ions stemming from

the sediment. Such effects were also observed in previous studies (Vogt et al., 2007).

The carbon isotope composition of CO_2 could only be determined in the ^{13}C-benzene cultivation due to the high amount of $^{13}CO_2$ in the ^{13}C-carbonate cultivation (Supplementary Figure S4). A nearly linear increase was observed in cultivations with ^{13}C-benzene, going up to $\delta^{13}C$ values of 17 500‰ (17.1% ^{13}C RIA; see Equations (2) and (3)). This increase was not observed in any control and can be regarded as a proof of the biotic degradation of ^{13}C-benzene to $^{13}CO_2$. A cessation in the increase near the end of the cultivation was in good agreement with the cessation of benzene degradation due to high sulfide concentrations. A comparison of the degradation and formation rates revealed that around 90% of the carbon of ^{13}C-benzene was degraded to $^{13}CO_2$. This agrees with previous findings, where also 87–95% of carbon derived from benzene was converted to CO_2 (Herrmann et al., 2010; Rakoczy et al., 2011):

$$^{13}C\,RIA = \left(\frac{1}{\left[\frac{\delta^{13}C}{1000} + 1 \right] \cdot R} + 1 \right)^{-1} \quad (2)$$

$$R = \frac{\text{standard } ^{13}C \text{ abundance}}{\text{standard } ^{12}C \text{ abundance}} = 1.2246 \cdot 10^{-2} \quad (3)$$

The formation of small amounts of methane was also observed. In ^{13}C-benzene and ^{13}C-carbonate cultivations, an increased ^{13}C RIA was found in the methane fraction compared to the corresponding ^{12}C cultivations, but methane amounts were too low for an exact quantification. Small amounts of labeled methane in ^{13}C-benzene-supplemented microcosms were also observed in a previous study (Herrmann et al., 2010), proving that a small part of the benzene-derived carbon was converted to methane.

Phylogenetic composition based on T-RFLP profiles

Community composition was monitored by T-RFLP fingerprinting targeting bacterial 16S rRNA genes. As shown in Supplementary Figure S5, the overall community structure did not change significantly over the experimental time, indicating that enrichment effects during the incubation on benzene were negligible. Therefore, the metagenome analyzed after 220 days of incubation was considered to be representative for the whole experimental period. The slightly different T-RFLP pattern of the 220 days sample (which also served for metagenome shotgun sequencing) might be explained by the different DNA purification procedure used for this sample. The T-RFLP pattern of the reactor sample directly taken from the Zeitz field site also points to a minor shift of the community composition, albeit the major phylotypes were identical. Regarding the specific conditions in the reactor columns, which are continuously percolated by the groundwater, minor

community shifts during batch cultivation in microcosms are not surprising.

Phylogenetic assignment of the T-RFs based on clone libraries revealed that the community was dominated by the Cryptanaerobacter/Pelotomaculum phylotype, which was previously identified as a key member of this benzene-degrading consortium (Kleinsteuber et al., 2008; Herrmann et al., 2010).

Protein identification

At several times within 300 days, samples from labeled and unlabeled microcosms were processed in parallel and extracted proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (see Supplementary Figures S6 and S7) before tryptic digest and analysis by UPLC-LTQ Orbitrap-MS/MS. Mass spectra were used for queries against the NCBInr database. The combined search of all samples yielded a total of 252 identified peptides belonging to 34 proteins.

The metagenome of the microbial community was sequenced by a shotgun approach. A 454 GS FLX Titanium sequencer run with subsequent sequence assembly resulted in 37 422 contigs ranging from 100 to 16 142 bp with an average length of 335 bp. The total size of the assembled data set was 24 120 113 bp. Automatic annotation using the MG-RAST platform (Meyer et al., 2008) resulted in 41 226 unique protein features and 4504 unique RNA features. The combined taxonomic profile of the annotation source databases used by MG-RAST revealed that 93.1% of all hits were assigned to Bacteria, 4.1% to Archaea and 1.9% to Eukaryota. The most abundant phyla were Proteobacteria with 22 876 hits and Firmicutes with 11 944 hits. According to the M5NR annotation source, the most abundant classes were Deltaproteobacteria and Clostridia, which were represented by 11 747 and 8453 hits, respectively. Among the Clostridia, Peptococcaceae was the predominant family.

Using the shotgun metagenome of the microbial community (see Supplementary Information for more details), a total of 2552 peptides from 688 proteins were identified as listed in the data set in the Supplementary Information. The significant rise of peptide/protein hits using a consortium-specific sequence database is an important aspect in metaproteomic approaches and was also described in other studies (Delmotte et al., 2009; Morris et al., 2010). For further analysis, only the identifications based on the shotgun metagenome data set were considered.

SIP results

The mass spectra of peptides identified in ^{13}C-benzene and ^{13}C-carbonate cultivations were analyzed for ^{13}C incorporation. In both experiments, specific groups of proteins showing different time-dependent RIAs were found and examples

are shown (Figure 1). In ^{13}C-benzene incubations, three different groups of peptides were found (Figures 1a–c). The first group showed a local
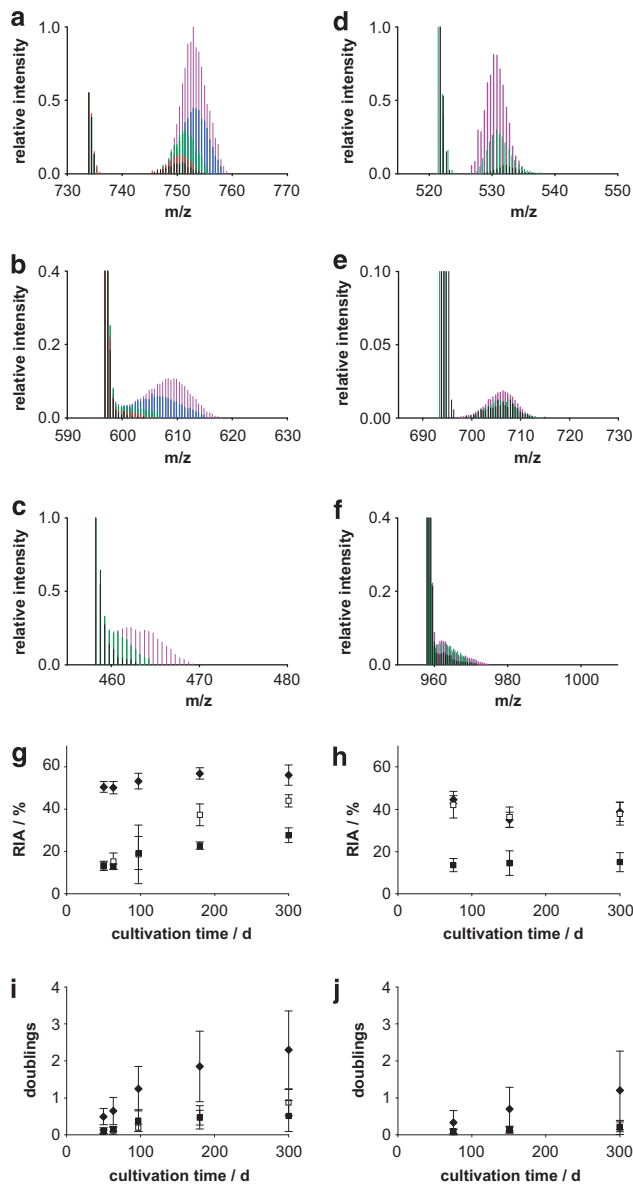

Figure 1 Exemplary peptide mass spectra with ^{13}C-incorporation patterns and development over time. (**a–c**) Examples of different ^{13}C-incorporation patterns of the ^{13}C-benzene cultivation at d_{50} (black), d_{63} (red), d_{97} (green), d_{180} (blue) and d_{300} (purple): Peptide SYQIVEGTSNIQK with local incorporation pattern at 60.1% RIA (**a**), peptide TLTAGQITPYK with incorporation range from 0% to 60% RIA (best fit at 43.6% RIA) (**b**) and peptide IATAQAANR with incorporation range from 0% to 40% RIA (best fit at 24.8% RIA) (**c**). (**d–f**) Examples of different ^{13}C-incorporation patterns of the ^{13}C-carbonate cultivation at d_{75} (black), d_{151} (green) and d_{300} (purple): Peptide APVVDDDGVR with local incorporation pattern at 40.6% RIA (**d**), peptide VFNIFGATGAD*m*K (*m* = oxidized methionine) with incorporation range from 25% to 45% RIA (best fit at 39.6% RIA) (**e**) and peptide SVLDNEAIVSDPILAGSSK with incorporation range from 0% to 25% RIA (best fit at 12% RIA) (**f**). Averaged RIA (**g**, **h**) and doublings of peptide amount (**i**, **j**) of all peptides with incorporation patterns as shown in **a** and **d** (black diamond), **b** and **e** (white square) or **c** and **f** (black square) from all time points of the ^{13}C carbonate and ^{13}C benzene incubations, respectively.

incorporation pattern with a fixed RIA (Figure 1a), indicating that benzene was used as a carbon source. The two remaining groups showed less focused, temporally shifting patterns, with low RIA close to the natural ^{13}C abundance at early time points and higher RIA up to 60 or 40% ^{13}C at later time points (Figures 1b and c). This can also be observed in the change of the average RIAs over time shown in Figure 1g. The shift is an indication of a carbon source changing its RIA over time, which is the case for metabolic products originating from an organism degrading a labeled carbon source.

In addition, in the ^{13}C-carbonate experiment, three groups of peptides were identified, one of which showed a local incorporation pattern (Figure 1d), which indicates the use of carbonate as a carbon source. The other patterns were again less focused, either around 40% RIA (Figure 1e) or not even exceeding 20% at late time points (Figure 1f). Furthermore, none of the groups showed a significant increase in RIA over time (Figure 1h).

Benzene mineralization to ^{13}CO$_2$ in the ^{13}C-benzene experiment and to non-labeled CO$_2$ in the ^{13}CO$_2$ experiment led to a respective enrichment or dilution of the CO$_2$ label. This can be clearly seen in both local patterns (Figures 1a and d) as a small shift of the RIA from early to late time points: to a higher RIA in the ^{13}C-benzene and to lower RIA in the ^{13}CO$_2$ experiment.

The doubling of the peptide amount, calculated from the labeling ratio, indicates an average synthesis rate more than twice as high for proteins with local incorporation patterns (Figures 1i and j). Despite having a high standard deviation due to variability in synthesis rates of the different proteins (see Supplementary Figure S8), averaged lr values correlated with the microbial growth rate, as was shown in a previous study (Taubert *et al.*, 2011b).

A comparison of the protein groups between the ^{13}C-benzene and the ^{13}CO$_2$ experiment revealed a strong correlation (Figure 2). Thus, based on the incorporation data three major groups of differentially labeled proteins were formed, putatively derived from three functional groups of organisms metabolizing different carbon sources.

Functional and phylogenetic annotation of identified proteins
On the basis of blastp and blastx searches, identified protein sequences were annotated in respect to their putative function and phylogenetic affiliation. An overview of the phylogenetic classification is given in Figure 3, showing all of the identified proteins (left) and the proteins of the three groups defined by ^{13}C incorporation. The total of proteins reflected the presence of bacteria from two major phyla, Firmicutes and Proteobacteria. The three groups defined by ^{13}C incorporation differed significantly from each other. While group I contained more than 60% of proteins affiliated to Clostridiales, group II
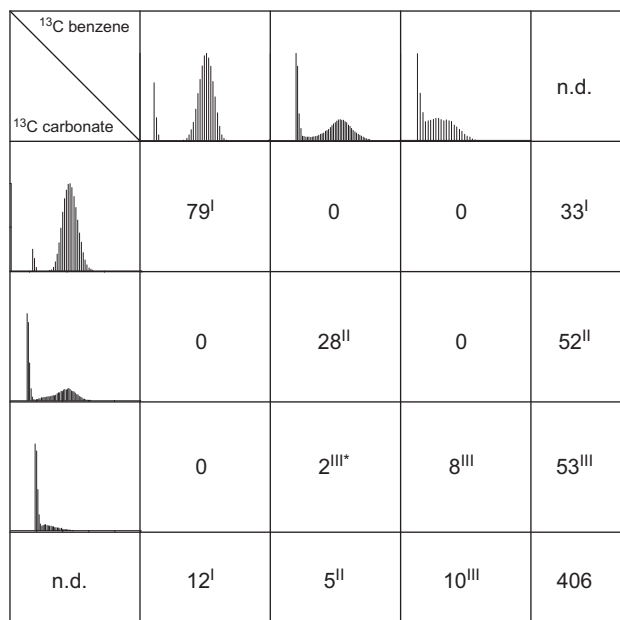
Figure 2 Abundance of ¹³C-incorporation patterns in identified proteins. The detected groups of incorporation patterns for the ¹³C-benzene experiment (top row) and the ¹³C-carbonate experiment (left column) are shown together with the number of identified proteins in each category. These were used to form three major groups of proteins from putative organisms utilizing different carbon sources. ND = no ¹³C-incorporation detectable. Roman numeral = group of proteins. *Proteins were classified as group III as differentiation of the patterns from the ¹³C-carbonate cultivation was more reliable than from the ¹³C-benzene cultivation.

Figure 3 Phylogenetic classification of identified proteins according to blast annotation of identified proteins. The broad bars show the phylogenetic groups present in the total amount of identified protein and in the three different groups of proteins that were formed based on incorporation patterns. The category 'uncertain/others' contains proteins assigned to various other phylogenetic groups than listed and proteins that produced significant blast results, but allowed no clear phylogenetic classification. The category 'unique' contains unique hypothetical proteins producing no significant blast results at all. The small bars show the phylogenetic distribution in selected phylogenetic marker proteins: chaperonines, subunits of F_0F_1 ATPases, ribosomal proteins and proteins involved in DSR. For group III, the number of phylogenetic marker proteins was too low and is not shown. The parenthetic numbers present the number of identified proteins in each group.

contained mostly proteins related to Deltaproteobacteria. This differentiation was also obvious in a subset of phylogenetic marker proteins (Gupta and Griffiths, 2002; Gevers et al., 2004). Group III, being more heterogeneous, showed a large proportion of proteins related to Bacteroidetes/Chlorobi, while most proteins could not be exactly classified. A more detailed phylogenetic classification based on the obtained blast data was possible for 209 of the identified proteins (see data set in the Supplementary Information). Within this subset, almost 95% of the group I Clostridiales affiliated proteins are related to the family Peptococcaceae with the genera Desulfotomaculum and Pelotomaculum (51 proteins). On the basis of the phylogenetic composition of the consortium according to T-RFLP profiles of bacterial 16S rRNA genes (see Supplementary Figure S5), it can be assumed that most of these proteins originate from the Cryptanaerobacter/Pelotomaculum sp. as it was the only abundant clostridial phylotype. More than 95% of the group II Deltaproteobacteria-related proteins were assigned to the genus Desulfobacca within the family Syntrophaceae (23 proteins). For the group III Bacteroidetes/Chlorobi related proteins, more than 60% were affiliated to the genus Rhodothermus within the Rhodothermaceae family (eight proteins).

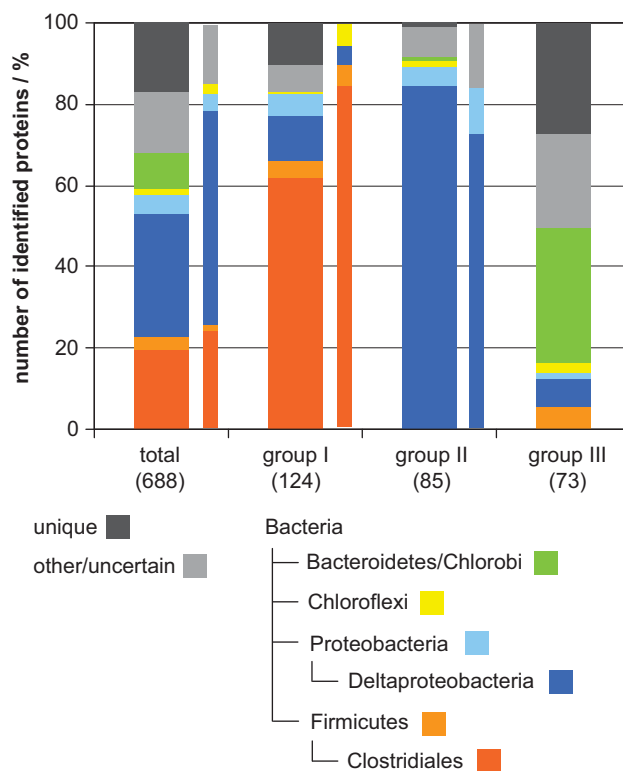The different protein groups based on distinct ¹³C incorporation patterns were clearly reflected in the

phylogenetic classification. The existing number of proteins from other phylogenetic groups can be explained by the lack of sequences for specific taxa in the NCBInr database used for the blast searches, leading to hits with a more distantly related protein. Moreover, horizontal gene transfer might have influenced protein distribution (Wiedenbeck and Cohan, 2011). In particular, syntrophy specialists such as *Pelotomaculum thermopropionicum* have been suggested to have evolved by genetic exchange with other niche-associated bacteria (Kosaka et al., 2008).

An overview of the functional categories of the identified proteins is given in Figure 4. Proteins related to dissimilatory sulfate reduction (DSR), aromatic compound degradation and C1 metabolism/CO₂ fixation are of particular interest to understand the ecophysiology of the benzene-degrading, sulfate-reducing community. While proteins involved in
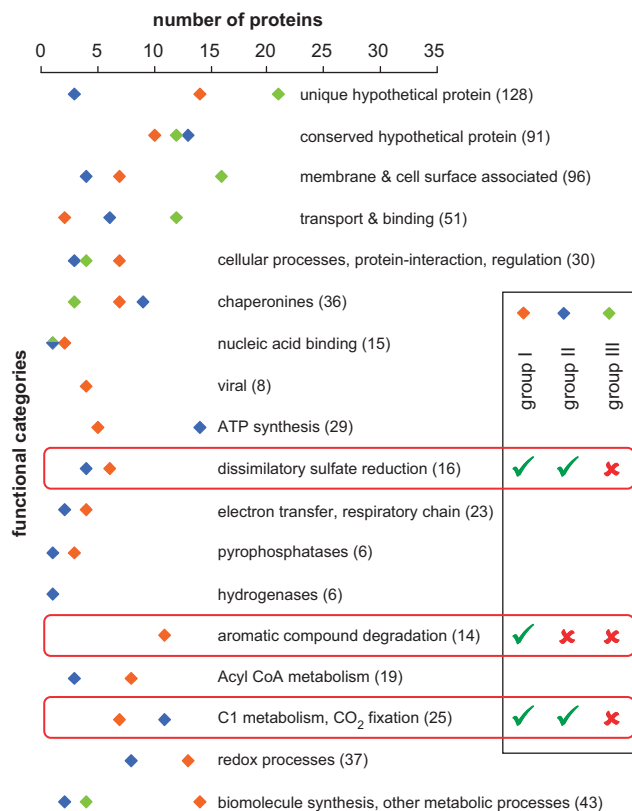
Figure 4 Functional classification of identified proteins. The numbers of proteins in different functional categories are shown for the protein groups with different ^{13}C-incorporation patterns. Numbers within parentheses show the total number of identified proteins in a category.
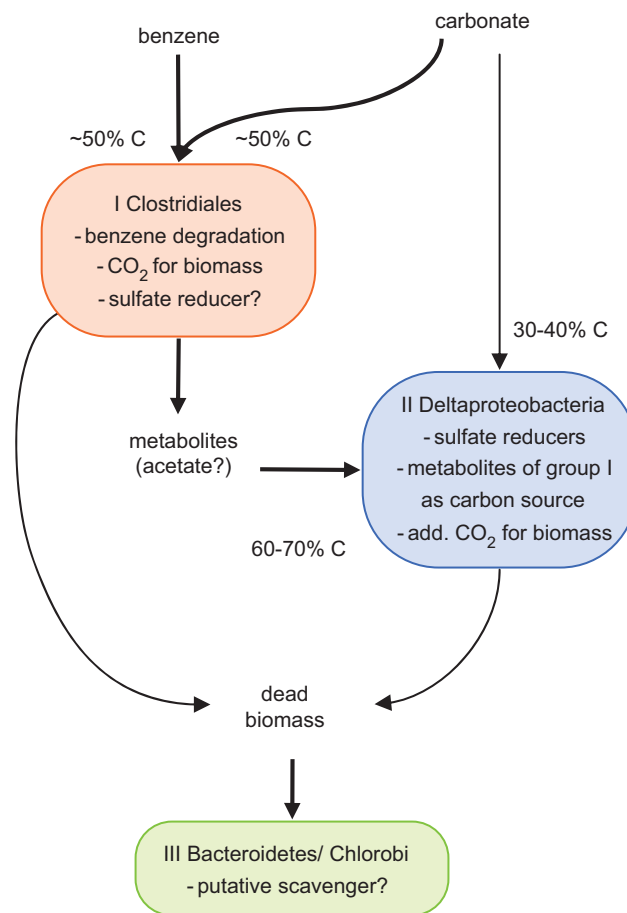

Figure 5 Hypothetical overview of the community. Carbon sources and carbon flux through the identified groups of organisms as conceived by the interpretation of the stable isotope probing results.

DSR and CO_2 fixation were found in both groups I and II, proteins involved in aromatic compound degradation were only found in group I (see also Supplementary Table S2).

Proteins involved in aromatic compound degradation cover nearly the whole benzoyl-CoA pathway, which is the proposed lower pathway of anaerobic benzene degradation (Vogt *et al.*, 2011). For the upper pathway with the initial step of benzene activation, only two proteins putatively involved in the β-oxidation pathway of benzylsuccinate were found, showing a low homology ($\sim 30\%$ identity) to BbsA and BbsB of Thauera aromatica. The presence of this pathway would favor an initial methylation of benzene over a carboxylation or hydroxylation reaction. However, a definitive annotation of the sequences is currently unfeasible due to the limited availability of closely related known proteins, so these proteins could also have different functions. Therefore, no further insights into the initial step of benzene degradation under sulfate-reducing conditions were obtained.

Community overview
On the basis of the ^{13}C incorporation data and the phylogenetic and functional protein annotation, a

scheme of the carbon flux within the community was developed (Figure 5). The group I organisms related to Clostridiales were found to be responsible for benzene degradation based on both ^{13}C incorporation and functional protein classification. This is in agreement with previous studies (Kleinsteuber *et al.*, 2008; Herrmann *et al.*, 2010), suggesting a Cryptanaerobacter/Pelotomaculum species to be the key organism for benzene degradation. Moreover, this organism showed a growth rate more than twice as high as other members of the community. This clearly highlights its primary role in carbon transformation. Furthermore, up to 50% of the assimilated carbon originated from carbonate. A high amount of heterotrophic carbon fixation is necessary, as species of the Peptococcaceae typically lack enzymes of the glyoxylate cycle (Kosaka *et al.*, 2006), which could be used for biomass accumulation from acetyl-CoA (Rabus, 2005). A heterotrophic carbon fixation rate of a similar level was observed in a toluene-degrading enrichment culture (Winderl *et al.*, 2010). The authors of that study proposed this to be a result of biomass accumulation from acetyl-CoA via carboxylation to pyruvate by the enzyme

pyruvate:ferredoxin oxidoreductase. Such an enzyme was also found in this study in one of the unlabeled cultures, so no information on ^{13}C incorporation was acquired. An alternative explanation of the high $^{13}CO_2$-derived label might be CO_2 fixation by the reductive acetyl-CoA pathway (Wood–Ljungdahl pathway) (Wood, 1991), of which several enzymes were identified in organisms belonging to group I.

For all proposed putative activation reactions of benzene to benzoyl-CoA, it is presumed that ATP is necessary to facilitate destabilization of the aromatic ring. The further degradation via the benzoyl-CoA pathway to acetyl-CoA yields 6 reducing equivalents, 1 equivalent of CO_2 and 3 equivalents of acetyl-CoA (Harwood et al., 1998), but no ATP. While the excess of reducing equivalents has to be eliminated by the organism, the initial activation step needs to be fueled by the further metabolization of acetyl-CoA. As the acetyl-CoA pathway is completely reversible, it could be used for the production of ATP from acetyl-CoA (Ragsdale, 1991). A much easier way for ATP production, however, would be acetate release via the ubiquitous enzymes phosphotransacetylase and acetate kinase (Wolfe, 2005). For the general carbon flux following the metabolization of benzene to 3 equivalents of acetyl-CoA, two different routes appear possible:

(1) The 3 equivalents of acetyl-CoA are further metabolized either to CO_2 or released as acetate by the pathways described above, yielding 3 equivalents of ATP, and in the case of the acetyl-CoA pathway, 3×8 additional reducing equivalents (Ragsdale, 1991). Thus, sufficient amounts of ATP are produced to fuel the benzene activation. Biomass synthesis can then proceed from acetyl-CoA via carboxylation to pyruvate. With ^{13}C-benzene, this could result in a biomass RIA of 50% to 60% (Winderl et al., 2010) and a nearly complete labeling of the released acetate. With ^{13}C-carbonate, the biomass RIA would be around 40% to 50%, while released acetate would be essentially unlabeled.

(2) In addition to the formation of acetyl-CoA from benzene degradation, acetyl-CoA is produced from CO_2, which requires both reducing equivalents and ATP. Acetyl-CoA is then released as acetate for ATP production. This would result in an ATP-neutral turnover of CO_2 to acetate, driven by benzene-derived reducing equivalents. Acetyl-CoA and released acetate should have an intermediate RIA between that of CO_2 and benzene. In this case, biomass accumulation has to proceed solely from acetyl-CoA to explain the observed RIA.

The high abundance of proteins from the reductive acetyl-CoA pathway could be explained by a preferential use of route 2. However, the detected proteins might as well be constitutively expressed and do not necessarily have to play an important role under conditions of benzene degradation. The data obtained from the SIP experiments and data from previous studies are in favor of route 1. The group I organism might be able to switch between the described pathways of metabolization of acetyl-CoA, depending on the level of reducing equivalents present. Assuming a syntrophic interaction with interspecies hydrogen transfer, the release of acetate would be necessary to prevent the accumulation of hydrogen that would otherwise inhibit a further elimination of reducing equivalents (Rakoczy et al., 2011). Strong indications for acetate and hydrogen as intermediate metabolites of benzene mineralization by the studied consortium have been recently gained in inhibition experiments, in which acetate and hydrogen supplementation transiently stopped benzene mineralization, even at low concentrations (Rakoczy et al., 2011). The hypothesis of a syntrophic interaction is also in accordance with a recent study of van der Zaan et al. (2012), who found that Peptococcaceae were involved in anaerobic degradation of benzene with different electron acceptors.

Group II organisms related to Deltaproteobacteria have been shown to use metabolic products of benzene degradation. For group II, the amount of labeling can be explained by the assimilation of fermentation products (for example, acetate) of group I organisms and heterotrophic CO_2 fixation. Furthermore, hydrogen released by group I organisms might be used as electron donor, and energy production by sulfate reduction seems plausible, as proteins involved in DSR have been found in group II. Such proteins (subunits of dissimilatory sulfite reductase and adenylyl sulfate reductase) have also been found in group I organisms; however, the results of recent research are inconclusive regarding the ability of Pelotomaculum species to perform DSR (Kosaka et al., 2008). As discussed by Plugge et al. (2011), the absence of genes encoding transmembrane electron transfer complexes in the genome of P. thermopropionicum is in accordance with the previously detected inability of Desulfotomaculum subcluster Ih strains to perform DSR. Therefore, it cannot be excluded that these proteins are evolutionary remnants or are actually involved in different processes (Plugge et al., 2011).

Group III organisms related to Bacteroidetes/ Chlorobi showed a lower RIA in both experiments that cannot be explained by the use of any of the discussed carbon sources. Nevertheless, incorporation of carbon derived from benzene was observed. Another carbon source with initially low RIA that slowly increased over time is the pool of lysed cells in the culture. In this instance, an increase in RIA can be explained by the lysis of partly labeled cells during cultivation. Thus, group III organisms might actually function as scavengers feeding on dead cells in the community.

2300

In summary, we were able to trace the flow of carbon derived from benzene and carbonate in a sulfate-reducing, benzene-degrading community by functional metaproteome analysis using protein-SIP. Three major functional groups of organisms were detected, each metabolizing different carbon sources. These groups were: (i) a clostridial group putatively fermenting benzene while fixing significant amounts of CO_2; (ii) a sulfate-reducing deltaproteobacterial group utilizing metabolites released during benzene fermentation; and (iii) a putative scavenger group belonging to Bacteroidetes/Chlorobi. The predicted pathway of benzene degradation is via metabolization to acetyl-CoA by the Clostridiales, initiated by a yet unknown activation step. Acetyl-CoA is then either further degraded to CO_2 or released as acetate by the Clostridiales, depending on the energy level. Deltaproteobacteria further metabolized the putatively released acetate, and interspecies hydrogen transfer possibly takes place between Clostridiales and Deltaproteobacteria as suggested previously (Rakoczy et al., 2011). Our study illustrates the power of protein-based SIP, in combination with a comprehensive metagenome data set, to decipher the carbon flow within complex microbial communities.

Acknowledgements

References

Bastida F, Rosell M, Franchini AG, Seifert J, Finsterbusch S, Jehmlich N et al. (2010). Elucidating MTBE degradation in a mixed consortium using a multidisciplinary approach. FEMS Microbiol Ecol 73: 370–384.

Benndorf D, Vogt C, Jehmlich N, Schmidt Y, Thomas H, Woffendin G et al. (2009). Improving protein extraction and separation methods for investigating the metaproteome of anaerobic benzene communities within sediments. Biodegradation 20: 737–750.

Boschker HTS, Nold SC, Wellsbury P, Bos D, de Graaf W, Pel R et al. (1998). Direct linking of microbial populations to specific biogeochemical processes by C-13-labelling of biomarkers. Nature 392: 801–805.

Chen Y, Murrell JC. (2010). When metagenomics meets stable-isotope probing: progress and perspectives. Trends Microbiol 18: 157–163.

Delmotte N, Knief C, Chaffron S, Innerebner G, Roschitzki B, Schlapbach R et al. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc Natl Acad Sci USA 106: 16428–16433.

Dumont MG, Murrell JC. (2005). Stable isotope probing–linking microbial identity to function. Nat Rev Microbiol 3: 499–504.

Gevers D, Vandepoele K, Simillon C, Van de Peer Y. (2004). Gene duplication and biased functional retention of paralogs in bacterial genomes. Trends Microbiol 12: 148–154.

Gödeke S, Richnow HH, Weiss H, Fischer A, Vogt C, Borsdorf H et al. (2006). Multi tracer test for the implementation of enhanced in-situ bioremediation at a BTEX-contaminated megasite. J Contam Hydrol 87: 211–236.

Gupta RS, Griffiths E. (2002). Critical issues in bacterial phylogeny. Theor Popul Biol 61: 423–434.

Harwood CS, Burchhardt G, Herrmann H, Fuchs G. (1998). Anaerobic metabolism of aromatic compounds via the benzoyl-CoA pathway. FEMS Microbiol Rev 22: 439–458.

Herrmann S, Kleinsteuber S, Chatzinotas A, Kuppardt S, Lueders T, Richnow HH et al. (2010). Functional characterization of an anaerobic benzene-degrading enrichment culture by DNA stable isotope probing. Environ Microbiol 12: 401–411.

Herrmann S, Kleinsteuber S, Neu TR, Richnow HH, Vogt C. (2008). Enrichment of anaerobic benzene-degrading microorganisms by in situ microcosms. FEMS Microbiol Ecol 63: 94–106.

Holmes DE, Risso C, Smith JA, Lovley DR. (2011). Anaerobic oxidation of benzene by the hyperthermophilic archaeon Ferroglobus placidus. Appl Environ Microbiol 77: 5926–5933.

Jehmlich N, Schmidt F, Hartwich M, von Bergen M, Richnow HH, Vogt C. (2008a). Incorporation of carbon and nitrogen atoms into proteins measured by protein-based stable isotope probing (Protein-SIP). Rapid Commun Mass Spectrom 22: 2889–2897.

Jehmlich N, Schmidt F, Taubert M, Seifert J, Bastida F, von Bergen M et al. (2010). Protein stable istope probing (Protein-SIP). Nat Protoc 5: 1957–1966.

Jehmlich N, Schmidt F, von Bergen M, Richnow HH, Vogt C. (2008b). Protein-based stable isotope probing (Protein-SIP) reveals active species within anoxic mixed cultures. ISME J 2: 1122–1133.

Kleinsteuber S, Schleinitz KM, Breitfeld J, Harms H, Richnow HH, Vogt C. (2008). Molecular characterization of bacterial communities mineralizing benzene under sulfate-reducing conditions. FEMS Microbiol Ecol 66: 143–157.

Kosaka T, Kato S, Shimoyama T, Ishii S, Abe T, Watanabe K. (2008). The genome of Pelotomaculum thermopropionicum reveals niche-associated evolution in anaerobic microbiota. Genome Res 18: 442–448.

Kosaka T, Uchiyama T, Ishii S, Enoki M, Imachi H, Kamagata Y et al. (2006). Reconstruction and regulation of the central catabolic pathway in the thermophilic propionate-oxidizing syntroph Pelotomaculum thermopropionicum. J Bacteriol 188: 202–210.

Laemmli UK. (1970). Cleavage of structural proteins during assembly of head of bacteriophage-T4. Nature 227: 680–685.

MacCoss MJ, Wu CC, Matthews DE, Yates JR. (2005). Measurement of the isotope enrichment of stable isotope-labeled proteins using high-resolution mass spectra of peptides. Anal Chem 77: 7646–7653.

Manefield M, Whiteley AS, Griffiths RI, Bailey MJ. (2002). RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. Appl Environ Microbiol 68: 5367–5373.

Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M et al. (2008). The metagenomics RAST

server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinform* **9**: 386.

Morris RM, Nunn BL, Frazar C, Goodlett DR, Ting YS, Rocap G. (2010). Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* **4**: 673–685.

Neufeld JD, Dumont MG, Vohra J, Murrell JC. (2007a). Methodological considerations for the use of stable isotope probing in microbial ecology. *Microbiol Ecol* **53**: 435–442.

Neufeld JD, Vohra J, Dumont MG, Lueders T, Manefield M, Friedrich MW *et al.* (2007b). DNA stable-isotope probing. *Nat Protoc* **2**: 860–866.

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**: 3551–3567.

Plugge CM, Zhang W, Scholten JC, Stams AJ. (2011). Metabolic flexibility of sulfate-reducing bacteria. *Front Microbiol* **2**: 81.

Rabus R. (2005). Functional genomics of an anaerobic aromatic-degrading denitrifying bacterium, strain EbN1. *Appl Microbiol Biotechnol* **68**: 580–587.

Ragsdale SW. (1991). Enzymology of the acetyl-CoA pathway of CO_2 fixation. *Crit Rev Biochem Mol Biol* **26**: 261–300.

Rakoczy J, Schleinitz KM, Müller N, Richnow HH, Vogt C. (2011). Effects of hydrogen and acetate on benzene mineralisation under sulphate-reducing conditions. *FEMS Microbiol Ecol* **77**: 238–247.

Schirmer M, Dahmke A, Dietrich P, Dietze M, Godeke S, Richnow HH *et al.* (2006). Natural attenuation research at the contaminated megasite Zeitz. *J Hydrol* **328**: 393–407.

Snijders APL, de Koning B, Wright PC. (2005). Perturbation and interpretation of nitrogen isotope distribution patterns in proteomics. *J Proteome Res* **4**: 2185–2191.

Taubert M, Baumann S, von Bergen M, Seifert J. (2011a). Exploring the limits of robust detection of incorporation of (13)C by mass spectrometry in protein-based stable isotope probing (protein-SIP). *Anal Bioanal Chem* **401**: 1975–1982.

Taubert M, Jehmlich N, Vogt C, Richnow HH, Schmidt F, von Bergen M *et al.* (2011b). Time resolved protein-based stable isotope probing (Protein-SIP) analysis allows quantification of induced proteins in substrate shift experiments. *Proteomics* **11**: 2265–2274.

van der Zaan BM, Talarico Saia F, Stams AJM, Plugge CM, de Vos WM, Smidt H *et al.* (2012). Anaerobic benzene degradation under denitrifying conditions: Peptococcaceae as dominant benzene degraders and evidence for a syntrophic process. *Environ Microbiol* **14**: 1171–1181.

Vogt C, Godeke S, Treutler HC, Weiss H, Schirmer M, Richnow HH. (2007). Benzene oxidation under sulfate-reducing conditions in columns simulating *in situ* conditions. *Biodegradation* **18**: 625–636.

Vogt C, Kleinsteuber S, Richnow HH. (2011). Anaerobic benzene degradation by bacteria. *Microb Biotechnol* **4**: 710–724.

Wiedenbeck J, Cohan FM. (2011). Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol Rev* **35**: 957–976.

Winderl C, Penning H, von Netzer F, Meckenstock RU, Lüders T. (2010). DNA-SIP identifies sulfate-reducing Clostridia as important toluene degraders in tar-oil-contaminated aquifer sediment. *ISME J* **4**: 1314–1325.

Wolfe AJ. (2005). The acetate switch. *Microbiol Mol Biol Rev* **69**: 12–50.

Wood HG. (1991). Life with CO or CO_2 and H_2 as a source of carbon and energy. *FASEB J* **5**: 156–163.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)