

Microbial metabolism in the deep subsurface: Case study of Opalinus Clay

THÈSE N° 6727 (2015)

PRÉSENTÉE LE 14 AOÛT 2015

À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE DE MICROBIOLOGIE ENVIRONNEMENTALE
PROGRAMME DOCTORAL EN GÉNIE CIVIL ET ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Alexandre BAGNOUD

acceptée sur proposition du jury:

Prof. S. Takahama, président du jury
Prof. R. Bernier-Latmani, directrice de thèse
Dr O. Leupin, rapporteur
Prof. J. R. van der Meer, rapporteur
Prof. T. Battin, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2015

Acknowledgements

This work would not have been possible without the help, the support, and the collaboration of many people, colleagues, and friends. The first person that I would like to acknowledge is Rizlan Bernier-Latmani. She is a great supervisor with many ideas, but she also listens to her students. She has an enormous passion and motivation for the projects she is handling. And because she easily transmits these feelings, she was really supportive when I was passing through a period of discouragement and doubt.

Then, I want to thank my other EML colleagues, which make the lab a great environment in which to work. It was a pleasure to come to work everyday with such nice people. In particular, Manon and Leia are acknowledged for their help during some of my numerous field trips to St-Ursanne. I also would like to acknowledge some other EPFL colleagues, including Pierre Rossi, Julien Maillard, Elena Rossel, Sylvain Coudret, Karine Vernez, and many others for their help with various analyses and/or for the interesting discussions I had with. I am also really grateful to Jacky and Marie from the magasin de chimie, who provide precious logistical help for our lab, always with a smile!

Then, I want to thank all of our collaborators from the Mont Terri project. Of course on the top this list is the Swisstopo crew from St-Ursanne, which works really hard to make the best working environment possible in the underground rock laboratory. Running a continuous in situ experiment would not have been possible without their support. Among these people, Thierry Theurillat and Nicola Kern helped a lot for our experimental sampling and monitoring, and I am really grateful to them. Then, I would like to thank our colleagues from Solexperts, namely Yanick Lettry and Thomas Fierz, who did a great job with the design, the building, and the installation of the bioreactor in the rock laboratory. Yanick even developed an excellent sense of patience while he was answering my innumerable questions and requests. Then, I would like to acknowledge Nagra for its financial support, and also for some of its collaborators who were really supportive for the scientific background of the project, including Bernhard Schwyn, Olivier Leupin, Nikitas Diomidis, and Herwig Müller. Finally, I want to thank all the participants in the annual technical discussion in St-Ursanne about geochemistry and microbiology experiments. All these discussions were really enriching, and it was great to share our findings with this involved audience.

Of course this project would not have been possible without some help and collaboration with other institutes. I am grateful to Anders Andersson, Ino de Bruijn and Johannes Alneberg in SciLifeLab in

Stockholm, which welcomed me for two memorable weeks, during which they introduced me to bioinformatics and helped with the analysis of the innumerable number of sequence data I received. I also want to thank Bob Hettich and Karuna Chourey from Oak Ridge National Laboratory for their work with the metaproteomic sequencing. And of course I have to acknowledge the Joint Genome Institute for their sequencing support, more particularly Tijana Glavina del Rio for the coordination.

And finally this long list would not be complete without all my friends and my family! Even though I am not sure they all understood what I was doing, they were of a great support, particularly my wife Mireille.

Lausanne, April 24th 2015

Abstract

Microbes have successfully colonized the deep subsurface, thanks to their small size and their diverse metabolic activities. This part of the biosphere remains a *terra incognita* for microbiologists, containing innumerable unknown microbial species and processes. We depend on it for many things, i. e. water supply, oil extraction and nuclear waste disposal. In Switzerland, Opalinus Clay will be used for deep geological repositories, because of its very low permeability. This rock has been studied for 20 years, however, the potential role of microbes has long been overlooked, even though they could play a major role in controlling geochemical conditions. It was showed that they were present in pristine rock, but were not active due to the lack of space. The potential for microbial activity in disturbed rock, under repository relevant conditions, remains unexplored. This is the focus of this thesis.

The first condition tested was the presence of space alone. This will occur when the rock is excavated to build the repository. The digging creates a network of large fractures in the rock zones near the galleries. This change is enough to promote microbial activity in a system, mainly composed of two species. The first one, a Peptococcaceae, oxidizes organic carbon to dioxide carbon, by reducing sulfate. The second organism is a *Pseudomonas* that seems to grow by fermenting organic carbon. From this model, it appears that the Peptococcaceae feeds on low molecular organic acids present in the Opalinus Clay, but also on fermentation products from *Pseudomonas*. What is less clear is the source of organic carbon for the latter. Is it only feeding on dead microbial cells, or is it also able to feed on reactive fossilized carbon contained in Opalinus Clay?

The second case study included an additional energy source, hydrogen. This gas will be produced by anoxic steel corrosion and could damage the repository if the pressure increases significantly. However, it can be oxidized *in situ* by autotrophic microbial species. The dominating microbe in this system is a sulfate-reducing Desulfobulbaceae. It produces organic carbon from carbon dioxide, which can later feed heterotrophic bacteria, which are also sulfate-reducing and that release carbon dioxide. A carbon loop was thus reconstructed for the first time in the deep subsurface. It includes a fermentation step that releases the organic substrate for heterotrophic sulfate-reducing bacteria that completely oxidize simple organic carbon molecules, i.e., acetate, to carbon dioxide.

In this system, hydrogen and sulfate consumption rates are 1.5 and 0.2 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$, respectively. This means that electrons derived from hydrogen reduce electron acceptors other than sulfate, confirming carbon fixation.

These two projects highlight the role that microbes can play for the safety of nuclear waste disposal. Hydrogen consumption is a beneficial process, but the overall impact of microbial activity can be negative, because it can promote corrosion and weathering of the different barriers surrounding the waste. Further work is needed in order to obtain a more complete picture of the role played by bacteria in a deep geological repository. Nonetheless, the results presented here represent a large improvement in our understanding of deep subsurface microbiology. Microbial processes were never described in this much detail in these environments, as is done here: from the metabolic pathway level, up to the ecosystem level.

Keywords

Deep subsurface microbiology, deep biosphere, Opalinus Clay, metagenomics, metaproteomics, sulfate reduction, hydrogen consumption

Résumé

Les microorganismes ont colonisé avec succès les souterrains profonds, grâce à leur petite taille et leur activité métaboliques diverses. Cette partie de la biosphère reste une *terra incognita* pour les microbiologistes, contenant des espèces et des processus microbiens inconnus. Nous y dépendons pour de nombreuses choses, telles que les réserves d'eau, l'extraction de pétrole, et le stockage de déchets nucléaires. En Suisse, les argiles à Opalinus vont être utilisés comme site de dépôts profonds, puisqu'ils ont une très faible perméabilité. Cette roche est étudiée depuis 20 ans bien que les microbes n'ont été inclus que plus tard, bien qu'ils soient connus pour jouer un rôle majeur, en contrôlant les conditions géochimiques. Il a été montré qu'ils sont présents dans la roche intacte, bien qu'inactifs à cause d'un manque d'espace. En revanche, rien n'est connu quant à l'activité microbienne dans la roche abîmée, qui sera présente dans les dépôts.

Les premières conditions testées concernent un unique apport d'espace dans la roche. Cela se produira lors de l'excavation du dépôt qui va créer un réseau de fractures plus grandes dans la zone de roche autour des galeries. Ce changement est suffisant pour promouvoir l'activité microbienne dans un système principalement composé de deux espèces. La première, un Peptococcaceae, couple l'oxydation du carbone organique en dioxyde de carbone avec une réduction du sulfate. Le deuxième organisme est un *Pseudomonas* qui semble pousser en fermentant le carbone organique. Il apparaît donc que le Peptococcaceae se nourrit des acides organiques de faible poids moléculaire produits par fermentation par le *Pseudomonas*. Ce qui est moins clair est la source de carbone organique de ce dernier. Se nourrit-il de cellules microbiennes mortes, ou également de la matière organique fossile et encore réactive présente dans les argiles à Opalinus ?

Le deuxième cas étudié inclut une source d'énergie supplémentaire, le dihydrogène. Ce gaz va être produit par la corrosion anoxique de l'acier et pourrait endommager le dépôt sur la pression augmente trop. Cependant, il peut être oxydé *in situ* par plusieurs espèces microbiennes autotrophes. Celle qui domine est un Desulfobulbaceae sulfato-réducteur. Elle produit depuis le dioxyde de carbone du carbone organique qui pourra plus tard nourrir des bactéries hétérotrophes, dont la grande majorité réduit également le sulfate et qui produit en retour du dioxyde de carbone. Un cycle du carbone a donc été décrit pour la première fois dans un sous-terrain profond. Il inclut une étape de fermentation qui libère des composés organiques pour les bactéries sulfato-réductrices hétérotrophes, qui oxydent complètement ces simples molécules de carbone, comme l'acétate par exemple, jusqu'au dioxyde de carbone.

Dans ce système, les taux de consommation de l'hydrogène et du sulfate sont de 1.5 et de 0.2 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$, respectivement. Cela signifie que les électrons provenant de l'hydrogène réduisent des accepteurs d'électrons autres que le sulfate, confirmant la fixation de carbone.

Ces deux projets mettent en valeur le rôle que les microorganismes jouent au niveau de la sécurité des dépôts de déchets nucléaires. La consommation du dihydrogène est un processus bénéfique, mais l'impact global d'une activité microbienne peut être négatif, puisqu'elle peut promouvoir la corrosion et l'altération des différentes barrières entourant les déchets. Bien sûr, d'autres conditions doivent être testées afin d'obtenir une image plus complète du rôle joué par les bactéries dans un dépôt en couches profondes. Ce travail est également une avancée significative dans la compréhension de la microbiologie sous-terrain profonde. Les processus microbiens n'y avaient jamais été décrits si précisément, depuis le niveau de la route métabolique jusqu'à celui de l'écosystème.

Mots-clés

Microbiologie du sous-terrain profond, biosphère profonde, Opalinus Clay, métagénomique, métaprotéomique, sulfato-réduction, consommation de l'hydrogène

Contents

ACKNOWLEDGEMENTS	I
ABSTRACT	III
KEYWORDS	IV
RÉSUMÉ	V
MOTS-CLÉS	VI
CONTENTS	VII
LIST OF FIGURES	XIII
LIST OF TABLES	XV
CHAPTER 1 INTRODUCTION TO DEEP SUBSURFACE MICROBIOLOGY	1
1.1 Definition	3
1.2 Interest in subsurface microbiology	4
1.3 Limitations of deep subsurface microbiology	5
1.4 General characteristics of terrestrial deep-subsurface environments	6
1.4.1 Pore size and interconnectivity	6
1.4.2 Absence of sunlight	6
1.4.3 Anoxicity	7
1.4.4 Water availability	7
1.4.5 Temperature	8
1.4.6 Pressure	8
1.5 Two major types of rock substrate	9
1.5.1 Sedimentary rocks	9
1.5.2 Igneous and metamorphic rocks	10
1.6 Microbial communities in the deep subsurface	11
1.6.1 Number of individual cells	11
1.6.1 Diversity of microbial populations	11
1.6.2 Metabolic state of microorganisms	11
1.6.3 Origin of microorganisms	12
1.7 Deep subsurface bioenergetics	13
1.7.1 Organic matter a source of energy	14
1.7.1.1 Degradation of organic matter in anoxic systems	14

1.7.1.2	Source of organic matter	17
1.7.2	Inorganic compounds as source of energy	18
1.7.2.1	The different sources of electron acceptors	18
1.7.2.2	Microbial energetic systems based on inorganic sources of electrons	18
CHAPTER 2 STATE OF THE ART: OPALINUS CLAY AS A DEEP GEOLOGICAL REPOSITORY HOST ROCK		21
2.1	General context of the study	23
2.1.1	Nuclear waste disposal in Switzerland	23
2.1.2	Underground rock laboratory of Mont Terri	24
2.1.3	A need for microbiological characterization of host rocks	24
2.2	The Opalinus Clay formation	26
2.2.1	Opalinus Clay diagenesis	26
2.2.2	Opalinus Clay solid phase	27
2.2.3	Opalinus Clay porewater	28
2.3	Microbiological study of the Opalinus Clay	28
2.3.1	First evidence of microbial activity in Opalinus Clay porewater	28
2.3.2	Impact of microbial activity in deep geological nuclear waste repository	29
2.3.3	Microbial activity in the Opalinus Clay porewater	29
2.3.4	Microbial activity in the Opalinus Clay rock	30
2.3.5	Conclusions of these studies and open questions	32
CHAPTER 3 MICROBIAL METABOLISMS IN THE POREWATER OF OPALINUS CLAY		35
Abstract		36
3.1	Introduction	37
3.2	Materials and Methods	38
3.2.1	Porewater sampling	38
3.2.2	Chemical analysis of porewater samples	39
3.2.3	DNA extraction	40
3.2.4	16S rRNA genes analysis	40
3.2.4.1	Libraries preparation and sequencing	40
3.2.4.2	Alpha-diversity analysis	41
3.2.4.3	Sample clustering analysis	41
3.2.5	Metagenomic analysis	41
3.2.5.1	Libraries preparation and sequencing	41
3.2.5.2	Reads assembly and contigs binning	41
3.2.5.3	Taxonomic and metabolic annotation of contigs	42
3.2.5.4	Clusters contribution to microbial community	42
3.2.5.5	Sequences deposition	43
3.3	Results	43
3.3.1	Redox potential assessment of porewater samples	43
3.3.2	Chemical characterization of samples	44
3.3.3	Microbial communities	44
3.3.3.1	Aerobic microbial communities	44
3.3.3.2	Anaerobic microbial communities	44
3.3.3.3	Alpha-diversity	46
3.3.3.4	Sample clustering	46
3.3.4	Metagenomic binning	47
3.3.4.1	Manual correction of binning	47
3.3.4.2	Draft genomes	49
3.3.4.1	ANI analysis	49
3.3.4.2	Annotation of dominating sulfate reducing bacteria	50

3.4	Discussion	52
3.4.1	Issue with line sampling	52
3.4.2	Anaerobic microbial communities	52
3.4.3	Metagenomic binning of BIC-A1 borehole	53
3.4.3.1	<i>Pseudomonas</i> c5	54
3.4.3.1	Peptococcaceae c7	55
3.4.3.2	Metabolic interactions	58
3.4.4	Outcome for the safety of Opalinus Clay repositories	60
3.5	Conclusion	60
3.6	Acknowledgements	61
CHAPTER 4	HYDROGEN CONSUMPTION BY MICROORGANISMS IN OPALINUS CLAY	63
Abstract		64
4.1	Introduction	65
4.2	Materials and Methods	65
4.2.1	Bioreactor	66
4.2.2	Chemical sampling and assays	69
4.2.2.1	Ion chromatography	69
4.2.2.2	Sulfide determination	69
4.2.2.3	Fe(II) determination	69
4.2.2.4	Alkalinity and pH	70
4.2.2.5	Dissolved gas determination	70
4.2.3	Planktonic cell density	70
4.2.4	DNA sampling and extraction	70
4.2.5	16S rRNA genes sequencing	71
4.2.6	Metagenomic analysis	71
4.2.6.1	Library preparation and sequencing	71
4.2.6.2	Contig binning	72
4.2.6.3	Taxonomic annotation of clusters	72
4.2.6.4	Pathway annotation of clusters	73
4.2.6.5	Clusters contribution to microbial community	73
4.2.6.6	Data deposition	73
4.2.7	Metaproteomics	73
4.3	Results	75
4.3.1	Chemical changes	75
4.3.2	Biological succession	76
4.3.1	Metagenomic analysis	77
4.3.1.1	Manual binning correction	77
4.3.1.2	Cluster quality assessment	78
4.3.1.3	Taxonomic annotation of clusters	78
4.3.1.4	Metabolic annotation of clusters	79
4.3.1.5	ANI analysis of good clusters	79
4.3.2	Metaproteomics analysis	80
4.4	Discussion	80
4.4.1	Microbial growth	80
4.4.2	Microbial community succession	81
4.4.3	Metabolic capabilities of draft genomes	83
4.4.4	Metabolic activities of draft genomes	85
4.4.4.1	Carbon cycling	87
4.4.1	Hydrogen concentration range for <i>Desulfobulbaceae</i> c16a	88
4.4.2	Putative Opalinus Clay indigenous bacteria	88
4.5	Conclusion	88

4.6	Acknowledgments	89
<hr/>		
CHAPTER 5	BIOLOGICAL HYDROGEN AND SULFATE CONSUMPTION RATE	91
<hr/>		
	Abstract	92
5.1	Introduction	93
5.2	Materials and Methods	94
5.2.1	Experimental set-up	94
5.2.1	Experimental procedures	95
5.2.2	Chemical analyses	97
5.2.3	Rate calculation	98
5.3	Results	99
5.3.1	Long-term change in anions concentration	99
5.3.2	Repeated sampling	99
5.3.3	Repeated injection	100
5.4	Discussion	100
5.4.1	Long-term change in sulfate concentration	100
5.4.2	Repeated sampling	102
5.4.3	Repeated injection	103
5.4.4	Rates of hydrogen and sulfate consumption	106
5.5	Conclusion	106
5.6	Acknowledgments	107
<hr/>		
CHAPTER 6	CONCLUSION AND OUTLOOK	109
<hr/>		
6.1	Microbial activity in Opalinus Clay pristine rock	111
6.1.1	State of the art	111
6.1.1	Attempts to extract DNA from pristine Opalinus Clay rock	111
6.1.1	Opalinus core microbial community	111
6.1.1	Future strategies	113
6.2	Microbial activity in Opalinus Clay porewater	113
6.2.1	Line sampling limitations	115
6.2.2	Microbial system in Opalinus Clay open space	115
6.2.3	Putative sources of energy	116
6.2.4	Future strategies	116
6.3	Microbial activity in Opalinus Clay porewater when an additional energy source is provided	117
6.3.1	Hydrogen as an energy source	117
6.3.2	Carbon cycling fuelled by hydrogen	117
6.3.3	Hydrogen consumption rate	118
6.3.4	Future strategies	118
6.4	Impact of microbial activity on nuclear waste repository	119
6.4.1	Radionuclide mobility	119
6.4.2	Gas production and consumption	119
6.4.3	Barrier alteration	120
<hr/>		
ANNEX 1	SUPPLEMENTARY INFORMATION TO CHAPTER 3	121
<hr/>		
A-1.1	Supplementary Materials and Methods	123
A-1.1.1	DNA extraction protocol using DNA Spin kit for Soil	123
A-1.1.2	JGI Illumina 16S rDNA amplicons (Itags) analysis pipeline	124

A-1.1.3	QIIME pipeline used for alpha-diversity analysis	126
A-1.1.4	DNA ethanol precipitation	127
A-1.1.5	Sequences deposition	127
A-1.2	Supplementary Tables	129
A-1.3	Supplementary Figures	140
ANNEX 2	SUPPLEMENTARY INFORMATION TO CHAPTER 4	142
A-2.1	Supplementary Materials and Methods	143
A-2.1.1	Calculation of dissolved gas concentration	143
A-2.1.2	Phenol-chloroform DNA extraction method	145
A-2.1.3	Sequences deposition	146
A-2.2	Supplementary Tables	148
A-2.3	Supplementary Figure	157
REFERENCES		159
CURRICULUM VITAE		169

List of Figures

Figure 1.1	The concept of biosphere, that includes all the places in Earth where life is possible	3
Figure 1.2	Local, intermediate and regional flow systems in the lithosphere.	4
Figure 1.3	Division rate of a barotolerant and two barophile microbes along a pressure gradient.....	9
Figure 1.4	An example of a sedimentary rock	10
Figure 1.5	An example of a metamorphic rock.....	10
Figure 1.6	Depth distribution of bacteria in sediments containing sapropels from the Mediterranean Sea.....	12
Figure 1.7	Oxidation of acetate by successive electron acceptors.....	15
Figure 1.8	Overall microbial metabolisms degrading organic matter, through the successive redox regime.....	16
Figure 1.9	Anaerobic pathway of organic matter mineralization	17
Figure 1.10	Energy production and carbon fixation in a chemolithoautotrophic sulfate reducing bacteria	19
Figure 1.11	Putative carbon loop in deep subsurface, based on the presence of H ₂ and SO ₄ ²⁻	19
Figure 2.1	The four different safety barriers surrounding high-level nuclear wastes	23
Figure 2.2	Drums containing solidified low- and intermediate-level.....	24
Figure 2.3	Situation of the Mont Terri underground rock laboratory in St-Ursanne	25
Figure 2.4.	Schematic of the geological history of the Opalinus Clay formation at Mont Terri	26
Figure 2.5	Excavation disturbed zone (EDZ) created by excavation works.....	33
Figure 3.1	Situation of the Underground Rock Laboratory of Mont Terri and borehole emplacements	39
Figure 3.2	Microbial community composition based on 16S rRNA gene sequencing	46
Figure 3.3	Diagram explaining how contigs were clustered	48
Figure 3.4	Putative energetic metabolism of <i>Pseudomonas</i> cluster c5 and Peptococcaceae cluster c7.....	56
Figure 3.5	Propionate fermentation pathway, as described by Stams <i>et al.</i> , (1984), and as depicted in this study.	58
Figure 3.6	Putative food webs between <i>Pseudomonas</i> c5 and Peptococcaceae c7	59
Figure 4.1	BRC-3 borehole equipment	66
Figure 4.2	Experimental set-up in recirculation mode with continuous H ₂ injection and in non-recirculation mode .	67
Figure 4.3	Chemical change in dissolved O ₂ , Fe(II) and S(-II) in borehole water	75
Figure 4.4	H ₂ consumption in the borehole over time.....	76
Figure 4.5	Bicarbonate concentration over time in borehole water.....	77
Figure 4.6	Change in pH and alkalinity of BRC-3 borehole water	78
Figure 4.7	Organic acid concentration over time in borehole water	79
Figure 4.8	Planktonic cell density over time in borehole water	81
Figure 4.9	Microbial community succession in the borehole water, after H ₂ injection.....	83

Figure 4.10	Metabolic pathways of <i>Desulfobulbaceae</i> c16a, according to proteomic and genomic data.....	86
Figure 4.11	Metabolic interactions between seven microorganisms, based on proteomic data	87
Figure 5.1	Borehole equipment of BRC-3	94
Figure 5.2	Timeline of the H ₂ injection experiment in BRC-3 borehole.....	95
Figure 5.3	Repetitive sampling experiment.	97
Figure 5.4	Change in borehole water concentration of sulfate, chloride and bromide over time	101
Figure 5.5	Sulfate, hydrogen and sulfide concentration from repeated sampling carried out at day 378	103
Figure 5.6	Chemical changes in borehole during the repetitive injection experiment.....	104
Figure 5.7	Change in the sulfide, sulfate and $\delta^{34}\text{S-SO}_4$ in an <i>in vitro</i> batch incubation of borehole water.	105
Figure A-1.1	Microbial community composition of all samples, based on the region V4 of the 16S rRNA gene.....	140
Figure A-1.2	Rarefaction curve of the number of observed species, based on the 16S rRNA gene.	141
Figure A-1.3	Sample clustering based on their microbial community	141
Figure A-2.1	Microbial community composition over time, based on the taxonomic affiliation of genomes.....	157

List of Tables

Table 1.1	Redox potential of some relevant half reactions	14
Table 2.1	Composition of the three main Opalinus Clay facies observed at the Mont Terri	27
Table 2.2	Chemical composition of the Opalinus porewater in comparison to seawater	28
Table 3.1	Summary of relevant characteristics for each sample	45
Table 3.2	OTUs shared by the four samples harboring an anaerobic microbial community	47
Table 3.3	Taxonomic affiliation of the selected clusters	49
Table 3.4	Summary of the pathway annotation of the selected clusters.....	50
Table 3.5	Genome comparison of Peptococcaceae c7 with two reference genomes.....	51
Table 4.1	Metabolic abilities of 15 selected genomes	84
Table 5.1	Major anions and cations composition of APWs and BRC-3 water samples.....	96
Table 5.2	Summary of all H ₂ oxidation and sulfate reduction rates calculated in this work.....	106
Table 6.1	OTUs from the 16S rRNA gene analysis that correspond to draft genomes	112
Table 6.2	OTUs similar between this project and other sequencing projects that took place in Mont Terri URL ...	114
Table A-1.1	ICP-MS and pH measurements of borehole water samples.....	129
Table A-1.2	IC measurements of borehole water samples.....	129
Table A-1.3	Number of observation of all 825 OTUs in the 23 samples	129
Table A-1.4	Genus-level microbial community composition based on 16S rRNA gene	130
Table A-1.5	Number of sequences, total number of species and number of observed species	131
Table A-1.6	Contig clustering, as outputted by CONCOCT and after a manual correction	131
Table A-1.7	Taxonomic affiliation of the metagenomic clusters	132
Table A-1.8	Single-copy gene plot of each metagenomic clusters	133
Table A-1.9	Results of average nucleotide identity between draft genomes of this study and known genome	134
Table A-1.10	Pathway annotation of the metagenomic clusters	135
Table A-1.11	Number of hydrogenases observed in each cluster	135
Table A-1.12	Number of genes degradation complex organic matter.....	136
Table A-1.13	Gene list from <i>Pseudomonas</i> c5 that might be involved in processes described in Figure 3.4.A	137
Table A-1.14	Gene list from Peptococcaceae c7 that might be involved in processes described in Figure 3.4.B.....	138
Table A-2.1	Number of OTUs observation for every sample	148

Table A-2.2	Clustering table indicating the raw CONCOCT output and the output that was manually corrected.	148
Table A-2.3	Single-copy gene analysis of clusters created by CONCOCT that were manually corrected.	149
Table A-2.4	Taxonomic affiliation of all clusters, based on 16S rRNA gene sequence.....	149
Table A-2.5	Analysis of 16S rRNA genes recovered from metagenomic binning	152
Table A-2.6	Pathway and functional annotation of the best 31 clusters.....	153
Table A-2.7	ANI analysis of the 15 selected draft genomes.	154
Table A-2.8	Protein profile of sample recovered at day 483.....	154
Table A-2.9	Protein profile of the first run of the sample recovered at day 483	154
Table A-2.10	Protein profile of the second run of sample recovered at day 483	154
Table A-2.11	List of proteins belonging to Desulfobulbaceae c16a	155
Table A-2.12	List of proteins identified in metaproteomic data that are mentioned in the main text	157

Chapter 1 Introduction to deep subsurface microbiology

This first chapter consists of a general introduction to deep subsurface microbiology. It does not include information that is directly related to the research project, but introduces key concepts that will be useful later for interpreting the results. If you are not familiar with environmental microbiology, or if you want to know more about deep subsurface microbiology, you may consider reading this chapter.

1.1 Definition

Eduard Suess, an Austrian geologist, introduced in 1888 the notion of ‘biosphere’, which is a very important ecological concept (Suess, 1888). In a broad sense, the biosphere is composed of all the places on Earth where life is possible, and is composed of the atmosphere, the hydrosphere and the lithosphere (Figure 1.1). The lithosphere, from a geological sense, is defined as the most superficial rock layer of Earth, which is solid. The lower limit of the lithosphere is defined by a maximal temperature of 1,300 °C, which is the melting point of olivine. Its depth varies between 200 km under old continents to 20 km under the Mid-ocean ridge (Skinner *et al.*, 2004) but it is clear that the lowest layers of the lithosphere do not belong to the biosphere. The lower boundary of the biosphere can be defined by a temperature threshold, which is 110°C, because no biological activity is possible above it this temperature (Furnes and Staudigel, 1999).

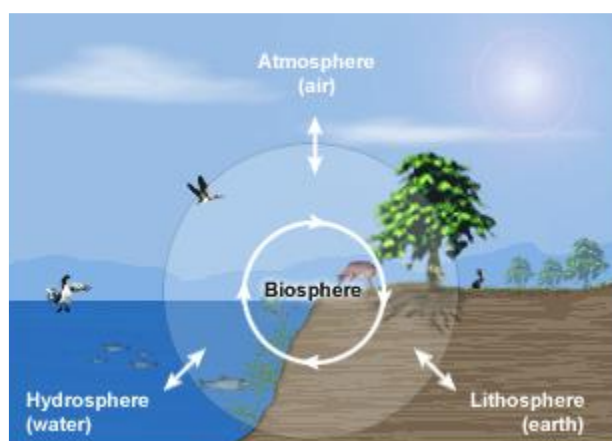


Figure 1.1 The concept of biosphere, that includes all the places in Earth where life is possible, thus overlapping with the hydrosphere, atmosphere and lithosphere. Modified from <https://www.ucar.edu/learn/1.htm>.

Conditions are harsher for living organisms in the subsurface because of the lack of sunlight and oxygen, and the small pore size, among other constraints. These parameters make prokaryotic organisms perfect for colonizing these environments, because of their diverse metabolic abilities, their small size, and their ability to survive under extreme conditions (Rothschild and Mancinelli, 2001).

The definition of the deep subsurface is not really clear, because the term ‘deep’ refers to a subjective notion. Here are different definitions of the ‘deep subsurface’ that can be found in the literature:

- Deep aquifer systems as those that are hundreds to thousands of meters below land surface (Ghiorse and Wilson, 1988) ;
- Sediments buried to depths greater than about 10 m (Fredrickson *et al.*, 1989) ;

- Zone beneath the seafloor and soil zones (Onstott *et al.*, 2009).

Because the great majority of deep subsurface microbiology studies deals with water-saturated environment, ‘deep subsurface’ can be defined by hydrological parameters. Lovley and Chapelle (Lovley and Chapelle, 1995) restricted this notion to the zones having intermediate or regional flow systems. In contrast to local flow systems, these systems are not firmly connected to the surface, leading to low rates of recharge and ground water flow, which means that they are not responding rapidly to individual precipitation events (Figure 1.2).

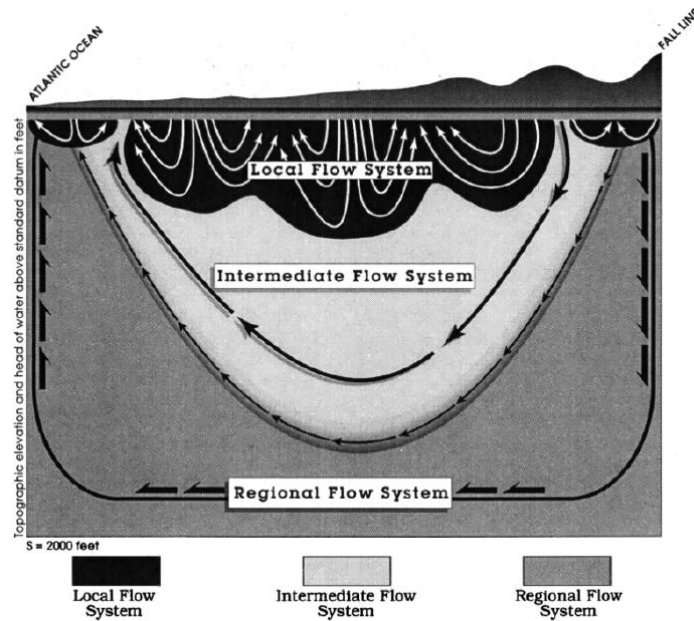


Figure 1.2 Local, intermediate and regional flow systems in the lithosphere. Deep subsurface refers to area where intermediate or regional flows occur. From (Lovley and Chapelle, 1995).

1.2 Interest in subsurface microbiology

Despite the physical discontinuity between our living world, the surface, and the deep subsurface, it is important to get a better understanding of these environments, mainly because:

1. The number of microorganisms living in the subsurface largely exceeds the number of those found in other ecosystems, even if the microbial density is generally low (10^3 to 10^8 cells / g or ml). But considering the viable volume of ocean floor, continental shelf and rock aquifer, the total biomass of these environments is potentially equal to the total plant biomass (Whitman *et al.*, 1998). A great part of living carbon biomass consists of microorganisms beneath the seafloor and soil zones.

2. Microbial activity in the deep subsurface is important, as it controls the organic and inorganic geochemistry of groundwater (Lovley and Chapelle, 1995).
3. Subsurface environments, as a biological *terra incognita*, are a reservoir of unknown microbial species and metabolisms.

There are of course also practical interests in deep subsurface microbiology, which have provided most of the opportunities for making progress in this research field. This can be attributed to the high cost of accessing the deep subsurface biosphere. Here are some industrial interests in deep subsurface microbiology:

1. Microbial activity can affect, positively or not, oil extraction.
2. Microorganisms can be useful in restoring contaminated groundwater from surface and underground disposal sites.
3. The storage of radioactive waste and heavy metals in deep geological formations requires a good knowledge of these environments, including the local microbial communities and their potential effect on waste storage.
4. There is large pool of methane gas hydrates, in sub-ocean floor sediments, which is expected to have a microbial origin and which could be an interesting energy resource.

1.3 Limitations of deep subsurface microbiology

The microbiological study of deep subsurface environments is a relatively young field of investigation, mainly due to the inaccessibility of such locations. The most widespread technique used to reach these environments is drilling and coring. The nature of these sampling methods makes it very difficult to maintain sterility during the procedure, even if many precautions are taken. This problem is amplified by the scarcity of microorganisms in such locations, making it challenging to determine whether a microbial population is indigenous (Fredrickson and Fletcher, 2001). Caves, tunnels and mines can also be useful to access the deep subsurface. However, those also often require drilling to sample undisturbed rock, which must be sufficiently far from the oxic atmosphere of the gallery (Russel, 1997). An additional challenge is that samples are difficult to handle by standard microbiological methods due to the hardness of the rock. The sampling of porewater is another possibility, but studies have shown that the planktonic microbial community differs from the sessile community (Lehman *et al.*, 2001).

Measures beyond traditional sterility precautions can be taken to assess the importance of contamination. The use of chemical tracers, mixed in with the drilling liquid, can help evaluate the penetration of

contaminating bacteria into the core, whose center is generally exempt of contamination (Amend and Teske, 2005). The results obtained from a rock sample can also be compared to a similar sample that was intentionally contaminated during the sampling procedure. Using this approach, Pedersen *et al.* (Pedersen *et al.*, 1997) showed that contaminating microbes are introduced during drilling but that the introduced biota do not establish themselves at detectable levels in aquifers due to the extreme conditions. Finally, the 16S SSU rRNA sequence of microorganisms can provide a good indication of their origin. Some microbial species are known contaminants, while sequences deviating from known sequences can indicate an undiscovered indigenous microorganism, even if such a result is difficult to establish beyond doubt (Santelli *et al.*, 2010).

1.4 General characteristics of terrestrial deep-subsurface environments

In this section are described the main environmental characteristics of deep subsurface zones that have a significant influence on bacterial growth and make such environments extreme for life.

1.4.1 Pore size and interconnectivity

Pore size and interconnectivity greatly influence microbial activity, by controlling the space available for life, as well as the mobility of microbes and water and nutrients. Considering that an average deep subsurface microbe size ranges between several tenths of microns and 10 μm , it is not surprising to find no biological activities in sediments with pore throats under 0.2 μm . Some studies showed that with increasing pressure, the macro-pores in clay materials do not automatically disappear, but, instead, their interconnectivity is reduced (Keller *et al.*, 2015), which results in decreased microbial activity (Pedersen, 2010). The results of Fredrickson *et al.* (Fredrickson *et al.*, 1997) suggest that subsurface microorganisms require interconnected pore throats larger than 0.2 μm in diameter for significant activity. Nonetheless, under the right conditions, viable microbes can be found and stimulated in poorly permeable rock. During clay rock diagenesis, because of the increased pressure due to sediment burial, porosity decreases to a level that inhibits microbial activity, resulting in a transition between biotic to abiotic diagenesis, as shown for Callovian–Oxfordian clayey formation of Bure (France; Lerouge *et al.*, 2011).

1.4.2 Absence of sunlight

There is of course no sunlight, and thus no photosynthesis, in deep subsurface environments. Such environments are considered aphotic, meaning that photosynthesis-dependent organic matter production does not occur. Therefore, four main sources provide low redox potential electrons for energetic metabolisms:

- The recharging groundwater carrying organic matter;
- The organic matter buried with the sediments;
- The rock itself, which may contain reduced compounds;
- The flow of reduced gases from the mantle that contains H₂ and CH₄ (Fredrickson and Balkwill, 2006);
- The radiolysis of water that is known to release H₂ (Barr and Allen, 1959).

The two first cases are discussed more in detail in section 1.7.1, and the three last cases, which are considered as independent from sunlight, are detailed in section 1.7.2.

1.4.3 Anoxicity

Due to the aphotic character of deep subsurface environments, oxygen production does not occur. Therefore, oxygen is mainly a result of advection by water. But, oxygen can be rapidly consumed by microorganisms if an appropriate electron donor (e.g., organic matter) is present, leading to the almost systematic anoxic character of deep subsurface systems. Only a few deep aquifers with no (or very limited) microbial activity and/or no organic carbon are still oxidic (Lovley and Chapelle, 1995). Thus, most of the organisms found in the deep subsurface are facultative or obligate anaerobes.

1.4.4 Water availability

In the deep subsurface settings, water is common, with a very high mineral surface area in comparison to the water volume. However, as there is relatively little space for water and for life only microorganisms can survive because pore size is too small for others organisms (Lovley and Chapelle, 1995) and most of the water movements are due to passive flows. Microorganisms are often in a dormant metabolic state due to limited nutrient availability. The water activity (a_w), a value ranging from 0 to 1 and expressing (among others) the accessibility of water to microorganisms in a given medium, can be a limiting factor for microbial growth. High solute concentrations and small porosities decrease this value. Most microorganisms can thrive at a_w values around 0.98, which is the value for seawater. Under 0.95, growth is not possible for most Gram-negative bacteria, and at a_w smaller than 0.65, life is not possible because DNA becomes permanently disordered (Brown, 1990). Some microorganisms have strategies to handle low water activities, such as survival in a non-vegetative state (e.g., endospores) but most cannot adapt to those conditions.

1.4.5 Temperature

Temperature and pressure increase with depth. Of the two, temperature is a more critical factor for life and seems to limit its development with depth. When temperature rises, chemical and enzymatic reactions occur more rapidly. However, above a given threshold, protein, nucleic acids and other components of the cell may be irreversibly damaged. Thermophilic and hyperthermophilic microorganisms have physiological adaptations to heat that include:

- Proteins and enzymes that are stable at high temperature, due to a few critical amino acids substitutions, having a large number of salt bridges that decreases the protein flexibility and a denser structure associated with a stronger internal hydrophobicity;
- Others cell components are adapted to high temperature, such as ribosomes and the cytoplasmic membrane, which is rich in saturated fatty acids for bacteria.

The upper temperature limit for life is thought to be approximately 113°C. This limit, in terms of depth, is highly variable. In hydrothermal vents, it reaches the ocean floor, but in shield rock, mountains and deep sediments, it can be deeper than 10'000 m (Jannasch and Taylor, 1984).

1.4.6 Pressure

Roughly, microbial activity decreases linearly with increasing pressure (Jannasch and Taylor, 1984). Three types of microorganisms have been studied as a function of their relation with pressure:

- The barotolerant microbes, which represents the most widespread type, includes all the organisms that grow better at atmospheric pressure but can withstand highest pressures. Most of those are insensitive to pressures in the range of 10 till 60 MPa (Figure 1.3).
- Barophilic microbes show faster growth rates at elevated hydrostatic pressures but can also thrive at atmospheric pressure (Figure 1.3).
- Finally, obligate barophile organisms thrive at higher pressure and cannot withstand any decompression to an atmospheric level.

Only a few proteins syntheses are controlled or induced by pressure in barophile organisms. The membrane, the cell wall and related structural proteins are the major components influenced by pressure. For example, more unsaturated fatty acids in the cytoplasmic membrane allow it to maintain its integrity at higher pressure. Moreover, enzymes must have higher affinity for their substrate, because a higher pressure tends to decrease this affinity (Jannasch and Taylor, 1984).

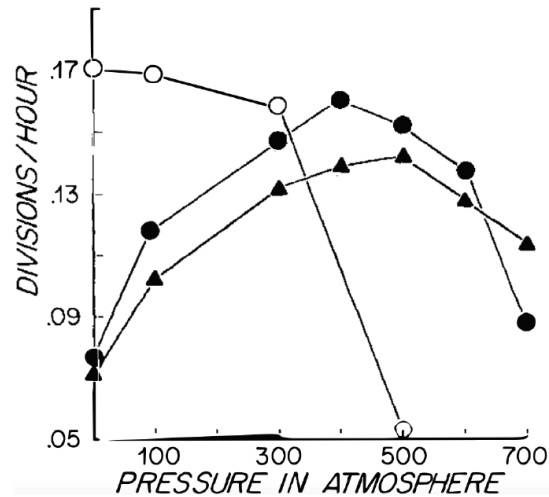


Figure 1.3 Division rate of a barotolerant (open circles) and two barophile (dark symbols) microbes along a pressure gradient. This plot clearly highlights the wide pressure range supported by microorganisms. From (Jannasch and Taylor, 1984).

1.5 Two major types of rock substrate

In deep subsurface microbiology, many different types of rock can be encountered. Here are presented the two end-members of this series, which are sedimentary rocks and metamorphic and igneous rocks.

1.5.1 Sedimentary rocks

The porosity of sedimentary rocks is mainly due to pores, which constitutes habitat for microorganisms (Figure 1.4). The water flow occurs via pore interconnections. So, pore size and pore interconnections are important parameters for the distribution of microorganisms and their metabolic activity. Typically, sedimentary rocks were not exposed to high temperatures and pressures during their formation, potentially allowing the persistence of buried microorganisms and organic carbon during sedimentation. Due to the low permeability of such environments, it is likely that indigenous microorganisms are as old as the rock itself (Fredrickson and Balkwill, 2006).



Figure 1.4 An example of a sedimentary rock: the Logan Formation (Mississippian) of Jackson County, Ohio, USA. From http://en.wikipedia.org/wiki/Sedimentary_rocks.

1.5.2 Igneous and metamorphic rocks

Igneous and metamorphic rocks usually do not contain organic matter, in contrast to sedimentary rocks, because they were subject to high temperature and/or pressure during their formation (Figure 1.5). Because of the absence of pores in crystalline rocks, fluid flows occur mainly via fractures, which is also the way microbial colonization proceeds. Metabolic substrates are also brought by fluid flow, via fractures (Fredrickson and Balkwill, 2006).



Figure 1.5 An example of a metamorphic rock: Precambrian metamorphic rock wall in Lamar Canyon, Yellowstone National Park, USA. From <http://www.nps.gov/features/yell/slidefile/geology/metamorphic/Page.htm>

1.6 Microbial communities in the deep subsurface

The goal of microbial ecology is typically to address questions about microorganisms found in a given environment: how many are there, who are they, are they active, and what type of activity are they carrying out? Herein are reviewed the more common findings for deep subsurface environments.

1.6.1 Number of individual cells

In deep subsurface environments, due to the general limited availability of space and the dearth of nutrients and water, the density of microbial cells is low (from 10^3 to 10^8 cells/g or ml). Generally, the bacterial population decreases with depth, but this distribution strongly depends on the local geochemical conditions that can increase the number of cells at specific depths (Figure 1.6; Parker & Wellsbury, 2004). In many consolidated rocks, microorganisms are undetectable. This is not only a question of detection limit, in some case microorganisms are simply absent.

1.6.1 Diversity of microbial populations

Diversity is expected to be high due to limited mixing in the rock substrate. Thus, microorganisms only a few meters away from each other can theoretically evolve completely independently. Moreover, the relative stability of the deep subsurface over time preserves microorganisms trapped in the sediment during its formation.

1.6.2 Metabolic state of microorganisms

The concentration of substrates for metabolism are often limiting, because they are renewed at a very slow rate due to very slow fluid movement. Thus, microbial metabolic rates are often very slow and microorganisms are frequently in a dormant state. This adaptation to extreme environments makes it difficult for them to thrive on classical media, which are too rich. Non-cultivability of deep subsurface inocula has been frequently observed, but it does not necessarily imply an absence of viable cells (Fredrickson *et al.*, 1997; Onstott *et al.*, 1998). For some aquifers between 200 and 400 m beneath the surface, the organic matter turnover was estimated to be in the range of decades or centuries (Onstott *et al.*, 2009). Parkes *et al.* (2000) have estimated a cell division time of more than 100'000 years for microorganisms in sapropel layers (rich in organic matter) of Mediterranean Sea sediments. Because only 8.3% of the bacteria were dividing, it meant that the population doubling time was 0.92 million years (myr). This rate was calculated by plotting the excess number of bacteria in these sapropel layers (in comparison to the surrounding inorganic layers) as a function of the sediment age (Figure 1.6). These findings also suggest that 4.7 myr-old organic matter is a substrate for subsurface microorganisms (Cragg *et al.*, 1998).

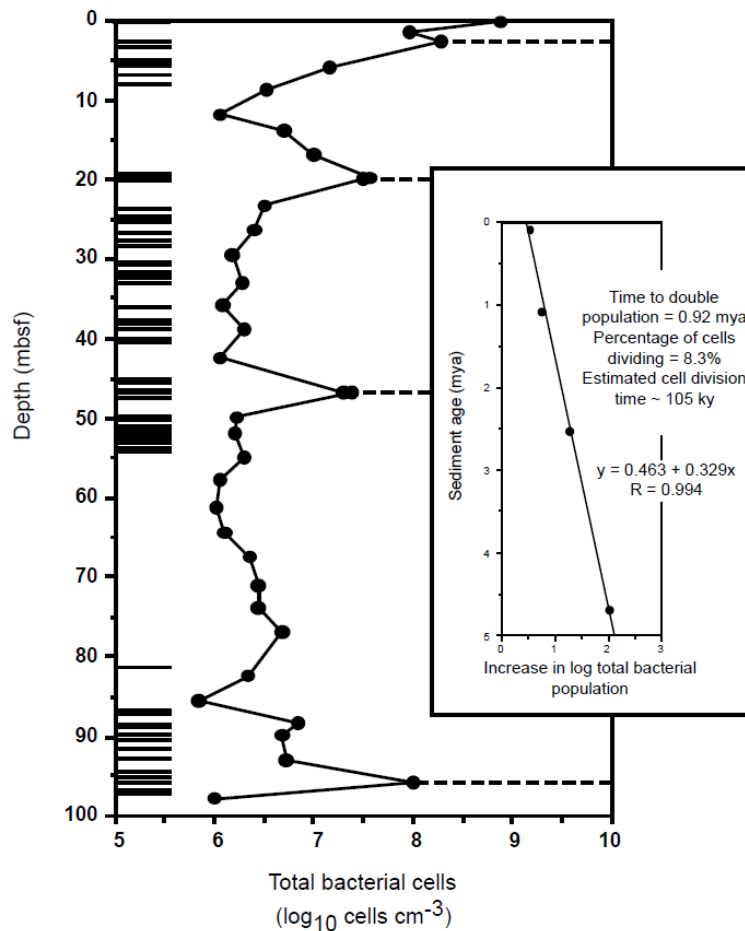


Figure 1.6 Depth distribution of bacteria in sediments containing sapropels from the Mediterranean Sea. Sapropel layers with high organic matter are shown as horizontal dark bands on left-hand axis. Significant increases in bacterial population numbers occurred in all four sapropels analyzed. Relative population increases within these sapropels are plotted against sediment age (inset), to provide an estimate of bacterial division time (Parkes *et al.*, 2000).

1.6.3 Origin of microorganisms

The question of the origin of microorganisms identified in the rock, which is commonly not addressed in the study of other environments, has great importance in the deep subsurface. First, as explained earlier, the low densities of *in situ* microbial cells and the nature of the sampling methods raise the question of whether the microbes found in the sample are indigenous or originate from contamination (for more details see section 1.3). Second, once the indigenous origin of a microorganism is established, the question remains of whether it originated from the vertical flow of groundwater or whether it was present in the rock at the time of its formation, making the observed microbial community as old as the rock. The hydrological conditions of the subsurface environment, in particular vertical water flow, can provide a good indication of the origin of *in situ* microbiota.

Vreeland *et al.* (2000) isolated a halotolerant bacterium from a 250 myr-old crystal salt. While the formation of endospores is an efficient way to survive unfavorable conditions, vegetative cells may also survive in less extreme environments where nutrients are limiting, such as deep subsurface locations. Indeed, the most commonly cultured and observed microbial cell forms in deep subsurface samples are vegetative cells. Because of the harsh subsurface conditions and the possible old age of microorganisms highlighted in many deep subsurface environments, it was necessary to devise a long-term starvation state of vegetative cells, which has been called ‘dormancy’, even if the exact nature of this metabolic state is unclear (Amy, 1997).

However, the survival of a microorganism in geological formation beyond 3 myr is currently being disputed in the literature (Willerslev, Hansen & Poinar, 2004; Willerslev, Hansen, Rønn, *et al.*, 2004). A microbe must be able to repair damage to its DNA and RNA, which can be due to spontaneous hydrolysis, oxidation and alkylation (Willerslev, Hansen, Rønn, *et al.*, 2004). Thus, dormant and spore-forming microorganisms, because of their inactivity, are doomed to die due to the inevitable degradation of their genetic material. This process could occur at geological time scales, depending on environmental conditions (Pedersen, 2000). Willerslev *et al.* (Willerslev, Hansen & Poinar, 2004; Willerslev, Hansen, Rønn, *et al.*, 2004) put forward a survival limit of 0.4 to 0.6 myr and showed that Gram-positive, non-spore forming bacteria survive the longest, surpassing spore-forming and Gram-negative bacteria. However, others authors have raised the possibility of a low-level cellular metabolic activity, increasing the survival time up to 1.5 myr for Johnson *et al.*, (2007) and up to 3 myr for Suzina *et al.* (2004). Schippers & Neretin (2006) proposed an inhibition of the DNA degradation, by sorption in minerals, preserving it beyond 1 myr after cell death. If this is true, this could cause false positive results when sequencing DNA extracted from rock.

1.7 Deep subsurface bioenergetics

In deep subsurface environments, energetic metabolisms used by microorganisms are likely to be identical to those found in shallow formations (Pedersen, 1993). Thus, soils and aquatic sediments represent a good model for the types of microbial metabolisms that might be found in the deep subsurface. Two types of energy sources, or low redox potential electron source, are expected to drive subsurface ecosystems: organic matter, linked to photosynthesis and ‘geogas’, which is composed of inorganic reduced compounds (H_2 , CH_4 , H_2S) that have a geogenic and thus abiotic origin.

1.7.1 Organic matter a source of energy

1.7.1.1 Degradation of organic matter in anoxic systems

Even if deep subsurface systems are aphotic, organic matter oxidation remains the primary way microorganisms gain energy. Different types of metabolisms, in function of the terminal electron acceptor, result in the oxidation of organic matter: O₂, NO₃⁻, Mn(IV), Fe(III), SO₄²⁻ and CO₂. These oxidized compounds are classified by their tendency to accept electron, which is expressed by the reduction potential E_0' at pH 7. The greater the E_0' difference between two redox couples, the more exergonic the reaction will be, thus providing more energy (Table 1.1).

Table 1.1 Redox potential (Eh; calculated for a temperature of 25°C and for pH = 7) of some relevant half reactions. Modified from Stumm & Morgan (1996).

Redox potentials of selected half-reactions at 25° C and pH 7	
Reaction	Eh (V)
$O_2 + 4H^+ + 4e^- = 2H_2O$	0.81
$NO_3^- + 6H^+ + 6e^- = \frac{1}{2}N_2 + 3H_2O$	0.75
$NO_3^- + 2H^+ + e^- = NO_2^- + H_2O$	0.42
$NO_3^- + 10H^+ + 8e^- = NH_4^+ + 3H_2O$	0.36
$Fe^{3+} + e^- = Fe^{2+}$	0.36
$NO_2^- + 8H^+ + 6e^- = NH_4^+ + 2H_2O$	0.34
$CH_3OH + 2H^+ + 2e^- = CH_4 + H_2O$	0.17
$CH_2O + 2H^+ + 2e^- = CH_3OH$	-0.18
$SO_4^{2-} + 8H^+ + 6e^- = S + 4H_2O$	-0.20
$SO_4^{2-} + 10H^+ + 8e^- = H_2S + 4H_2O$	-0.21
$CO_2 + 8H^+ + 8e^- = CH_4 + 2H_2O$	-0.24
$N_2 + 8H^+ + 6e^- = 2NH_4^+$	-0.28
$H^+ + e^- = \frac{1}{2}H_2$	-0.41
$CO_2 + 4H^+ + 4e^- = \frac{1}{6}C_6H_{12}O_6 + H_2O$	-0.43
$CO_2 + 4H^+ + 4e^- = CH_2O + H_2O$	-0.48
$Fe^{2+} + 2e^- = Fe$	-0.85

This rule helps us to understand why terminal electron acceptors are used sequentially in closed systems: the most exergonic reactions is favored until the most reactive electron acceptor is depleted. Thus, respecting this sequence, the sequence of respiration metabolisms is: aerobic respiration (with O₂ reduction), denitrification and the dissimilatory nitrate reduction (with NO₃⁻ reduction), Mn(IV) reduction, Fe(III) reduction, sulfate reduction (with SO₄²⁻) and methanogenesis (with CO₂ reduction; Figure 1.7). The concentration of H₂, which reflects the redox potential of a given environment, can indicate which type of anaerobic respiration is currently occurring (Figure 1.7).

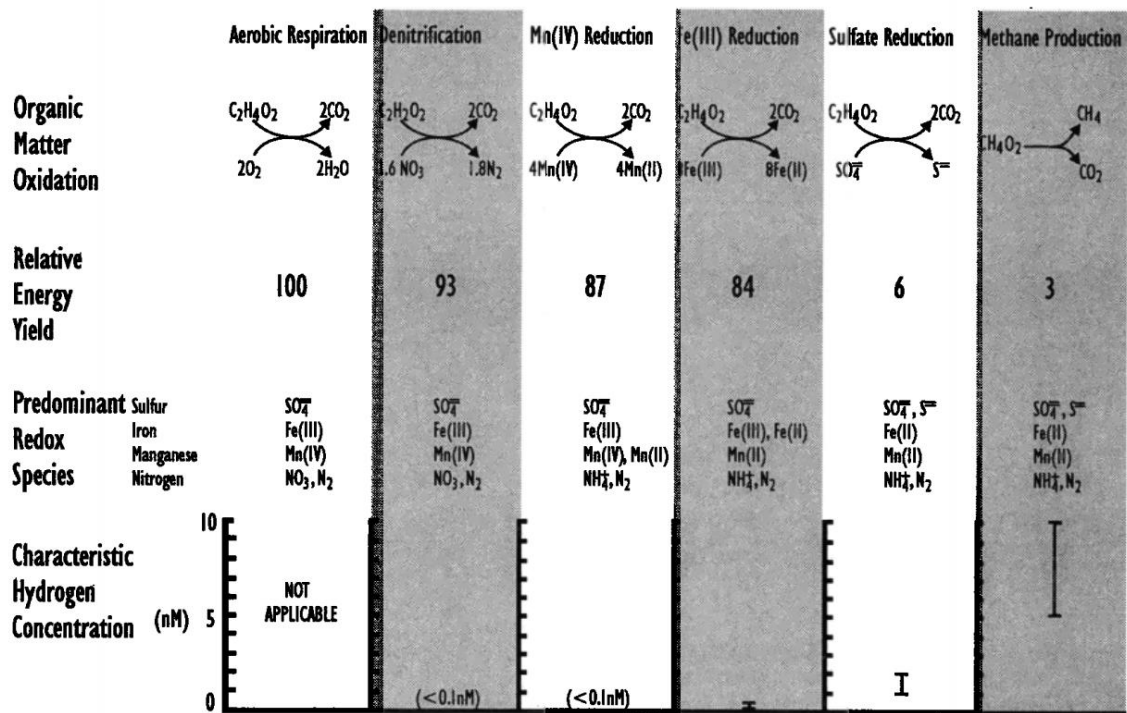


Figure 1.7 Oxidation of acetate by successive electron acceptors. Acetate has been chosen as model compound because it is the most organic compound used by all type of organic matter respiring metabolisms. From Lovley & Chapelle, 1995.

Microorganisms using oxygen and nitrate as terminal electron acceptors can completely degrade organic matter to CO_2 on their own. This is partly due to the secretion of extracellular enzymes hydrolyzing complex organic matter, liberating mainly sugars and amino acids but also aromatic compounds and long chain fatty acids. All of these compounds, directly derived from biomass, are defined as primary substrate (Lovley and Chapelle, 1995; Gobat *et al.*, 2004). But due to the anoxic character of deep subsurface, aerobic respiration and denitrification are expected to be a minor factor in such environments.

For anaerobic respiring microorganisms (reducing Mn(IV), Fe(III) or SO_4^{2-}), the ability to degrade organic matter is less sophisticated. Indeed, they can only oxidize a very small range of organic compounds, which are simpler from a chemical point of view. These compounds, generated by the fermentation of primary substrates, are defined as secondary substrates (Gobat *et al.*, 2004) and include organic acids (such as acetate, propionate, butyrate, lactate, or succinate), alcohols (such as ethanol (Leschine, 1995)). Even H_2 can be used as electron donor by Mn(IV), Fe(III) and SO_4^{2-} reducing bacteria (Kashefi and Lovley, 2000; Muyzer and Stams, 2008). Most of bacteria oxidizing hydrogen in that way are chemolithoautotrophic organisms, reducing CO_2 to produce their own organic matter (Lovley, 1991). Anaerobic respiring microorganisms are also able to directly oxidize some primary substrates, such as aromatics and long chain fatty acids (Lovley and Chapelle, 1995).

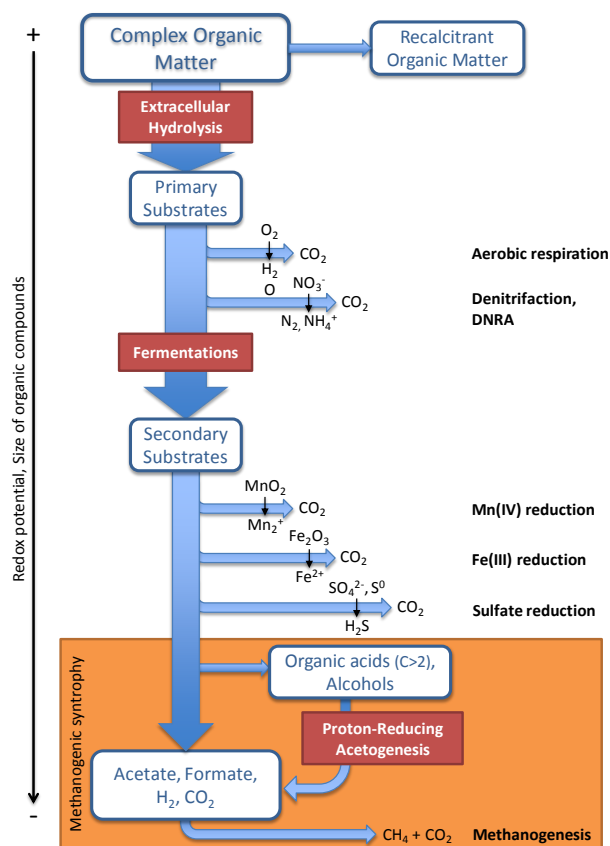


Figure 1.8 Overall microbial metabolisms degrading organic matter, through the successive redox regime. Modified from Gobat *et al.*, (2004).

When conditions become more reducing so that Mn(IV), Fe(III) and sulfate are depleted, some microorganisms can still degrade the rest of organic matter to CO_2 and CH_4 by a metabolism called methanogenesis. The methanogenic microbes all belong to the Archaea domain and can only degrade very small organic compounds such as formate and acetate. Organic acids larger than two carbon and alcohols have to be degraded to acetate by proton-reducing acetogenesis, a secondary fermentation process (by contrast to primary fermentation) that oxidizes organic compounds by reducing protons (H^+) to hydrogen. But because the energetic gain provided by this reaction is too low to sustain life (Table 1.1), it can only be carried out by microorganisms if the H_2 concentration is kept very low, which is done by methanogenic partners. This process is a syntrophy because it is a common energetic strategy between two metabolisms that are mutually dependent: the first one (the proton-reducing acetogen) is providing the carbon substrate for the second one (the methanogen) and the second one is providing the optimal chemical conditions making the metabolic reaction of the first one exergonic (Schink, 1997). The overall microbial reactions for degrading organic matter are presented in Figure 1.8

The situation is different in anoxic sulfate bearing environments. Because sulfate-reducing bacteria have a better affinity for hydrogen than methanogenic microorganisms, they will outcompete them. Moreover, because sulfate-reducing bacteria are more metabolically versatile than methanogenic microbes, the complete

oxidation of organic matter to carbon dioxide does not require secondary fermentation anymore (Schink, 1997; Figure 1.9).

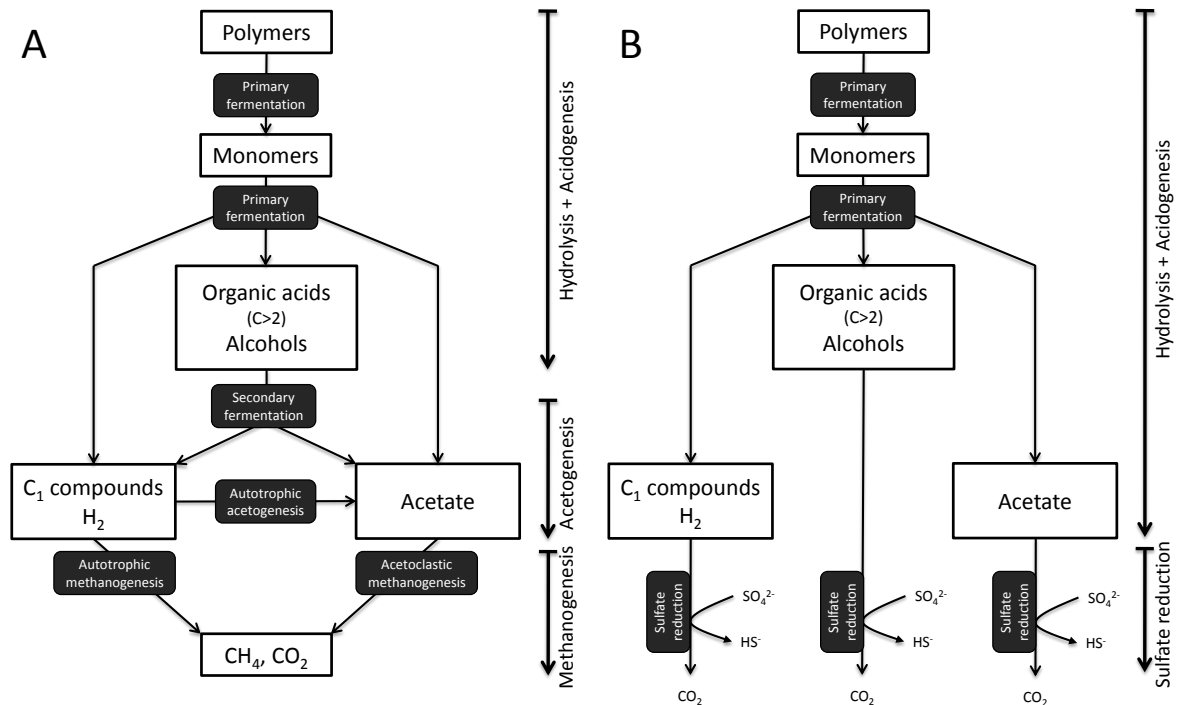


Figure 1.9 Anaerobic pathway of organic matter mineralization (A) in sulfate-free environments and (B) in sulfate-bearing environments. Modified from (Schink, 1997).

1.7.1.2 Source of organic matter

The supply of organic matter derived from photosynthesis generally decreases with depth and with decreasing pore sizes, which reduces water flows. In fractured crystalline rocks, the supply of fresh organic carbon from the surface can sustain heterotrophic metabolisms. However, organic matter content is often limiting in deep subsurface ecosystems and the proportion of chemolithoautotrophic organisms increases with depth.

Of course they are exceptions, which are petroleum reservoirs and organic-rich sediments and sedimentary rocks. To 'survive' over time, organic matter must be protected against biological degradation, which is caused by the lack of electron acceptor, physical inaccessibility, or the recalcitrant nature of organic matter itself. This implies very slow activity of heterotrophic metabolisms in these environments (Fredrickson and Balkwill, 2006).

1.7.2 Inorganic compounds as source of energy

1.7.2.1 The different sources of electron acceptors

Inorganic reduced compounds, mainly H₂, but also H₂S and CH₄, can serve as a primary electron source. They have mainly three different origins:

1. Outgassing of the Earth's mantle, which releases H₂, CH₄ and H₂S in neutral or slightly acidic fluids (Gold, 1992).
2. Serpentinization, which is an abiotic process consisting on the anoxic weathering of ultramafic ferrous silicates, can release H₂ and can occur at room temperature (Stevens and McKinley, 1995).
3. The radiolysis of water can produce H₂ (Barr and Allen, 1959).

1.7.2.2 Microbial energetic systems based on inorganic sources of electrons

A completely photosynthesis-independent subsurface ecosystem has been an intriguing prospect for microbial ecologists for more than ten years, because it could potentially persist indefinitely. It would also represent a good model for describing the first ecosystems that existed before the evolution of photosynthesis. It also represents the most likely ecosystem that could be found outside our planet, such as on Mars. Indeed, the distribution of subsurface lithoautotrophic microbial ecosystems (SLiMEs) could be widespread because they only required appropriate mineralogical and physical conditions.

To be considered a SLiME, both electron source and electron acceptor should be completely dissociated from photosynthesis and must have a geological origin. This implies that the electron acceptor cannot be O₂, NO₃⁻, Fe(III) and SO₄²⁻, because all of these compounds are oxidic species that are a consequence of the oxygenation due to photosynthesis. Consequently, the only electron acceptor in a SLiME is CO₂. The electron source can be more diverse but mainly restricted to H₂, H₂S and CH₄. The microbial community of SLiMEs should be mainly composed of autotrophic organisms, such as methanogenic Archaea, even though heterotrophic organisms are likely to be present, feeding on dead biomass.

It seems that it is really difficult to find such a place on Earth (Stevens and McKinley, 1995). In fact, not a single SLiMEs has been found without any strong controversy so far. A subsurface environment completely free from direct and indirect photosynthesis products, with a porosity allowing life, and which is neither too warm or too reduced so CO₂ is present has proven difficult to find (Nealson *et al.*, 2005).

However, carbon cycling can occur in the deep subsurface in non-SLiME systems, when Fe(III) or SO₄²⁻ are present. So far, no such system has been described. But theoretically, if an inorganic reduced compound is present, organic carbon can be synthesized by a chemolithoautotrophic microorganism, oxidizing it using Fe(III) or SO₄²⁻ as terminal electron acceptors, which produces energy. This energy can then be invested to

transfer electron to CO_2 , reducing it to organic carbon (Madigan *et al.*, 2010). Figure 1.10 shows this process in the case of a sulfate reducing bacterium, using H_2 as the electron and energy source.

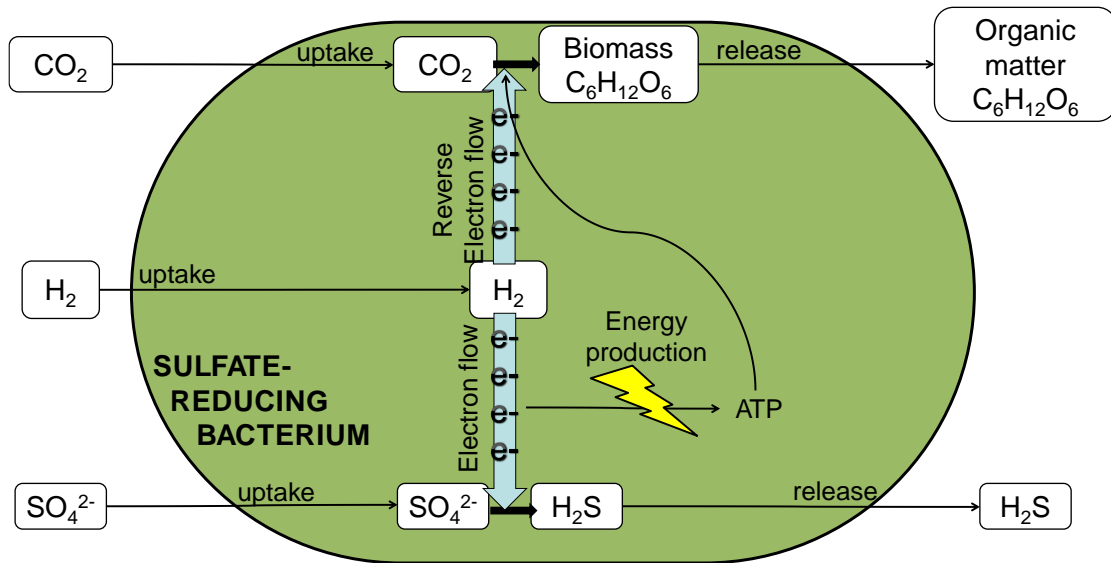


Figure 1.10 Energy production and carbon fixation in a chemolithoautotrophic sulfate reducing bacteria, using H_2 as electron and energy source.

Once organic carbon has been produced, all other metabolisms degrading it (presented in section 1.7.1.1) can take place. In the case of a deep subsurface environment bearing SO_4^{2-} , fermentative microorganisms are expected to be the ones feeding on complex biomass, i.e. dead autotrophic cells. Finally, fermentation products can be oxidized to CO_2 by heterotrophic sulfate reducers, closing the loop of a carbon cycle (Figure 1.11).

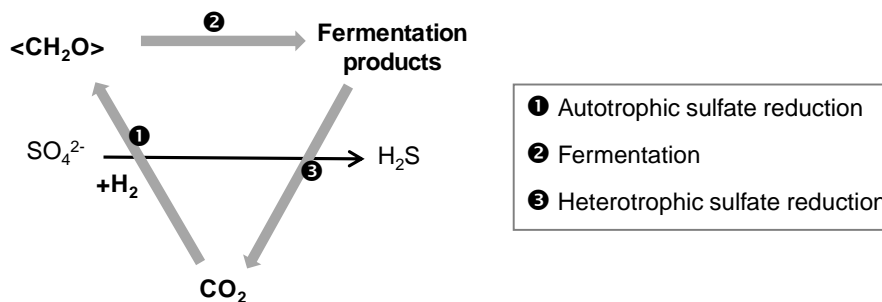


Figure 1.11 Putative carbon loop in deep subsurface, based on the presence of H_2 and SO_4^{2-} . Each grey arrow represents a metabolism category, whose name is indicated in the box. $\langle\text{CH}_2\text{O}\rangle$ is a generic formula representing organic carbon.

Chapter 2 State of the art: Opalinus Clay as a deep geological repository host rock

This second chapter develops the concept of deep geological repository for the disposal of radioactive waste that is envisaged in Switzerland. It explains why microbiology should be considered in this undertaking. It also describes the main findings of the previous microbiological studies that were carried out in Opalinus Clay.

2.1 General context of the study

2.1.1 Nuclear waste disposal in Switzerland

As many other developed countries, Switzerland uses nuclear power as an important source of its energy. Since the 70's, the fraction of the consumed energy produced by nuclear fuel has grown steadily, to reach almost 40% today. Nuclear power plants produce high-level and long-lived intermediate-level nuclear waste that must be stored for hundreds of thousands of years. This duration corresponds to the amount of time needed to reduce their radioactivity. The storage is envisioned in deep geological repositories designed to protect people and the environment, according to the Swiss Nuclear Energy Act of 2003. Nuclear waste is currently being stored in facilities at the surface, which permits ready monitoring. However, surface storage is a temporary solution due to the potential for release. The long-term solution is to bury nuclear waste in stable geological formation (Nagra, 2004, 2008).

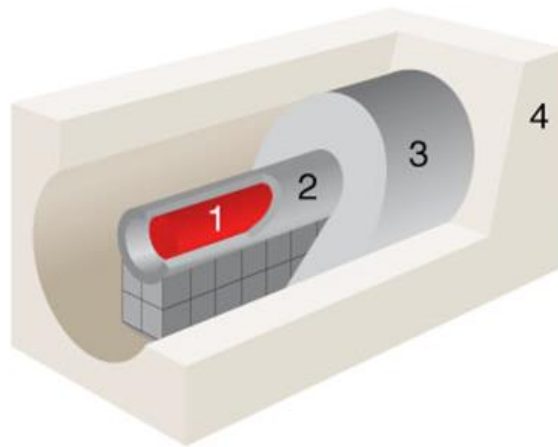


Figure 2.1 The four different safety barriers surrounding high-level nuclear wastes in a deep underground repository: 1) the cladding or the glass matrix, 2) the steel container, 3) the bentonite backfill and 4) the host rock (Nagra, 2013).

This mission has been assigned to the National Cooperative for the Disposal of Radioactive Waste (NAGRA), which was created by the operators of nuclear power plants and the Swiss Confederation and is in charge of nuclear waste produced by medicine, industry and research. The high-level waste is to be stored underground surrounded by four safety barriers. The first barrier is the cladding containing uranium pellets in the case of spent fuel or a vitrified glass matrix in the case of fission products. The second barrier is a thick-walled steel canister and the gallery in which the canister is placed is backfilled with bentonite -a colloidal

swelling clay with significant water retention capacity that represents the third barrier. Finally, the host rock is the fourth and last barrier (Figure 2.1; Nagra, 2013).

The repository of low- and intermediate-level waste exhibits a different configuration. The first barrier consists of a drum that will contain the solidified waste. Then several drums will be placed in a concrete canister and embedded in mortar (second barrier). The two last barriers will be another mortar that will backfill the gallery, and the host-rock (Figure 2.2; Nagra, 2013).



Figure 2.2 Drums containing solidified low- and intermediate-level will embedded by a mortar in a cement canister. These structure will be then place in deep geological repositories, after having backfilled galleries with a special type of mortar (Nagra, 2013).

2.1.2 Underground rock laboratory of Mont Terri

A good host-rock is argillite, a sedimentary rock mainly made of swelling clays, which has very low permeability and a good capacity to bind dissolved substances over geological times. The Opalinus Clay rock, which has similar characteristics and is found in the Jura mountains and in the northern part of the Switzerland, is currently the option supported by Nagra to host nuclear waste (Nagra, 2015). In order to evaluate the suitability of this formation for nuclear waste storage, an underground rock laboratory has been built under Mont Terri (St-Ursanne, Switzerland) taking advantage of the security gallery of a motorway tunnel crossing an Opalinus Clay rock layer (Figure 2.3). Since 1996, the Mont Terri Project, an international research program, aims at assessing the hydrogeological, the geochemical and the geotechnical characteristics of the Opalinus Clay rock (Thury and Bossart, 1999; <http://www.mont-terri.ch>).

2.1.3 A need for microbiological characterization of host rocks

It seems evident that microbiological studies should be part of this research project, because the geochemistry of a given environment depends greatly on the diversity and the activity of microorganisms. However, during the first few years of the Opalinus Clay characterization effort, this aspect was neglected. Only recently has it become increasingly evident that microorganisms may have an important role to play in Opalinus Clay. Indeed, microbial growth has been observed in some Mont Terri experiments, modifying the

chemistry of the porewater (Stroes-Gascoyne *et al.*, 2011), as detailed in section 2.3.1. Since this discovery, microbiological characterizations are fully integrated in the study of the Opalinus Clay rock at the Mont Terri Project. However, only limited findings have emerged from these investigations. The low microbial density and activity in such a rock environment could be responsible for the difficulties encountered by previous studies. A more complete microbiological study of the Opalinus Clay rock at Mont Terri Rock Laboratory - one that would use molecular biology tools - is called for in order to assess the diversity and the metabolic potential of microorganisms and their influence on long-term nuclear waste storage.

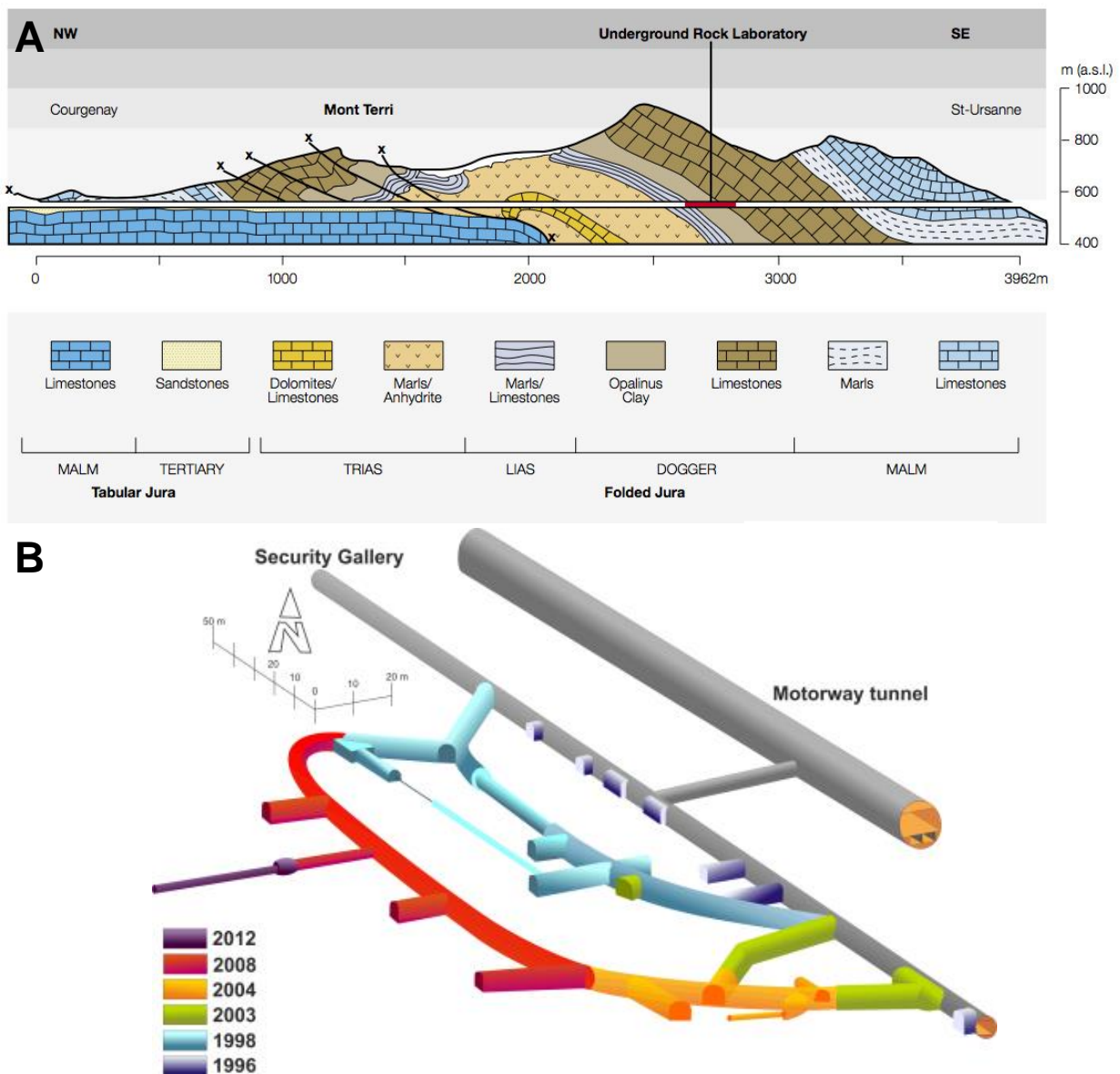


Figure 2.3 A) Geological section of the motorway tunnel going through Mont Terri. B) Map of the underground rock laboratory, which has been sequentially build from 1996 to 2012. Modified from (Bossart and Thury, 2008).

2.2 The Opalinus Clay formation

2.2.1 Opalinus Clay diagenesis

The Opalinus Clay formation, which is likely to host Switzerland's high-level and long-lived intermediate nuclear waste, is a sedimentary rock deposited 174 myr ago in a shallow marine environment of the Jurassic sea. The name 'Opalinus' come from ammonite fossils of *Leioceras opalinum*, which are found in this geological formation. The opalescent luster of their shell has been preserved thanks to the stability provided by the argillite (Gautschi, 2001). The sedimentary detritus, which are at the origin of that formation, come from the erosion of isolated massifs, which were the only part of the land outside the water in the mid-Jurassic. Later, during about 120 myr, the Opalinus Clay formation was covered by and buried under more than 1,000 m of marine limestone, marls/clays and others sediments (now eroded away), which led to its compaction. For the next 30 myr, the disappearance of the sea allowed the erosion of the upper geological formations and the infiltration of freshwater. Finally a last marine stage brought tertiary sediments over the geological sequence until 10 myr ago (Pearson *et al.*, 2003). All these steps are synthetized in Figure 2.4.

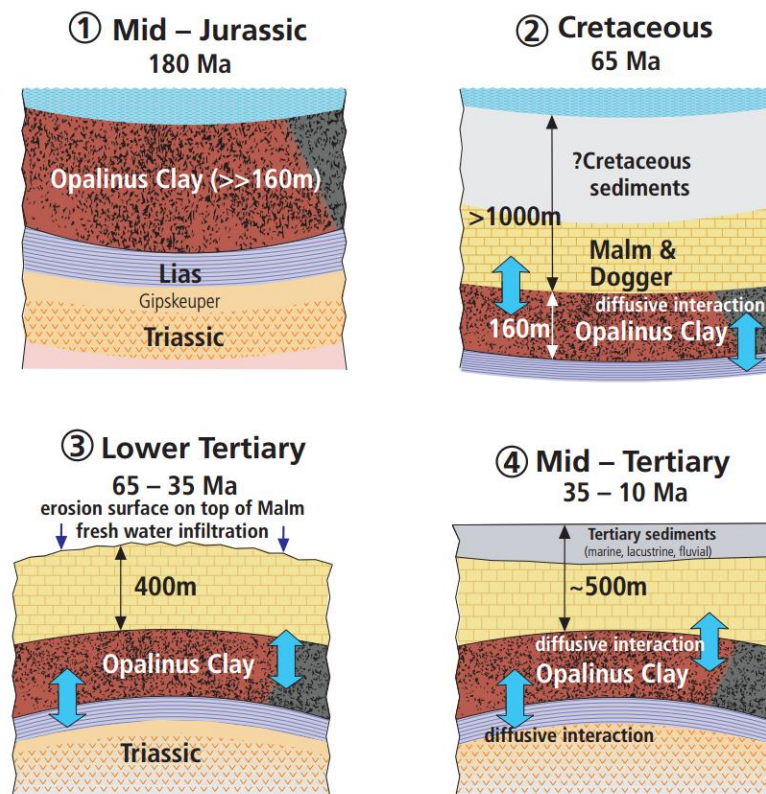


Figure 2.4. Schematic of the geological history of the Opalinus Clay formation at Mont Terri: (1) Deposition of sediments forming the Opalinus Clay ; (2) Deposition of upper sedimentary layers ; (3) Sea retreat and erosion of the upper layers ; (4) Last deposit of sediment due to another sea stage. From Pearson *et al.* (2003).

2.2.2 Opalinus Clay solid phase

The Opalinus Clay rock is a well compacted and moderately consolidated clay stone mainly made of clay minerals, quartz and carbonates. The poor crystallinity of the clay phase indicates a maximal temperature of 80 °C during diagenesis. Three different facies, corresponding to three different sedimentation conditions, were described: the shaly facies, which is the most widespread, the shaly carbonate-rich facies and the sandy facies. The mineral composition of these sub-units is qualitatively similar but has quantitative differences, which are summarized in Table 2.1 (Pearson *et al.*, 2003).

Carbonate and silicate are the dominant minerals in the Opalinus Clay formation. Beside cation exchange ability, silicate minerals can be considered unreactive (Mazurek, 1999). The main redox-active elements present in the Opalinus Clay are C, Fe, Mn and S. Fe(II) is present in siderite, ankerite, pyrite and illite, and Fe(III) in chlorite and silicates. Total iron accounts for 5% (w/w) of the rock. Mn(II) can be found in carbonates but it represents only 0.05 % (w/w) of the rock. S is present in the sulfate contained in the porewater and in sulfur minerals, principally celestite (SrSO₄) and pyrite (FeS₂). Celestite is mainly present in veins in undisturbed Opalinus Clay samples, as well as in the bulk formation, but only at trace levels. S represents 0.5 % (w/w) of the rock (Pearson *et al.*, 2003). This mineralogical composition explains the reducing redox conditions and the near neutral pH conditions of the rock (Stroes-Gascoyne *et al.*, 2007).

Table 2.1 Composition of the three main Opalinus Clay facies observed at the Mont Terri. From Pearson *et al.* (2003).

	Shaly facies (% w/w)	Sandy carbonate-rich facies (% w/w)	Sandy facies (% w/w)
Calcite	5-28	36-42	7.17
Dolomite/ ankerite	0.2-2	2.1-3.3	0.3-2
Siderite	1-4	1.1-2.4	1.1-3
Pyrite	0.6-2	0.2-1.3	0.7-3.2
Organic C	<0.1-1.5	0.2	0.2-0.5
Quartz	6-24	27-31	16-32
K-feldspar	1-3.1	3-5	2.5-5
Albite	0.6-2.2	1.1-2.4	0.8-2.2
Clay minerals	58-76	25	45-70
- Illite	16-40	6-9	15-35
- Illite-smectite	5-20	3-6	5-20
- Chlorite	4-20	2.2-4	4.4-15
- Kaolinite	15-33	8-13	13-35

Organic carbon represents up to 1% (w/w) of the rock in the shaly facies. The greatest part of this organic matter is strongly bound to mineral particles but a small fraction is dissolved in the porewater (ranging from 1 to 16 mg/L), which one third are small organic acids like acetate (203 µM), propionate (27 µM), lactate (9 µM) and formate (2 µM; Courdouan *et al.*, 2007).

The main feature that makes Opalinus Clay a good host rock for long-term nuclear storage is its low hydraulic conductivity, typically in the range of 10⁻¹³ to 10⁻¹⁴ m/s. Therefore, diffusion is the dominating

transport process and the residence time for porewater is in the range of several myr (Stroes-Gascoyne *et al.*, 2007). This low hydraulic conductivity accounts for the preservation of a seawater signature in the chemical composition of the Opalinus Clay porewater, as it is shown in Table 2.2.

2.2.3 Opalinus Clay porewater

Opalinus Clay rock is fully water saturated, which correspond to 5 to 8 % (w/w) of the rock, corresponding to 12 to 18 % of the total porosity (Thury and Bossart, 1999; Pearson *et al.*, 2003). The average pore throat diameter of pore is very small, between 0.01 to 0.02 μm (Mais, 1998), significantly below typical microbial sizes. A more recent study showed that 97% of the Opalinus Clay pore volume is composed of pores whose radius is smaller than 15 nm (Keller *et al.*, 2011). This, added to poor pore interconnectivity and restricted availability of nutrients and water ($a_w=0.93$ to 0.95), makes Opalinus Clay an inhospitable environment, unsuitable for microbial occurrence, activity and mobility (Stroes-Gascoyne *et al.*, 2007).

Table 2.2 Chemical composition of the Opalinus porewater in comparison to seawater. From Pearson *et al.* (2003).

	Opalinus porewater	Seawater
pH	7.96	8.1
Na ⁺ (mM)	245	470
K ⁺ (mM)	1.1	10.1
Ca ²⁺ (mM)	15.2	10.3
Mg ²⁺ (mM)	17.1	53.8
NH ₄ ⁺ (mM)	0.57	
Cl ⁻ (mM)	287	546
SO ₄ ²⁻ (mM)	13.7	28.2
HCO ₃ ⁻ (mM)	0.8	2.25
TOC (mM)	1.1	0.04

2.3 Microbiological study of the Opalinus Clay

2.3.1 First evidence of microbial activity in Opalinus Clay porewater

The porewater chemistry (PC) experiment, whose aim was to follow the chemical changes of artificial porewater put into contact with the Opalinus Clay in a borehole (Wersin *et al.*, 2011), highlighted anomalies in the pattern of change of chemical parameters, that were attributable to microbial activity (Battaglia and Gaucher, 2003). In particular, a decrease in the sulfate concentration and an increase in the inorganic dissolved carbon content in the water were observed. The statistical most probable number method revealed that the water contained more than $1.1 \cdot 10^5$ microbial cells per mL. The presence of microbial aggregates, which could lead to an underestimation of the concentration of living cells, suggests that these organisms likely form

biofilms that cover the pipes and the surfaces of the borehole. Hydrogen sulfide was detected during sampling, indicating sulfate-reducing activity. This microbial activity was made possible first by the presence of space, brought by the borehole drilling, and second by glycerol that leaked from a pH probe, serving as the carbon and energy source for heterotrophic sulfate reducing bacteria (De Cannière *et al.*, 2011).

2.3.2 Impact of microbial activity in deep geological nuclear waste repository

From this point on, it became necessary to assess the role of microorganisms in the Opalinus Clay formation, because they can positively or negatively influence the disposal of nuclear waste (Stroes-Gascoyne *et al.*, 2007):

1. Microorganisms can impact radionuclide mobility. They can reduce it by sorption for instance. Microbes can also, by their activity, modify the pH and Eh of the environment, which, in turn, influences radionuclide sorption, solubility and thus mobility.
2. Microorganisms can produce or consume gases such as CO₂, CH₄ and H₂, depending on their metabolism. A gas production has a negative impact, because it can build up pressure in vaults and cause physical damage, which could impede the rock impermeability. This means that gas consumption is beneficial for the integrity of the repository.
3. Microorganisms can enhance steel canister corrosion by modifying pH and Eh and/or by producing H₂S. This last point is directly relevant because the Opalinus Clay porewater contains SO₄²⁻, which is likely to serve as an electron acceptor in case of anaerobic respiration when an electron donor is available.
4. Microorganisms can enhance the weathering of the host rock itself by reducing Fe(III) from Fe-bearing minerals, and by secreting organic acids, which reduce the pH and complex metal ions (Uroz *et al.*, 2009).

Hence, the Mont Terri Project is keen on investigating the role of microorganisms in the Opalinus Clay formation in more detail to better understand the potential impact on indigenous and introduced microorganisms on long-term nuclear waste storage.

2.3.3 Microbial activity in the Opalinus Clay porewater

More microbial studies were carried out to better understand the biological processes occurring in Opalinus Clay formation porewater, where the lack of space is not limiting for microbial activity. Stroes-Gascoyne *et al.* (2011) summarized analyses performed by several microbiological laboratories carried out on PC porewater. A concentration of 10⁸ cells per mL was measured by DAPI and SYBRGreen staining, which is

higher than in oligotrophic groundwater that contains between 10^3 and 10^5 cell per mL. The metabolic diversity was found to be high and the community that was probed by culturing methods followed by molecular identification (through 16S rRNA gene sequencing) included aerobic and anaerobic heterotrophic, nitrate-reducing, iron-reducing, sulfate-reducing and methanogenic microorganisms. The occurrence of sulfate-reducing bacteria and methanogens is supported by the presence of H_2S , pyrite and CH_4 . Some microorganisms were identified by extracting and amplifying 16S rRNA genes directly from PC porewater. They include a *Desulfosporosinus* sp., a sulfate-reducing bacterium (Spring and Rosenzweig, 2006), a *Hyphomonas* sp., a facultative anaerobe (Moore *et al.*, 1984), and Clostridiaceae and Peptococcaceae organisms, which are strict anaerobes (Rogosa, 1971; Wiegel *et al.*, 2006). These metabolic annotations perfectly match environmental conditions encountered in PC borehole and highlight the importance of Gram-positive bacteria in Opalinus Clay, because all of the identified microorganisms, but the *Hyphomonas* sp., belong to the Firmicutes phylum.

The importance of *Desulfosporosinus* genus was shown by glycerol enrichments of porewater coming from PC experiment, but also from ‘Diffusion and Retention’ (DR) experiment, also contaminated with glycerol from a leaking probe (Gimmi *et al.*, 2014). Enriched microorganisms were all belonging to *Desulfosporosinus* genus, the majority being *Desulfosporosinus lacus* (Fruttschi and Bernier-Latmani, 2010a, 2010b), which is able to grow on glycerol (Ramamoorthy *et al.*, 2006a).

2.3.4 Microbial activity in the Opalinus Clay rock

A separate series of microbial experiments considered in more detail the microbiology of Opalinus Clay rock samples. Because drilling techniques imply microbial contamination even with the most stringent precautions, a particle tracer test was performed and showed that microspheres (0.4 μm in diameter) never penetrated more than 5 mm inside the core rock. Thus, the central part of cores can be considered pristine (Mauclaire *et al.*, 2007). Drilling with compressed air, as compared to nitrogen drilling, did not have an obvious effect on the microbial community of the core. Direct observation of microbes in Opalinus rock sample using DAPI staining and CARD-FISH failed, due to too low cell concentrations and interference with clay particle (Stroes-Gascoyne *et al.*, 2007).

Four different laboratories attempted to extract DNA from Opalinus Clay rock, but none succeeded. This is likely due to the low viable cell content and to the strong DNA binding effect of clay minerals (Stroes-Gascoyne *et al.*, 2007). DNA extraction from added spore samples indicated that the spore detection limit was below 10^3 spores per g of rock, which indicates that Opalinus Clay contains less than 10^3 cells per g. In this analysis, PCR amplifications were most efficient by using non-diluted DNA extracts, indicating that Opalinus Clay did not contain high concentration of DNA polymerase inhibitors (Poulain *et al.*, 2008).

Phospholipids fatty acids (PLFA) analysis revealed that the Opalinus Clay contains about 10^5 or 10^6 viable cells per g of rock, which is small in comparison to soil samples. Because PLFA degradation can be strongly

decreased in environments with little free water and by association with certain minerals (Onstott *et al.*, 1998; Mauclaire *et al.*, 2007), PLFA analysis can lead to an overestimation of the viable cells concentration in Opalinus Clay. The analysis also suggested little nutrient turnover due to weak biological activity. Indeed, there was 14 times more microbial debris than PLFA indicating viable cells. The specific biomarkers indicated mostly the presence of Eubacteria in the rock. Among them, a weak percentage corresponds to sulfate-reducing bacteria, such as *Desulfobulbus*, *Desulfovibrio* and *Desulfobacter*, all Gram-negative, while the microbial community mainly consisted of anaerobic Gram-negative bacteria.

Out of the four laboratories, two were able to enrich microorganisms. The difficulty and the long incubation time needed to obtain microbial enrichments suggested that no contaminating bacteria had been introduced during the sampling procedure. They also indicate that viable cells are present in Opalinus Clay in small populations well adapted to the extreme conditions in the clay. Mauclaire *et al.* (2007) cultured bacteria in media for sulfate-reducing bacteria containing lactate or malate after three and seven months of incubation. Growth was confirmed by DAPI staining and microscope observation. The concentration of cultured microbes was low, between 1 and $18 \cdot 10^3$ cells/mL, and their number was larger in the lactate medium, which is less selective. The average size of the recovered microorganisms was quite small, around 0.15 μm . Poulain *et al.*, (2008) obtained positive enrichments in media for aerobic heterotrophs, for anaerobic heterotrophs, for anaerobic heterotrophs using inorganic electron acceptors, and for methanogens, after 1 to 35 weeks of incubation time. Growth was detected by amplifying the 16S rRNA gene by PCR. For the methanogen enrichment, no Archaea were detected by PCR, indicating that another type of microorganism was enriched. Incubations with the introduction of non-indigenous microorganisms showed that Opalinus Clay contains enough nutrients to sustain the presence of a microbial population when water and the pore size are not limiting factors (Stroes-Gascoyne *et al.*, 2007).

Even if nearly all the retrieved cultured microorganisms come from extreme or nutrient-starved environments, seemingly a confirmation of their indigenous origin, it is quite surprising to enrich many aerobic microorganisms from the Opalinus Clay rock. One explanation could be that the Opalinus Clay rock has trapped all the microbial community present 174 myr ago in the Jurassic sea, including aerobic bacteria. Then, during the anoxification of the sediment and its diagenesis, only microorganisms adapted to extreme or nutrient-starved environments have survived until today.

The question of survival beyond three million years in a geologic formation is currently being disputed in the literature, as explained in sections 1.6.2 and 1.6.3. However, it is possible that the low water activity of the Opalinus Clay (a_w between 0.93 and 0.95) has decreased the rate of spontaneous hydrolysis of DNA, increasing the survival time of microorganisms. Another possible alternative is a recent microbial intrusion by groundwater via the major fault zones occurring less than 3 myr, after the last sea stage. However, it is difficult in that case to explain how microorganisms have left the bigger fractures to colonize the smaller pore zones where they were recovered (Stroes-Gascoyne *et al.*, 2007). It is why a sedimentary origin of the

microorganisms seems more likely. Those were deposited during the Jurassic together with the sediments that later formed Opalinus Clay. They kept living, thriving and evolving at a very low rate in the few macropores available in the rock, adapting themselves to extreme or nutrient-starved environments, extracting energy from the exergonic oxidation of organic compounds of Opalinus Clay, or of hydrogen putatively produced by radiolysis of water, owing to the presence of radioactive elements such as U, Th and ^{40}K , and to the accumulation of ^4He derived from alpha particles (Pearson *et al.*, 2003).

In Opalinus Clay, the average pore size is smaller than $0.02\ \mu\text{m}$ but there exist macropores that are in the micron range. Thus, it is likely that microorganisms survived in the macropores from the deposition of the Opalinus sediment, 170 myr ago. It seems also possible than microbes have survived in fractures, but this kind of environment was not targeted by any study to date (Stroes-Gascoyne *et al.*, 2007). But so far, no one has put forward the hypothesis that, because microorganisms have likely the same origin of the sedimented material composing the rock and were entrapped in this geological formation, they could have survived in a non-measurable porosity, being totally surrounded by rock and resisted the pressure of the geological formation. It would mean that, even if the measurable porosity is too small to support microbial life, living microorganisms could be present within the rock.

2.3.5 Conclusions of these studies and open questions

The results to date suggest that Opalinus Clay contains a small indigenous microbial community, which is probably almost inactive or in a dormant state. This can be explained by space and water limitations. However, any disturbance, by providing space, water, but also nutrients, can revive these microorganisms, as is seen in porewater experiments.

Building an underground nuclear waste repository will greatly impact the properties of Opalinus Clay, creating better conditions for microbial activity. Contemporary microorganisms and organic matter will be brought during the excavation and be present in some low- and-intermediate level waste. Pore size in Opalinus Clay will be also increased, as a consequence of drilling. Bossart *et al.* (2002) showed that such an undertaking leads to the creation of a network of larger fractures surrounding the gallery, the excavation disturbed zone (EDZ; Figure 2.5). Another positive disturbance for microbial activity is anoxic steel corrosion, known to produce hydrogen gas, which can be used as a source of energy by microorganisms (see section 1.7.2.2). This process will occur in the repository, because canisters for high-level waste and drums for low- and intermediate-level waste are made of carbon steel, and is well known to stimulate sulfate reduction activity (Enning and Garrelfs, 2014).

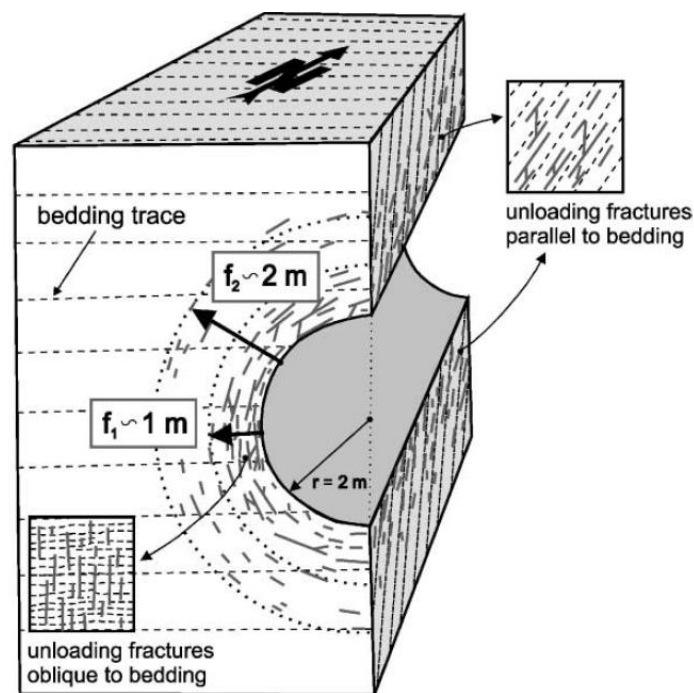


Figure 2.5 Excavation disturbed zone (EDZ) created by excavation works, consists of a network of bigger fractures surrounding the gallery. From Bossart *et al.* (2002).

However, nothing is known about microbial activity in these two particular environments of Opalinus Clay, when space alone is available, and when space and hydrogen are available, both of which are relevant to a deep geological repository. The aim of this Ph.D. thesis is thus to better characterize microbial communities, in terms of composition and microbial activity, in these two conditions, in order to better evaluate their impact on the safety of underground nuclear waste disposal. The first case, presented in Chapter 3, describes microbial communities in borehole porewater where no organic carbon was added, which is a perfect analogue of microbial activity taking place in the EDZ and in the zones with a bigger porosity distribution of Opalinus Clay. The second case, presented in Chapter 4 and Chapter 5, describes the microbial response to an addition of hydrogen in borehole, in order to simulate hydrogen produced by anoxic steel corrosion.

Chapter 3 Microbial metabolisms in the porewater of Opalinus Clay

This third chapter describes the first research project whose initial aim was to describe the microbial community in Opalinus Clay borehole waters. As the great majority of samples collected were contaminated with aerobic biomass that was likely released from biofilms growing in the sampling lines, this project focuses on a single sample that can be considered as an analogue to the zones of Opalinus Clay containing more available space for microorganisms, such as in the excavation disturbed zones.

This chapter was submitted in a shorter version to FEMS journal, under the title “A minimalistic microbial food web in an excavated deep subsurface clay rock”.

All the work presented in this chapter was carried out by the student, except the 16S rRNA sequencing and analysis that was performed by the DOE Joint Genome Institute (Walnut Creek, CA, USA), the metagenomic library preparation and sequencing that were performed by the Genomic Technologies Facility of the University of Lausanne (Switzerland), and the metagenomic read assembly that was performed by Ino de Bruijn at the Bioinformatics Infrastructure for Life Sciences (Stockholm, Sweden).

Abstract

Clay rocks are being considered for radioactive waste disposal but little is known about the potential impact of microbes on the long-term safety of geological disposal systems. Thus, a more complete understanding of the microbial community structure and function in these environments would help increase the sustainability of the geological disposal of the waste. Additionally, it would provide a unique glimpse into a poorly studied deep subsurface microbial ecosystem. Previous studies concluded that, in a pristine Clay rock (Opalinus Clay), the microbial community was inactive due to the pore size distribution favoring small pores and the resulting lack of space. This study shows that an increase in pore size alone is sufficient to stimulate bacterial growth. Metagenomic sequencing and genome-binning of a porewater sample revealed a remarkably simple heterotrophic microbial community, fueled by sedimentary organic carbon, mainly composed of two organisms: a *Pseudomonas* sp. fermenting bacterium growing on organic macromolecules and releasing organic acids and H₂, and a sulfate-reducing Peptococcaceae able to oxidize organic molecules to carbon dioxide. This microbial system likely thrives in fractured zones of the Opalinus Clay formation and may have beneficial impacts on the stability of a repository (H₂ consumption) as well as negative impacts (steel corrosion and rock weathering). The study also reveals ubiquitous operational taxonomic units that likely constitute a core microbial community indigenous to the Clay rock.

3.1 Introduction

The existence of microbial life in the deep terrestrial subsurface has been documented (Pedersen, 2000; Colwell and Smith, 2004) but little is known about the diversity and functioning of the microbial communities in these systems. We depend on the subsurface for several critical services such as groundwater supply, resource recovery, gas storage and nuclear waste disposal. Hence, understanding the distribution and diversity of microbial communities as well as their metabolic potential and activity in the subsurface is an active area of research. Unfortunately, it is hampered by the logistical difficulties associated with access to this environment (Edwards *et al.*, 2012).

Clay formations such as Opalinus, Boom, or Callovo-Oxfordian Clay are low-permeability rocks that are being considered in several European countries for the disposal of radioactive waste. Such disposal sites, named geological repositories, represent an international consensus on the favored option to contain the waste until radionuclides decay to an acceptable level of radioactivity. Repository layouts and engineered barrier systems are specific to each country, but clay rocks are generally preferred due to their reliable barrier function. Extensive study of the physical and chemical characteristics of potential host clay rocks has been conducted over the past 30 years but the extant microbial communities and their associated metabolic potential remain poorly described. The rationale for investigating the metabolic potential of repository rock microbial communities is to contribute to the scientific basis for the evaluation of their possible impact on the safety of engineered barriers. There are a number of microbial processes that could impact these barriers either positively or negatively. For instance, gas production could lead to pressure build-up in the vaults and enhancement of the weathering of the barriers around the waste (Stroes-Gascoyne *et al.*, 2007). Positive effects include the consumption of gas produced by steel corrosion in order to avoid critical pressure build-up in repositories, and the reduction of radionuclide mobility when microorganisms precipitate and/or assimilate these compounds. Delineating which of the above metabolic processes are significant in these environments is the first step of a more complete microbial characterization of repository rocks and a more exhaustive safety case.

In Switzerland, the currently planned strategy is to dispose of radioactive waste in the Opalinus Clay formation, a sedimentary rock that was formed about 170 million years ago. Salient features of this rock formation are its very small pore size, with a mean pore size of 10 to 20 nm, a very low hydraulic conductivity ($< 10^{-13}$ m/s) and a strongly reducing environment, which is due to the presence of minerals such as siderite and pyrite (Mazurek, 1999; Thury and Bossart, 1999). The Opalinus Clay also contains up to 1.5% (w/w) organic matter that is mainly associated with the solid phase (Pearson *et al.*, 2003; Courdouan *et al.*, 2007). The slow diffusion of water occurring in this formation results in the porewater remaining in chemical equilibrium with the rock, and thus retaining the marine signature of its depositional environment, which explains sulfate concentrations around 10 to 20 mM (Pearson *et al.*, 2003). The pore size distribution, favoring

very small pores, also results in limited microbial activity in the rock, simply due to the lack of physical space (Stroes-Gascoyne *et al.*, 2007). Previous studies attempting to characterize microorganisms in Opalinus Clay using traditional culturing and 16S rRNA gene-based methods (Stroes-Gascoyne *et al.*, 2007; Poulain *et al.*, 2008) concluded that a small number of microorganisms were present in the Opalinus Clay but that they were inactive. However, the introduction of space in the rock through e.g. drilling could enhance microbial activity. It is expected that during the construction of a repository, a network of micro- and macro-scale fractures will develop as a result of excavation, leading to the formation of an excavation-disturbed zone (EDZ) around the galleries, that is characterized by slightly higher hydraulic conductivity (Bossart *et al.*, 2002).

The present study aims to determine whether microbial activity in Opalinus Clay can be stimulated by the addition of space alone and to characterize active metabolisms extant in this space.

3.2 Materials and Methods

3.2.1 Porewater sampling

The Underground Rock Laboratory (URL) of Mont Terri (St-Ursanne, Switzerland, Figure 3.1), located about 300 m below the ground surface, was the location of this study. This facility allows access to the Opalinus Clay via a number of boreholes. This work mainly focused on a single borehole (BIC-A1) that has a vertical descending orientation, is 15.6 meters long and was drilled in Gallery 08 on 19th of March 2012 (Figure 3.1.C). The borehole was maintained under anoxic conditions until sampling by immediately flushing it with argon and by installing a hydraulic inflatable packer within one day of the drilling, sealing off the rock formation from the atmosphere of the tunnel. Porewater from the formation accumulated into the borehole over time at an average rate of 44 mL/day and provided a possible environment for microbial growth. A water sample was recovered from the BIC-A1 borehole 10 months after its drilling. Sampling was performed using a sterile glass bottle attached to an aluminum rod sterilized by ethanol and flaming. The water samples were immediately filtered using a sterile filtration device and a sterile 0.22 μm polycarbonate membrane. The membrane was stored in LifeGuard Soil Preservation Solution (Mobio, Carlsbad, USA) and placed at -20°C within a few hours.

Twenty-two other samples were recovered from 7 other boreholes (Figure 3.1.C), using three approaches, including the one described for BIC-A1 for open boreholes. Most boreholes are sealed with a packer, a large hydraulic stopper that separates borehole water from the oxic atmosphere of the gallery, and long tubing lines that allow borehole sampling from the gallery. For boreholes with sufficiently high internal pressures, water was sampled directly into a sterile and anoxic bottle after discarding the dead volume of the lines (second approach). Where boreholes were not significantly over-pressurized, a sterile and anoxic gas-sampling bag

was connected to one of the lines, and the water was recovered over time at rates between 10 and 25 mL per day (third approach). Because these bags remained in the gallery for several weeks, oxygen probably diffused through the bag wall, causing slight oxygen contamination in the water sample. All water samples were filtered within hours of recovery, using a sterile filtration device and a sterile 0.22 μm polycarbonate membrane. The membrane was frozen at -20°C until DNA extraction.

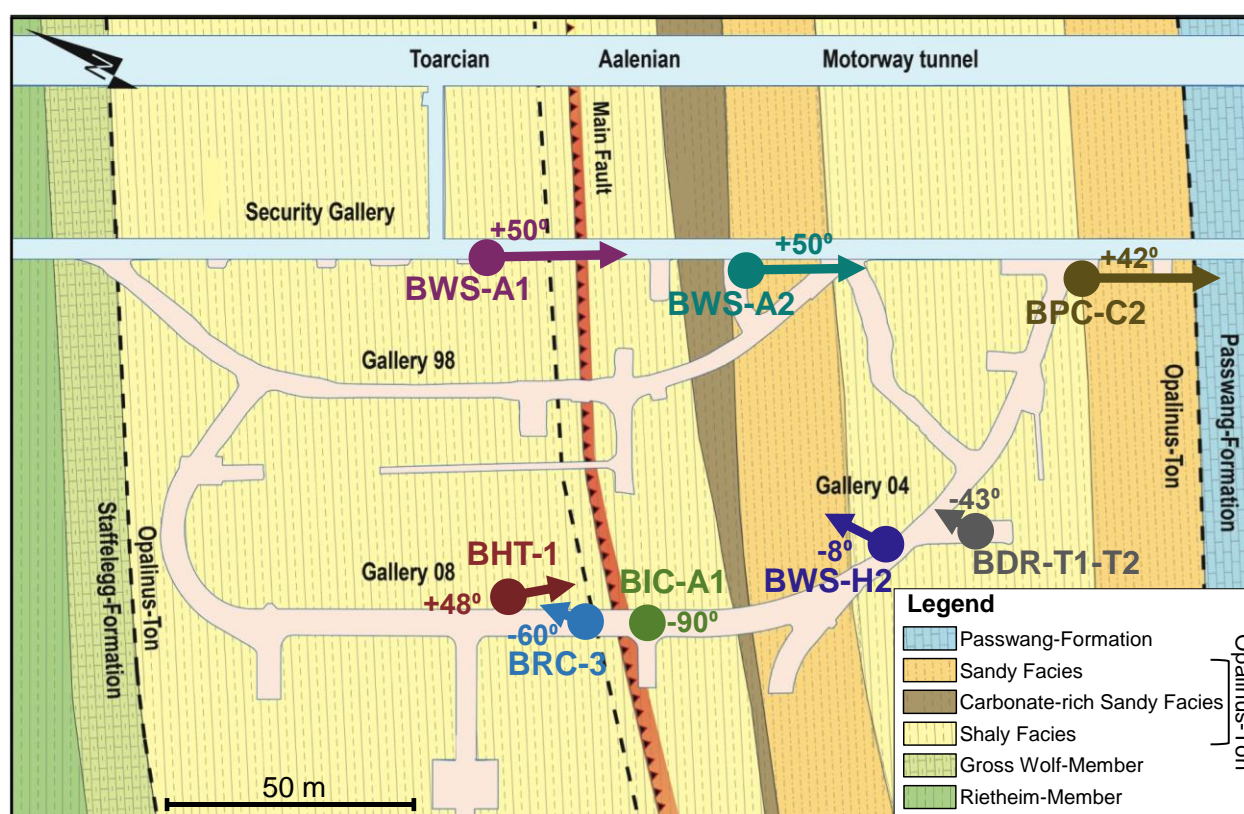


Figure 3.1 Situation of the Underground Rock Laboratory (URL) of Mont Terri and borehole emplacements. A) Location of the Mont Terri URL in Switzerland. B) Cross-section of the motorway tunnel of Mont Terri, whose security gallery was used for excavating the URL. C) Location of the boreholes sampled in this study. The circles indicate the position of the boreholes opening and the angles indicate the boreholes orientation compared to the ground (a negative angle means a descending borehole). Modified from Swisstopo (Bossart and Thury, 2008).

3.2.2 Chemical analysis of porewater samples

Filtered water samples were characterized chemically for cation and anion composition using ion chromatography (DX-3000, Dionex, Sunnyvale, USA). Major anions were separated with an IonPac AS11-HC column, using a KOH gradient from 0.5 to 30 mM for elution while major cations were separated with an IonPac CS16 cation-exchange column, using 40 mM of methanesulfonic acid for elution. Metals (Mn, Cr, Zn, Si, Al, Fe, Co, Cu, Sr and Ni) were measured using inductively coupled plasma–optical emission spectroscopy (ICP-OES, Shimadzu ICPE 9000), and the samples and standards were prepared in 0.1M HNO_3 (final

concentration). The pH was measured in unfiltered water using an Orion™ 8102BNUWP ROSS Ultra™ pH Electrode (Thermo Scientific).

When possible, oxidation-reduction potential (ORP) was measured using a combined Pt-ring electrode (ref. 6.0451.100, Metrohm, Switzerland) equipped with a flow-through cell, and Fe(II) and S(-II) were measured on 0.2 µm filtered water sample, using assays described by Stookey (Stookey, 1970) and Cline (Cline, 1969), respectively.

3.2.3 DNA extraction

DNA was extracted from particulate matter retained by the filtration membranes using a slightly modified protocol of the DNA Spin Kit for Soil (MP Biomedicals, Santa Ana, USA), including (i) a 5 minute incubation of the biomass in the lysis buffer at 60°C before bead-beating, and (ii) a second extraction with new lysis buffer and another bead-beating step. Extracted DNA was further purified using the standard protocol of Genomic DNA Clean & Concentrator purification kit (Zymoresearch, Irvine, USA). The complete protocol is described in section A-1.1.1.

3.2.4 16S rRNA genes analysis

3.2.4.1 Libraries preparation and sequencing

16S rRNA sequencing was done by the Joint Genome Institute through a community sequencing program project (CSP 1505). Libraries were prepared for Illumina MiSeq sequencing using PCR primers for the hypervariable region V4 of the 16S rRNA gene, which are 515F (5'-GTG CCA GCM GCC GCG GTAA-3') and 806R (5'-GGA CTA CHV GGG TWT CTA AT-3'; Caporaso *et al.*, 2011), using 30 cycles with an annealing temperature of 30°C. Sequencing was done using a 2x250 bp reads configuration. After sequencing, amplicons were analyzed using the iTagger version 1.1 pipeline. The QC filtered reads were scanned with Duk version 1.05 for removing contaminating sequences (Li, 2011), trimmed to 165 bp, merged with FLASH version 1.2.6 (Magoč and Salzberg, 2011), and poor quality merged reads were discarded or trimmed. USEARCH version 7.0.959 (Edgar, 2010) was used for merging libraries between samples using a similarity threshold of 97%, forming Operational Taxonomic Units (OTUs). Chimeric sequences were filtered by UCHIME (Edgar *et al.*, 2011). RD Classifier version 2.5 (Wang *et al.*, 2007) was used for taxonomic assignments, comparing sequences to a Greengenes V4 database (DeSantis *et al.*, 2006). Finally, QIIME version 1.7.0 (Caporaso *et al.*, 2010) was run to generate the 'biom' file and the phylum-level table. The detailed pipeline is described in section A-1.1.2.

3.2.4.2 Alpha-diversity analysis

All the steps described here were carried out using QIIME (Caporaso *et al.*, 2010). The pipeline is available in the section A-1.1.3.

3.2.4.3 Sample clustering analysis

The sample cluster analysis was carried out using R software (R Development Core Team, 2008), using scripts found Numerical Ecology with R (Borcard *et al.*, 2011). Sample clustering based on the microbial community composition was carried out using 9 methods, using three distance matrices (Bray-Curtis, Hellinger and Chord) and three linkage methods (UPMGA, complete and Ward). According to the cophenetic correlation test and to the Gower distance test, the best methods are the Hellinger and Bray-Curtis distance, both with UPGMA distances. They give the same sample clustering, but the first one was selected because the dendrogram was easier to read.

3.2.5 Metagenomic analysis

3.2.5.1 Libraries preparation and sequencing

For metagenomic sequencing, DNA was extracted from a second filtration membrane with the method described above. However, the purification step of the DNA extract was done by ethanol precipitation (see section A-1.1.4 for more information). TruSeq Nano DNA Sample Prep Kit (Illumina) was used for library preparation, using a fragment length of 400 – 800 bp, before Illumina paired-end HiSeq sequencing.

3.2.5.2 Reads assembly and contigs binning

First, the Illumina metagenomic reads were post-processed by Casava version 1.8.2 (http://support.illumina.com/sequencing/sequencing_software/casava.html). The assembly was performed with Ray version 2.3.1 (Boisvert *et al.*, 2012) on a Cray XE6 system using 1,024 cores in 4.5 hours and a kmer length of 31. Reads were mapped to the contigs in order to calculate their coverage, using Bowtie 2 version 2.1.0 (Langmead and Salzberg, 2012) with default parameters. Duplicate reads were removed subsequently using MarkDuplicates from Picard tools version 1.77 (<http://picard.sourceforge.net/>). Reads were also mapped to contigs belonging to 16 others metagenomic samples, all originating from an adjacent borehole, that had been amended with hydrogen and that contained similar microbial species according to 16S rRNA gene sequencing (refer to Chapter 3). For binning, CONCOCT version 0.2 was run (Alneberg *et al.*, 2014) using a contig threshold size of 5000 bp. Prodigal version 2.60 was used to identify genes in the contigs and to translate them into protein sequences (Hyatt *et al.*, 2010). The proteins were annotated using WebMGA (Wu *et al.*, 2011) and the COG database. This annotation was only used for single-copy gene analysis, in order to evaluate the completeness and purity of each cluster (Alneberg *et al.*, 2014) using a custom Perl script.

3.2.5.3 Taxonomic and metabolic annotation of contigs

In parallel, contigs were annotated using the JGI IMG pipeline (Markowitz *et al.*, 2012). Additionally, pathways were identified manually using KO, COG, EC annotations, the KEGG database (Kanehisa *et al.*, 2014), the MetaCyc database (Caspi *et al.*, 2014) and textbook biochemical pathways (Kim and Gadd, 2008). PSORTb version 3.0.2 was used for predicting subcellular localization of genes products (Yu *et al.*, 2010). The prediction of ribosomal RNA genes in the metagenome was done by Barrnap version 0.4.2 (<https://github.com/Victorian-Bioinformatics-Consortium/barrnap>) and their taxonomic affiliation with the RDP Classifier version 2.7 (Wang *et al.*, 2007), using a confidence threshold of 0.8. NCBI BLASTN platform (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was also used for taxonomic annotation of 16S rRNA genes. MLTreeMap version 2.061 was used for taxonomic annotation of the bins (Stark *et al.*, 2010), as well as phylogenetic distribution output from the JGI IMG pipeline. Taxonomic annotation of clusters considered as draft genomes was also done using average nucleotide identity (ANI) with references genomes (Goris *et al.*, 2007), using the Kostas Lab platform (<http://enve-omics.ce.gatech.edu/ani/index>). Uniprot was also used to annotate some of the genes (The Uniprot Consortium, 2014). To compare clusters of genes involved in sulfate reduction, hydrogen oxidation and acetyl-CoA pathway of a microorganism abundant in the BIC-A1 borehole with that of reference genomes, BLAST version 2.2.28 (Altschul *et al.*, 1990) and Mauve version 2.3.1 (Darling *et al.*, 2009) were used. Fasta files and protein table files of reference genomes were downloaded from NCBI database (<http://www.ncbi.nlm.nih.gov/genome>).

3.2.5.4 Clusters contribution to microbial community

The abundance $a(C_i)$ of each cluster C_i compared the abundance of the entire microbial community M is defined as follows:

$$M = \sum_{i=1}^m a(C_i)$$

where m is the number of clusters and

$$a(C_i) = \frac{\sum_{j=1}^n l(C_j) \cdot c(C_j)}{\sum_{j=1}^n l(C_j)}$$

where n is the number of contigs in cluster C_i , $l(C_j)$ is the length of contig C_j and $c(C_j)$ the mean coverage of C_j . That is the abundance of a cluster is expressed as its mean coverage. The relative abundance $r(C_i)$, or contribution, of each cluster is defined as

$$r(C_i) = \frac{a(C_i)}{M}$$

3.2.5.5 Sequences deposition

All sequencing data, 16S rRNA genes, draft genomes, unclassified contigs, genes annotations and raw reads were deposited in NCBI and IMG databases. Accession numbers can be found in section A-1.1.5.

3.3 Results

3.3.1 Redox potential assessment of porewater samples

The 23 samples considered are listed in Table 3.1 along with general information about their sampling, including ORP assessments. The borehole water samples whose ORP assessment is considered as low were all sampled in “one shot” and are:

- BPC-2/1, whose reduction potential measured using a flow-through cell was found to be -287 mV, and which contained about 100 μM of Fe(II);
- BWS-A2/1, whose reduction potential measured using a flow-through cell was found to be -164.7 mV, and which contained about 50 μM of Fe(II);
- BHT-1/3, which contained 9.2 μM of soluble Fe(II), and which was amended with hydrogen prior to sampling;
- BDR-T1 samples, which smelled like sulfide when sampled, and were contaminated with glycerol through a pH probe prior to sampling;
- and BIC-A1 sample, which contained 7.5 μM of soluble S(-II).

Some other samples are clearly oxic because they were recovered:

- through sampling bags that were left in place several weeks (BWS-H2/0 to BWS-H2/3, BRC-3/2, BRC-3/4, BRC-3/15);
- from an open canister (BHT-1/1);
- or from an open descending borehole (BRC-3/1).

Finally, for the others samples (BHT-1/2, BRC-3/5, BRC-3/13, BRC-3/17, BRC-3/18, BWS-A1/1, BWS-A1/4 and BDR-T2/1), there is no direct evidence of oxygen contamination because they were sampled in “one shot”, but no evidence of low ORP either.

3.3.2 Chemical characterization of samples

Chemical analysis of sampled porewater carried out by IC and ICP-MS obtained values similar to those previously reported for Mont Terri borehole waters (Pearson *et al.*, 2003; Table A-1.1 and Table A-1.2). Only Si measurements differ significantly, as most of the values measured in this study are zero, while the concentrations reported by (Pearson *et al.*, 2003) are mostly greater than 1000 ppb. Cr, Zn and Al are below detection limit. Fe is higher in samples containing soluble Fe(II), which are BPC-2/2 and BWS-A2/2. Nitrate concentrations greater than 0.03 mM indicate an oxygen contamination in samples BHT-1/1, BRC-3/1, BRC-3/4 and BWS-H2/1, which are all considered as oxic, due to their sampling methods, as explained in section 3.3.1. Some high Br values in BRC-3/5 to BRC-3/18 are higher due to Br injections for a tracer test. Finally, sample BPC-2/2 has a chemical signature that is significantly different from the other samples, highlighting the fact that this borehole water originates from another rock formation (Passwang rock formation; Figure 3.1).

3.3.3 Microbial communities

The Illumina sequencing produced 2 574 000 good quality reads of the V4 region of the 16S rRNA gene. Among the 23 samples, 825 OTUs were identified (Table A-1.3). The fraction of strict anaerobes and strict aerobes identified by 16S rRNA gene sequencing is listed in Table 3.1 (which is based on Table A-1.4) and shows that out of the 23, only four lacked strict aerobic bacteria as is expected from the Opalinus Clay rock environment. Paradoxically, among the samples in which an aerobic microbial community was identified, two exhibited chemical characteristics associated with anoxic conditions: BHT-1/3 and BWS-A2/1.

3.3.3.1 Aerobic microbial communities

A histogram of the microbial community composition based on the region V4 of 16S rRNA gene of all samples can be found in Figure A-1.1. Microbial communities of the 19 samples that do not harbor predominantly strict anaerobes are dominated by Alpha-, Gamma- and Betaproteobacteria. The most frequent microbial groups observed contain strict aerobes, such as Xanthomonadaceae and Phaeobacter, metabolically diverse members of Methyloversatilis and Rhodospirillaceae, and facultative anaerobic bacteria such as Sphingomonadaceae and *Pseudomonas*. The metabolism type and oxygen affinity of the most important microbial groups are presented in Table A-1.4.

3.3.3.2 Anaerobic microbial communities

The three boreholes from which the four samples were found to harbor exclusively strict or facultative anaerobic bacteria were: (a) BPC-2 whose reduction potential measured using a flow-through cell was -287 mV; (b) BIC-A1, which contained 7.5 μM of soluble sulfide; and (c) BDR-1 that gave off a strong hydrogen sulfide smell. These boreholes differed significantly in their characteristics. BPC-2 not only drains Opalinus Clay but also the adjacent rock formation (the limestones of the Passwang formation; Figure 3.1) and thus

produces a significantly greater flow-rate than boreholes that only intersect Opalinus Clay. BDR-T1 was the site of an experiment in which porewater was recirculated and a pH probe slowly leaked glycerol into the system. BIC-A1 is the newest borehole sampled and was sealed for a period of 10 months immediately after drilling.

Table 3.1 Summary of relevant characteristics for each sample: borehole origin, sampling date, sampling method (direct or bag), evidence for high or low ORP, presence of an electron donor, and proportion of strict anaerobes and strict aerobes (based on V4 region of 16S rRNA gene taxonomic affiliation). When low ORP samples harbor a strictly aerobic microbial community (samples BHT-1/3 and BWS-A2/1), most of the sampled biomass was likely flushed from the sampling lines that were contaminated with oxygen.

Sample	Borehole	Sampling date	Sampling method	Evidence for high ORP	Evidence for low ORP	e- donor	Proportion of strict anaerobes	Proportion of strict aerobes
BPC-2/1	BPC-2	12.12.2012	Direct		Eh = - 287 mV		84.8%	0.1%
BHT-1/1	BHT-1	29.06.2010	Direct	Contact with O ₂			0.1%	48.1%
BHT-1/2	BHT-1	08.06.2011	Direct				0.3%	9.2%
BHT-1/3	BHT-1	10.01.2013	Direct		[soluble Fe] = 9.2 µM	Hydrogen	5.0%	39.6%
BIC-A1/1	BIC-A1	10.01.2013	Direct		[S(-II)] = 7.5 µM		73.0%	0.0%
BRC-3/1	BRC-3	13.09.2010	Direct	Contact with O ₂			2.3%	13.7%
BRC-3/2	BRC-3	13.01.2011	Bag	Contact with O ₂			7.1%	10.0%
BRC-3/4	BRC-3	03.04.2012	Bag	Contact with O ₂			1.9%	50.9%
BRC-3/5	BRC-3	22.03.2012	Direct				30.1%	8.8%
BRC-3/13	BRC-3	05.07.2012	Direct				7.5%	32.9%
BRC-3/15	BRC-3	31.07.2012	Bag	Contact with O ₂			2.8%	9.9%
BRC-3/17	BRC-3	18.10.2012	Direct				18.2%	24.7%
BRC-3/18	BRC-3	09.11.2012	Direct				2.1%	65.1%
BWS-A1/1	BWS-A1	29.06.2010	Direct				7.9%	11.7%
BWS-A1/4	BWS-A1	16.11.2011	Direct				0.3%	62.1%
BWS-A2/1	BWS-A2	28.09.2012	Direct		Eh = - 164.7 mV		3.4%	22.0%
BWS-H2/0	BWS-H2	17.07.2009	Bag	Contact with O ₂			1.1%	20.2%
BWS-H2/1	BWS-H2	29.06.2010	Bag	Contact with O ₂			9.2%	12.7%
BWS-H2/2	BWS-H2	26.11.2010	Bag	Contact with O ₂			1.2%	2.5%
BWS-H2/3	BWS-H2	01.02.2011	Bag	Contact with O ₂			3.4%	16.8%
BDR-T1/1	BDR-1-int1	05.03.2009	Direct		Sulfide smell	Glycerol	32.5%	0.0%
BDR-T1/2	BDR-1-int1	27.01.2010	Direct		Sulfide smell	Glycerol	35.4%	0.0%
BDR-T2/1	BDR-1-int2	27.01.2010	Direct				2.2%	16.9%

The microbial communities from the three boreholes corresponding to the four aforementioned samples are presented in Figure 3.2. First, it is striking to observe the low microbial diversity obtained from the four samples: only 5 OTUs (OTU 6, OTU 9, OTU 2, OTU 5 and OTU 0) compose more than 75% of the microbial communities (Table 3.2). Second, 6 OTUs were observed in all 23 samples (OTU 6, OTU 2, OTU 0, OTU 1, OTU 3 and OTU 15; 23/23 samples in Table 3.2). There is significant overlap between the 5 OTUs dominating anoxic samples and these 6 ubiquitous OTUs. More specifically, OTU 6, OTU 2 and OTU 0 represent dominant OTUs in at least some of the four aforementioned samples and are present in all 23 samples whereas OTU 1, OTU 3 and OTU 15 are only present in low abundance in all 23 samples.

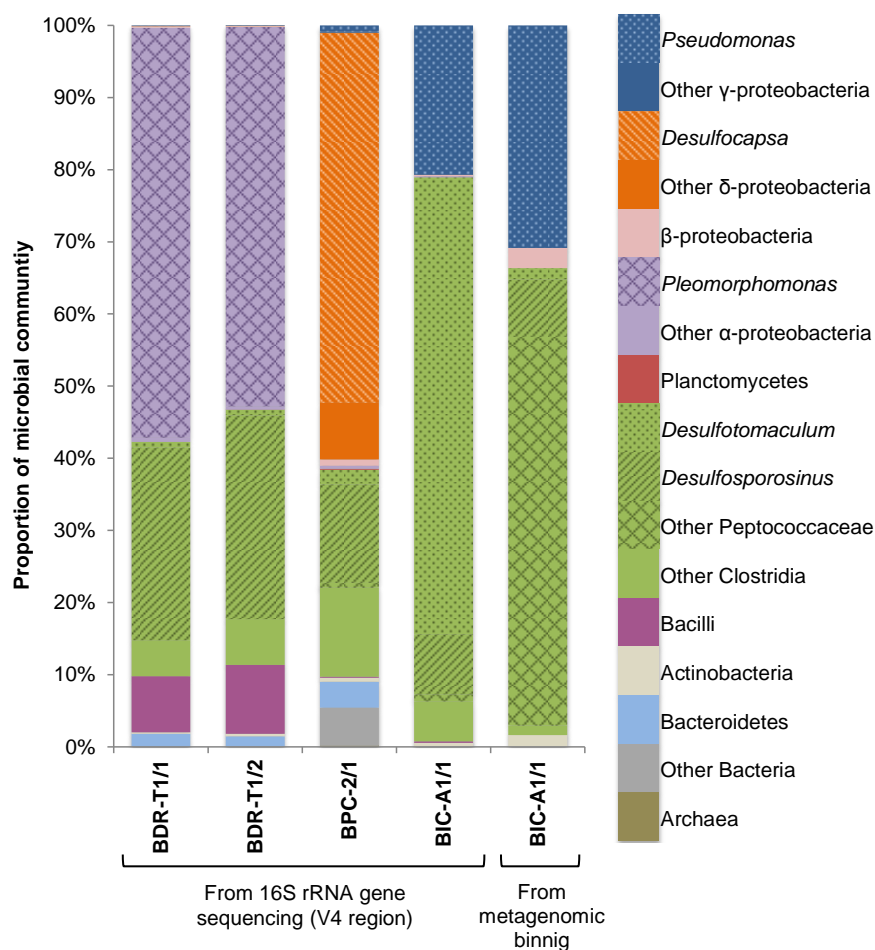


Figure 3.2 Microbial community composition based on 16S rRNA gene sequencing (V4 region) of samples BDR-T1/1, BDR-T1/2, BIC-A1/1 and BPC-2/1, and based on the average contig coverage of the metagenome binning of sample BIC-A1/1.

3.3.3.3 Alpha-diversity

The alpha-diversity varies greatly across sample (Figure A-1.2 and Table A-1.5). True anoxic samples generally have low diversity. For sample BDR-T1, which was inadvertently contaminated with glycerol, exhibits a low diversity community with few microbial species that were strongly selected for their ability to utilize this compound. The low diversity observed in sample BIC-A1/1 can be due to the presence of low molecular weight organic acids (LMWOAs) in Opalinus Clay porewater (Courdouan *et al.*, 2007) in this newly drilled borehole. However, the last anoxic sample, BPC-2/1, harbors one of the greatest diversity observed in this study. This could be due to oligotrophic conditions, as the borehole drains the Passwang rock formation.

3.3.3.4 Sample clustering

Sample clustering according to OTU composition clearly indicates that spatial variations are more significant than temporal variations, because each borehole forms a specific cluster, with the only exception

being sample BRC-3/5 (Figure A-1.3). Samples from BDR-T1, which were contaminated with glycerol, and sample BPC-2/1, coming from another rock formation, are the two that differ the most from the other borehole water samples.

Table 3.2 OTUs shared by the four samples harboring an anaerobic microbial community, including their proportion in each sample, the total number of observations among the original 23 samples, the taxonomic affiliation and the expected metabolism type. OTUs present in all 23 samples are highlighted with grey shading.

OTU #	Proportion in samples				Number of observations	Taxonomic affiliation	Metabolism type	Reference
	BIC-A1/1	BPC-2/1	BDR-T1/1	BDR-T1/2				
6	0.0416	0.0095	0.2667	0.2824	23 / 23	<i>Desulfosporosinus</i>	Sulfate-reducing bacteria	Spring and Rosenzweig, 2006
9	0.5717	0.0184	0.0083	0.0079	20 / 23	<i>Desulfotomaculum</i>	Sulfate-reducing bacteria	Auello <i>et al.</i> , 2013
2	0.2594	0.0056	0.0007	0.0006	23 / 23	<i>Pseudomonas</i>	O ₂ and NO ₃ ⁻ respiring bacteria	Moore <i>et al.</i> , 2006
5	1.57E-05	0.0001	0.5684	0.5258	22 / 23	<i>Pleomorphomonas</i>	O ₂ respiring or fermenting bacteria	Xie and Yokota, 2005
0	0.0001	0.5124	0.0001	0.0001	23 / 23	<i>Desulfocapsa</i>	Sulfate-reducing bacteria	Kuever <i>et al.</i> , 2005
117	0.0001	0.0005	0.0016	0.0026	18 / 23	<i>Actinotalea</i>	O ₂ respiring or fermenting bacteria	Yi <i>et al.</i> , 2007
15	3.92E-05	0.0033	0.0001	2.87E-05	23 / 23	<i>Pseudomonas</i>	O ₂ and NO ₃ ⁻ respiring bacteria	Moore <i>et al.</i> , 2006
1	0.0001	0.0001	0.0001	0.0001	23 / 23	<i>Novispirillum</i>	O ₂ and NO ₃ ⁻ respiring bacteria	Yoon <i>et al.</i> , 2007
3	0.0001	0.0001	4.43E-05	0.0001	23 / 23	Rhodobacteraceae	Metabolically diverse	Garrity <i>et al.</i> , 2005
145	0.0001	3.52E-05	4.43E-05	2.87E-05	18 / 23	<i>Desulfosporosinus</i>	Sulfate-reducing bacteria	Spring and Rosenzweig, 2006
17	7.84E-06	0.0018	1.48E-05	7.19E-06	20 / 23	<i>Thiobacillus</i>	O ₂ and NO ₃ ⁻ respiring bacteria	Kelly <i>et al.</i> , 2005
4	1.57E-05	0.0001	7.38E-06	4.31E-05	22 / 23	Xanthomonadaceae	O ₂ respiring bacteria	Saddler <i>et al.</i> , 2005
8	2.35E-05	1.76E-05	2.21E-05	2.87E-05	21 / 23	Sphingomonadaceae	O ₂ respiring or fermenting bacteria	Yabuuchi and Kosako, 2005
13	7.84E-06	1.76E-05	7.38E-06	2.87E-05	22 / 23	<i>Methyloversatilis</i>	Metabolically diverse	Kalyuzhnaya <i>et al.</i> , 2006



3.3.4 Metagenomic binning

3.3.4.1 Manual correction of binning

Illumina metagenomic sequencing of the BIC-A1 sample produced more than 60 million high quality reads, assembled into 2,246 contigs longer than 5,000 bp (whose N50 length is 45,685 bp). These were binned into 14 genome-clusters using CONCOCT (Alneberg *et al.*, 2014) that combines information on coverage across samples and tetranucleotide composition (Table A-1.6 and Figure 3.3). Two contigs were considered to be contaminating sequences, because they matched *Arabidopsis thaliana* sequences perfectly, and were removed (Table A-1.6). Five of the 14 clusters harbored a 16S SSU rRNA gene (Figure 3.3). One of them, cluster c7, contained multiple and non-identical versions of the 16S rRNA gene. Thus, two contigs from this cluster were reassigned to other clusters (c4 and c14c) as described in Figure 3.3. The inaccurate binning of contigs containing the 16S rRNA gene results from the highly conserved regions in this gene, which leads to incorrect coverage assessments. Also, because they are structural RNA genes, their nucleotide signatures are distinct from other parts of the same genome. Other taxonomic annotations, such as those provided by MLTreeMap and by the IMG pipeline, allowed the manual re-assignment of these inaccurately clustered 16S rRNA gene-bearing contigs to other bins, and enabled the taxonomic annotation of clusters missing a 16S rRNA gene (Table A-1.6, Table A-1.7 and Figure 3.3). Based on this taxonomic annotation along with single-copy gene analysis and the 16S rRNA-derived community profile (Figure 3.2 and Table A-1.3), clusters c10 and c13, both classified as *Streptomyces*, were merged in a newly defined cluster c14c (Figure 3.3). Three clusters were

not considered further: cluster c11, which consists of a single contig, 5733 bp in size, and clusters c0 and c2, both small in size (Figure 3.3) and annotated as *Pseudomonas*. The latter were ignored because only one dominating *Pseudomonas* OTU was detected by 16S rRNA gene sequencing, and it corresponds to cluster c5 (Figure 3.2 and Table A-1.3). We could have merged these three *Pseudomonas* clusters, but this would have created an 8.15 Mbp draft genome, which is significantly larger than the largest known *Pseudomonas* genome. One explanation for the existence of these two small *Pseudomonas* clusters could be that they are constituted by misassembled and chimeric contigs containing some *Pseudomonas* sequences. Thus, after manual correction, 10 clusters were identified (Figure 3.3), suggesting the dominance of 10 microorganisms in the porewater. The final binning is presented in Table A-1.6.

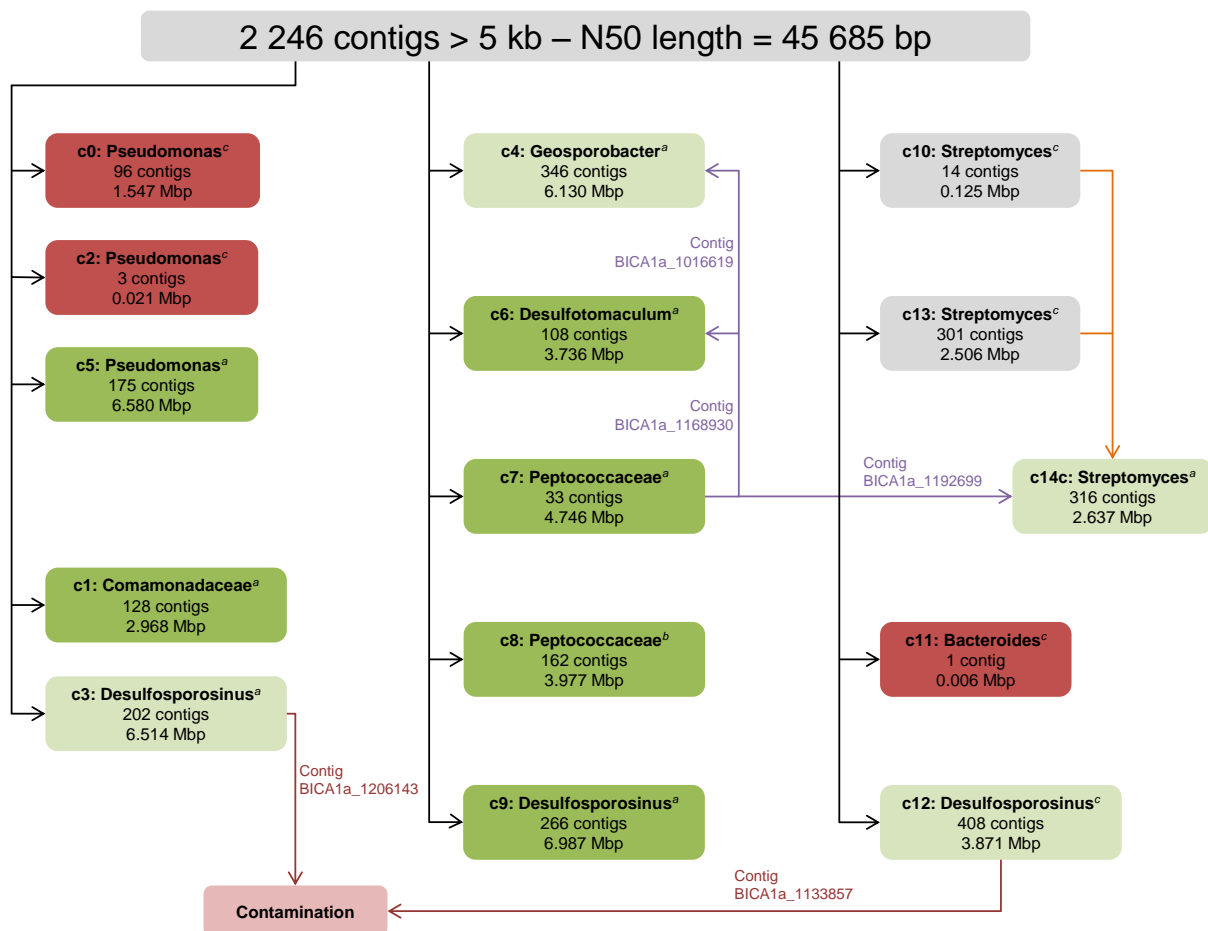








Figure 3.3 Diagram explaining how contigs were clustered. Black arrows indicate clustering carried out by CONCOCT. Later, some clusters were manually merged (orange arrows), some 16S rRNA gene bearing contigs were manually transferred to other or new clusters (purple arrows), or removed (red arrows) because they were considered to be contaminating sequence. Clusters in red boxes were ignored because they are likely due to metagenomic artifacts. Clusters in green boxes were considered as real, but only the ones in dark green boxes were considered as draft genomes. Taxonomic affiliation was based on ^a16S rRNA gene, ^bMLTreeMap or ^cIMG phylogenetic distribution. See Table A-1.7 for more information about taxonomic affiliation.

3.3.4.2 Draft genomes

Out of the ten clusters identified, six clusters were nearly pure and complete (clusters c1, c5, and c6 to c9; Table 3.3 and Table A-1.8) based on the single-copy gene analysis, and thus can be considered draft genomes. They all possess every single-copy gene and have no more than 5 extra copies of single-copy genes. These clusters represent four Peptococcaceae species (among them a *Desulfosporosinus* (c9) and a *Desulfotomaculum* species (c6)), a *Pseudomonas* (c5) and a Comamonadaceae species (c1).

Cluster Peptococcaceae c7 harbors three 16S rRNA-bearing contigs. One of them, which is identical to OTU 7 from 16S rRNA sequencing, was reassigned to cluster c6, based on the matching 16S rRNA genes and metagenome contribution to the microbial community (Table 3.3, Figure 3.2, and Table A-1.3). Cluster c6 could thus be annotated as a *Desulfotomaculum* sp.. The other two 16S rRNA gene sequences of Peptococcaceae c7 are partial and don't span the V4 region covered by the 16S rRNA analysis. However, this cluster can be assigned to OTU 9 because both entities have a similar contribution to the microbial community (Table 3.3 and Table A-1.3). Perfect BLAST alignment of 16S rRNA genes of *Pseudomonas* c5 and OTU 2 indicate that these entities are similar. The overall accuracy of the taxonomic assignment for each cluster is evident from Table A-1.7, in which the majority of the genes in a given cluster correspond to the assigned genus or family. The BIC-A1 microbial community profile obtained here is qualitatively similar to the one obtained with the 16S rRNA sequencing, as shown in Figure 3.2.

Table 3.3 Taxonomic affiliation, size, number of contigs, proportion in the microbial community, average coverage, number of single-copy genes and corresponding OTU from 16S rRNA analysis (if any) for the selected clusters. For single-copy genes, green indicates 1 copy, yellow 2 copies and orange 3 copies of the single-copy gene. For the list of single-copy genes, refer to Table A-1.8.

Clusters	Taxonomic affiliation	cluster length (Mbp)	Number of contigs	Proportion	Coverage	Single copy genes	Corresponding OTU from itag analysis
c1	Comamonadaceae	2.97	128	0.03	52.52x		
c5	<i>Pseudomonas</i>	8.15	274	0.35	691.41x		OTU 2
c6	<i>Desulfotomaculum</i>	3.73	107	0.02	29.99x		OTU 7
c7	Peptococcaceae	4.76	36	0.50	999.06x		OTU 9
c8	Peptococcaceae	3.98	162	0.01	23.76x		
c9	<i>Desulfosporosinus</i>	6.99	266	0.03	60.48x		

3.3.4.1 ANI analysis

The similarity of the six selected clusters to known genomes was calculated using their average nucleotide identity (ANI). Clusters c6 to c9 were compared to all known Peptococcaceae genomes, while clusters c1 and c5 were contrasted with genomes harboring the highest 16S rRNA identity according to BLASTN. The results show that five of the six clusters, represent at minimum a new species. For clusters c6 to c9, all known

Peptococcaceae genomes were used as reference genomes for ANI analysis. Clusters c7 and c9 might also constitute new genera because their highest identities with known Peptococcaceae members were smaller than 85% (Goris *et al.*, 2007). However, clusters c6 and c8 have an ANI higher than 93% with *Desulfosporosinus* sp. OT and *Desulfosporosinus* sp. Tol-M respectively, indicating that they belong to this genus. But because all other known *Desulfosporosinus* genomes were not similar enough to these clusters for measuring ANI (because of insufficient hit numbers), we suggest that both *Desulfosporosinus* sp. OT and Tol-M were wrongly classified as *Desulfosporosinus*. Both genomes were only published as genome announcements and little physiological data is available although one organism (strain Tol-M) is unable to grow with sulfate as a terminal electron acceptor, providing fodder for the argument that perhaps, it should not be included in *Desulfosporosinus* (Abicht *et al.*, 2011; Laban *et al.*, 2015). This means that clusters c6 and c8 can only be classified as Peptococcaceae (Table A-1.9).

Table 3.4 Summary of the pathway annotation of the selected clusters. Numbers indicate the number of occurrences of relevant genes in the corresponding cluster. For a more complete list of annotated pathways, refer to Table A-1.10, Table A-1.11 and Table A-1.12.

Clusters	Taxonomic affiliation	Respiration										Central carbon metabolism			Hydrogenases				Fermentation						Carbon oxidation by SRB												
		Disimilatory sulfate reduction	Disimilatory nitrate reduction	Denitrification	electron transferring flavoproteins	electron transferring quinones	electron transferring [Fe-S] proteins (excluding hydrogenases)	electron transferring cytochromes	ATP synthase	Glycolysis (EMP pathway)	Entner-Doudoroff pathway	Citrate cycle (TCA cycle)	[NiFe] group 1	[NiFe] group 3a	[NiFe] group 3b	[NiFe] group 4	[FeFe]	succinate production	formate production	ethanol production	butanol production	butyrate production	acetate production	lactate production	propionate production	acetoin production	butanediol production	modified TCA cycle	-Acetyl-CoA pathway	Key enz.							
c1	Comamonadaceae	18	24	7	7							0	1	2	0	2																					
c5	<i>Pseudomonas</i>	31	42	24	23							0	1	0	0	2																					
c6	<i>Desulfotomaculum</i>	14	21	38	5							1	8	0	0	7																					
c7	Peptococcaceae	22	22	41	8							1	9	0	1	7																					yes
c8	Peptococcaceae	25	20	34	10							3	5	0	0	17																					yes
c9	<i>Desulfosporosinus</i>	36	38	50	15							4	6	0	3	33																					yes

For *Pseudomonas* c5, reference genomes were also selected based on BLASTN of the complete 16S rRNA gene. Organisms showing a similarity of 99% were selected (*Pseudomonas chloritidismutans* AW-1, *Pseudomonas xanthomarina* S11 and *Pseudomonas knackmussii* B13), plus two *Pseudomonas stutzeri* that were the second and the third best alignment. The *Pseudomonas* c5 could not be inferred by ANI analysis, because it is matching with a similarity greater than 97% two different genomes: *Pseudomonas chloritidismutans* and *Pseudomonas xanthomarina* (Kim *et al.*, 2014; Table A-1.9).

3.3.4.2 Annotation of dominating sulfate reducing bacteria

Because the dominant cluster (c7) is phylogenetically close to the genus *Desulfotomaculum* (Table A-1.7) and because several species of this genus have been sequenced, its genome was compared to that of *Desulfotomaculum acetoxidans*, a sulfate reducing bacteria (SRB) capable of complete oxidation but unable to ferment, and *Desulfotomaculum reducens*, a fermenting SRB for which the mechanism of sulfate reduction

was studied in some detail. Overall, the genes involved in sulfate reduction, hydrogen oxidation and complete oxidation of acetate are remarkably conserved both in terms of identity and synteny (Table 3.5).

Table 3.5 Genome comparison of Peptococcaceae *c7* with two reference genomes: *Desulfotomaculum acetoxidans* for the reductive acetyl-CoA pathway and hydrogenases involved in respiratory processes, and *Desulfotomaculum reducens* for genes involved in sulfate reduction. Grey boxes represent putative gene clusters in Peptococcaceae *c7*. The comparison was carried out using BLASTN. All e-values are significant ($<4 \times 10^{-20}$).

Ref. genome	Metabolic function	Ref. gene ID	<i>Desulfotomaculum c7</i>		BLAST alignment			
			Gene ID	IMG gene product annotation	Identity (%)	Ref. gene coverage	c7 gene coverage	
Comparison to <i>Desulfotomaculum reducens</i> M1-1 genome	NADH oxidation, H ⁺ translocation, quinone reduction	Dred_2036	BICA1a_1184301213	NADH:ubiquinone oxidoreductase subunit 2 (chain N)	64.36	0.997	0.997	
		Dred_2037	BICA1a_1184301212	NADH:ubiquinone oxidoreductase subunit 4 (chain M)	72.48	0.989	0.985	
		Dred_2038	BICA1a_1184301211	NADH:ubiquinone oxidoreductase subunit 5 (chain L)	71.1	0.996	0.997	
		Dred_2039	BICA1a_1184301210	NADH:ubiquinone oxidoreductase subunit 11 or 4L	72.16	0.909	0.892	
		Dred_2040	BICA1a_1184301209	NADH:ubiquinone oxidoreductase subunit 6 (chain J)	47.9	0.961	0.929	
		Dred_2041	BICA1a_1184301208	NADH:ubiquinone oxidoreductase 23 kD subunit (chain I)	48.36	0.730	0.772	
		Dred_2042	BICA1a_1184301207	NADH:ubiquinone oxidoreductase subunit 1 (chain H)	67.24	0.992	0.984	
		Dred_2043	BICA1a_1184301206	NADH:ubiquinone oxidoreductase 49 kD subunit 7	73.35	0.990	0.995	
		Dred_2044	BICA1a_1184301205	NADH:ubiquinone oxidoreductase 27 kD subunit	51.72	0.948	0.937	
		Dred_2045	BICA1a_1184301204	NADH:ubiquinone oxidoreductase 20 kD subunit	76.05	0.844	0.924	
	Dred_2046	BICA1a_1184301203	NADH:ubiquinone oxidoreductase subunit 3 (chain A)	64.96	0.986	0.978		
	e ⁻ transfer from quinone pool to DsrAB	Dred_3197	BICA1a_102506429	Dissimilatory sulfite reductase, gamma subunit	73.33	0.984	0.984	
		Dred_3198	BICA1a_102506430	Fe-S oxidoreductase	64.75	0.968	0.979	
		Dred_3199	BICA1a_102506431	Nitrate reductase gamma subunit	54.24	0.990	0.994	
	Sulfite reduction to sulfide	Dred_3186	BICA1a_102506439	Dissimilatory sulfite reductase, alpha and beta subunits	58.5	0.981	0.898	
		Dred_3187	BICA1a_102506440	Dissimilatory sulfite reductase, alpha and beta subunits	54.55	0.973	0.874	
	Sulfate activation and reduction to sulfite	Dred_0635	BICA1a_118430176	ATP sulfonylase (sulfate adenylyltransferase)	87.86	0.991	0.986	
		Dred_0636	BICA1a_118430177	Adenosine-5'-phosphosulfate reductase beta subunit	75.54	0.935	0.941	
		Dred_0637	BICA1a_118430178	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	75.12	0.997	0.994	
		Dred_0638	BICA1a_118430179	No gene product annotation	66.59	0.996	0.991	
	Dred_0639	BICA1a_118430180	Methyl-viologen-reducing hydrogenase, delta subunit	66.35	0.998	0.998		
	Qmo complex reduction by heterodisulfide reductase	Dred_1325	BICA1a_1172615263	No gene product annotation	50.3	0.869	0.855	
		Dred_1326	BICA1a_1172615264	Heterodisulfide reductase, subunit B	50.18	0.987	0.980	
		Dred_1327	BICA1a_1172615269	NADPH-dependent glutamate synthase beta chain	57.48	0.993	0.991	
		Dred_1328	BICA1a_1172615265	Pyridine nucleotide-disulphide oxidoreductase	53.55	0.988	0.991	
		Dred_1329	BICA1a_1172615275	Coenzyme F420-reducing hydrogenase, delta subunit	62.5	0.418	0.959	
		Dred_1329	BICA1a_1172615268	Coenzyme F420-reducing hydrogenase, delta subunit	42.27	0.637	0.837	
Qmo complex reduction by flavoproteins	Dred_1330	BICA1a_1172615273	Coenzyme F420-reducing hydrogenase, beta subunit	25.49	0.780	0.784		
	Dred_1778	BICA1a_1075448205	Electron transfer flavoprotein, beta subunit	61.09	0.993	0.994		
Dred_1779	BICA1a_1075448204	Electron transfer flavoprotein, alpha subunit	69.33	0.995	0.995			
H ⁺ translocation coupled to PPI hydrolysis	Dred_2985	BICA1a_118430132	Inorganic pyrophosphatase	79.61	0.995	0.972		
Comparison to <i>Desulfotomaculum acetoxidans</i> DSM 771 genome	[NiFe]-hydrogenase	Dtox_0791	BICA1a_120211075	Formate hydrogenlyase subunit 3	68.68	0.996	0.997	
		Dtox_0792	BICA1a_120211074	Formate hydrogenlyase subunit 4	74.74	0.948	0.948	
		Dtox_0793	BICA1a_120211073	Hydrogenase 4 membrane component (E)	63.03	0.974	0.983	
		Dtox_0794	BICA1a_120211072	Formate hydrogenlyase subunit 3	72.41	0.982	0.980	
		Dtox_0795	BICA1a_120211071	Ni,Fe-hydrogenase III large subunit	64.67	0.984	0.984	
		Dtox_0796	BICA1a_120211070	Ni,Fe-hydrogenase III small subunit	74.58	0.993	0.993	
	[FeFe]-hydrogenase	Dtox_0168	BICA1a_1094883155	Iron only hydrogenase large subunit, C-terminal domain	75.82	0.991	0.983	
		Dtox_0169	BICA1a_1094883154	Fe-S-cluster-containing hydrogenase components 1	65.77	0.986	0.990	
		Dtox_0172	BICA1a_1094883150	Iron only hydrogenase large subunit, C-terminal domain	79.37	0.995	0.995	
		Dtox_0173	BICA1a_1094883149	NADH:ubiquinone oxidoreductase, NADH-binding (51 kD) subunit	80.74	0.997	0.997	
		Dtox_0174	BICA1a_1094883148	Ferredoxin	53.91	0.953	0.969	
		Dtox_0175	BICA1a_1094883147	Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase	53.89	0.964	0.929	
		Dtox_0176	BICA1a_1094883146	NADH:ubiquinone oxidoreductase 24 kD subunit	60	0.990	0.990	
		Dtox_0177	BICA1a_1094883145	DRTGG domain	64.86	0.943	0.959	
		Dtox_0178	BICA1a_1094883144	Iron only hydrogenase large subunit, C-terminal domain	72.58	0.974	0.969	
		Acetate oxidation to C ₄ (acetyl-CoA pathway)	Dtox_1285	BICA1a_1172615266	5,10-methylenetetrahydrofolate reductase	59.68	0.979	0.963
			Dtox_1276	BICA1a_1172615277	Pterin binding enzyme	72.83	0.994	0.994
	Dtox_1275		BICA1a_1172615278	CO dehydrogenase/acetyl-CoA synthase delta subunit	69.62	0.995	0.995	
	Dtox_1272		BICA1a_1172615280	CO dehydrogenase/acetyl-CoA synthase gamma subunit	68.31	0.996	0.996	
	Dtox_1271		BICA1a_1172615281	CO dehydrogenase/acetyl-CoA synthase beta subunit	80.22	0.998	0.998	
	Dtox_1270		BICA1a_1172615282	6Fe-6S prismatic cluster-containing protein	61.52	0.998	0.998	
	CO ₂ reduction to formate	Dtox_3503	BICA1a_1089845305	Uncharacterized anaerobic dehydrogenase	54.56	0.580	0.995	
		Dtox_3507	BICA1a_1089845306	Uncharacterized anaerobic dehydrogenase	68.18	0.392	0.992	
		H ₄ F + formate → formyl-H ₄ F	Dtox_0231	BICA1a_109488329	Formyltetrahydrofolate synthetase	44.81	0.986	0.982
	formyl-H ₄ F → methylene-H ₄ F	Dtox_2589	BICA1a_1172615153	5,10-methylene-tetrahydrofolate dehydrogenase	42.76	0.980	0.987	

3.4 Discussion

3.4.1 Issue with line sampling

A few of the samples collected exhibited chemistry consistent with the absence of oxygen but harbored a microbial community dominated by strict aerobes. These are BHT-1/3 and BWSA-2/1, which include 39.6% and 22% of strict aerobes respectively (Table 3.1 and Table A-1.4). The discrepancy between the sample's chemistry and its microbial community is likely attributable to the phylogenetic signature from biofilm growing inside the lines and sloughed off by pumping. This biomass developed by oxidizing LMWOAs contained in Opalinus Clay porewater using oxygen as a terminal electron acceptor. These biofilms can develop when O₂ diffusing in from the gallery comes into contact with borehole water in the line. This hypothesis is supported by the numerous observations of *Pseudomonas* and Xanthomonadaceae, which are well known for their abilities to form biofilms (Saddler and Bradbury, 2005; Moore *et al.*, 2006). Our efforts to flush these pipes, discarding the first part of the available borehole water, were not sufficient to fully dislodge this biomass. These findings suggest that great care must be taken when sampling borehole water for microorganisms at the Mont Terri URL. Sampling from lines should be avoided if possible. Alternatively, they must be flushed for long periods of time to fully remove the biofilm prior to sampling.

3.4.2 Anaerobic microbial communities

Nonetheless, based on the four samples of trusted quality as well as a number of the oxidized samples, it was possible to identify a subset of microorganisms present in every sample (Table 3.2). The metabolic potential of these common microbes includes fermentation (e.g., Rhodobacteraceae and *Pseudomonas*), denitrification (e.g., Rhodobacteraceae, *Novispirillum*, *Pseudomonas*), sulfate reduction and inorganic sulfur disproportionation (e.g., *Desulfocapsa* and *Desulfotomaculum*; Garrity *et al.*, 2005a; Moore *et al.*, 2006; Yoon *et al.*, 2007; Kuever *et al.*, 2005b; Widdel, 2006). It is remarkable that this subset of microbes is consistently detectable in all the boreholes probed independent of drilling conditions, source rock, geochemical conditions, flow rate, sampling procedure, and microbial contamination sources. Such uniformity suggests a core microbial community in the Mont Terri URL that is present across the Opalinus Clay rock, but that may develop into distinct communities in response to disturbances. The variable overall microbial community obtained at each borehole may reflect the adaptation of the core community to the specific conditions at that location.

The microbial OTUs identified in consistently anoxic samples BDR-T1/1, BDR-T1/2, BIC-A1/1 and BPC-2/1 mainly comprise sulfate-reducing bacteria and facultative anaerobes that can ferment or denitrify (Figure 3.2; Table 3.2). Samples from the BDR-T1 borehole, which was inadvertently contaminated with glycerol, are dominated by a *Pleomorphomonas* species. This nitrogen-fixing genus usually respire oxygen, but some

members are known to ferment glucose, producing organic acids (Xie and Yokota, 2005). Glycerol could be metabolized to pyruvate through glycolysis prior to fermentation, as was shown for *Klebsiella pneumonia* (Wang *et al.*, 2003). Organic acids, the products of this putative fermentation, could be oxidized by *Desulfosporosinus* sp., an incomplete oxidizer, acetate-producing, sulfate-reducing bacterium (Spring and Rosenzweig, 2006) also identified in this sample. Another scenario could be the direct oxidation of glycerol by the *Desulfosporosinus* species, as it was reported that some members of this genus only grow on this carbon source (Ramamoorthy *et al.*, 2006b).

The microbial energy source of sample BPC-2/1 is unknown, because the water originates from the adjacent Passwang rock formation (Figure 3.1), which is not as well characterized as Opalinus Clay porewater. No hydrogen sulfide was detected in this sample. Nonetheless, the dominating microbial process is thought to be sulfate reduction, according to the high proportion of *Desulfocapsa* sp. in the sample. Hydrogen sulfide may not be detectable due to its precipitation with Fe(II) that is present at a concentration of about 100 µM. Sample BPC-2/1 is the only sample in this study that harbors an important proportion of Gram-negative SRB (*Desulfocapsa* sp.). All other samples mainly contain *Desulfosporosinus* and *Desulfotomaculum* spp. that are classified as Gram-positive SRB (Figure 3.2).

The BIC-A1 borehole received no amendment and is only different from undisturbed Opalinus Clay in that more space is available. Hence, it is an accurate model of fractured Opalinus Clay. The availability of space alone was sufficient to support microbial growth in this rock. Evidence of microbial growth includes the presence of soluble hydrogen sulfide (Table 3.1), which is ordinarily absent from pristine porewater (Pearson *et al.*, 2003), and results from microbial sulfate reduction fueled by organic compounds naturally present in this environment (Pearson *et al.*, 2003; Courdouan *et al.*, 2007). Further, the pH, which is higher than typical Opalinus Clay pH (Pearson *et al.*, 2003), is also indicative for a sulfate reduction activity, as shown by the following equation:



Finally, the absence of PCR inhibitors in Opalinus Clay rock combined with the inability to extract genomic DNA from the rock (as opposed to the porewater; Poulain *et al.*, 2008), indicate that microbial density was increased in porewater relative to the rock, reflecting a higher microbial activity when space is available.

3.4.3 Metagenomic binning of BIC-A1 borehole

Metagenomic sequencing and binning enabled the identification of the metabolic capabilities of six microorganisms (table 1): one Comamonadaceae (c1), one *Pseudomonas* (c5), and four Peptococcaceae (c6 to c9), together representing 93% of the microbial community. Thus, this community is of remarkably low diversity. This low diversity highlights the challenges of microbial survival in this oligotrophic environment.

It is particularly striking in comparison to the microbial community from acid mine drainage, which includes more than 100 genera (Chen *et al.*, 2014).

We can place these organisms into two groups according to their metabolic capabilities. Clusters c1 and c5 are capable of respiration using nitrate as the final electron acceptor, reducing it either to ammonium or dinitrogen (Table 3.4). This respiration can be coupled to the oxidation of complex organic compounds to carbon dioxide through the Embden–Meyerhof–Parnas (EMP) or the Entner-Doudoroff (ED) pathways. However, no nitrate was detected in the BIC-A1 borehole (Table A-1.2), suggesting that an alternative metabolism is underway. We propose that the organisms corresponding to these two clusters ferment rather than respire. Even though the great majority of *Pseudomonas* species are not fermenters, some are known to ferment pyruvate, producing acetate (Moore *et al.*, 2006; Schreiber *et al.*, 2006). For example, it was shown that *Pseudomonas aeruginosa* is able to live using this pathway after oxygen and nitrate are depleted (Eschbach *et al.*, 2004).

The second group of organisms (c6, c7, c8 and c9), all belonging to Peptococcaceae family, are similar with respect to their metabolic abilities (Table 3.4). Based on their individual genome sequences, they are all able to use sulfate as a terminal electron acceptor, have an elaborate electron transfer chain and include ATP synthase. All sulfate-reducing bacteria identified in this sample are complete oxidizers, meaning that they include the complete set of enzymes required to oxidize acetate to carbon dioxide, including the catalytic subunit of carbon monoxide dehydrogenase. In addition to sulfate reduction, the organisms harboring these genomes appear to ferment organic carbon, through glycolysis, producing succinate, formate, ethanol, butyrate, butanol and acetate. The four sulfate-reducing clusters all have genes coding for group 1 [NiFe]-hydrogenases, suggesting that they can use dihydrogen as a source of electrons for respiration. As they all possess the enzymes for the acetyl-CoA pathway, which is reversible, they are predicted to be autotrophic in the absence of organic carbon and in the presence of dihydrogen.

3.4.3.1 *Pseudomonas* c5

Here we will further describe the metabolic potential of the two dominant clusters (which, when combined, correspond to 85% of the microbial community), as they are representative of the two groups of organisms identified.

The *Pseudomonas* genome (c5), which represent 35% of the microbial community of sample BIC-A1 (Table 3.3), includes genes encoding for proteins involved in extracellular phosphate hydrolysis (Table A-1.12): phytases, specific and non-specific organic phosphatases (Rodríguez and Fraga, 1999), polyphosphatase, inorganic pyrophosphatases (Gobat *et al.*, 2004) and the machinery to produce gluconic acid that solubilizes inorganic phosphate from phosphate-bearing minerals in Opalinus Clay (apatite or monazite; Pearson *et al.*, 2003). Extracellular degradation of complex organic matter is also part of the capabilities in this cluster (Table A-1.12): genes encoding proteins for lipid degradation (ABC transport, chemotaxis and

esterase/lipase), protein degradation (membrane-bound, secreted, ABC transport), nucleic acid degradation (others than restriction nuclease, ribonucleases, polymerases and repair nucleases), amino-acid mineralization (asparaginases and glutaminases), urea mineralization (Gobat *et al.*, 2004), and starch and amorphous cellulose degradation (Kim and Gadd, 2008) are all identifiable in the cluster. Figure 3.4.A shows the putative metabolic pathways of this cluster (c5) in the BIC-A1 porewater. From these observations, it appears that this organism can use complex organic carbon compounds that are broken down extracellularly prior to transport inside the cell and catabolized through diverse fermentation pathways. *Pseudomonas* c5 can generate dihydrogen gas using [FeFe]-hydrogenases or by group 3a [NiFe]-hydrogenases, or it can produce lactate, ethanol, butanol or succinate through fermentation.

3.4.3.1 Peptococcaceae c7

The dominant cluster in the Peptococcaceae family is cluster c7, which represents 50% of the microbial community (Table 3.3). As the other clusters belonging to this family, it represents a SRB capable to complete oxidize organic carbon, but also, in presence of H₂ capable of autotrophic growth.

If a carbon source is available, Peptococcaceae c7 can generate reducing equivalents from the oxidation of acetate, formate, lactate, ethanol, butanol, butyrate and propionate, using reverse fermentation pathways (Figure 3.4.B). Interestingly, the propionate oxidation pathway is only found in sulfate-reducing bacteria in this study, but all of them are missing methylmalonyl-CoA carboxytransferase (EC 2.1.3.1), which can generate pyruvate and methylmalonyl-CoA from oxaloacetate and propionyl-CoA (Figure 3.5.A). This metabolism appears to be catalyzed instead by propionyl-CoA carboxylase (EC 6.4.1.3) for which both clusters c6 and c7 encode a gene. This alternate reaction can transform propionyl-CoA to methylmalonyl-CoA and pyruvate can be produced from oxaloacetate by an oxaloacetate decarboxylase (EC 4.1.1.3; Figure 3.5.B). Additionally, *Desulfotomaculum* c7 is able to catabolize glucose via the EMP pathway followed by the oxidative acetyl-coA pathway for complete oxidation of the acetyl-coA produced (Figure 3.4.B).

If H₂ is available, Peptococcaceae c7 can use it for ATP generation by reducing SO₄²⁻, and can use the reductive acetyl-CoA to produce biomass from CO₂. As shown in Figure 3.4.B, the c7 genome possess genes coding for enzyme able to reduce ferredoxin and NADP⁺ from H₂, which will in a second step transfer electrons to the enzymes of the reductive acetyl-CoA.

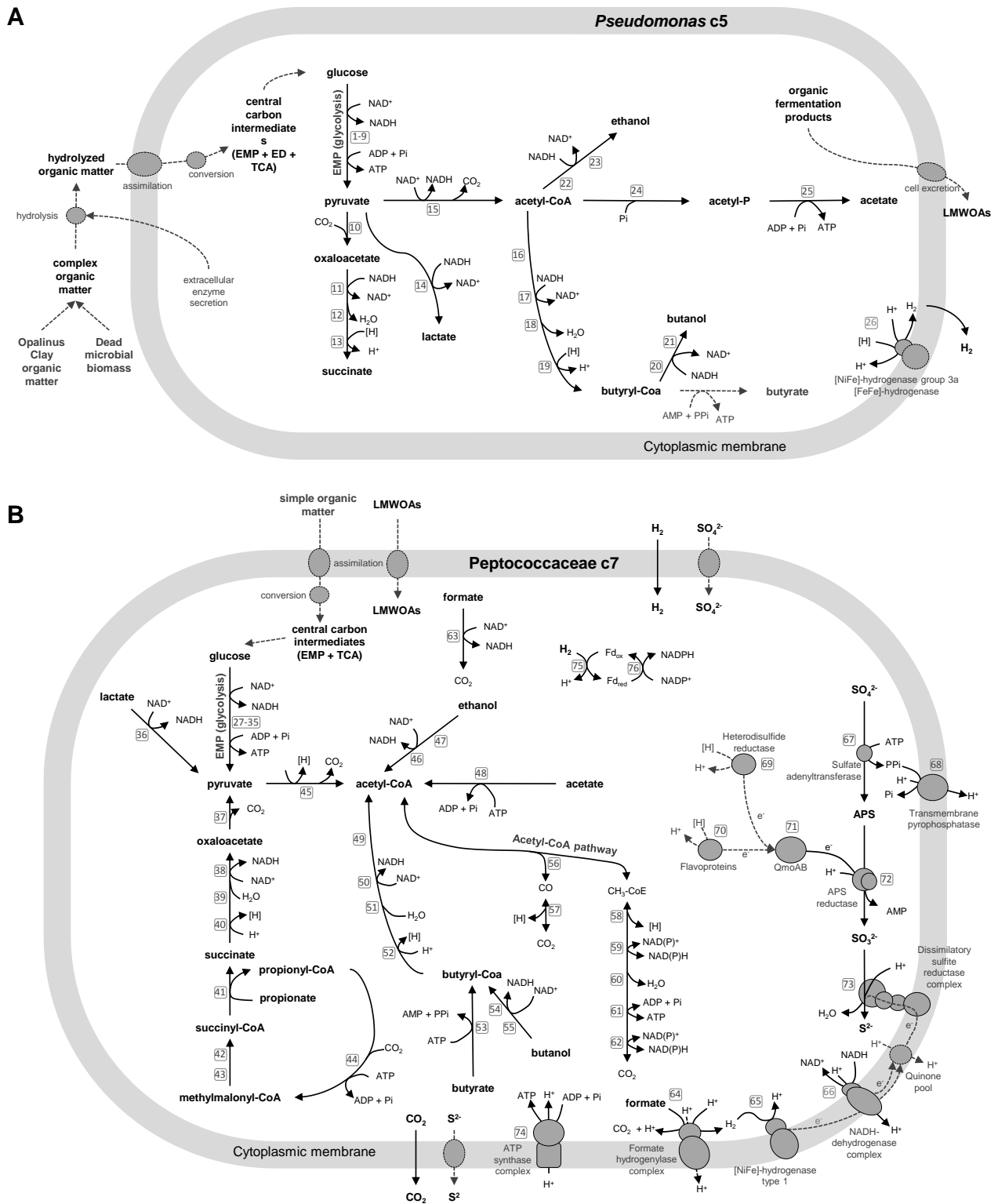


Figure 3.4 Putative energetic metabolism of (A) *Pseudomonas* cluster c5 and (B) Peptococcaceae cluster c7. All dashed elements represent biological processes that have not been formally identified in these two metagenome clusters. For more information concerning genes coding for proteins involved in each of these reactions, refer to Table A-1.13 (for *Pseudomonas* c5) and Table A-1.14 (for Peptococcaceae c7). (A) *Pseudomonas* c5 is thought to hydrolyze complex organic matter extracellularly, and to metabolize the products through fermentation, producing acetate, lactate, ethanol, succinate, and dihydrogen gas. (B)

Peptococcaceae cluster c7 is thought to oxidize low-molecular weight organic acids (LMWOAs) such as formate, acetate, lactate, propionate and butyrate, and alcohols such as ethanol and butanol all the way to CO₂. Also, H₂ oxidation is coupled to sulfate reduction to produce extra ATP. Finally, if the H₂ are high enough, the Peptococcaceae can be autotrophic, using the reductive acetyl-CoA pathway to fix CO₂.

As indicated above, acetyl-CoA is a key pathway for, Peptococcaceae c7, because it can use it in an oxidative or reductive way depending on the conditions. A BLAST comparative analysis to the genome of *Desulfotomaculum acetoxidans* (Spring *et al.*, 2009) showed that genes coding for enzymes of this pathway are strongly conserved at the gene cluster level (Table 3.5). This genome comparison also revealed that genes coding for two membrane-bound hydrogenase complexes (a [NiFe]- and an [FeFe]-hydrogenase) were similar between these organisms (Table 3.5). The [NiFe]-hydrogenase belongs to group 3 hydrogenases, and is part of a membrane-bound complex that also contains formate hydrogenlyase (Bagramyan and Trchounian, 2003). Therefore, it is likely a hydrogenase that transfers reducing equivalents from formate to dihydrogen, which subsequently is oxidized by a membrane-bound [NiFe]-hydrogenase type 1 (Figure 3.4.B).

The machinery for reducing sulfate was compared to *Desulfotomaculum reducens* (Junier *et al.*, 2010) and all its constituents were strongly conserved at the gene cluster level (Table 3.5). Adenosine triphosphate (ATP) consumed by sulfate adenylyltransferase (EC 2.7.7.4) for sulfate activation can be partially recovered through a transmembrane pyrophosphatase that translocates a proton while hydrolyzing PPi (Figure 3.4.B). Unlike *Desulfotomaculum reducens*, Peptococcaceae c7 does not possess a membrane-bound adenylyl-sulfate reductase (EC 1.8.99.2), but a cytoplasmic one. Before reducing adenylyl-sulfate, electrons coming from acetate oxidation are transferred to APS reductase, through flavoproteins and/or heterodisulfide reductase, and then through the QmoAB complex, without transiting into the cytoplasmic membrane (Figure 3.4.B). Similarly to *Desulfotomaculum reducens* and in contrast to Gram-negative SRB, no membrane-bound Qmo subunit could be identified. Electrons coming from dihydrogen, formate and NADH oxidation are transferred in the cytoplasmic membrane, via the quinone pool, generating a proton motive force, and are used by dissimilatory sulfite reductase (EC 1.8.99.3).

Thus, Peptococcaceae c7 is able to use sulfate as a terminal electron acceptor, producing hydrogen sulfide. It oxidizes fermentation products as well as monomers of others organic molecules to carbon dioxide, and is capable of autotrophic grow when carbon dioxide and H₂ are available. The genome annotation of Peptococcaceae c7 clearly shows that key processes for energy conservation are highly conserved among Gram-positive SRB, because they were consistent with the ones described *Desulfotomaculum reducens* (Junier *et al.*, 2010) and *Desulfotomaculum acetoxidans* (Spring *et al.*, 2009), as shown in Table 3.5. It is remarkable that although Peptococcaceae c7 is not similar to these two others Peptococcaceae members according to ANI (Table A-1.9), genes involved in sulfate reduction, hydrogen oxidation and the acetyl-CoA pathway are almost perfectly conserved at the gene cluster level among these three organisms. This suggests that the genes coding

for these metabolic capabilities are orthologous and were all present in the common ancestor of these organisms.

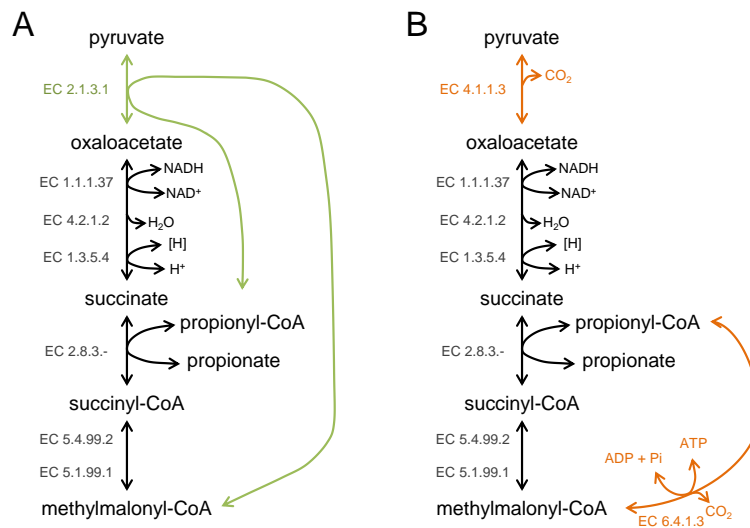


Figure 3.5 Propionate fermentation pathway, (A) as described by Stams *et al.*, (1984), and (B) as depicted in this study in Peptococcaceae clusters 6 and 7. The common reactions shared by both variants are displayed in black. In the genomes of Peptococcaceae clusters 6 and 7, gene coding for methylmalonyl-CoA carboxytransferase (EC 2.1.3.1) is missing. However, this enzyme can be replaced by propionyl-CoA carboxylase (EC 6.4.1.3) and oxaloacetate decarboxylase (EC 4.1.1.3), whose coding genes were identified in these two genomes.

3.4.3.2 Metabolic interactions

From the annotations of the *Pseudomonas* c5 and Peptococcaceae c7 clusters, we reconstructed potential interactions between these two species in the BIC-A1 borehole. Depending on the conditions, two different model can be depicted.

The first one corresponds to an absence of H₂ and is shown in Figure 3.6.A. As shown previously (Courdouan *et al.*, 2007), acetate is an organic acid commonly found in Opalinus Clay porewater, at concentrations that can reach 200 μM. Acetate can serve as a carbon and energy source for the Peptococcaceae c7 for sulfate reduction to hydrogen sulfide. After the death of Peptococcaceae c7 cells, complex organic matter from the dead biomass is available to the *Pseudomonas* cluster for energy conservation through fermentative pathways, likely producing acetate and dihydrogen, as these products generate the highest yield of ATP. In turn, these compounds can serve as electron donors and energy sources by the Peptococcaceae cluster c7. However, this putative carbon loop is not sustainable if the only energy source is acetate found in the Opalinus Clay porewater. Over time, after each cycle, because energy conversion efficiency in heterotrophic bacteria is typically 60% (Calow, 1977), acetate concentration would continuously decrease.

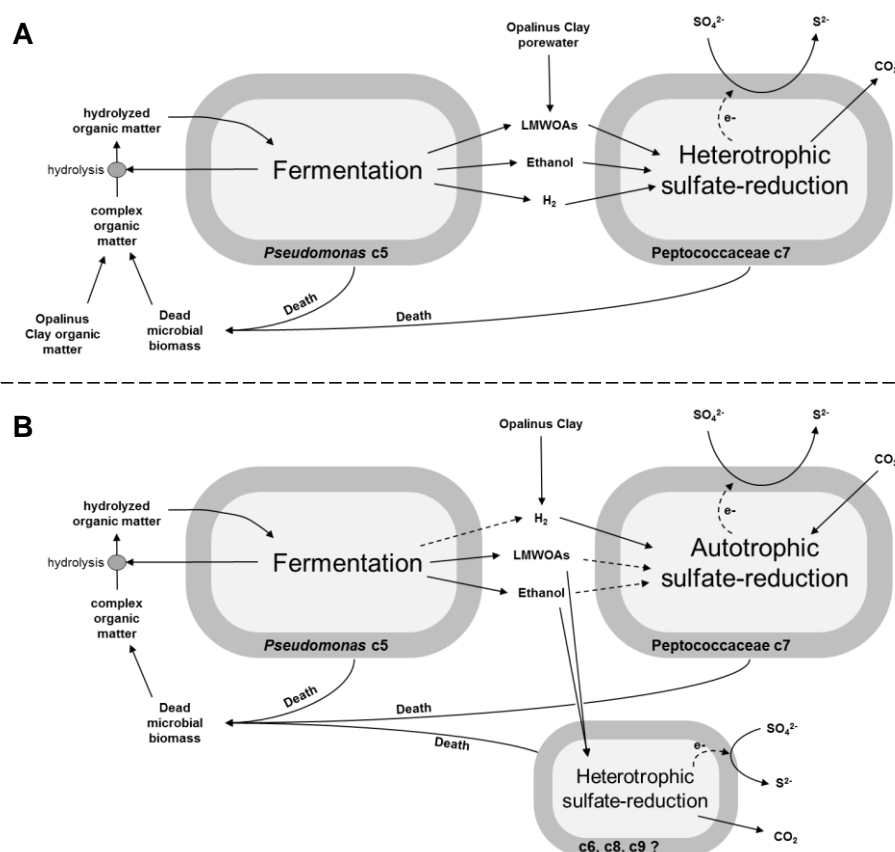


Figure 3.6 Putative food webs between *Pseudomonas* c5 and Peptococcaceae c7, when the main energy source is (A) organic matter contained in Opalinus Clay, or (B) H_2 produced *in situ* by an unknown process.

Another potential energy source that can feed this carbon loop and render it sustainable would be the buried organic matter contained in Opalinus Clay. The question is whether *Pseudomonas* c5 is able to utilize this organic carbon. Organic matter in Opalinus Clay is thermally immature and largely composed of humic and fulvic acids, while other organic compounds such as kerogen type III (humic type) and n-alkanes C_{20} - C_{33} were also identified (Jones and Tipping, 1998). *Pseudomonas* species are known to utilize this type of organic carbon for respiration (Tikhonov *et al.*, 2010; De Haan, 1974) but it remains unclear whether these carbon sources can support fermentation.

The second model is based on the presence of H_2 in Opalinus Clay. A study characterizing the Opalinus Clay gas phase measured hydrogen in several boreholes 2 years after their closure. However, in subsequent samplings, 1 or 2 years later, no H_2 was detected. Thus, the authors, unable to explain the origin of H_2 in the boreholes, concluded that this compound was not part of the natural Opalinus Clay gas phase (Pearson *et al.*, 2003). This means that an unknown process in Opalinus Clay could sustain microbial activity through production of H_2 . In this case, the microbial system would be organized differently. Primary production could be carried out by the dominating Peptococcaceae c7, transferring the electrons from H_2 to SO_4^{2-} and producing energy. This energy would be invested to reduce CO_2 to biomass, via the reductive acetyl-CoA pathway, using

the same electron source. The complex organic matter can later be fermented by *Pseudomonas* c5, which produces LMWOAs and ethanol. In turn, these compounds could serve as carbon and energy sources for an unidentified heterotrophic SRB, which could correspond to cluster c6, c8 or c9. This ecosystem model is present in Figure 3.6.B.

3.4.4 Outcome for the safety of Opalinus Clay repositories

The microbial metabolisms described above will likely occur in Opalinus Clay nuclear waste repositories and they should be considered when assessing the long-term safety of such an undertaking. For instance, SRB can produce CO₂ from organic carbon oxidation that is, depending where it is produced in a geological repository, likely to dissolve in the porewater, precipitate while reacting with the cement or less likely, diffuse as a gas phase. Sulfide production by SRB could also increase the rate of anoxic steel corrosion of the drums containing waste. Also, higher microbial activity could lead to local rock weathering, because active organisms may gather nutrients from the rock itself by secreting protons, organic acids and siderophores that will complex metal ions (Uroz *et al.*, 2009). On the other hand, beneficial impacts of microbial activity include the consumption of H₂ gas that will be produced in the repository through anoxic steel corrosion.

3.5 Conclusion

Despite technical limitations that led to sampling artifacts, this survey revealed that a few microbial OTUs were observed in every sample, suggesting they originate from unperturbed Opalinus Clay rock. This study clearly showed that, by providing space in the rock, microorganisms were able to thrive with no additional amendment, indicating that the pore size distribution in Opalinus Clay is the main limiting factor for microbial growth. Once space is available, the organic compounds and sulfate found in the rock can sustain microbial life, albeit of limited diversity. According to reconstructed genomes, pathway annotation, and observed environmental parameters, Gram-positive, complete-oxidizer sulfate-reducing bacteria are able to grow using organic compounds present in Opalinus Clay porewater (e.g., acetate) and/or H₂ to respire sulfate. Gram-negative bacteria breakdown and ferment dead biomass from the SRB and potentially the Opalinus Clay organic matter to produce low-molecular weight organic acids that further sustain the sulfate-reducing bacteria. The reconstruction of a food chain dominated by two organisms and able to thrive in drilled but otherwise undisturbed rock represents a significant step in furthering our understanding of microbial processes in the subsurface.

3.6 Acknowledgements

The work conducted by the U.S. Department of Energy Joint Genome Institute (JGI) is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. JGI is acknowledged for providing resources for the sequencing analysis through the grant CSP 1505. A.F.A. is supported by a grant from the Swedish Research Council VR (grant 2011-5689). Computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC). The Swisstopo team of the Mont Terri Rock Laboratory provided the best possible conditions for scientific research in the URL and we are deeply indebted to them for their help. We thank Manon Frutschi for her invaluable help during sampling. Elena Rossel, Karin Vernez Thomas, Sylvain Coudret and Pierre Rossi are acknowledged for their help for laboratory analyses.

Chapter 4 Hydrogen consumption by microorganisms in Opalinus Clay

This fourth chapter describes the second research project that focused on microbial hydrogen consumption in Opalinus Clay porewater. This is a crucial question regarding the success of nuclear waste disposal, because anoxic steel corrosion will release this gas, causing pressure build up, which might threaten the integrity of the repository.

This chapter was submitted for publication in a shorter version, under the title “Hydrogen-driven carbon cycling in the deep subsurface”.

All the work presented in this chapter was carried out by the student, except the 16S rRNA sequencing and analysis, and the metagenomic library preparation and sequencing of 6 samples that were performed by the DOE Joint Genome Institute (Walnut Creek, CA, USA), the metagenomic sequencing of 10 other samples that was performed by the Genomic Technologies Facility of the University of Lausanne (Switzerland), the metagenomic read assembly that was performed by Ino de Bruijn at the Bioinformatics Infrastructure for Life Sciences (Stockholm, Sweden), and the metaproteomic extraction, sequencing and analysis that were performed by Karuna Chourey at Oak Ridge National Laboratory (TN, USA).

Abstract

Microbial metabolic processes in the deep subsurface are poorly understood. Here we show that H₂, typically derived from serpentinization or radiolysis of water, supports a complex microbial community in Opalinus Clay, 300 meters below the surface. Metagenomic binning and metaproteomic analysis reveal a complete carbon cycle, driven by autotrophic hydrogen oxidizers. Dead biomass from these organisms is a substrate for a fermenting bacterium that produces acetate as a product. In turn, complete oxidizer heterotrophic sulfate-reducing bacteria utilize acetate and oxidize it to CO₂, closing the cycle. This metabolic reconstruction sheds light onto a hydrogen-based, carbon cycling, and sunlight-independent ecosystem in the deep subsurface.

4.1 Introduction

The vast majority of microbial life is found in the subsurface, in the water trapped within igneous, metamorphic and sedimentary rock formations (Whitman *et al.*, 1998). Life in those environments is independent of photosynthesis and relies on geogenic energy sources, mostly CH₄ and H₂. The latter is produced by the serpentinization (anoxic oxidation) of mafic and ultramafic rocks (Stevens and McKinley, 1995) and the radiolysis of water (Barr and Allen, 1959). Only two studies have investigated the metabolic potential of deep subsurface microbial communities (Chivian *et al.*, 2008; Dong *et al.*, 2014) but they both uncovered a single-organism ecosystem, suggesting minimal complexity in biogeochemical cycling in such environments. This highlights the fact that our current knowledge of the biology of deep subsurface remains in its infancy.

For many reasons, we depend on the deep subsurface, even though we live far removed from this environment. One reason is the use of geological formations as nuclear waste repositories. In Switzerland, the current option is to use Opalinus Clay as the host rock for such an undertaking, in a design that includes steel materials in the form of drums or canisters. Because this environment is anoxic, steel corrosion will release hydrogen gas that can ultimately threaten repository integrity in the case its pressure reaches excessive levels. On the other hand, as explained in the previous paragraph, H₂ is a compound that can be easily oxidized by microorganisms in order to sustain their growth. Thus, this project aims at investigation whether this gas could be consumed by microbial activity in Opalinus Clay and at deciphering the responsible metabolic processes.

To achieve these objectives, we investigated carbon cycling in a deep surface community stimulated by the addition of H₂, using a combination of metagenomic sequencing, genome binning, and metaproteomic analysis. We reconstructed a metabolic network that includes autotrophic and heterotrophic bacteria and that represents several metabolisms including sulfate reduction and fermentation. Thus, we show that H₂ can be consumed biologically in Opalinus Clay and that a light-independent microbial metabolic network is able to cycle carbon in the deep subsurface.

4.2 Materials and Methods

H₂ was injected into porewater in BRC-3 borehole in the Opalinus Clay formation in Mont Terri underground rock laboratory (URL; St-Ursanne, Switzerland), located about 300 m below the surface (Figure 3.1), in an experiment that lasted more than 500 days, in BRC-3 borehole.

4.2.1 Bioreactor

BRC-3, a 25 m long borehole, was drilled from the gallery floor. A hydraulic neoprene packer was installed at the bottom, in order to create a 2.74 m long chamber isolated from oxic gallery atmosphere, where porewater constantly produced by the borehole (at a rate of 20 mL/day) accumulated. Multiple polyamide lines were placed for connecting this chamber to surface equipment, allowing water recirculation and sampling. To avoid lines clogging with particles, a PVC screen was also installed in the chamber. An artist's rendering and a schematic drawing of the borehole equipment in presented in Figure 4.1.

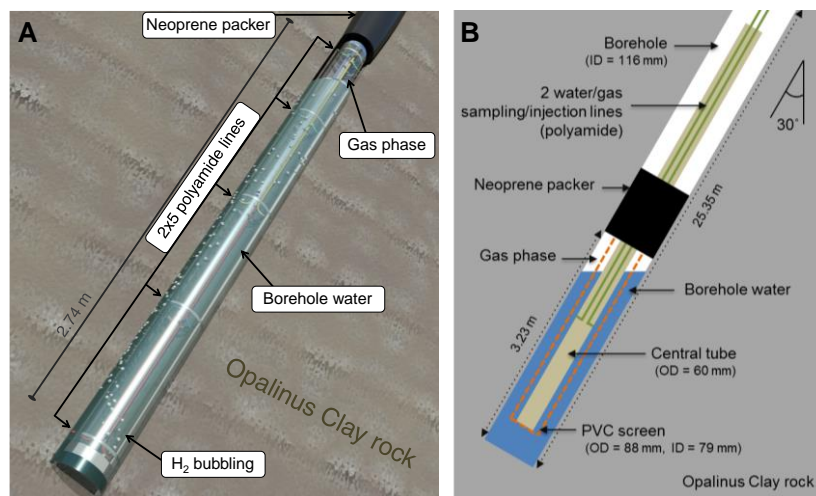


Figure 4.1 BRC-3 borehole equipment. (A) An artist's rendering of borehole equipment. For better clarity, the PVC screen is not shown. (B) Schematic drawing of the same borehole equipment, which has a vertical descending orientation, forming a 30° angle with the zenith. For better clarity, only a set of two polyamide lines are indicated, instead of 5 sets of two lines.

The surface equipment, through which borehole water was recirculated, consists of PEEK lines connected in a circulation loop to a Plexiglas sediment trap (originally designed as sampling cylinders), a peristaltic pump (with Pharmed BPT tubing), a flow-meter, pH and dissolved oxygen probes, a gas permeable membrane connected to a 500 mL H₂ reservoir and 2 needle valves, the first one placed right after borehole and second one right before the borehole, in the direction of water flow (Figure 4.2). In order to protect this experiment from oxygen contamination when borehole water was recirculated, a Plexiglas cabinet was installed and was regularly flushed with argon. In addition to the circulation loop, another line was dedicated to borehole pressure monitoring, releasing water when pressure was above 0.5 bars (relative pressure). H₂ was later directly and non-continuously injected into the borehole chamber, thus creating a gas phase.

It was not possible to install this experimental setting under truly sterile conditions. However, great care was taken to limit contamination. These steps include, depending on the material, autoclaving,

ethanol flaming, rinsing with ethanol 70 %, HCl 1 M (overnight), or bleach 7 %, before rinsing with sterile water.

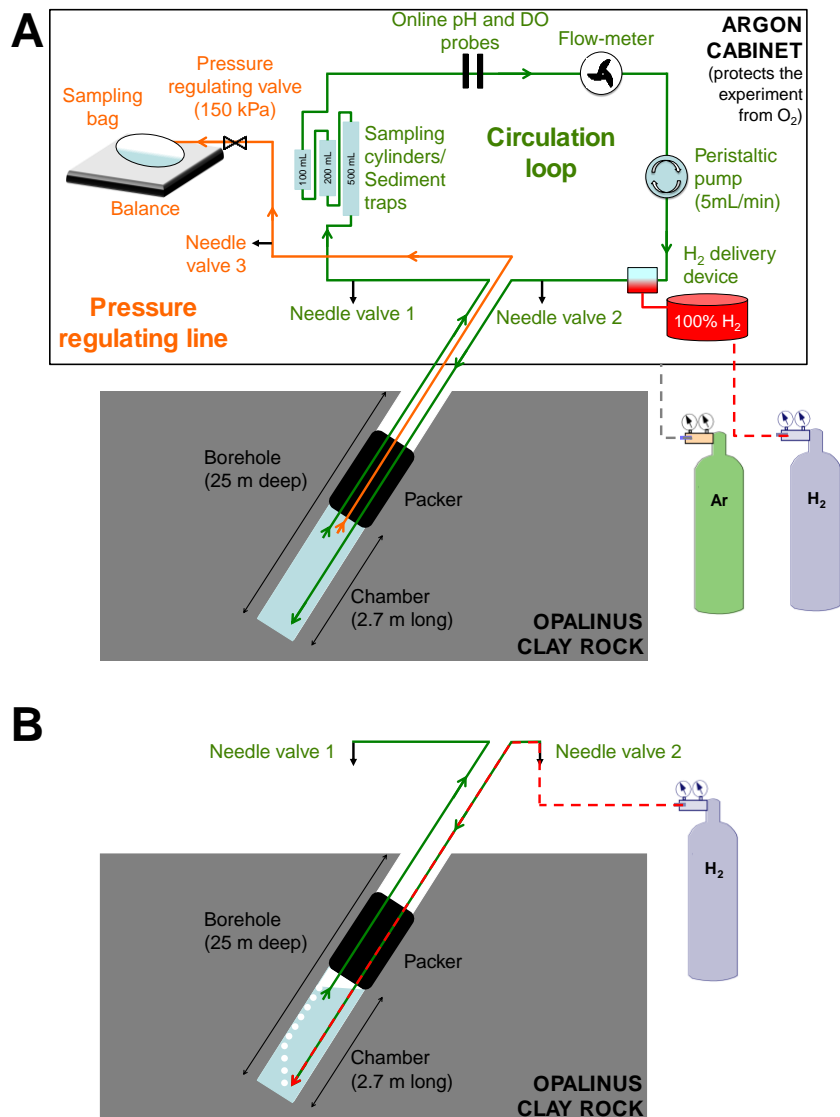


Figure 4.2 Experimental set-up: (A) in recirculation mode with continuous H₂ injection and (B) in non-recirculation mode with discrete H₂ injections into the borehole. Dashed lines represent temporary connections between gas tanks and elements of the experiment. (A) In circulation mode, water is withdrawn from the borehole by the upper green line and goes through a first needle valve (needle valve 1), 5 cylinders that act as sediment traps (only three are indicated in the cartoon), pH and DO probes, a peristaltic pump (5 mL/min), a flow-meter, a gas-permeable membrane in contact with a H₂ reservoir, and a second needle-valve (needle valve 2), before being re-injected into the bottom of the borehole. The third line (in orange) serves as a pressure valve, in order to keep a stable pressure of 0.5 relative bars in the borehole. It allows additional water produced by the borehole to be removed and collected into a bag. Here, surface equipment is kept under anoxic conditions in an Ar-flushed cabinet in order to prevent O₂ from contaminating borehole water. (B) In non-recirculation mode, only needle valves are used for borehole water sampling, artificial water injection (not shown in the cartoon) and H₂ injection that created a gas phase at the top of borehole chamber.

During recirculation mode, from day 0 to day 139, H₂ was constantly delivered to porewater through the gas permeable membrane connected to the H₂ reservoir. Most of samples were recovered by connecting a sterile and anoxic 1 L bottle to needle valve 1 when the pump was shut down. This sample, which consists of several hundreds mL of borehole water, was mainly used for DNA extraction, but also served for some chemical analyses. For some other analysis, borehole water was directly sampled from needle valves. More information about this particular type of sampling can be found in section 4.2.2. Borehole overpressure was used as a driving force for water sampling. But, in order to keep borehole pressure constant, sterile and anoxic artificial porewater (APW), amended with HCO₃⁻ right before use, was injected into needle valve 2, from a sealed bottle connected to a sterile Ar flux whose pressure was 0.5 relative bar. In this way, pressure and water volume in borehole stayed constant during sampling. In order to avoid sampling this artificial porewater, all connections between needles valves 1 and 2 in surface equipment were closed, forcing the water flow to go from artificial porewater bottle, to the interval and to the sampling bottle. Tracer tests were carried out to figure out how much borehole water can be sampled before injected artificial porewater is sampled.

During non-recirculation mode, starting at day 177, in order not to sample the water contained in the lines dead volume, pump was either started for 20 min at 5 mL/min, or the first 100 mL of water were discarded. The higher borehole pressure caused by the injection of H₂ allowed for a larger sampling volume even before any artificial porewater injection. However, to sample the last several hundreds milliliters, when borehole pressure got closer to atmospheric pressure, artificial porewater was injected, as described earlier, to increase borehole pressure. This injection continued until after the sampling was done, in order to replace all the sampled water. After this injection, 1 to 3 liters of sterile H₂ was injected to borehole through needle valve 2.

Between days 139 and 177, we used a third mode, which consisted of a transition between the two modes described above: borehole water was recirculated but H₂ was non-continuously injected into the borehole. In term of sampling and monitoring, this mode is similar to the recirculation mode.

APW composition is given in Table 5.1. The composition of APW I, used until day 324, was based on porewater modeling for another borehole located nearby in the underground rock laboratory (Wersin *et al.*, 2011). Starting on day 324, it was decided to use an APW matching the ionic composition of natural porewater from BRC-3 better. The APW II composition was based on ion chromatograph measurements from BRC-3 porewater that was sampled before the experiment started.

4.2.2 Chemical sampling and assays

4.2.2.1 Ion chromatography

To measure organic acids, a 1.5 to 6 mL aliquot was withdrawn from the 1 L sampling bottle, filtered with a 0.2 µm pore size filter and stored at -20 °C until analysis. Before analysis, a volume of 4.5 mL of APW II was added (if needed) to complete the sample volume to 6 mL. 1 M BaCl₂ solution (150 µL) was added, in order to precipitate SO₄²⁻. After a 10 min incubation, samples were filtered through a 0.2 µm filter and passed through OnGuard Ag cartridges 1cc and OnGuard H cartridges 1cc (Dionex, Sunnyvale, USA) according to manufacturer's protocol, in order to remove chloride ions. 1M NaOH solution (3 µL) was then added to 1.5 mL of treated samples. Finally, samples were measured by ion chromatography (DX-3000, Dionex, Sunnyvale, USA), using an IonPac AS11-HC column and 0.5 mM KOH solution for the elution.

4.2.2.2 Sulfide determination

For each sampling session, S(-II) was measured by averaging values obtained for several sub-samples from the 1 L sampling bottle or directly from needle valve 1 or 2, using the method described by Cline (Cline, 1969). For each sample, about 1.6 mL of borehole water was placed in a 2 mL tube containing 0.4 mL of 50 g/L zinc acetate solution. The exact amount of water sampled was determined by difference in tube weight before and after sampling. Samples were stored at 4 °C if they were analyzed within 1 or 2 days, or at -20 °C. 80 µL of Cline reagent, which is a 50 % (v/v) HCl solution containing 4 g/L of *N,N*-Dimethyl-1,4-phenylenediamine oxalate and 6 g/L of iron(III) chloride hexahydrate, was added to 1 mL of sample. Absorbance was read at 664 nm using a spectrophotometer after an incubation period of 30 min in the dark.

4.2.2.3 Fe(II) determination

For each sampling session, Fe(II) was measured by averaging values obtained by several sub-samples from the main sampling bottle or directly from needle valve 1 or 2, using the method described by Stookey (Stookey, 1970). For each sample, about 1 mL of borehole water was placed in a 2 mL tube containing 1 mL of 1M HCl solution. The exact amount of water sampled was determined by the difference in tube weight before and after sampling. Samples were stored at 4 °C if they were analyzed within 1 or 2 days, or at -20 °C. 900 µL of 50 mM HEPES solution (pH 7.0) containing 1 g/L of ferrozine was added to 100 µL of sample. Absorbance was read at 562 nm using a spectrophotometer immediately.

4.2.2.4 Alkalinity and pH

Alkalinity and pH were measured from unfiltered porewater sub-sampled from the 1 L sampling bottle, using an Orion™ 8102BNUWP ROSS Ultra™ pH Electrode (Thermo Scientific). Alkalinity was determined by titrating porewater with a solution of 0.02N H₂SO₄ till pH reaches 4.5.

4.2.2.5 Dissolved gas determination

For measuring dissolved gases (H₂, CO₂ and CH₄), 5 to 10 mL of borehole samples were recovered in a pre-prepared sealed serum bottle. The bottle contained 3.7 mg of mercury(II) chloride, was flushed with N₂, and its internal volume and pressure (null or slightly positive relative to atmospheric pressure) was measured. The samples were directly recovered from needle valves, using a needle. The exact amount of water sampled was determined by the difference in bottle weight before and after sampling. Each sample was stored on its side at 37 °C between 1 and 2 days for gas phase equilibration. After this incubation, the gas phase was sampled using a gas-tight syringe and 1 mL was injected in a GC-FID (Varian 450-GC, Agilent, Santa Clara, USA). Initial concentration of H₂, CO₂ and CH₄ can then be back calculated. For this part, more details can be found in section A-2.1.1.

4.2.3 Planktonic cell density

Samples for planktonic cell density were recovered from needle valve 1 in a sterile and anoxic serum bottle. The sample was measured within five hours. 1 mL was filtered at 0.2 µm on a black polycarbonate membrane, and then rinsed 3 times with PBS buffer (8 g/L of NaCl, 0.2 g/L of KCl, 1.44 g/L of Na₂HPO₄, 0.24 g/L of KH₂PO₄, pH = 7.4). The membrane was then stained with SYBRGreen I as described by Lunau *et al.* (Lunau *et al.*, 2005). For each samples, 15 fields were observed with an epifluorescence Nikon Eclipse E800 microscope at 1,000x magnification and a B-2A filter. The standard deviation was calculated by averaging 3 groups of 5 samples. Planktonic cell density was calculated with this formula:

$$Total\ cell\ concentration\ \left[\frac{cell}{mL}\right] = \frac{Mean\ number\ of\ cells \times filter\ area\ [\mu m^2]}{counting\ field\ area\ [\mu m^2] \times sample\ volume\ [mL]}$$

4.2.4 DNA sampling and extraction

Several hundreds milliliters of borehole water were first recovered in anoxic and sterile bottles, before filtration using sterile 0.2 µm polycarbonate filters, which was immediately placed in a 1.5 mL sterile tube containing 0.4 mL of LifeGuard Soil Preservation Solution (MO BIO Laboratories Inc., Carlsbad, USA), prior to being frozen at -20 °C. The DNA of a first set of samples was extracted using a modified version of FastDNA SPIN Kit for Soil (MP Biomedicals, Santa Ana, USA). A detailed version of the protocol can be find in A-1.1.1.

An extra purification step was carried out subsequently, using the standard protocol of Genomic DNA Clean & Concentrator purification kit (Zymoresearch, Irvine, USA). DNA from samples recovered at 2, 42, 48, 56, 101 and 134 days were extracted using this method. However, because this method recovered poor DNA quality (in term of fragment length) on samples containing S(-II) and black precipitates, it was decided to use a method that doesn't involve a bead beating step. It is why a phenol-chloroform DNA extraction was carried out for all other samples. The method starts with the recovery of biomass from filtration membranes. The membrane was placed in a 60 mL sterile bag containing 0.6 mL of TE buffer pH 7.5-8.0. The bag was closed and biomass was transferred to the TE buffer by rubbing the membrane with one's fingers on the outside of the bag. The TE buffer containing the filtrate was transferred to a new tube, and combined with the pellet obtained by centrifuging the LifeGuard solution (after having removed the membrane) at 7'000 xg for 10 minutes. For DNA extraction, a phenol-chloroform extraction was carried out. A detailed protocol can be found in section A-2.1.2.

4.2.5 16S rRNA genes sequencing

16S rRNA genes sequencing of the 32 samples of the H₂ injection experiment was done by the Joint Genome Institute (Walnut Creek, USA) through a community sequencing program project (CSP 1505), together with the sequencing of the 23 samples from the diversity survey presented in Chapter 3. The protocol used is the same one described in section 3.2.4.1.

4.2.6 Metagenomic analysis

4.2.6.1 Library preparation and sequencing

For samples recovered at days 181, 188, 195, 202, 206, 209, 214, 238, 246 and 250, True seq DNA LT protocol (Illumina, San Diego, USA) was used for library preparation, with 0.5 µg of DNA as a starting material, and with a targeted fragment length of 500 bp. DNA was sequenced using Illumina HiSeq 2500, generating 100 bp paired-end reads at the Lausanne Genomic Technologies Facility. The reads were then quality trimmed with sickle version 1.210 (<https://github.com/najoshi/sickle>). To reduce the computational demands of the assembly, the ~2 billion reads were subsequently subsampled to 25 percent with seqtk version 1.0-r32 (<https://github.com/lh3/seqtk>). These 10 samples were co-assembled with Ray version 2.3.1 (<http://denovoassembler.sourceforge.net/>) on a Cray XE6 system using 1,024 cores in 2.5 hours and a kmer of 41.

Samples recovered at days 14, 48, 101, 122, 134 and 233 were sequenced at the Joint Genome Institute (Walnut Creek, USA). For preparing libraries, 100ng of DNA was sheared to 270 bp using the covaris E210 (Covaris, Woburn, USA) and size selected using SPRI beads (Beckman Coulter,

Fullerton, USA). The fragments were treated with end-repair, A-tailing, and ligation of Illumina compatible adapters (Integrative DNA Technologies, Coralville, USA) using the KAPA-Illumina library creation kit (KAPA biosystems, Wilmington, USA). Quantitative PCR was used to determine the concentration of the libraries and were sequenced on the Illumina HiSeq. For sequencing, the libraries were quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified libraries were then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v3, and Illumina's cBot instrument to generate a clustered flowcell for sequencing. Sequencing of the flowcell was performed on the Illumina HiSeq2000 sequencer using TruSeq SBS sequencing kits, v3, following a 2x150 indexed run recipe. For QC filtering, Raw Illumina metagenomic reads were screened against Illumina artifacts with a sliding window with a kmer size of 28, step size of 1. Screen reads were trimmed from both ends using a minimum quality cutoff of 3, reads with 3 or more s or with average quality score of less than Q20 were removed. In addition, reads with a minimum sequence length of <50 bps were removed.

4.2.6.2 Contig binning

Bowtie 2 version 2.1.0 (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) and MarkDuplicates from Picard tools version 1.77 (<http://picard.sourceforge.net/>) were used (with default parameters) to map quality trimmed read from all 16 samples on contigs from co-assembly, in order to calculate contig coverage. Only contigs > 5000 bp were used as input for CONCOCT version 0.2 (Alneberg *et al.*, 2014).

4.2.6.3 Taxonomic annotation of clusters

For taxonomic affiliation of genomes, the following method was applied. If a 16S rRNA gene was detected using Barrnap version 0.4.2 (<http://www.vicbioinformatics.com/software.barrnap.shtml>), the annotation of RDP classifier version 2.7 (<http://sourceforge.net/projects/rdp-classifier/>) was used with a confidence threshold of 0.8. If no 16S rRNA gene could be detected, MLTreeMap version 2.061 (<http://mltreemap.org/>) annotations were used instead, choosing the taxonomic level that contains at least 66.67% of all marker genes. When the default and geba tree produced results with different taxonomic levels, the highest one was chosen. If all marker genes were affiliated to a single species, the highest taxonomic level that is not shared by adjacent species on the tree was chosen. IMG phylogenetic distributions (<https://img.jgi.doe.gov/mer/>) were not used for assessing the taxonomic annotation of genomes.

The average nucleotide identity (ANI) between the selected draft genomes and references genomes found in NCBI genome database (<http://www.ncbi.nlm.nih.gov/genome/>) was calculated using the default parameters of the Kostas Lab's ANI calculator (<http://enve-omics.ce.gatech.edu/ani/>). For

selecting reference genomes, a BLASTN analysis of the 16S rRNA gene of the draft genomes was carried out on NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), using as database the 16S ribosomal RNA sequences database. Depending on the outcome, either the 15 top BLASTN hits, or all genomes belonging to a genus or family, were used as references genomes.

4.2.6.4 Pathway annotation of clusters

Genes were annotated with the IMG pipeline (<http://img.jgi.doe.gov/>). Metabolic pathways were manually annotated using in-house databases based on KEGG (<http://www.genome.jp/kegg/>) and MetaCyc (<http://metacyc.org/>) databases, and textbook biochemical pathways (Kim and Gadd, 2008). In order to determine the subcellular localization of proteins, PSORTb version 3.0.2 (<http://www.psort.org/psortb/>) was used. For comparing some gene cluster at the synteny level, BLAST version 2.2.28 (<http://www.ncbi.nlm.nih.gov/books/NBK1762/>) and Mauve version 2.3.1 (<http://asap.genetics.wisc.edu/software/mauve/>) were used.

4.2.6.5 Clusters contribution to microbial community

Draft genome contribution to microbial community was calculated as described in section 3.2.5.4.

4.2.6.6 Data deposition

16S DNA sequences, the 15 draft genomes, the unclassified contigs, genes annotations, raw reads from V4 region and metagenomic sequencing were deposited to NCBI and IMG databases. Accession numbers can be found in section A-2.1.3.

4.2.7 Metaproteomics

Borehole water (0.5-1 L) was filtered by Sterivex 0.22 µm polyethersulfone membrane (Millipore, Billerica, USA) and directly frozen in dry ice. The frozen filters was cut into small pieces and immersed in detergent based lysis buffer, described by (Chourey *et al.*, 2010). The cells trapped on the filters were heat-lysed and processed as described earlier (Chourey *et al.*, 2013). The cell lysate (supernatant) was aliquoted into fresh tubes and amended with chilled 100% trichloroacetic acid (TCA) to a final concentration of 25% (vol/vol) and kept at -20°C overnight. The filter pieces were discarded. Following TCA precipitation, the cell lysate was centrifuged at 21,000 x g for 20 min to obtain a protein pellet. The pellet was retained and the supernatant discarded. The protein pellet was washed twice with chilled acetone as described earlier (Chourey *et al.*, 2013). The protein pellet was dried at room temperature and solubilized in 6 M guanidine buffer (6 M guanidine; 10 mM dithiothreitol [DTT] in Tris-CaCl₂ buffer (50mM Tris; 10mM CaCl₂, pH 7.8) and incubated at 60°C for three hours with intermittent vortexing. An aliquot of 25 µl was utilized for protein estimation, carried out using the RC/DC protein

estimation kit (Bio-Rad Laboratories, Hercules, USA) as per the manufacturer's instructions. Bovine serum albumin (supplied with the kit) was used as standard for the assay. Following protein estimation results, rest of the protein sample was digested overnight at 37, peptides desalted using seppak column and solvent exchanged as described earlier (Thompson *et al.*, 2007). Peptides were stored at -80°C until MS analysis.

All chemicals used for sample processing, cleanup and mass spectrometry analysis were obtained from Sigma Chemical Co. (St. Louis, USA). Sequencing-grade trypsin was acquired from Promega (Madison, USA). Any other sources used to obtain the chemicals have been mentioned where required. High performance liquid chromatography- (HPLC-) grade water and other solvents were obtained from Burdick & Jackson (Muskegon, USA), 99% formic acid was purchased from EM Science (Darmstadt, Germany).

A single aliquot of 75ug peptide mix was loaded onto a biphasic resin packed column [SCX (Luna, Phenomenex, Torrance, USA) and C18 (Aqua, Phenomenex, Torrance, USA)] as described earlier (Brown *et al.*, 2006; Thompson *et al.*, 2007). Following sample loading, the column was washed for 15 min, offline as described by Sharma (Sharma *et al.*, 2012) and connected to the C18 packed nanospray tip (New Objective, Woburn, USA) mounted on Proxeon (Odense, Denmark) nanospray source as described earlier (Sharma *et al.*, 2012). Peptides were subjected to 24h multi-step chromatographic separation via the Ultimate 3000 HPLC system (Dionex, Sunnyvale, USA) connected to the mass spectrometer and measurements done using the Multi-Dimensional Protein Identification Technology (MuDPIT) approach as described earlier (Brown *et al.*, 2006; Thompson *et al.*, 2007; Sharma *et al.*, 2012). The peptide fragmentation was executed and recorded via Orbi Trap Elite mass spectrometer (Thermo Fisher Scientific, Germany) operated in data dependent mode, via Thermo Xcalibur software V2.1.0. Each full scan (1 microscan) was followed by collision-activated dissociation (CID) based fragmentation using 35% collision energy of 20 most abundant parent ions (1 microscan) with a mass exclusion width of 0.2 m/z and dynamic exclusion duration of 60 s.

For protein identification, the raw spectra were searched against selected proteins database generated by IMG annotation pipeline, via Myrimatch v2.1 algorithm (Tabb *et al.*, 2007) set to parameters described by (Xiong *et al.*, 2015) with minor modifications such as omission of static cysteine and dynamic oxidation modifications. Identification of at least two peptides per protein (one unique and one non-unique) sequence was set as a prerequisite for protein identification. Common contaminant peptide sequences from trypsin and keratin were concatenated to the database. Reverse database sequences were also included in the database as decoy sequences to calculate false discovery rate (FDR). False discovery rate (FDR) cutoff for peptide to spectrum identification was maintained at < 1%. For downstream data analysis, spectral counts of identified peptides was normalized as described before (Paoletti *et al.*, 2006) to obtain the normalized spectral abundance factor (NSAF), also referred to as

normalized spectral counts (nSpc). Average of nSpc from duplicate runs was used for further statistical analysis of the data.

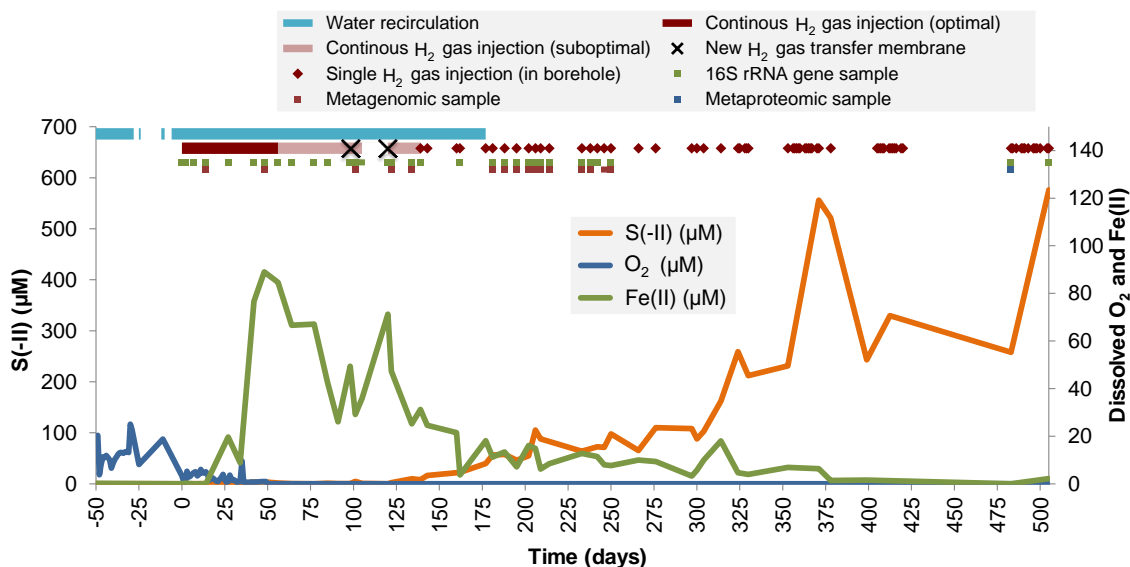


Figure 4.3 Chemical change in dissolved O₂ (blue line, Y-right axis), Fe(II) (green line, Y-right axis) and S(-II) (orange line, Y-left axis) in borehole water, starting 50 days before the first H₂ injection that occurs at day 0. The experiment procedure is indicated at the top of the plot: the light blue line shows when water was recirculated, the dark red line shows when H₂ was optimally delivered from the surface, the light red line shows when H₂ was sub-optimally delivered from surface, the dark crosses show when the H₂-permeable membrane was changed, and red diamonds show when H₂ was injected directly into the borehole. Samples are also indicated at the top of the plot: green squares show samples for 16S rRNA gene sequencing, red squares show samples for metagenomic sequencing, and blue square show the sample for metaproteomic analysis.

4.3 Results

4.3.1 Chemical changes

During the recirculation mode, O₂ was consumed rapidly (within 1 month), leading to anoxic conditions (Figure 4.3). Later, starting at day 25, Fe(II) accumulated into the porewater. Concomitantly, H₂ was consumed in the borehole water, as shown by the difference in H₂ concentrations between the water going to and coming from borehole (Figure 4.4.A). Later, when Fe(II) concentration starts to decrease, H₂ input stopped abruptly (Figure 4.4.A). This likely resulted from clogging of the gas permeable membrane. At days 98 and 120, the membrane was replaced, but this change had practically no impact on H₂ input. It is why it was decided to inject H₂ pulses regularly and directly into the borehole, starting at day 139. From that point on, S(-II) increased in the porewater, overtaking Fe(II) at day 209, before reaching a concentration of about 600 μM (Figure 4.3). During the H₂ discrete injections

phase, H₂ consumption in the borehole was clearly occurring, highlighted by the decreasing concentration of H₂ between each sampling point (Figure 4.4.B).

4.3.2 Biological succession

Planktonic cell density increased during the three distinct phases, (i) when O₂ was reduced, (ii) when Fe(II) was produced (Figure 4.8.A) and (iii) when S(-II) was produced (Figure 4.8.B). Planktonic cell density is missing for the sampling event of day 276 because, unlike all other samples that were measured right after they were recovered, this sample was stored two days at 4 °C, leading to biased results. The 16S rRNA gene analysis of 32 samples produced 3,803,075 reads, classified in 606 OTUs (Table A-2.1). It shows that microbial community drastically changed over time, as pictured in Figure 4.9. The oxic phase was dominated by Gammaproteobacteria, mainly Xanthomonadaceae and *Pseudomonas* organisms, which can use O₂ or NO₃⁻ as electrons acceptor, or ferment (Saddler and Bradbury, 2005; Moore *et al.*, 2006). Later, once O₂ was totally consumed, other bacteria took over. A member of *Desulfocapsa*, a Gram-negative genus known to reduce sulfate (Kuever *et al.*, 2005b), strongly dominated the microbial community. Other SRB belonging to Gram-positive genera *Desulfotomaculum* and *Desulfosporosinus* also took a significant place in this microbial community (Spring and Rosenzweig, 2006; Widdel, 2006). Throughout the whole experiment, members of Alphaproteobacteria, belonging to the genera *Novispirillum* and *Hyphomonas*, and to the family of Rhodobacteraceae, represented an important part of the microbial community. These organisms are known to have diverse metabolic abilities (Moore *et al.*, 1984; Garrity *et al.*, 2005a; Yoon *et al.*, 2007).

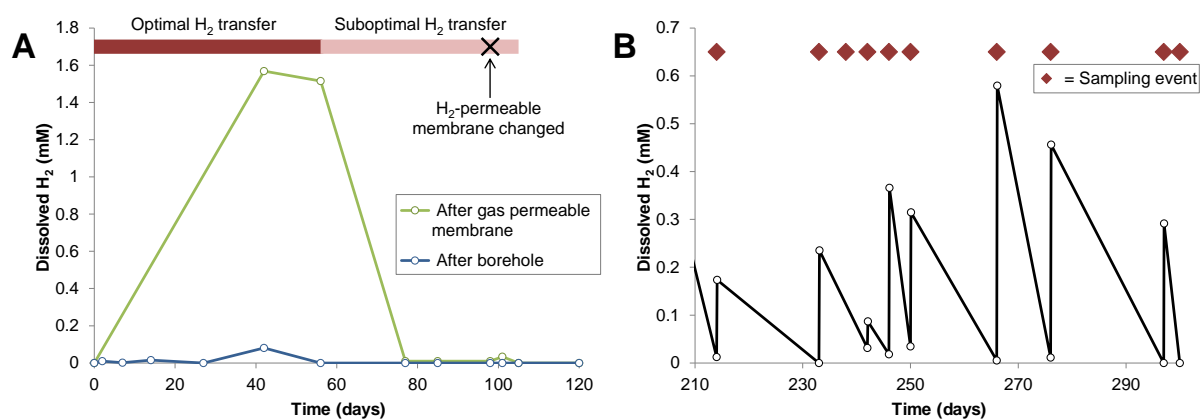


Figure 4.4 H₂ consumption in the borehole over time (A) during the phase when water was recirculated and H₂ continuously injected and (B) during the phase when water was not recirculated and H₂ delivered through discrete injections into the borehole. (A) At the top of the plot, the dark red line indicates the period during which H₂ was optimally delivered through the gas permeable membrane, while the light red line indicates the period during which H₂ was sub-optimally delivered. Gas membrane was changed H₂ at day 98 as indicated by the cross. Starting at day 105, H₂ injection was stopped (till day 120). (B) Red diamonds indicate the sampling events, during which H₂ was injected to the borehole and significant amount of borehole water was replaced by sterile water.

SO_4^{2-} concentration decreased till day 324, before starting to increase (Figure 5.4), while CO_2 concentration kept decreasing (Figure 4.5). Alkalinity of the borehole water samples shows the same trend, while pH increased (Figure 4.6). Organic acid shows higher concentration at day 0, with more than 70 μM of acetate and more than 40 μM of lactate. Later these concentrations decreased and stayed low, these two compounds being still detectable, together with formate, propionate and butyrate (Figure 4.7).

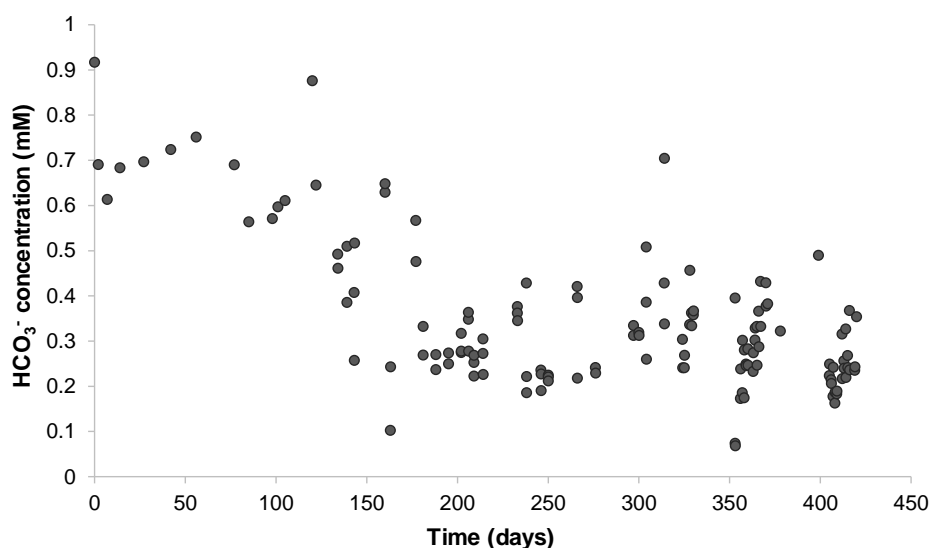


Figure 4.5 Bicarbonate concentration over time in borehole water.

4.3.1 Metagenomic analysis

4.3.1.1 Manual binning correction

Metagenomic binning allowed the identification of 65 putative genomes. The output binning of CONCOCT, was manually corrected. Some bins were subdivided based on their coverage pattern across samples (Table A-2.2). This concerns cluster c4 divided in c4a, c4b and c4c, cluster c7 divided in c7a and c7b, cluster c8 divided in c8a, c8b and c8c, cluster c16 divided in c16a and c16b, cluster c20 divided in c20a and c20b and cluster c40 divided in c40 and c40x (composed of only one contig). Because clusters c40x and c53 have complementary single-copy genes, and have similar taxonomic annotations (results not shown), it was decided to merge them into a newly defined cluster c58. Cluster c15 contained 10 different 16S rRNA genes. This can be explain by the difficulty of binning contigs harboring these conserved genes, because of specific tetranucleotides frequency patterns and because of biased coverage patterns. However, based on taxonomic annotation and coverage pattern comparison, these incorrectly annotated contigs could be transferred to their correct cluster. This concerns contig BRHa_1003273 placed in cluster c11, contig BRHa_1000755 placed in cluster c35, contig BRHa_1003404 placed in cluster c41, contig BRHa_1000455 placed in cluster c46, contig

BRHa_1007118 placed in cluster c52, contig BRHa_1004490 placed in cluster c20a, and contig BRHa_1004053 placed in cluster c8a. Contig BRHa_1000758 was classified as contaminating sequence because it perfectly matches PhIX phage genome. The first 40 nucleotides of BRHa_1006152 were trimmed off because they match an adaptor used for sequencing processes. The final binning and the one output by CONCOCT are presented in Table A-2.2.

4.3.1.2 Cluster quality assessment

Purity and completeness of each cluster was then assessed using a single-copy gene analysis (Table A-2.3). Clusters that lack less than three single-copy genes and whose average number of single-copy genes is contained between 0.97 and 1.03 are considered as draft genomes that are pure and nearly complete. Based on this criterion, 15 were found to be pure and nearly complete, and together represented more than 75% of the microbial community (Table A-2.3).

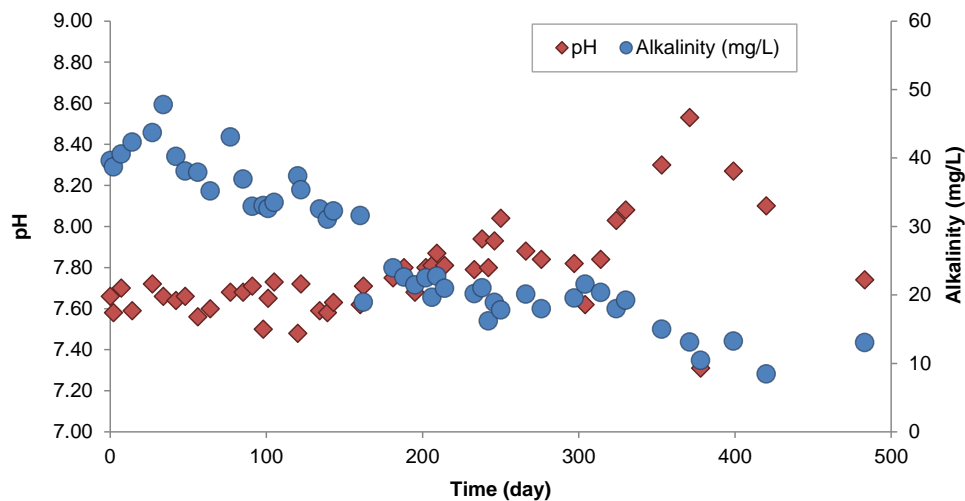


Figure 4.6 Change in pH (blue circles, left Y-axis) and alkalinity (red diamonds, right Y-axis) of BRC-3 borehole water along the course of the H₂ injection experiment.

4.3.1.3 Taxonomic annotation of clusters

The good quality of this binning is also highlighted by their homogeneous taxonomic annotation, based on separate methods (Table A-2.4). A similar picture of the microbial community to the one from 16S rRNA analysis was obtained (Figure A-2.1).

Among the 40 16S rRNA genes identified in contigs, 20 are complete and 23 can be linked to OTUs issued from the 16S rRNA gene analysis (Table A-2.5). The taxonomic annotations of the corresponding clusters and OTUs perfectly match, even though the taxonomic level can be different. Moreover, because the same 16S rRNA gene analysis was also carried out on the samples from Chapter 3, it allows to determine if genomes harboring a completed 16S rRNA gene were detected in other boreholes. This is the case for all of them, as shown in Table A-2.5. Moreover, *Desulfobulbaceae* c16a,

Pseudomonas c35, *Roseovarius* c41 and Rhodospirillaceae c57 were also observed in all 23 samples presented in Chapter 3.

4.3.1.4 Metabolic annotation of clusters

The pathway annotation of the 15 selected draft genomes resulted in their classification into five categories: (i) autotrophic SRB able to use H₂ as an energy and electron source to reduce CO₂ through the reductive acetyl-CoA pathway; (ii) SRB unable to oxidize H₂ expected to exhibit heterotrophic growth; (iii) autotrophic bacteria unable to reduce sulfate, but able to oxidize H₂; (iv) autotrophic bacteria unable to reduce sulfate and to oxidize H₂ expected to exhibit heterotrophic growth and (v) heterotrophic bacteria unable to reduce sulfate or to oxidize H₂. The metabolic abilities of these 15 genomes were compiled in Table 4.1. The pathway annotation of the 31 best genomes, those having an average single-copy gene number comprised between 1.11 and 0.66 (Table A-2.3), is presented in Table A-2.6. The same metabolic capabilities were highlighted here than with the 15 selected genomes.

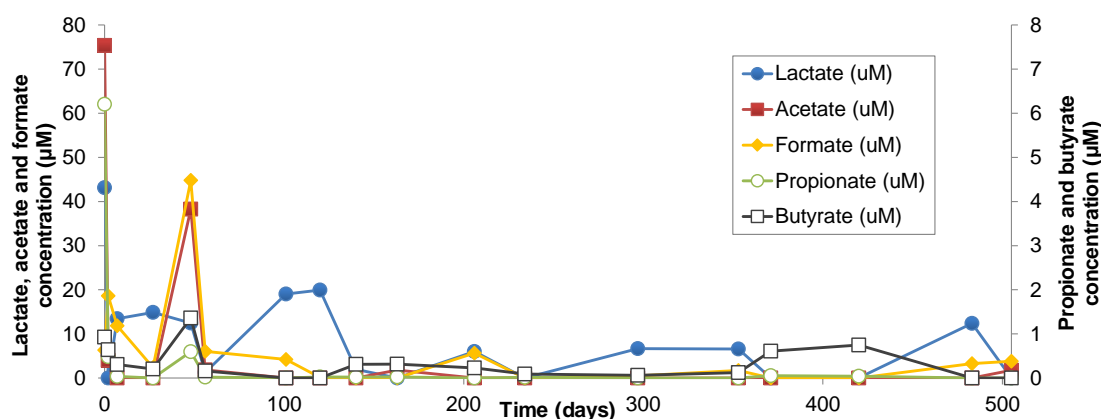


Figure 4.7 Organic acid concentration over time in borehole water: lactate (solid blue circles), acetate (solid red squares), formate (solid yellow diamonds), propionate (open green circles) and butyrate (open black squares). Solid symbols belong to the left Y-axis, while open symbols belong to the right Y-axis.

4.3.1.5 ANI analysis of good clusters

As shown in Table A-2.7, the ANI analysis of the 15 selected clusters indicates that most of them are significantly different from the sequenced genomes that have been deposited to NCBI. The identities are about 80%, indicating that the genomes issued from this study might represent new microbial families (Goris *et al.*, 2007). This is not consistent with the taxonomic annotations of the clusters, which often match a genus level (Table A-2.4). But this latter was based on only one or a set of marker genes, which might not reflect the characteristics of the whole genomes. Two selected draft genomes have very high ANI with genomes deposited to NCBI, and should belong to the same microbial species. They

are clusters c35 and c41, which can be considered as *Pseudomonas xanthomarina* and *Roseovarius* sp., respectively.

4.3.2 Metaproteomics analysis

Metaproteomic analysis, combined with the reconstructed individual genomes, enabled the observations of 1087 proteins in the sample recovered at day 483 (Table A-2.8). The result of each run is presented in a separate table (Table A-2.9 and Table A-2.10). The majority of these proteins belong to seven clusters, all considered as draft genomes, and allows the reconstruction of their active metabolic routes. They are Desulfobulbaceae c16a (Figure 4.10 and Table A-2.11), three Peptococcaceae, c4a, c8a and c23, *Desulfatitalea* c12, Rhodospirillaceae c57 and *Hyphomonas* c22. Together, these organisms represent more than 60% of the microbial community (Figure 4.11). All proteins used for annotating the active pathways of these 7 clusters are listed in Table A-2.12, together with their IMG annotation.

4.4 Discussion

H₂ promotes the successive reduction of O₂, Fe(III) and SO₄²⁻ in BRC-3 borehole water, as shown in Figure 4.3. This finding perfectly matches the electron tower theory stating that electron acceptors with higher redox potential are used first by microorganisms because their reduction liberates more energy, as detailed in section 1.7.1.1. The biological nature of these reactions can be demonstrated by the increasing number of planktonic cells, in each of these 3 phases (Figure 4.8).

4.4.1 Microbial growth

In Figure 4.8.A, the first planktonic cell density peak (at day 14) corresponds to the suboxic phase, when O₂ was still present in borehole, and when the Xanthomonadaceae and *Pseudomonas* populations are at their maximum (Figure 4.9). After 27 days, O₂ concentration dropped to zero (Figure 4.3), causing the fall of microaerobic microorganisms. At 56 days, a second peak coincides with the maximal Fe(II) concentration measured (Figure 4.3), and highlights the growth of anaerobic microorganisms. However, after 56 days, H₂ transferred stopped suddenly (Figure 4.4), impacting microbial growth. A third peak appears immediately after replacing the H₂-permeable membrane (black cross in Figure 4.8.A) but because it was clogged again soon thereafter, planktonic cell density rapidly dropped.

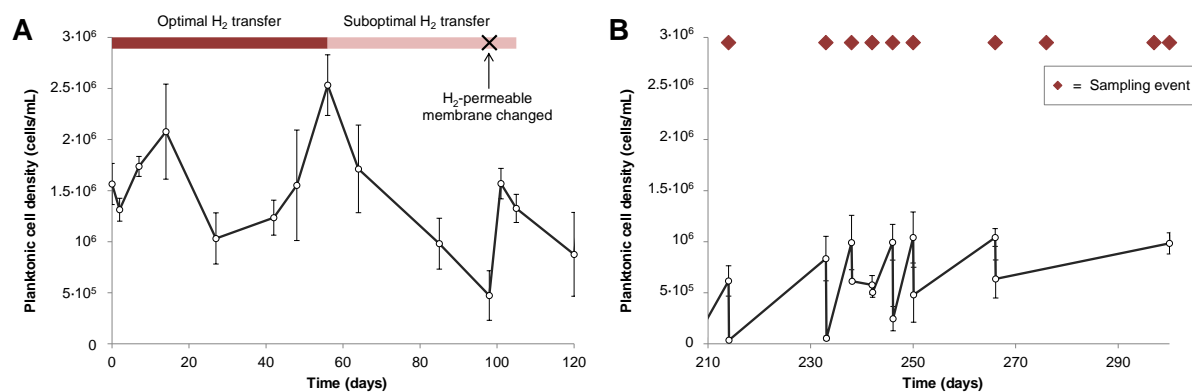


Figure 4.8 Planktonic cell density (black lines and open circles) over time in borehole water (A) during the phase when water was recirculated and H₂ continuously injected (B) during the phase when water was not recirculated and H₂ delivered through discrete injections into the borehole. (A) At the top of the plot, the dark red line indicates the period during which H₂ was optimally delivered through the gas permeable membrane, while the light red line indicates the period during which H₂ was sub-optimally delivered. The gas membrane was changed on day 98 as indicated by the cross. Starting at day 105, H₂ injection was stopped (until day 120). (B) Red diamonds indicate the sampling events, during which H₂ was injected to the borehole and a significant amount of borehole water was replaced by sterile artificial porewater.

During the non-recirculating phase, for each sampling event, sampled borehole water was replaced with synthetic, sterile and anoxic porewater (Figure 4.8.B). This is the reason why the second measurement of planktonic cell density is lower, because it was carried out after having replaced the borehole water. But after a few days, planktonic cell density always increased (except one time, between 238 and 242 days), indicating that microorganisms were growing during sulfate reduction phase (Figure 4.3).

The decreasing alkalinity and HCO₃⁻ concentrations clearly indicate biological consumption of this compound in the porewater (Figure 4.5 and Figure 4.6). This decrease is not due to a decrease in pH value, because the latter tends to increase during the course of the experiment (Figure 4.6). These observations can then only be explained by a microbial HCO₃⁻ uptake for carbon fixation.

4.4.2 Microbial community succession

The microbial community succession based on the 16S rRNA gene is consistent with the redox regimes (Figure 4.3). Aerobic Gammaproteobacteria (Moore *et al.*, 2006; Saddler and Bradbury, 2005) initially represented about 60% of the microbial community when O₂ is still present. When the latter vanished, Gammaproteobacteria significantly decrease. Only *Pseudomonas* members remain during the entire duration of the experiment. Among them is *Pseudomonas* c35, which is the same organism than the *Pseudomonas* c5 described in Chapter 3 (Table A-2.5), a fermenter that can break down complex organic matter in Opalinus Clay porewater (Figure 3.4.A). Indeed both genomes harbor an identical

16S gene, and have an ANI higher than 98% (results not shown). Its ability to ferment can explain why this organism does not vanish once O₂ is depleted.

When this event occurs, Desulfobulbaceae c16a, which corresponds to *Desulfocapsa* OTU 0, takes a dominating place in the microbial community (Figure 4.9). This organism, even though it is considered to be an SRB (Kuever *et al.*, 2005a), is thought to be responsible for Fe(II) production, as some Desulfobulbaceae members were shown to reduced Fe(III) (Holmes *et al.*, 2004). Iron reduction is also thought to be carried out by some Alphaproteobacteria members. Indeed, Rhodobacteraceae, whose most important OTU correspond to *Roseovarius* c41, reach their highest proportion during that period.

Starting at day 48, the Fe(II) concentration starts to decrease (Figure 4.9), highlighting a production of S(-II), which likely precipitates with Fe(II) to form FeS. This means that even though S(-II) is not yet detectable, SO₄²⁻ reduction already started. Desulfobulbaceae c16a remains the dominating organism, but other SRB develop, mainly Peptococcaceae members, which correspond to clusters c4a, c4b, c8a and c23.

The two sharp peaks of Novispirillum members in Figure 4.9, whose most important OTU is Rhodospirillaceae c57, corresponds to two phases where H₂ input was stopped, due to complete membrane clogging at day 97, and due to a deliberate pause in H₂ injection at day 120. This highlights the ability of this organism to switch between autotrophic and heterotrophic metabolic activity.

The sample recovered at day 139 is the first one collected following the first H₂ injection directly into the borehole. For unclear reasons, this new mode of H₂ delivery affected the microbial community composition as SRB Peptococcaceae and Alphaproteobacteria increased their proportion (Figure 4.9). But after 42 days, the situation naturally returned as it was before H₂ delivery into the borehole. It is a perfect illustration of the resilience capability of this microbial ecosystem.

At day 483, the microbial community is significantly different than what was observed in the previous samples (Figure 4.9). This can be explained by an interval of two months between this sampling and the last one. This allowed the building of lower redox conditions in the borehole water. Indeed, even if great care was taken during each sampling not to contaminate the porewater with O₂, a small amount of air is inevitably injected into the borehole together with new artificial porewater, after each sampling. Thus, the important proportion of Peptococcaceae and Anaerolinaceae families at day 483 can be explained by more stringent anoxic conditions that can build when APW is not injected for a long time. In this sample, Alphaproteobacteria are significantly in lower abundance than in the previous samples.

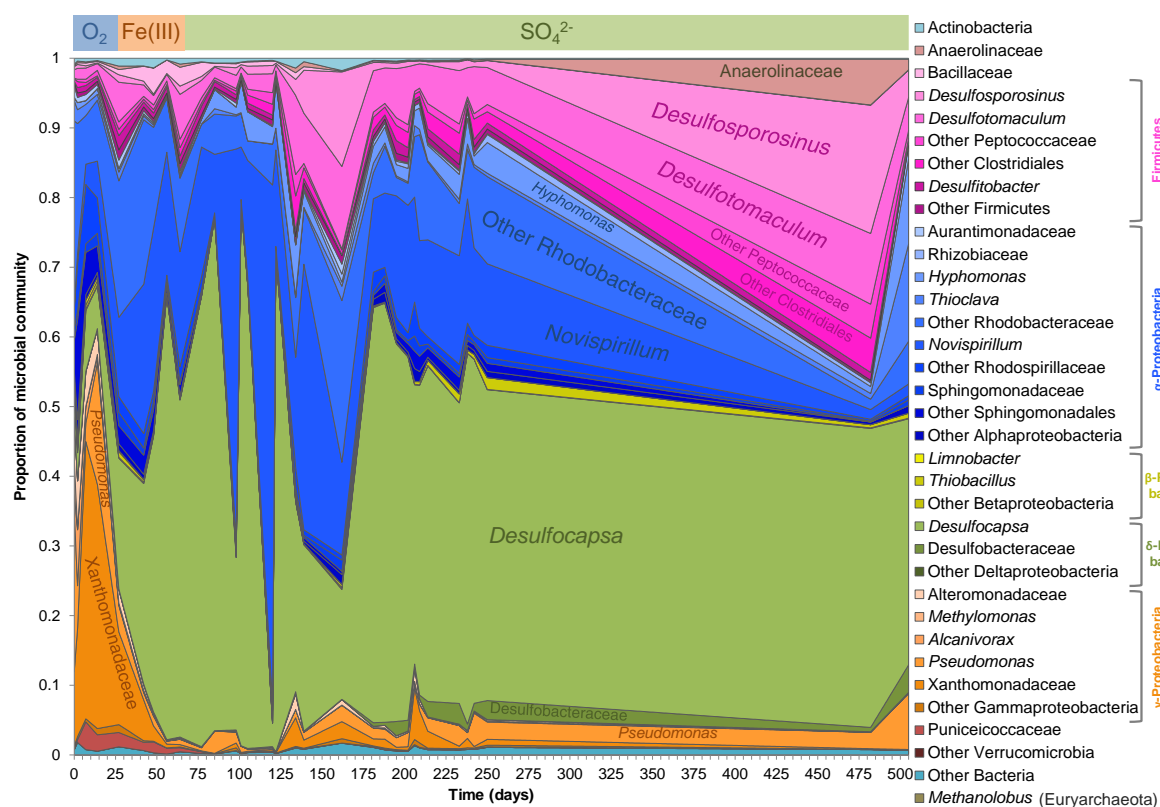


Figure 4.9 Microbial community succession in the borehole water, after H₂ injection, based on sequencing of region V4 of the 16S rRNA gene. The colored bar at the top of the plot indicates the final electron acceptor in the different redox regimes (Figure 4.3).

An explanation for this decrease is the ability of some organisms of this phylum to use SO_3^{2-} as terminal electron acceptor. This compound is usually not present in environments where sulfate-reduction takes place, as all forms of sulfur are reduced to S(-II) (Zopfi *et al.*, 2004). A way to produce SO_3^{2-} in this system is to oxidize S(-II), which can be carried out by O₂ brought by the APW injections. A metaproteomic analysis of a porewater sample recovered at 420 days showed that Rhodospirillaceae c57 and *Roseovarius* c41 were composing a significant part of the microbial community, and that they were using SO_3^{2-} as a terminal electron acceptor (results not shown). Thus the decrease of Alphaproteobacteria populations at day 420 can be explained by the lower SO_3^{2-} concentration, due to a longer period with no O₂ contamination.

4.4.3 Metabolic capabilities of draft genomes

Among the 15 selected draft genomes, most have the capability to growth autotrophically, mainly fixing carbon via the reductive acetyl-CoA (Table 4.1). However, Rhodospirillaceae c57 and *Hoeflea* c9 can use the Calvin cycle instead. Some of those putative autotrophs, Peptococcaceae c4a and c4b and *Hoeflea* c9, cannot use H₂ as a source of electron for respiratory processes, having no group-1 [NiFe]-hydrogenases (Vignais and Billoud, 2007). This means that they are not able to obtain energy

from H₂ oxidation, and that they are expected to grow heterotrophically in this system, oxidizing or fermenting organic carbon.

Cluster c4b, annotated as an SRB Peptococcaceae, includes both the acetyl-CoA pathway and the Calvin cycle (Table 4.1). The facts that two different carbon fixation pathways coexist in a single genome, and that a SRB harbors the Calvin cycle are remarkable. Indeed, no other known SRB uses a carbon fixation pathway different than the acetyl-CoA pathway or the reverse tricarboxylic acid (TCA) cycle (Hansen, 1993). However, the presence two carbon fixation pathways in a single genome can possibly be due to a binning error, even though the single-copy gene profile of Peptococcaceae c4b is perfect (Table 4.1) and the genes of these two pathways are found in numerous contigs. *Hoeflea* c9 also possesses key genes for both carbon fixation pathways (Table A-2.3). However, the low completeness of acetyl-CoA and the 9 occurrences of its key enzyme, the carbon monoxide dehydrogenase (EC:1.2.99.2) in this genome indicate that this key enzyme, and thus this pathway, are absent from *Hoeflea* c9. This latter is then thought to only possess the Calvin cycle as autotrophic pathway.

Table 4.1 Metabolic abilities of 15 selected genomes, including respiratory processes, central carbon metabolism, H₂ oxidation linked to respiratory processes, carbon fixation and fermentation pathways. The table also includes, for each genome, its average proportion and its proportion within each sample. The same table for a larger set of draft genomes is presented in Table A-2.6.

Metabolic ability	Cluster	Taxonomic affiliation	Proportion heatmap (across the 16 metagenomic samples)																							
			Average proportion (%)	Respiration	Central carbon metabolism			Carbon fixation				Fermentation pathways														
				Dissimilatory sulfate reduction	Dissimilatory nitrate reduction	Denitrification	Glycolysis	Pentose phosphate pathway	Entner-Doudoroff pathway	Citric acid cycle	Wood-Ljungdahl pathway	Key enzyme*	Calvin cycle	Key enzyme**	Hydroxypropionate-hydroxybutyrate cycle	3-Hydroxypropionate bi-cycle	Reductive citric acid cycle	Dicarboxylate-hydroxybutyrate cycle	Number of group-1 [NiFe]-hydrogenases	Succinate production	Formate production	Ethanol production	Butanol production	Butyrate production	Acetate production	Lactate production
Autotrophic H ₂ oxidizing SRB	c12	<i>Desulfatitalea</i>	1.4%									yes							4							
	c16a	<i>Desulfobulbaceae</i>	42.3%									yes							5							
	c23	<i>Peptococcaceae</i>	0.5%									yes							11							
	c8a	<i>Peptococcaceae</i>	1.6%									yes							2							
Autotrophic non-H ₂ oxidizing SRB	c4a	<i>Peptococcaceae</i>	0.7%									yes							0							
	c4b	<i>Peptococcaceae</i>	0.5%									yes	yes						0							
Autotrophic H ₂ oxidizing non-SRB	c57	<i>Rhodospirillaceae</i>	10.2%										yes						5							
	c41	<i>Roseovarius</i>	8.3%									yes							2							
Autotrophic non-H ₂ oxidizing non-SRB	c20a	<i>Clostridiaceae</i>	0.6%									yes							5							
	c9	<i>Hoeflea</i>	1.1%									yes	yes						0							
Heterotrophic non-H ₂ oxidizing non-SRB	c0	<i>Gammaproteobacteria</i>	0.3%																0							
	c22	<i>Hyphomonas</i>	3.7%																0							
	c29	<i>Hyphomonadaceae</i>	6.6%																0							
	c35	<i>Pseudomonas</i>	1.2%																0							
	c54	<i>Flavobacteriales</i>	0.2%																0							

*Key enzyme for Wood-Ljungdahl pathway is carbon monoxide dehydrogenase (EC 1.2.99.2)

**Key enzyme for Calvin cycle is ribulose-bisphosphate carboxylase (EC 4.1.1.39)

Organisms that cannot reduce SO₄²⁻ are expected to use another electron acceptor, or to use a fermentative mode of growth. The other putative electron acceptors does not include NO₃⁻ that was never measured in any porewater sample withdrawn from the borehole (results not shown), but include Fe(III) and intermediate sulfur species, as SO₃²⁻. Ferric minerals were never reported in Opalinus Clay,

except after oxidation of pyrite (Pearson *et al.*, 2003). This indicates that the presence of Fe(III) in Opalinus Clay porewater is an experimental artefact due to exposure to air.

In our system, Fe(III) is more likely to be used as electron acceptor in the early anoxic phase, when S(-II) is not present. Indeed S(-II) can abiotically reduce Fe(III)-bearing minerals, which produce S⁰ (Mortimer *et al.*, 2011), making Fe(III) less available for bacteria. This means that this process is another source of intermediate valence sulfur species from S(-II) oxidation, besides the one caused by O₂ injection as described in section 4.4.2. Overall, these considerations mean that the presence of both Fe(III) and intermediate valence sulfur species as electron acceptors ensue from a contamination of the borehole with O₂, and thus that their biological use as electron acceptor is not expected to take place in undisturbed Opalinus Clay porewater that was never in contact with air.

4.4.4 Metabolic activities of draft genomes

Metaproteomic analysis allowed the reconstruction of the active metabolic routes for the most abundant bacteria. Two organisms are responsible for primary production in this ecosystem: Desulfobulbaceae (cluster c16a) and Rhodospirillaceae (cluster c57). Proteomic results for Desulfobulbaceae confirmed its autotrophic capabilities by providing evidence for all the proteins in the reductive acetyl-CoA pathway (Figure 4.10). The ability of Rhodospirillaceae to fix carbon was also evidenced by its proteome (Table A-2.12), but unlike Desulfobulbaceae, it uses the Calvin cycle. These two autotrophic microorganisms both utilize H₂ as an electron donor, as they both harbor group 1 [NiFe]-hydrogenases (Table A-2.12), but they differ in their electron accepting process. Desulfobulbaceae is a sulfate-reducing bacterium and includes the entire dissimilatory sulfate-reducing pathway in its proteome (Figure 4.10). In contrast, in Rhodospirillaceae, a sulfite reductase was identified in the proteome (Table A-2.12). Thus, this organism is likely to use intermediate valence sulfur species as terminal electron acceptors. As explained in section 4.4.3, this means that Rhodospirillaceae c57 is not expected to use the electron acceptor in undisturbed Opalinus Clay, where O₂ contamination never occurred.

The five other microorganisms are heterotrophic as surmised by the absence of [NiFe]-type 1 hydrogenases and/or of carbon fixation pathway from their proteome, even though some were defined as autotrophic based on their genomes (clusters c8a, c12 and c23; Table 4.1 and Table A-2.12). *Hyphomonas* (cluster c22) is the only organism that can degrade organic macromolecules, such as proteins and RNA. We propose that *Hyphomonas* sp. serves to ferment organic compounds derived from dead biomass, producing acetate. This conclusion was reached based on the absence of a clear electron acceptor and based on the presence of proteins involved in acetate production in its proteome. This organism harbors proteins involved in respiratory processes, such as ATP synthase and electron transport proteins (Table A-2.12). This means that this bacterium has the capability to respire, retains

respiratory proteins even when it switches to fermentation because no electron acceptor is present. This strategy makes sense, as it allows the bacterium to switch back instantly to a respiratory metabolism when an electron acceptor is available. Remarkably, the latter contains 17 different iron complex outer membrane receptor proteins, suggesting an important role for iron in biosynthesis (Table A-2.12).

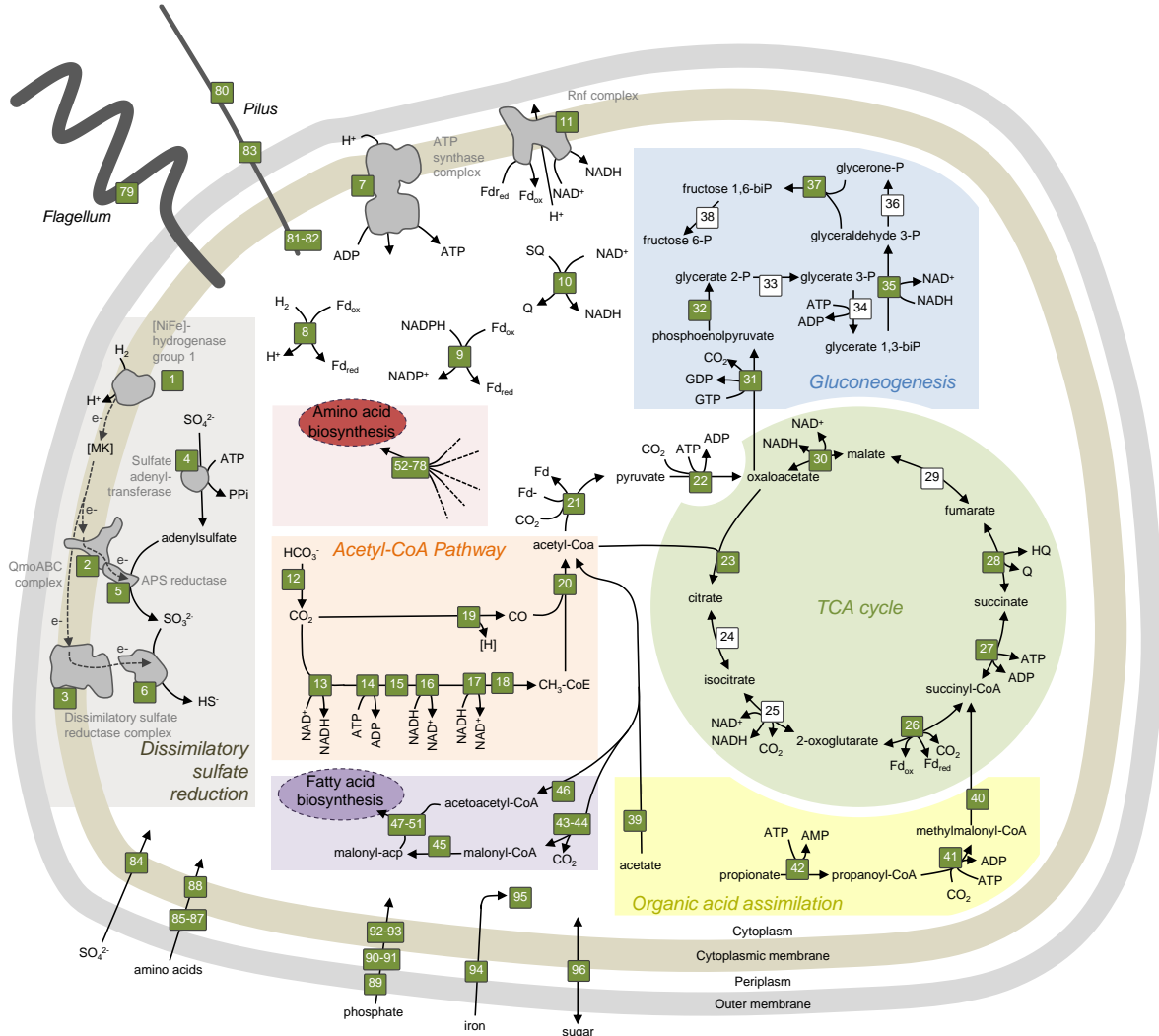


Figure 4.10 Metabolic pathways of *Desulfobulbaceae* c16a, according to proteomic (green boxes) and genomic (white boxes) data. This microorganism reduces SO_4^{2-} to HS^- using electrons from H_2 (grey box), thus producing a proton gradient across the cytoplasmic membrane. The intracellular proton translocation is coupled to ATP generation. ATP and electrons (the latter being transferred via reduced ferredoxin and NADH/NADPH) can be used to reduce CO_2 to acetyl-CoA via the reductive acetyl-CoA pathway (orange box). Biomass biosynthesis takes place through the TCA cycle (green box), and gluconeogenesis (blue box). Fatty acids and amino acids biosynthesis is not described in detail, but the proteins involved are listed in violet and red boxes, respectively. Finally, it can also use acetate and propionate as a source of carbon (yellow box). All proteins of this figure are listed in Table A-2.11.

Furthermore, the proteome of cluster c4a, a member of the *Peptococcaceae* family, shows an abundance of aromatic compound degradation proteins suggesting that it uses complex organic

compounds derived from dead biomass, in addition to ethanol, butyrate and formate, as electron donors and carbon sources (Table A-2.12). The other three genomes represent Gram-positive SRB with proteomic evidence supporting utilization of acetate and other organic acids and/or ethanol as electron donors (clusters c8a, c12 and c23; Table A-2.12). These four heterotrophic SRB can oxidize carbon to CO_2 using the oxidative acetyl-CoA pathway (Table A-2.12).

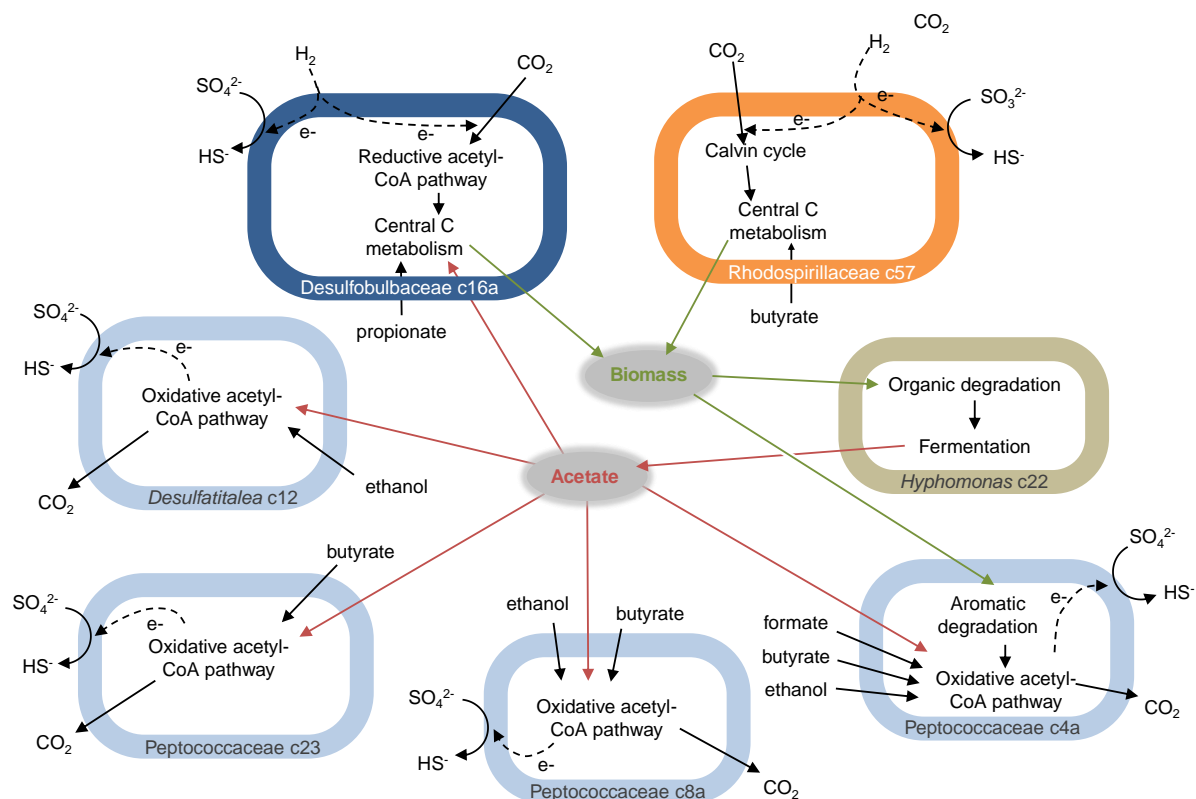


Figure 4.11 Metabolic interactions between seven microorganisms, based on proteomic data. Carbon fixation is carried out by a sulfate-reducing bacterium (dark blue) and by a non-sulfate reducing bacterium (orange). Both use H_2 as an electron source. All other organisms are heterotrophic. A fermenting bacterium can oxidize organic macromolecules to acetate (brown), while sulfate-reducing bacteria oxidize acetate to CO_2 (light blue). All protein data used to build this metabolic interaction are listed in Table A-2.12. The processes linked to complex organic compounds are indicated in green, and the processes linked to fermentation products are indicated in red.

4.4.4.1 Carbon cycling

From careful biochemical pathway annotation of all seven genomes, we were able to identify whole community carbon cycling (Figure 4.11 and Table A-2.12). The two autotrophic organisms fix CO_2 and produce biomass. *Hyphomonas* sp. degrades dead biomass and produces acetate. However, we observe no net accumulation of acetate in the water (Figure 4.7), presumably due to its oxidation to CO_2 by the four heterotrophic SRB (clusters c4a, c8a, c12 and c23) via the oxidative acetyl-coA pathway. Other fermentation products (i.e., ethanol, butyrate, and formate) are also oxidized by these SRB, suggesting that the fermentation pathways of this system were not all identified. The two autotrophic bacteria also

appear to use fermentation products as a carbon source, because they harbor proteins able to metabolize acetate and butyrate (Table A-2.12).

4.4.1 Hydrogen concentration range for Desulfobulbaceae c16a

Desulfobulbaceae c16a is adapted to a wide range of H₂ concentrations. Indeed, this organism dominates the microbial community in the early anoxic phase, when H₂ input, delivered through the gas permeable membrane, was only limited by gas/water equilibrium (Figure 4.9, Figure 4.4). Later, starting at day 77, the membrane was clogged and the H₂ input dropped drastically. As a result, the overall concentration of planktonic cells decreased but, as shown by 16S rRNA sequencing, Desulfobulbaceae remained the most abundant microorganism. These data suggest that this organism is able to use H₂ as an electron donor at concentrations ranging from saturated to non-detectable.

4.4.2 Putative Opalinus Clay indigenous bacteria

A salient limitation of deep subsurface microbial sampling is the possibility of contamination. In order to address this possibility, 16S rRNA genes from genomes were compared to OTUs observed in 23 porewater samples presented in Chapter 3 (Table A-1.5) and originated from 8 other boreholes across the Mont Terri URL (Figure 2.3.A). We could detect the two autotrophic bacteria (Desulfobulbaceae c16a and Rhodospirillaceae c57) in all 23 samples (Table A-2.5). Thus, they are surmised to be indigenous to Opalinus Clay rock. Average nucleotide analysis (Table A-2.7) of these two genomes revealed that they are distinct from known genomes, likely constituting new families, which support an Opalinus Clay origin.

These findings suggest that autotrophic hydrogen oxidizing bacteria are extant in undisturbed Opalinus Clay, without any H₂ amendment. This is not surprising, considering reports of the presence of H₂ in pristine boreholes (Pearson *et al.*, 2003). The origin of H₂ in this formation is uncertain but it serves as an energy source for this microbial ecosystem.

4.5 Conclusion

The present study clearly shows that H₂ consumption occurs in Opalinus Clay, and that it is a biologically driven reaction. This process has a significant positive impact on the safety of nuclear waste storage, because it limits the H₂ pressure build-up that could eventually threaten the repository integrity.

This work also provides detailed insights into the metabolic interactions of microorganisms in a photosynthesis-independent subsurface ecosystem where H₂ is the primary energy source. Most

remarkably, we uncovered carbon cycling from autotrophic organisms, acting as primary producers, through a fermenting organism degrading dead biomass and producing acetate, to heterotrophic SRB oxidizing acetate back to CO₂. The current study provides a unique understanding of carbon cycling in subsurface microbial systems and a glimpse into microbial metabolism on early Earth.

4.6 Acknowledgments

We thank Manon Fruttschi for her invaluable help during sampling. We are grateful to the Swisstopo team of the Mont Terri Rock Laboratory that provides the best possible conditions for scientific research in the URL. Thierry Theurillat and Nicola Kern, from Swisstopo, performed some samplings and maintenance tasks, we are indebted to them for their help. We thank also Yanick Lettry, Thomas Fierz and Herwig Mueller for their help in the design and functioning of the sampling and recirculation system. The work conducted by the U.S. Department of Energy Joint Genome Institute (JGI) is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. JGI is acknowledged for providing resources for the sequencing analysis through the grant CSP 1505. A.F.A. is supported by a grant from the Swedish Research Council VR (grant 2011-5689). Computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC).

Chapter 5 Biological hydrogen and sulfate consumption rate

This fifth chapter describes in details how sulfate and hydrogen consumption rates have been estimated in the hydrogen injection experiment, through two specific *in situ* experiments.

This chapter will be submitted to Applied Geochemistry journal.

All the work presented in this chapter was carried out by the student, except the $\delta^{34}\text{S}\text{-SO}_4$ analysis that was performed by Hydroisotop GmbH (Schweitenkirchen, Germany).

Abstract

Microbial activity can impact the safety of deep geological nuclear repositories in several ways. Recent studies showed that H₂ is consumed primarily by autotrophic sulfate-reducing bacteria in Opalinus Clay, a rock formation studied as a model for clayey host rocks for nuclear waste. Anoxic steel corrosion of canisters is expected to release gaseous hydrogen, potentially causing overpressurization of disposal vaults. Thus, microbial consumption of H₂ is globally beneficial. However, to fully consider this biological process in models of repository change over time, it is crucial to determine the *in situ* rates of microbial hydrogen oxidation and sulfate reduction. These rates were estimated through two distinct *in situ* experiments, using several measurement and calculation methods. Volumetric consumption rates are about 1.48 μmol·cm⁻³·day⁻¹ for H₂, and 0.17 μmol·cm⁻³·day⁻¹ for sulfate. This means that electrons derived from H₂ reduce electron acceptors other than sulfate, confirming previous findings of CO₂ fixation and Fe(III) reduction. These rate estimates are critical to evaluate whether biological H₂ consumption can negate H₂ production in repositories, and to determine whether sulfate reduction can consume sulfate faster than it is replenished by diffusion, which could lead to methanogenic conditions.

5.1 Introduction

The disposal of radioactive waste requires long-term confinement to allow for radionuclide decay, which can last up to a 1 million years. For that reason, in several countries, nuclear waste is to be placed in deep geological repositories that are disconnected from the surface (Wildi *et al.*, 2000). Such an undertaking implies extensive knowledge of the host rock and its behavior under the variety of conditions that are relevant for the repository, in order to have a predictive understanding of the change in the waste over time.

A few European countries have opted for a repository design that consists of carbon steel containers in which the waste is sealed, and that are placed in a low-permeability Clay rock formation. At the time scale of the repository, the steel is expected to corrode quickly, producing Fe(II) and gaseous H₂ (Marsh and Taylor, 1988). Another expected source of H₂ is the radiolysis of water (Le Caër, 2011). H₂ build-up will lead to an increase in pressure that can threaten the success of the disposal, by damaging the containment barriers (Bonin *et al.*, 2000). However, H₂ is also known to be an excellent source of energy that can support microbial growth and activity (Nealson *et al.*, 2005). In that sense, microbial activity could be beneficial to the long-term safety of the repository (Libert *et al.*, 2011).

In Opalinus Clay, which is the envisaged host rock for deep geological repository in Switzerland, H₂ is presumed to be consumed via biological processes because it disappeared 20 times faster than accounted for by abiotic processes (Vinsot *et al.*, 2014). Additionally, an *in situ* experiment, presented in Chapter 4, clearly showed that H₂ was oxidized microbiologically through a complex metabolic network. Thus, it is expected that hydrogen gas will be depleted by biological processes in the repository, but its rate of consumption remains uncertain. Determination of the *in situ* kinetics of H₂ oxidation is essential in order to complete the geochemical modeling of the repository and to account for the biological contributions to H₂ consumption.

Furthermore, the rate of sulfate reduction is a key parameter in evaluating whether sulfate depletion via hydrogen-fueled microbial sulfate reduction will exceed the influx of sulfate via diffusion from the rock. The depletion of sulfate, if it occurs, may enable the growth of methane-producing microorganisms, which would negatively impact repository safety because its accumulation, just like that of H₂, may lead to pressure build-up (Libert *et al.*, 2011).

Thus, the focus of the present work is to assess the rates of H₂ oxidation and sulfate reduction *in situ* in the Opalinus Clay formation.

5.2 Materials and Methods

5.2.1 Experimental set-up

The experimental set-up is described in section 4.2.1. Briefly, a vertical descending borehole (borehole BRC-3) at the Mont Terri Underground Rock Laboratory (URL) was equipped with a hydraulic stopper in order to isolate borehole porewater from the oxic atmosphere of the gallery (Figure 5.1). Polyamide lines allowed water or gas sampling and injection from the surface. The set-up in the gallery (referred to as surface equipment) enabled water recirculation using a peristaltic pump (at a flow rate of 5 mL/min), water sampling through needle valves, and hydrogen gas injection using a gas reservoir connected to the water circuit through a gas permeable membrane. A schematic representing H₂ injection as a function of time is presented in Figure 5.2.

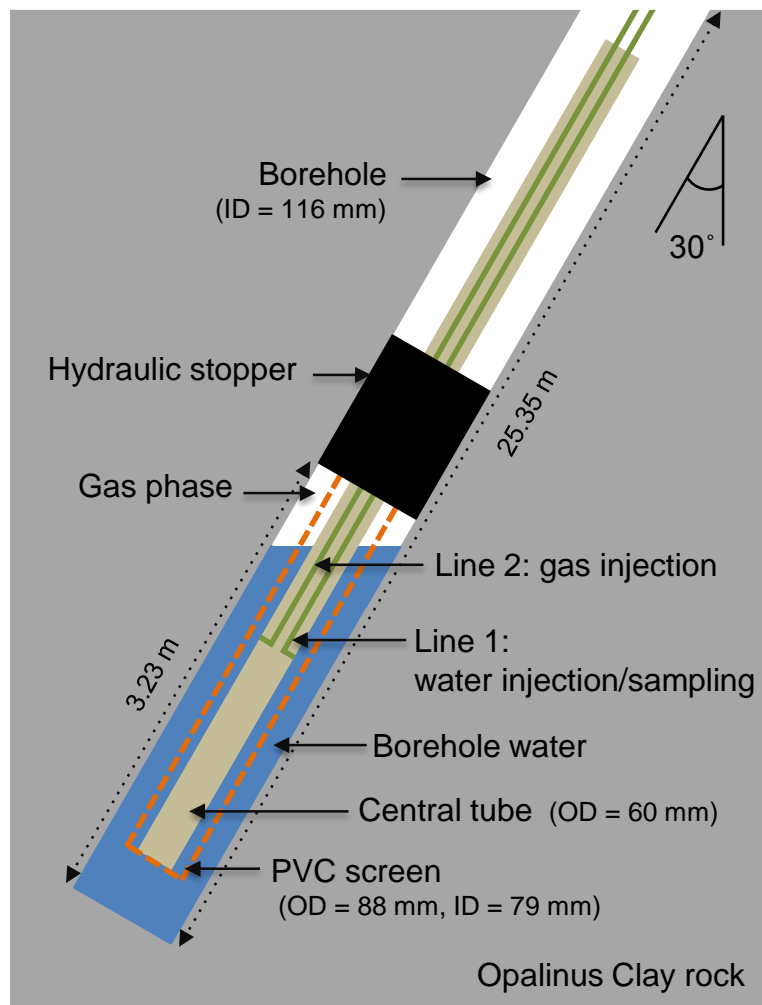


Figure 5.1 Borehole equipment of BRC-3, which has a vertical descending orientation, forming a 30° angle with the zenith. Downhole equipment includes a central tube containing sampling lines (line 1 for water sampling and injection; line 2 for gas injection), a hydraulic packer and a PVC screen.

At the start of the experiment, bromide was amended to the borehole water as a conservative tracer. Hydrogen gas was supplied to the borehole for over 500 days. During the first 139 days, this amendment was made via the gas permeable membrane. Later, H₂ was directly injected into the borehole, creating a gas phase. Water circulation was discontinued starting at day 178. Approximately once a week, water (between 400 and 1000 mL) was sampled and replaced by artificial porewater (APW). APW I was the replacement solution until day 324, after which APW II was used (Table 5.1). These large sample volumes were collected for molecular analysis of the microbial community, as described in section 4.2.4, but a small volume was also preserved from each sample to quantify major anion concentrations.

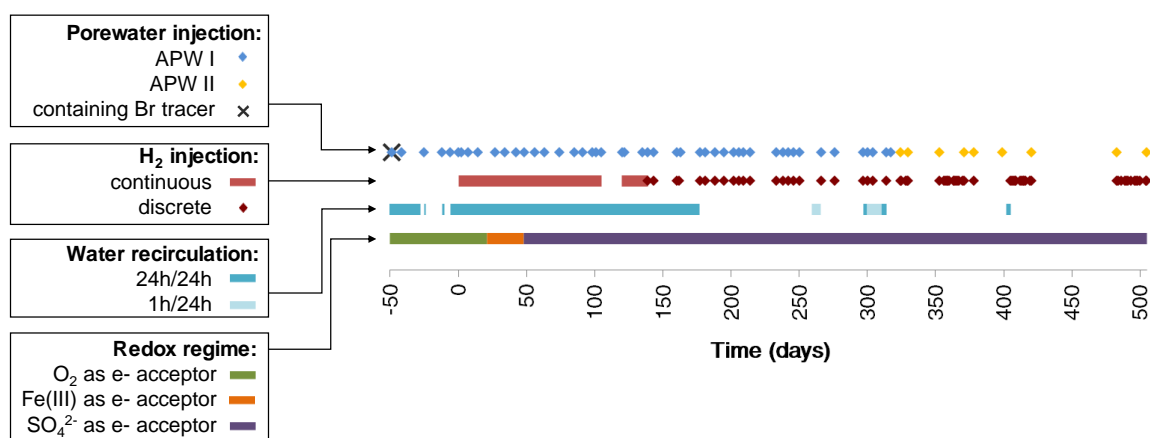


Figure 5.2 Timeline of the H₂ injection experiment in BRC-3 borehole. Day 0 corresponds to the first injection of H₂ in borehole water. On the first line, the black cross indicates when the Br tracer was injected, the blue squares when APW I was injected and the yellow squares when APW II was injected in the borehole. On the second line, the red line indicates when H₂ was continuously injected to the borehole water, and the red diamonds indicate individual H₂ injections into the borehole. On the third line, the dark blue line indicates when borehole water was continuously recirculated, and the light blue line when borehole water was recirculated one hour per day. On the fourth line, the green line indicates that O₂ when is the electron acceptor used by the microbial community, the orange line when Fe(III) is used, and the purple line when SO₄²⁻ is used (Figure 4.3).

5.2.1 Experimental procedures

During the 500 days of hydrogen gas injection, three experiments were carried out specifically to evaluate the rates of hydrogen oxidation and of sulfate reduction in the borehole.

The first experiment (repeated sampling) was carried out by repeated sampling of the borehole on day 378, one week after the last APW I injection, and 24 hours after the last injection of 1L of hydrogen gas. Borehole water was withdrawn without pumping, simply using the fact that the borehole pressure was in excess of gallery pressure. The water was withdrawn in 100 mL aliquots, to investigate the changing composition of the water as a function of distance from the gas phase (Figure 5.3). Each 100

mL sample was filtered through a 0.2 μm filter, and analyzed for SO_4^{2-} , S(-II) and dissolved H_2 . In total, 10 samples were collected before the gas phase reached the sampling line.

The second experiment (repeated injection) was conducted on day 405, 3 days after the injection of APW II and recirculation of borehole water. The pump was shut down, and water collected using borehole pressure. After discarding the equivalent of the volume of the line, water was analyzed for SO_4^{2-} , Br^- , S(-II), dissolved H_2 and $\delta^{34}\text{S-SO}_4$. Immediately after sampling, 1 L of H_2 gas was injected into the borehole, and a new sample collected to quantify the concentration of dissolved H_2 resulting from this injection. The combination of hydrogen injection, followed by water sampling, was repeated in total 11 times within a period of 15 days.

Table 5.1 Major anions and cations composition of APW I, APW II, three BRC-3 porewater samples before any H_2 injection (sample 4 was analyzed two times), and 5 selected BRC-3 porewater samples after H_2 was injected.

Sample	Na^+ (mM)	Mg^{2+} (mM)	K^+ (mM)	Ca^{2+} (mM)	Cl^- (mM)	SO_4^{2-} (mM)	HCO_3^- (mM)	
Artificial porewater	APW I	255.59	17.89	1.46	15.79	294.58	14.70	0.57
	APW II	270.00	23.02	2.12	18.88	307.76	23.83	0.57
Initial BRC-3 borehole water (before any H_2 injection)	sample 1	329.46	26.32	2.43	21.51	288.84	25.37	n. d.
	sample 2	289.79	24.66	2.01	19.06	261.60	26.85	n. d.
	sample 3	234.80	16.61	1.73	17.76	243.02	16.60	n. d.
	sample 4a	284.26	24.05	2.32	17.05	302.07	24.90	n. d.
	sample 4b	284.02	23.44	2.13	19.01	304.00	25.43	n. d.
BRC-3 borehole water (after H_2 was injected)	at day 0	320.14	23.25	2.15	18.79	299.69	17.88	n. d.
	at day 101	285.31	22.70	1.05	16.53	311.02	16.58	n. d.
	at day 238	269.69	10.08	0.48	15.22	273.41	12.93	n. d.
	at day 304	239.59	16.41	0.56	16.11	258.20	14.61	n. d.
	at day 420	n. d.	n. d.	n. d.	n. d.	317.10	5.88	n. d.

'n. d.' stands for 'not determined'

The third experiment (laboratory incubation) took place in the laboratory, using 950 mL of borehole water collected anoxically and under sterile conditions on day 314. A one-liter bottle containing the borehole water was amended with 1.5 mL of 150 mM KH_2PO_4 (pH 7.2), and its headspace was flushed with 100% H_2 to a pressure of 0.85 relative bar. The bottle was then incubated at 17.5°C and sampled periodically over a period of 107 days (10 samples were collected) to quantify SO_4^{2-} , S(-II) and $\delta^{34}\text{S-SO}_4$.

5.2.2 Chemical analyses

For major cation and anion quantification, a sample of 2 mL of water was filtered using a 0.2 μm filter, and stored at $-20\text{ }^{\circ}\text{C}$ until analysis. Immediately before analysis, each sample was re-filtered and diluted 10, 100 and 1,000 times with MilliQ water. Samples were analyzed by ion chromatography (DX-3000, Dionex, Sunnyvale, USA) with an IonPac CS16 Cation-Exchange column and 40 mM of methanesulfonic acid for elution, to measure major cations, and with an IonPac AS11-HC column and KOH gradient from 0.5 to 30 mM for the elution, to measure major anions.

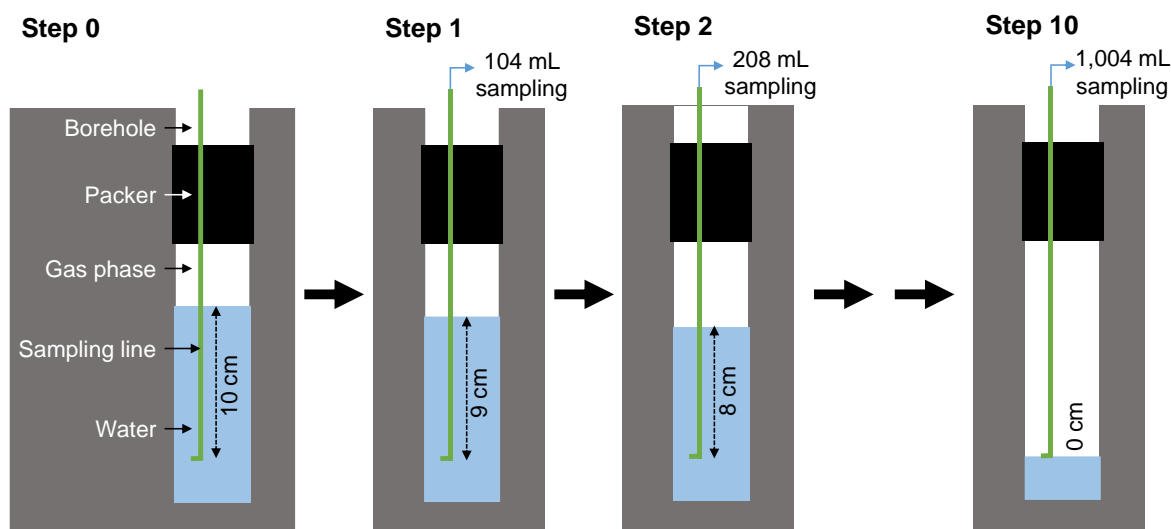


Figure 5.3 Repeated sampling experiment: because the surface area of water/gas interface in the borehole is 104 cm^2 , the water level goes down by 1 cm every time 104 mL are sampled. This means that by repeated sampling, we can access water that is closer to the headspace, where hydrogen gas is dissolved in the water. This figure is a simplified view that does not account for borehole incline and for the presence of a central tube and a screen, as shown in Figure 5.1.

$\delta^{34}\text{S}\text{-SO}_4$ was measured on water samples filtered through 0.2 μm filters. Prior to analysis, S(-II) was removed by reaction with Zn acetate at a concentration of 1% (w/v) and incubation for 2 hours at room temperature. Subsequently, the solution was centrifuged at $4,500 \times g$ for 10 minutes and the pellet discarded. Analysis of $\delta^{34}\text{S}\text{-SO}_4$ was carried out by Hydroisotop (Schweitenkirchen, Germany). Briefly, BaSO_4 was precipitated by the addition of BaCl_2 and 0.2 mg of the precipitate mixed with NiO_4 into tin cups, and the mixture placed into an auto-sampler for analysis by a Carlo-Erba Elemental Analyzer (CHN EA1110) for combustion to SO_2 . The clean SO_2 was carried by helium to a Thermo Finnigan DELTA plus XL Continuous Flow Stable Isotope Mass Spectrometer for sulfur isotope ratio analysis. Results were corrected to sulfate standards IAEA-SO-5, IAEA-SO-6 and NIST-127 (all BaSO_4), and/or IT2 substandard IT2-42 (BaSO_4). 'IT2' denotes Internal Standards with values calibrated to NIST International Standards. The 1-sigma error for clean standard material is $\pm 0.3\text{‰}$ for sulfur.

S(-II) was determined by the method described by Cline (Cline, 1969) on samples filtered through 0.2 μm filters. Dissolved hydrogen was measured using the procedure described in section 4.2.2.5. A volume of about 10 mL of water was recovered in a serum bottle that was previously flushed with N_2 , at a known final pressure, and contained HgCl_2 . After sampling, the bottle was incubated horizontally at 37°C for at least 24 hours to equilibrate the gas phase with the aqueous phase. The gas phase composition was analyzed using a GC-FID (Varian 450-GC, Agilent, Santa Clara, USA). The dissolved H_2 concentration was calculated using the bottle's initial pressure, the bottle volume, the sampled water volume, the incubation temperature and Henry's Law constant for H_2 .

5.2.3 Rate calculation

Hydrogen and sulfate consumption rates (R) were assessed in several ways.

1. First, the rate R of hydrogen or sulfate consumption was obtained simply by calculating the concentration C difference during a time interval t .

$$R [\mu\text{mol} \cdot \text{cm}^{-3} \cdot \text{d}^{-1}] = \frac{dC [\mu\text{mol} \cdot \text{cm}^{-3}]}{dt [\text{d}]}$$

2. The rate of hydrogen oxidation was also assessed using Fick's first law of diffusion. When hydrogen diffuses from the gas phase and is consumed in the water phase, assuming a steady state, the flux of hydrogen F can be calculated according to this equation:

$$F [\mu\text{mol} \cdot \text{cm}^{-2} \cdot \text{s}^{-1}] = D_w [\text{cm}^2 \cdot \text{s}^{-1}] \cdot \frac{dC [\mu\text{mol} \cdot \text{cm}^{-3}]}{dz [\text{cm}]}$$

in which $D_w = 5.13 \cdot 10^{-5} \text{ cm}^2 \cdot \text{s}^{-1}$ and z represents the sample height (Jähne *et al.*, 1987). The volumetric rate of hydrogen consumption R can be calculated by dividing the hydrogen flux by the height:

$$R [\mu\text{mol} \cdot \text{cm}^{-3} \cdot \text{d}^{-1}] = \frac{F [\mu\text{mol} \cdot \text{cm}^{-2} \cdot \text{d}^{-1}]}{dz [\text{cm}]}$$

3. Additionally, the rate of sulfate reduction was also calculated using the sulfate to bromide concentration ratio. The amount of sulfate S_r consumed by sulfate-reducing bacteria per unit volume can be calculated according to this equation:

$$S_{r_{t=i}} [\text{mM}] = \frac{C_{t=0}^{\text{SO}_4}}{C_{t=0}^{\text{Br}}} \cdot C_{t=i}^{\text{Br}} - C_{t=0}^{\text{SO}_4}$$

Then equation 1 can be used for calculating the sulfate reduction rate.

4. Finally, the rate of sulfate reduction was evaluated by measuring the change in $\delta^{34}\text{S-SO}_4$. Indeed, sulfate-reducing bacteria will preferentially reduce the lighter S isotope, resulting in an increase in $\delta^{34}\text{S-SO}_4$ (Canfield, 2001). First, the ratio between the change in $\delta^{34}\text{S-SO}_4$ and the change in sulfate concentration over time R_{fr} was determined in the laboratory and calculated using this equation:

$$R_{fr} [\text{‰} \cdot \text{mM}^{-1}] = \frac{d(\delta^{34}\text{S} - \text{SO}_4) [\text{‰}]}{dC [\text{mM}]}$$

5. In a second step, the measured change in $\delta^{34}\text{S-SO}_4$ *in situ* is translated to a sulfate concentration reduced by bacteria Sr , using this equation, before calculating the sulfate reduction rate with the first equation shown in this section.

$$Sr [\text{mM}] = \frac{d(\delta^{34}\text{S-SO}_4) [\text{‰}]}{R_{fr} [\text{mM}^{-1}]}$$

5.3 Results

5.3.1 Long-term change in anions concentration

The concentration of sulfate decreases during the first 276 days of hydrogen injection, then increases till day 330 before dropping again at day 420 (Figure 5.4). In contrast, the concentration of chloride is stable and the concentration of bromide decreases to near zero (Figure 5.4).

5.3.2 Repeated sampling

The results of the repeated sampling experiment are presented in Figure 5.5. To determine the position of samples in the water column relative to the sampling line, the volume of water withdrawn prior to sampling was divided by the surface area of the water-gas interface. The latter was calculated to be 104 cm^2 using the geometry of the borehole introduced in Figure 5.1. The sulfate concentration decreased and the sulfide concentration increased as the samples were collected from volumes progressively closer to the gas phase. Based on S(-II) production, the rate of sulfate reduction is estimated to be $0.08 \text{ } \mu\text{mol} \cdot \text{cm}^{-3} \cdot \text{day}^{-1}$ (Table 5.2). The concentration of dissolved hydrogen was stable in the first 6 samples, and then steadily increased as the samples approached the gas phase. The rate of hydrogen consumption was estimated for these two phases, using the change in concentration over time for the first one, and Fick's first law of diffusion for the second. Rates are presented in Table 5.2.

5.3.3 Repeated injection

Results of the repeated injection experiment are shown in Figure 5.6. The sulfide concentration steadily increased while the hydrogen concentration fluctuated up and down, depending on whether the sample was recovered immediately or one day after the hydrogen injection (Figure 5.6.A). Both sulfate and bromide concentrations decreased abruptly two days into the experiment, before slowly increasing. The change in their ratio shows steady sulfate depletion relative to bromide. $\delta^{34}\text{S-SO}_4$ appears to be stable over the course of the experiment, except for the last data point, which is almost 2 ‰ greater than the last measurements (Figure 5.6.B). Hydrogen and sulfate reduction rates based on hydrogen and sulfide concentration changes over time, and on sulfate bromide ratios are presented in Figure 5.6.C and Table 5.2. Finally, the *in vitro* experiment, during which 1.4 mM of sulfate was reduced and $\delta^{34}\text{S-SO}_4$ was increased by 0.9 ‰, allowed the estimation of R_{fr} (0.64 ‰·mM⁻¹; Figure 5.7). The rate of sulfate reduction could then be back-calculated from *in situ* $\delta^{34}\text{S-SO}_4$ measurement described above and was found to be 0.20 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$ (Table 5.2).

5.4 Discussion

5.4.1 Long-term change in sulfate concentration

It is challenging to measure the rate of sulfate reduction in a borehole. This is because there is the potential for continued input of sulfate from the formation. In this particular experiment, the system was further complicated by the replacement of the large water samples collected for DNA analysis with artificial porewater (APW), as explained in section 4.2.4. In the first part of the experiment (from day 0 to day 324), the composition of APW (APW I) was based on another borehole in the URL (Table 5.1; Pearson, 2002), where extensive characterization of porewater geochemistry took place. However, the sulfate concentration at that location is lower than that at the BRC-3 borehole (Table 5.1). To correct this mismatch between APW and borehole water sulfate concentration, a second APW (APW II) was prepared with a higher sulfate concentration (Table 5.1). Another challenge was related to the fact that mixing was limited in the borehole, resulting in significant concentration gradients, particularly for H₂ because the only source of H₂ was the gas phase at the top of the water column. Additionally, the gas/water interface, where sulfate reduction is likely to take place, was located above the water sampling line, making it difficult to sample this volume directly (Figure 5.1).

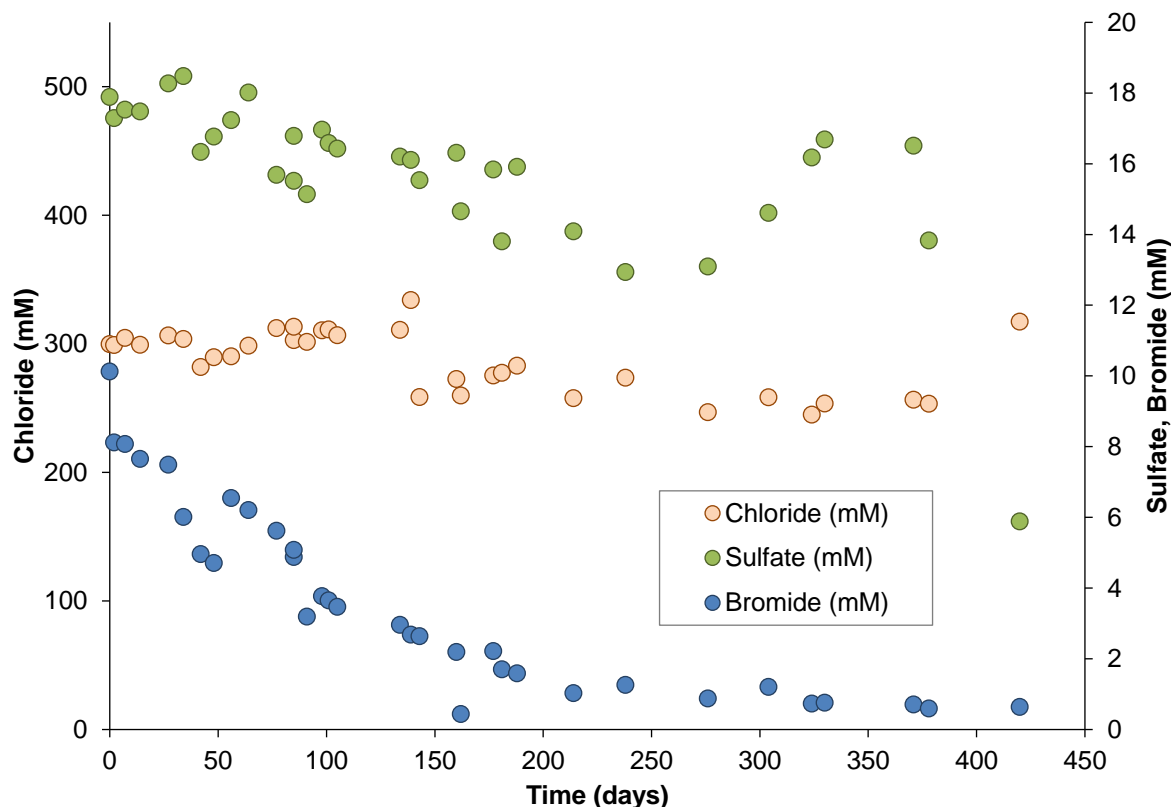


Figure 5.4 Change in borehole water concentration of sulfate (green circles), chloride (light orange circles) and bromide (blue circles) over time. The decrease in sulfate concentration observed during the first ~300 days is mainly caused by borehole water dilution with APW I that has a lower sulfate concentration. Later, after day 324, APW II composition was adapted to initial borehole water composition, with a higher sulfate concentration, explaining why sulfate concentrations increase (Table 5.1). Bromide concentrations decreased over time because this tracer was only injected as the start of the experiment and was diluted by each APW injection.

The concentration of sulfate decreased continuously overtime until day 276 (Figure 5.4). This decrease is due primarily to the injection of replacement APW with a lower sulfate concentration. After day 324, the sulfate concentration tended to increase due to amendment the new APW containing a higher sulfate concentration (APW II, Table 5.1). The last sulfate measurement is significantly lower, less than half of the last one (Figure 5.4). This decrease is likely due to the coupled effect of sulfate reduction and borehole homogenization. Between days 353 and 378, H_2 was injected intensively followed by a continuous borehole water circulation between days 402 and 405 (Figure 5.2). Because of the large influx of H_2 , the sulfate concentration in the water volume in close proximity to the gas phase decreased significantly. However, this consumption of sulfate consumption is not detectable in the samples obtained between days 353 and 402, because the sampling line was below the location where sulfate reduction took place. But after borehole water circulation for 3 days, mixing with low-sulfate water from above resulted in a lower concentration of sulfate. This observation perfectly illustrates the borehole water heterogeneity, and the difficulty in homogenizing it.

5.4.2 Repeated sampling

The repeated sampling experiment was designed to capture the chemical heterogeneity of the water column caused by hydrogen gas diffusion from the headspace into water. At each sampling, the water level went down, moving the gas/water interface closer to the sampling line (Figure 5.3). Using the geometry of the system, we linked the amount of water withdrawn before each sample with the original elevation of the sample relative to the sampling line. Because the sampling line is oriented horizontally, each sample is composed of water coming from above and below the line. This means that a sample recovered after withdrawing 500 mL will consist of 250 mL of water originally located above the sampling line, and 250 mL of water originally located below the sampling line, mixed at a 1:1 ratio.

The slope of the concentration of H₂ along the water column shows two different behaviors (Figure 5.5). The first six samples exhibit a near-constant hydrogen concentration, around 600 μM, and indicate that H₂ injected a few hours before is still being consumed in this part of the water column. Thus, it is possible to calculate the rate of oxidation of hydrogen by dividing the drop in concentration by the time elapsed. Depending on whether the initial H₂ concentration is taken to be the highest H₂ measurement or, to correspond to the solubility limit of H₂ under *in situ* conditions (4.24 mM), we obtain a low-end or a high-end estimate of the rate: 0.72 or 3.30 μmol·cm⁻³·day⁻¹. A better assessment of the initial H₂ concentration may be obtained by multiplying the highest measured H₂ concentration value of by two. This is because this sample, which is closest to the gas phase, has been diluted in a 1:1 ratio with water from deeper in the borehole and containing no H₂. This value gives an intermediate hydrogen oxidation rate of 1.93 μmol·cm⁻³·day⁻¹ (Table 5.2).

The last four water samples exhibit a steady increase in the concentration of H₂ (Figure 5.5). This trend suggests that H₂ diffusion from the gas phase and H₂ consumption are competing processes in this samples. If we assume steady state, we can calculate the rate of H₂ oxidation by computing the flux of H₂ from gas phase using Fick's first law of diffusion. For this calculation, we used H₂ concentrations for the 7th and 10th samples, both multiplied by two to account for the dilution by H₂-free porewater, and a diffusion coefficient of 5.13·10⁻⁵ cm²·s⁻¹ (Jähne *et al.*, 1987). The H₂ flux was determined to be 2.19 μmol·cm⁻²·day⁻¹, which corresponds to an H₂ oxidation rate of 1.13 μmol·cm⁻³·day⁻¹ (Table 5.2).

In the same experiment, we also calculated the rate of sulfate reduction by dividing the highest concentration of S(-II) reached during the experiment by the time elapsed since APW injection. Unlike for H₂, it is not necessary to multiply the S(-II) concentration by two. Indeed, this compound is produced by microbial activity in the borehole and is thus distributed more homogeneously than H₂. Nonetheless, this estimate is likely to under-estimate the rate of sulfate reduction because some dilution of sulfide is expected and soluble sulfide only represents a fraction of the total sulfide produced (due to precipitation with ferrous iron). The sulfate reduction rate obtained based on this calculation is 0.08 μmol·cm⁻³·day⁻¹ (Table 5.2).

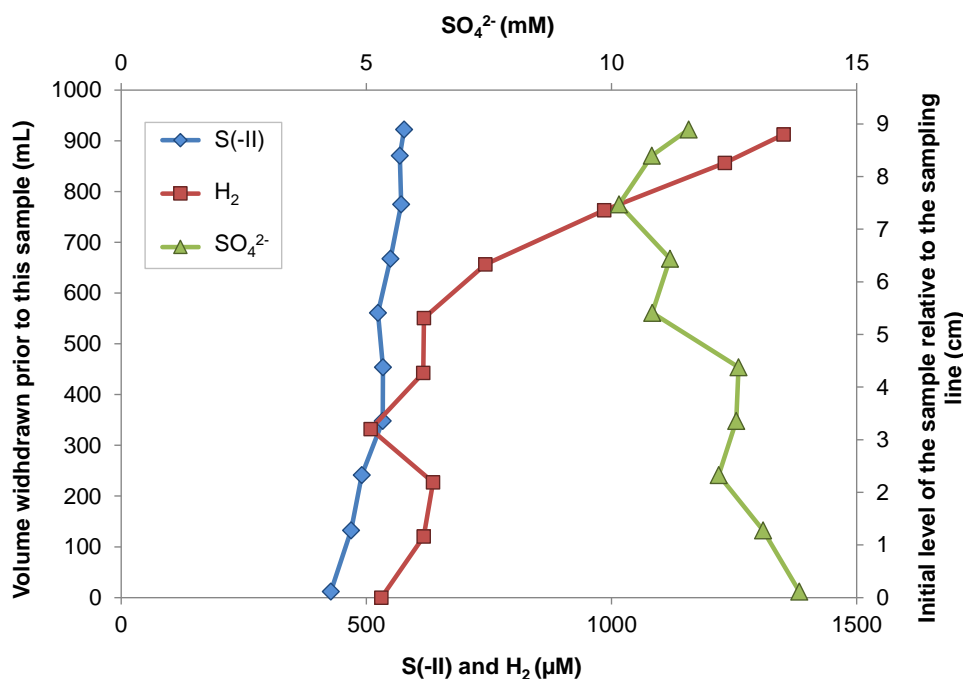


Figure 5.5 Sulfate (green triangles), hydrogen (red squares) and sulfide (blue diamonds) concentrations from repeated sampling carried out at day 378. For each sample, 100 mL of borehole water was withdrawn. As sampling proceeded, the level of the water in the borehole progressively moved closer to the sampling line, allowing collection of water closer to the headspace.

5.4.3 Repeated injection

The second experiment, involving repeated H₂ injection in a borehole where water was mixed by recirculation for 3 days, was designed to capture the rate of sulfate reduction. Every day during 2 weeks, about 150 mL of water was withdrawn for sampling, moving the gas phase progressively closer to the sampling line. Each sampling was followed by the injection of 1 L of hydrogen gas, likely homogenizing the H₂ concentration in the water above the sampling line (Figure 5.6.A).

The rate of sulfate reduction was evaluated in three different ways:

1. Using soluble S(-II) production: after 15 days, the rate of sulfate reduction is estimated to be $0.04 \mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$ (Figure 5.6.C; Table 5.2). As discussed in section 5.4.2, soluble S(-II) does not represent the total S(-II) produced by sulfate reduction, meaning that this calculation is an under-estimate.

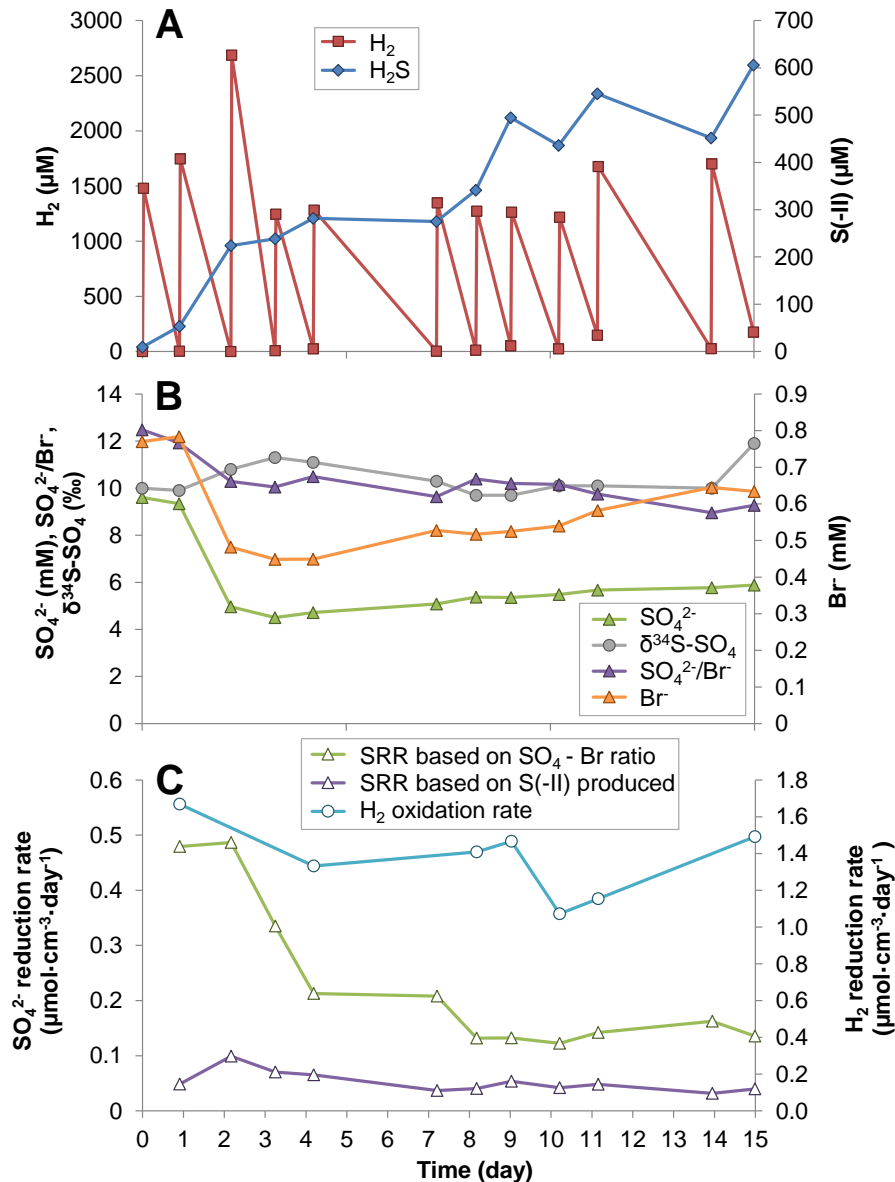


Figure 5.6 Chemical changes in the borehole for the two weeks during which H_2 was injected a total of 10 times. (A) Concentration of H_2 (red squares) and S(-II) (blue diamonds) used to calculate H_2 and SO_4^{2-} consumption rates. (B) Concentration of SO_4^{2-} (green triangles), Br^- (orange triangles), ratio between SO_4^{2-} and Br^- concentrations (purple triangles), and $\delta^{34}\text{S-SO}_4$ (blue circles). The latter two were used to evaluate the rate of SO_4^{2-} consumption. (C) Sulfate reduction rate (SSR) based on the ratio between SO_4^{2-} and Br^- (green triangles) or on sulfide production (purple triangles, and H_2 oxidation rate (blue circles).

2. Comparing sulfate and bromide concentrations: bromide serves as a conservative tracer. Hence, the steady decrease of the ratio between sulfate and bromide concentrations indicates that sulfate is consumed (Figure 5.6.B). Sulfate consumption is estimated by calculating the decrease in sulfate relative to a constant sulfate to bromide concentration ratio. After 15 days, the rate of sulfate reduction stabilized. By averaging the 6 last values, we could calculate a rate of $0.14 \pm 0.01 \mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$ (Figure 5.6.C; Table 5.2).

3. Using $\delta^{34}\text{S-SO}_4$ measurements: porewater was incubated in the laboratory in the presence of hydrogen gas, in a homogenized and closed system in order to estimate the isotopic fractionation of sulfur ($\delta^{34}\text{S-SO}_4$) during sulfate reduction by the borehole microbial community. The fractionation per mole of sulfate reduced was measured to be 0.64‰ (Figure 5.7). Throughout the *in situ* experiment, $\delta^{34}\text{S-SO}_4$ doesn't change very much except when the sample obtained is close enough to the borehole headspace to register microbial sulfate reduction and a higher $\delta^{34}\text{S-SO}_4$ value (Figure 5.6.C). Thus, the change of 1.9‰ in $\delta^{34}\text{S-SO}_4$ corresponds to 2.96 mM of sulfate reduced, which represents a rate of sulfate reduction $0.20 \mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$ (Table 5.2).

Sulfate reduction rates estimated by the bromide concentration and by $\delta^{34}\text{S-SO}_4$ are comparable. The fact that these estimates originate from two different analytical techniques and match reasonable well suggests the robustness of the approach. Moreover, by comparing these two rates with the one obtained with S(-II) concentration, we calculate that roughly 75 % of the S(-II) produced precipitates with ferrous iron within a time period corresponding to the experimental duration (15 days).

Estimating the rate of hydrogen gas oxidation is also challenging in this system due to potential losses of the gas to the formation and via small leaks through the sampling lines. However, we used the experiment last described to calculate a hydrogen oxidation rate of $1.37 \pm 0.2 \mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$ by averaging the rates obtained from several day-long incubations (excluding the second and third incubations on account of inaccurate H_2 measurements at those time points; Figure 5.6.A; Table 5.2).

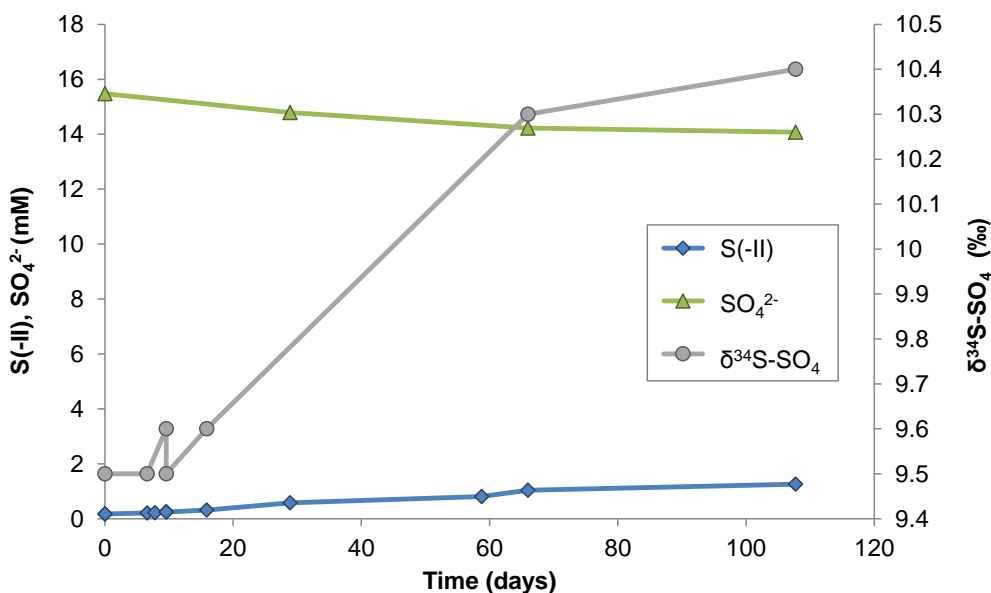
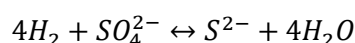


Figure 5.7 Change in the sulfide concentration, sulfate concentration, and $\delta^{34}\text{S-SO}_4$ over time in an *in vitro* batch incubation of borehole water.

5.4.4 Rates of hydrogen and sulfate consumption

Overall, hydrogen and sulfate consumption rates can be calculated by averaging obtained values from both experiments (Table 5.2), excluding low and high-end estimates. The rates of hydrogen oxidation and sulfate reduction were determined to be $1.48 \pm 0.41 \mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$ and $0.17 \pm 0.04 \mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$, respectively.

Based on the stoichiometry of hydrogen oxidation coupled to sulfate reduction, given by this equation, each mole of sulfate reduced, requires 4 moles of hydrogen oxidized.



This result suggests that, in Opalinus Clay, electrons derived from hydrogen oxidation not only reduce sulfate, but also carbon dioxide, via carbon fixation pathways, and probably also iron(III), through alternate respiratory processes, as presented in section 4.4.4. This ratio between H_2 and sulfate consumed is comparable to the one reported by Vinsot et al. (Vinsot *et al.*, 2014) in another experiment where H_2 consumption in Opalinus Clay was monitored at the Mont Terri URL. In that study, they estimated that 6 H_2 molecules were oxidized per sulfate reduced. Unfortunately, in the Vinsot et al. study (Vinsot *et al.*, 2014), no information is available on the volumetric rates of H_2 and sulfate consumption in the porewater.

Table 5.2 Summary of all H_2 oxidation and sulfate reduction rates calculated in this work.

	Value	Method	Experiment	Remark
H ₂ oxidation rate	0.72 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$	dC/dt	repeated sampling	Low-end estimate
	3.30 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$	dC/dt	repeated sampling	High-end estimate
	1.93 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$	dC/dt	repeated sampling	Used for final estimate
	1.13 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$	Fick's first law	repeated sampling	Used for final estimate
	1.37 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$	dC/dt	repeated injection	Used for final estimate
SO ₄ ²⁻ reduction rate	0.08 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$	dC/dt of S(-II)	repeated sampling	Low-end estimate
	0.04 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$	dC/dt of S(-II)	repeated injection	Low-end estimate
	0.14 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$	SO ₄ ²⁻ /Br ⁻ ratio	repeated injection	Used for final estimate
	0.20 $\mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$	δ ³⁴ S-SO ₄	repeated injection	Used for final estimate

5.5 Conclusion

The present work demonstrates that microorganisms, especially sulfate-reducing organisms, are able to rapidly oxidize hydrogen in Opalinus Clay. This biological process is beneficial for the geological disposal of nuclear waste. Indeed, hydrogen produced by anoxic corrosion of steel canisters will likely

not accumulate in the rock due to these processes, thus avoiding a pressure build up that could damage the host rock and increase its permeability.

Nonetheless, further studies are needed to fully assess the safety of repositories with respect to H₂ consumption. For instance, the geometry of the repository and how it impacts the surface area of the gas/water interface, is an important parameter controlling the rate of hydrogen oxidation. Additionally, the availability of sulfate, the main electron sink for hydrogen oxidation, will significantly impact the consumption of H₂. The concentration of sulfate in Opalinus Clay porewater ranges between 15 to 20 mM and could be locally depleted due to low hydraulic conductivity of the rock. The rate of hydrogen oxidation in the absence of sulfate, linked for example to methanogenesis, also remains an open question.

5.6 Acknowledgments

The Swisstopo crew of St-Ursanne is acknowledged for providing excellent working conditions in the Mont Terri underground laboratory. Among these people, Thierry Theurillat is thanked for his precious help for samplings. We also thank Yanick Lettry and Thomas Fierz from Solexperts AG for their help in designing and installing the borehole and surface equipment.

Chapter 6 **Conclusion and outlook**

This last chapter summarizes this entire thesis and provides direction for future developments. It covers three different conditions for microbial growth: pristine Opalinus Clay, Opalinus Clay porewater, and Opalinus Clay porewater amended with hydrogen. It also tries to briefly address the impact of the observed microbial activity on the safety of nuclear waste disposal in the deep subsurface.

6.1 Microbial activity in Opalinus Clay pristine rock

6.1.1 State of the art

As described in section 2.3.4, the small pore size distribution in Opalinus Clay inhibits its biological activity. Previous studies were able to cultivate bacteria from the rock using standard media for specific metabolic groups, which indicates that this material contains living organisms (Stroes-Gascoyne *et al.*, 2007). This task was tedious, as a small number of cultures showed evidence of microbial growth, and as most of the cultivated organisms grew very slowly. The authors of the study also tried to describe the *in situ* microbial community by DNA extraction followed by 16S rRNA genes sequencing. Due to a low microbial density, they were unsuccessful at amplify this gene, which clearly indicates that microbial density is low in undisturbed Opalinus Clay (Poulain *et al.*, 2008).

6.1.1 Attempts to extract DNA from pristine Opalinus Clay rock

One of the research project of this thesis also focused on the characterization of the microbial community in Opalinus Clay rock, using a molecular approach, namely 16S rRNA sequencing. Because we did not make significant progress in this project, it was not included in the present manuscript as a chapter, but is briefly discussed in this section. In order to overcome the problem of low microbial density, we tried to use a DNA extraction protocol developed for difficult soil samples, and that uses 50 g of rock as starting material, significantly increasing the sample size (Selenska and Klingmüller, 1991). The method already worked successfully on Opalinus Clay rock samples (Selenska-Pobell & Bachvarova, pers. comm.). We were able to reproduce it on a single Opalinus Clay rock sample that was contaminated with O₂ during its storage, which led to microbial growth. When applying the same method to samples stored in proper conditions, this extraction was very difficult to reproduce. A small amount of DNA could be extracted, with great difficulty, and 16S rRNA sequences retrieved. These results will be compared to the ones obtained by other laboratories working in parallel in order to rule out microbial contamination during the procedure. This inter-laboratory comparison has not come to fruition yet and is still in progress.

6.1.1 Opalinus core microbial community

As explained in section 3.4.2, the microbial communities of 23 porewater samples collected from 8 different boreholes share similar OTUs. Some are found in all 23 samples, and others in most of the samples. In total 32 OTUs are found in at least 75% of the samples (Table A-1.3). This comparison is made possible by the high throughput sequencing data that enables the detection of low abundance OTUs in samples. These results

suggest the presence of a core microbial community in Opalinus Clay, because the 8 boreholes have a different history and were all drilled in different conditions, precluding the possibility that the same contaminating microbes. Among them is BHT-1 that was drilled under sterile conditions (Vinsot *et al.*, 2014). Some of these ubiquitous OTUs correspond to draft genomes obtained from the BIC-A1 borehole and the hydrogen experiment using similarity criteria of 98% identity covering 98% of the shorter sequence compared. The draft genomes of 7 OTUs belonging to this core microbial community were recovered, two in duplicate versions (one for the BIC-A1 borehole and one for the hydrogen experiment; Table 6.1).

Table 6.1 OTUs from the 16S rRNA gene analysis that correspond to draft genomes obtained from the metagenomic analysis of BIC-A1 borehole (fourth column) and the hydrogen experiment (fifth column). The third column indicates in how many porewater sample OTUs were observed. The color of the draft genomes indicate their quality, green for ‘good’, blue for ‘acceptable’ and red for ‘bad’, according to the criteria explained in sections 3.3.4.2, 4.3.1.2 and 4.3.1.4. The last column indicates the average nucleotide identity between draft genomes obtained in duplicate.

OTU	Taxonomy	Number of observation	BIC-A1 bin	BRH bin	ANI
OTU_2	<i>Pseudomonas</i>	23/23	c5	c35	98%
OTU_7	<i>Desulfotomaculum</i>	19/23	c6	c8a	98.84%
OTU_0	<i>Desulfocapsa</i>	23/23		c16a	
OTU_1	<i>Novispirillum</i>	23/23		c57	
OTU_3	Rhodobacteraceae	23/23		c41	
OTU_4	Xanthomonadaceae	22/23		c13	
OTU_11	Rhodospirillaceae	20/23		c5	
OTU_12	<i>Hyphomonas</i>	21/23		c29	
OTU_24	Peptococcaceae	19/23		c15	
OTU_26	Clostridiales	11/23		c20a	
OTU_29	<i>Desulfotomaculum</i>	13/23		c4b	
OTU_34	<i>Hyphomonas</i>	18/23		c22	
OTU_42	Flavobacteriaceae	16/23		c52	
OTU_46	Desulfobacteraceae	13/23		c12	
OTU_48	<i>Brevundimonas</i>	21/23		c15	
OTU_49	Puniceococcaceae	11/23		c24	
OTU_58	<i>Natronincola</i>	18/23	c4		
OTU_60	Rhodobacterales	8/23		c46	
OTU_65	Rhizobiaceae	10/23		c11	
OTU_70	Clostridiales	10/23		c7a	
OTU_88	Rhodobacteraceae	14/23		c2	
OTU_89	<i>Sphingobium</i>	18/23		c58	
OTU_139	<i>Streptomyces</i>	8/23	c14		
OTU_188	Flavobacteria	6/23		c54	
OTU_328	<i>Malikia</i>	2/23	c1		
OTU_590	<i>Escherichia</i>	3/23		c28	

OTU sequences were also compared to other 16S rRNA gene sequences obtained from other sequencing projects that took place in the Mont Terri URL, including the BN experiment (Leys and Valcke, in prep.), the PC experiment (Stroes-Gascoyne *et al.*, 2011; Frutschi and Bernier-Latmani, 2010b), the IC experiment (Daumas and Necib, 2014), the PC-C experiment (Vinsot *et al.*, 2008; founded by Andra, Nagra, CEN-SCK

and BGR), the HT experiment (Daumas *et al.*, 2014; Vinsot *et al.*, 2014; founded by NMWO), the DR experiment (Frutschi and Bernier-Latmani, 2010a), and microorganisms isolated from Opalinus Clay rock samples (Poulain *et al.*, 2008). Because most of these sequences were never published, they were merged in a single fasta file (except sequences from the BN experiment) that is accessible from this webpage: dx.doi.org/10.5075/epfl-thesis-6727. The analysis revealed that 70 OTUs identified in the current project were also observed in these other projects (Table 6.2). This number could have been greater knowing that some sequencing projects only output a small number of sequences and that they did not detect small OTUs. Overall, these comparisons highlight the presence of omnipresent microbial species in the Mont Terri porewaters that seems to indicate an Opalinus Clay origin. This finding requires confirmation through DNA extraction from pristine rock samples.

6.1.1 Future strategies

Some improvements should be made in order to characterize the microbial community present in pristine Opalinus Clay rock. The key parameter limiting DNA extraction is the concentration of DNA in the rock. In order to make this procedure more reproducible, the same extraction method should be applied to a larger sample, i.e., 200 grams. Also, it appears that extraction success depends on the sample, which reflects an inherent heterogeneity in microbial density within the Opalinus Clay. This means that a large number of samples should be investigated, and unsuccessful samples should be discarded. A complementary approach to characterize the microbial community would be to cultivate microorganisms. To increase chances of successful cultivation, a large inoculum should be used. Also, we propose to use a large volume of medium, allowing repeated sub-sampling, which is useful to monitor slow microbial growth. This could be done using molecular approaches as well, such as real-time PCR. The Opalinus Clay contains all needed nutrients for microbial growth, including phosphate and nitrogen, so it is only necessary to add to the medium, which should have a composition as close as possible to Opalinus Clay porewater, as well as an electron donor favor growth.

6.2 Microbial activity in Opalinus Clay porewater

The first signs of microbial activity in Opalinus Clay were observed in borehole water, as described in section 2.3.1. However, the results were not informative, because this activity was fueled by glycerol, which is not a relevant electron donor in a deep subsurface environment or in a nuclear waste repository.

Table 6.2 OTUs similar between this project and other sequencing projects that took place in Mont Terri URL. The grey row indicates if sequences were obtained from porewater (PW) or from enrichment (Enr.), and the number of sequences (seq) or OTUs that resulted from these studies. The third column indicates in how many porewater samples from this study OTUs were observed.

This study		Number of observation	BN exp (Leys and Valcke, in prep.)	PC porewater (Stroes-Gascoyne et al., 2011)	PC porewater (Stroes-Gascoyne et al., 2011)	PC exp (Fruttschi and Bernier-Latmani, 2010b)	IC exp (Daumas and Necib, 2014)	PC-C exp (Vinsot et al., 2008)	HT exp (Daumas et al., 2014)	DR exp (Fruttschi and Bernier-Latmani, 2010a)	Opalinus rock (Poullain et al., 2008)	Source
OTU	Taxonomy		PW 62 OTUs	PW 4 seq	Enr. 32 seq	Enr. 94 seq	PW 70 seq	PW? 70 seq	PW 61 seq	Enr. 92 seq	Enr. 7 seq	Seq/OTUs #
PW 1065 OTUs												
OTU_6	<i>Desulfosporosinus</i>	23/23	x		x	x				x		
OTU_2	<i>Pseudomonas</i>	23/23	x		x			x	x			
OTU_19	<i>Limnobacter</i>	19/23	x					x	x			
OTU_48	<i>Brevundimonas</i>	21/23	x					x	x			
OTU_12	<i>Hyphomonas</i>	21/23		x				x				
OTU_15	<i>Pseudomonas</i>	23/23			x			x				
OTU_22	<i>Phenylobacterium</i>	17/23	x					x				
OTU_58	<i>Natronincola</i>	18/23	x		x							
OTU_73	Anaerolinaceae	7/23	x				x					
OTU_104	<i>Agrobacterium</i>	6/23	x					x				
OTU_236	<i>Brevundimonas</i>	18/23	x					x				
OTU_339	Marinilabiaceae	1/23						x	x			
OTU_3	Rhodobacteraceae	23/23						x				
OTU_7	<i>Desulfotomaculum</i>	19/23					x					
OTU_11	Rhodospirillaceae	20/23						x				
OTU_13	<i>Methyloversatilis</i>	22/23	x									
OTU_16	<i>Thioclava</i>	20/23						x				
OTU_17	<i>Thiobacillus</i>	20/23								x		
OTU_20	<i>Xenophilus</i>	13/23	x									
OTU_24	Peptococcaceae	19/23	x									
OTU_27	Rhodospirillaceae	15/23						x				
OTU_32	Erysipelotrichaceae	8/23	x									
OTU_33	<i>Xanthobacter</i>	12/23								x		
OTU_34	<i>Hyphomonas</i>	18/23		x								
OTU_36	<i>Desulfosporosinus</i>	22/23	x									
OTU_44	<i>Trichococcus</i>	13/23			x							
OTU_45	<i>Cupriavidus</i>	9/23	x									
OTU_46	Desulfobacteraceae	13/23	x									
OTU_59	Rhodobacteraceae	17/23						x				
OTU_62	<i>Limnohabitans</i>	18/23	x									
OTU_70	Clostridiales	10/23	x									
OTU_72	<i>Hyphomicrobium</i>	12/23						x				
OTU_77	<i>Variovorax</i>	8/23	x									
OTU_79	Rhodobacteraceae	12/23						x				
OTU_88	Rhodobacteraceae	14/23						x				
OTU_89	<i>Sphingobium</i>	18/23	x									
OTU_94	<i>Sphingosinicella</i>	15/23	x									
OTU_102	<i>Aminobacter</i>	16/23	x									
OTU_110	<i>Ralstonia</i>	14/23	x									
OTU_117	<i>Actinotalea</i>	18/23	x									
OTU_132	Clostridiales	16/23					x					
OTU_143	<i>Thermincola</i>	6/23	x									
OTU_147	<i>Mycoplana</i>	15/23						x				
OTU_156	<i>Pseudomonas</i>	16/23	x									
OTU_167	<i>Arenimonas</i>	4/23						x				
OTU_176	Rhodospirillaceae	8/23						x				
OTU_179	<i>Variovorax</i>	6/23	x									
OTU_182	Chloroflexi	10/23						x				
OTU_195	<i>Stenotrophomonas</i>	12/23	x									
OTU_201	<i>Pseudomonas</i>	5/23	x									
OTU_244	<i>Brevundimonas</i>	15/23						x				
OTU_254	<i>Bosea</i>	5/23	x									
OTU_279	<i>Rhodopseudomonas</i>	10/23	x									
OTU_359	<i>Nocardioiodes</i>	10/23									x	
OTU_365	<i>Marinobacter</i>	10/23						x				
OTU_367	<i>Exiguobacterium</i>	3/23									x	
OTU_380	<i>Staphylococcus</i>	6/23	x									
OTU_394	<i>Achromobacter</i>	6/23	x									
OTU_397	Sphingobacteriaceae	8/23	x									
OTU_410	Erythrobacteraceae	4/23						x				
OTU_438	Ectothiorhodospiraceae	2/23	x									
OTU_466	Rhodospirillaceae	5/23						x				
OTU_471	<i>Antarctobacter</i>	4/23						x				
OTU_500	Rhodospirillaceae	4/23	x									
OTU_517	Corynebacteriaceae	2/23	x									
OTU_546	<i>Marinobacter</i>	1/23							x			
OTU_570	Desulfobulbaceae	4/23					x					
OTU_589	<i>Streptaphyta</i>	5/23	x									
OTU_660	<i>Streptococcus</i>	7/23	x									
OTU_990	<i>Anaerococcus</i>	2/23	x									

6.2.1 Line sampling limitations

By trying to investigate the microbial diversity encountered in Opalinus Clay porewater, we realized that many samples, even ones chemically described as anoxic, harbored an aerobic microbial community. This discrepancy can be explained by an issue involving sampling lines: subsequent to previous sampling events, the sampling lines contained residual porewater (with low concentration of organic acids present in the water), which combined with O₂ diffusing from the gallery, allowed for the development of a biofilms in the lines. These biofilms were sloughed off during sampling, contaminating the sample with aerobic bacteria. This is an important fact that should be considered when sampling borehole waters at the Mont Terri URL (or other facilities) for microbial analysis. We suggest avoiding the use of lines for sampling if possible. For example, in open boreholes with a vertical descending orientation, a bottle attached to a sterile aluminum rod can be used, as described in section 3.2.1. For other boreholes equipped with a packer that cannot be removed, we propose:

- If using an old line, to try to clean it as much as possible, by recirculating artificial, anoxic and sterile porewater at high speed and for a long time;
- If using a new or a clean line, to maintain it under anoxic conditions between sampling events, by flushing it using a sterile and inert gas, such as nitrogen or argon. For this purpose, a second line could be useful for flushing.

6.2.2 Microbial system in Opalinus Clay open space

Among the 23 samples investigated, only one was characterized as anoxic and was not sampled using a line. This precious sample, originating from the BIC-A1 borehole, is thought to be representative of conditions occurring in the disturbed zone of Opalinus Clay, when enough space is available to allow microbial activity. We could describe quite precisely the microbial community of this system, using a metagenomic approach. This microbial ecosystem is very simple, composed of a few microbial species, the two most abundant of which represent 85% of the community. We could also characterize the putative metabolisms of the individual populations. A set of organisms, which is capable of respiring NO₃⁻, is thought to use a fermentative metabolism in this environment, because this putative electron acceptor was not present in this borehole water. Within this group, a *Pseudomonas* sp., is very well equipped to degrade organic macromolecules, likely feeding on dead biomass and bioavailable organic carbon in Opalinus Clay. The second set of microorganisms includes sulfate-reducing bacteria and comprises the large majority of the microbial community. They are able to feed on fermentation products and on H₂. This means that they can be either heterotrophic or autotrophic, depending on the available electron donor. Among them, a Peptococcaceae represents 50% of the bacteria composing this microbial community.

6.2.3 Putative sources of energy

The electron donors feeding this microbial system were not strictly identified. Opalinus Clay porewater contains a significant amount of low molecular weight organic acids (Courdouan *et al.*, 2007), which are very likely to feed sulfate-reducing bacteria. However, two other putative energy sources can play an important role in this borehole water, making this carbon loop sustainable. Firstly, a fraction of the fossilized organic matter of the Opalinus Clay, even though it was described as recalcitrant (Jones and Tipping, 1998), could be available to microorganisms, i. e. to the dominating *Pseudomonas*, as an energy and carbon source. Secondly, because all sulfate-reducing bacteria can use hydrogen as an electron source, this compound could be used as an energy source for carbon fixation. Hydrogen was detected in the water of various boreholes, before disappearing over time (Pearson *et al.*, 2003). The authors of the study could not explain this phenomenon and proposed that hydrogen was produced by anoxic corrosion of stainless steel present in the borehole. But it could be due to an unidentified abiotic process releasing hydrogen from Opalinus Clay, similarly to serpentinization of water radiolysis. BIC-A1 borehole contains stainless steel (316L), which was chosen for its resistance for corrosion in marine water, and thus is not expected to release any hydrogen, even though this possibility cannot be entirely ruled out.

6.2.4 Future strategies

More detailed monitoring of the microbial community thriving in Opalinus Clay porewater could help bring deeper understanding of its functioning. First, any material that could release a putative energy source for a microbial metabolism should be avoided, i.e., stainless steel. The microbial biomass, along with organic acids and dissolved hydrogen, should be measured over time, in order to understand the kinetics of growth. Such data would be helpful in deciphering whether organic acids contained in porewater are the only source of energy. If that were the case, we would expect a rapid increase in biomass soon after space is made available, followed by a slow decrease in concentration over time. In that case, spiking porewater with (labeled) acetate and observing the microbial community could test this hypothesis. If the microbial biomass is stable or increases slowly over time, it would suggest that another source of energy is available to microorganisms. In this case, lab experiment should be carried out to determine whether:

- Hydrogen is released from the rock in anoxic conditions. This should be carried out in abiotic conditions, by the addition of sodium azide for instance, otherwise hydrogen concentrations would remain too low, being consumed biotically.
- Organic matter contained in Opalinus Clay can be used as a carbon source. To answer that question, organic matter could be extracted and characterized by Rock-Eval analysis. The biological degradability of extracted organic matter could then be measured during a biological oxygen demand test, using as inoculum aerobic bacteria degrading organic carbon. If this is

conclusive, this can be repeated in Opalinus Clay conditions, using porewater as medium and inoculum.

In addition, determining *in situ* conditions in which genes are active and transcribed to proteins would provide useful information to describe more precisely the microbial processes. Metaproteomic sequencing would be perfect tool to identify active enzymes, but it can be very challenging due to the very low microbial density expected in this environment. Another solution can be the metatranscriptomics, which can be more sensitive.

6.3 Microbial activity in Opalinus Clay porewater when an additional energy source is provided

Prior to this work, it was shown that the addition of an electron donor, such as glycerol leaking from pH probes, can boost microbial activity in the Opalinus Clay porewater, as described in section 2.3.1. Other electron donors are more likely to be present in a deep geological repository, such as hydrogen, known to be released from the anoxic corrosion of steel drums and canisters that will seal in nuclear waste. Excessive buildup of hydrogen could lead to pressure increases that endanger vault integrity. On the other hand, hydrogen is one of the best electron donors for respiratory metabolism, as shown in Table 1.1, hence it is likely to be consumed by bacteria *in situ*.

6.3.1 Hydrogen as an energy source

By injecting hydrogen in a borehole water for more than 500 days, we clearly showed that this compound was oxidized by a microbial process. Hydrogen sulfide production, increasing number of cells and decreasing dissolved carbon dioxide indicated that the dominating microbial metabolism was autotrophic sulfate-reduction. This was confirmed by 16S rRNA gene sequencing, highlighting a complex microbial community dominated by a member of Desulfobulbaceae, a microbial family known to be capable of autotrophic growth when reducing sulfate.

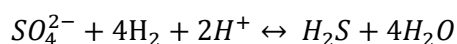
6.3.2 Carbon cycling fuelled by hydrogen

Metagenomic binning and metaproteomic analysis allowed the detailed description of carbon cycling by the microbial community. Carbon fixation was carried out by two separate autotrophic organisms, one reducing sulfate, which is the dominating Desulfobulbaceae described earlier (in section 6.3.1), and the other reducing sulfide, which is a Rhodospirillaceae. Organic macromolecules are then released when cells die, and are catabolized by *Hyphomonas* sp., which oxidizes these compounds to acetate via fermentation. Finally, acetate

and other fermentation products are oxidized to carbon dioxide via the oxidative acetyl-CoA pathway by heterotrophic sulfate-reducing bacteria, closing this carbon loop. Four of the heterotrophic SRB were identified: three Peptococcaceae members and one *Desulfatitalea* sp.. It is interesting to notice that autotrophic members of this community also uptake fermentation products, as an extra carbon source. This carbon loop perfectly fits the model described in section 1.7.2.2.

6.3.3 Hydrogen consumption rate

In situ rates of hydrogen and sulfate consumption were estimated to be $1.37 \mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$ and $0.16 \mu\text{mol}\cdot\text{cm}^{-3}\cdot\text{day}^{-1}$, respectively. Approximately 8 molecules of hydrogen are oxidized per molecule of sulfate reduced, which is double what is dictated by the stoichiometry of the redox reaction. Indeed, according to this equation, four hydrogen molecules are needed to reduce one sulfate molecule:



This means that electrons from hydrogen reduce an electron acceptor other sulfate. Other electron acceptors could be used, such as sulfide and Fe(III), but because this microbial system is autotrophic, the second most important electron sink to hydrogen is thought to be carbon dioxide. Another project in Mont Terri URL considered hydrogen consumption in Opalinus Clay (Vinsot *et al.*, 2014). Unfortunately, because its focus was the gas phase and not the aqueous phase, it is not possible to compare the rates of hydrogen consumption obtained in these two independent experiments. This other project measured continuously hydrogen after its injection. It appears that this compound is immediately consumed, even after a first injection, reaching non-measurable level after a few days. The lack of lag time seems to indicate that bacteria oxidizing hydrogen are naturally present and active in Opalinus Clay porewater. Hydrogen seems to be an electron donor for bacteria, probably coming from organic matter fermentation and/or from an unidentified abiotic process taking place in Opalinus Clay, as discussed in sections 3.4.3.2 and 6.2.3.

6.3.4 Future strategies

The biological hydrogen reduction rate is rapid enough to exceed the production rate through steel corrosion. Sulfate, used as the electron acceptor in this process, could be depleted in specific environments, if its replenishment through diffusion from the formation is not sufficiently rapid. Such a scenario would support a different microbial community, based on autotrophic methanogenesis. Archaea is the only domain of life that harbors this particular metabolism, and representatives of this domain are exceedingly scarce in Opalinus Clay porewater. Among more than 3.8 million 16S rRNA sequences obtained by sequencing samples from the hydrogen experiment, only 2 were affiliated to archeal organisms, more precisely to the genus *Methanlobus*, which is methanogenic (Table A-2.1). Under the right conditions, such organisms may be able to thrive and produce methane gas.

During the course of the hydrogen amendment experiment, sulfate concentration did not decrease significantly, primarily because porewater was regularly replaced with artificial porewater containing sulfate. In order to deplete sulfate in borehole water, one should avoid repeatedly injecting porewater containing sulfate. To achieve methanogenesis in the borehole, we propose the replacement of porewater by an artificial medium devoid of sulfate.

Hydrogen is not the only energy source that will be delivered to a deep geological repository. low- and intermediate-level waste contains organic carbon. Thus, it would be informative to stimulate microbial growth using this type of organic carbon to evaluate if the anoxic microbial degradation of organic carbon observed in soils and sediments, as described in section 1.7.1.1, can be reproduced in Opalinus Clay porewater.

6.4 Impact of microbial activity on nuclear waste repository

As explained in section 2.3.2, microbial activity can impact the disposal of nuclear waste in the subsurface either positively or negatively. The goal of this work was partly to address this question, and from the results obtained, it is possible to draw general conclusions regarding this issue.

6.4.1 Radionuclide mobility

The few studies that looked at interactions between actinides and bacteria isolated from Opalinus Clay concluded that radioactive elements like Cm(III) and U(VI) were strongly complexed at cells surfaces (Lütke *et al.*, 2013; Moll *et al.*, 2014). Moreover, it is well establish that sulfate-reducing bacteria can reduce U(VI) to U(IV), significantly decreasing its mobility (Wall and Krumholz, 2006). Overall, this means that microbial activity will likely have a benefic impact, rendering radionuclides less mobile in repositories.

6.4.2 Gas production and consumption

As described in Chapter 3 and in section 6.3 of this Chapter, hydrogen will be rapidly consumed by microbial activity, limiting a pressure build-up that could damage the repository. This process is clearly beneficial for a deep geological repository. Dissolved CH₄ was never detected at concentrations higher than 15 µM during the hydrogen injection experiment (results not shown), because the sulfate concentration was always sufficiently high to inhibit methanogenesis. However, an evaluation of whether a transition to methanogenesis could occur once sulfate is depleted, as suggest in section 6.3.4, is critical. Finally, carbon dioxide produced by oxidation of the organic matter contained in Opalinus Clay or in nuclear waste would only have a limited impact on repository pressure. Indeed, it is expected from geochemical modelling that this gas will either precipitate in the cement surrounding the low- and intermediate-level waste, or dissolve in the

Opalinus Clay porewater, which is a carbonate-rich system. Moreover, autotrophic activity fueled by hydrogen will decrease the carbon dioxide partial pressure, by reducing it to organic carbon.

6.4.3 Barrier alteration

Sulfate-reduction activity will very likely enhance the alteration of the barriers surrounding the waste. Indeed, it is well established that this metabolism promotes the corrosion of steel via two mechanisms that modify the chemical equilibrium in the direction of steel dissolution (Enning and Garrelfs, 2014). First, hydrogen sulfide produced by sulfate-reducing bacteria precipitates Fe(II) cations released from steel oxidation, producing FeS. Second, sulfate-reducing bacteria also consume hydrogen, which is another product of anoxic steel corrosion, as described in Chapter 3. Thus, hydrogen consumption can be considered as a negative process. However, steel canisters are expected to only last for a short time, on the time scale of the lifetime of the repository, even without bacterial activity. Moreover, in the current reference concept of waste disposal, microbial activity near steel compounds will be inhibited by a high-density bentonite or a high-pH backfill. This implies that autotrophic sulfate reduction will be physically separated from steel canisters and drums and that H₂ and S(-II) will have to diffuse from their source to their sink for this process to be effective. Thus, these current designs strongly limit it. Overall, this means that biological hydrogen consumption remains a positive process for the safety of nuclear waste disposal, even if it accelerates drum and canister corrosion.

More generally, microbial activity is well known for enhancing the weathering of rocks and minerals. Indeed, bacteria have several strategies for obtaining the nutrients they need to build their biomass. These include secretion of protons, organic acids, or siderophores that will liberate and gather the elements needed from the solid matrix (Uroz *et al.*, 2009). This may include the mining of phosphorus, present in Opalinus Clay as apatite and monazite (Pearson *et al.*, 2003).

For clearly quantifying the effect of microbial alteration of repository materials, experiments in which they are placed in contact with an active microbial community in Opalinus Clay porewater, under *in situ* conditions, could be envisioned. The steel and the bentonite alteration are already studied in the IC-A experiment, but a system designed for easier monitoring and more frequent sampling of the biomass responsible for this process would allow better characterization of such phenomena.

Annex 1 Supplementary information to
Chapter 3

A-1.1 Supplementary Materials and Methods

A-1.1.1 DNA extraction protocol using DNA Spin kit for Soil

This method was adapted method from DNA Spin kit for Soil (MP Biomedicals). Here is the detailed protocol:

1. Add the filter to a Lysing Matrix E tube.
2. Add 650 μ l Sodium Phosphate Buffer to sample in Lysing Matrix E tube.
3. Add 80 μ l MT Buffer.
4. Gently mix the tube by invert it by hand.
5. Incubate the tube 5 minutes at 60°C.
6. Homogenize in the FastPrep Instrument for 45 seconds at a speed of 5.5 m/s.
7. Centrifuge at 13'000 rpm for 2 minutes to pellet debris.
8. Transfer 400 μ l supernatant to a clean 2.0 ml microcentrifuge tube.
9. Add 400 μ l Sodium Phosphate Buffer to sample in Lysing Matrix E tube.
10. Add 50 μ l MT Buffer.
11. Gently mix the tube by slowly turning it over.
12. Incubate the tube 5 minutes at 60°C.
13. Vigorously shake the Lysing Matrix E tube for 45 seconds.
14. Centrifuge at 13'000 rpm for 2 minutes.
15. Transfer 500 μ l supernatant to the 2.0 ml microcentrifuge tube.
16. Add 250 μ l PPS (Protein Precipitation Solution).
17. Mix by shaking the tube by hand 10 times.
18. Centrifuge at 13'000 rpm for 3 minutes to pellet precipitate.
19. Transfer supernatant to a clean 15 ml tube.

20. Resuspend Binding Matrix suspension and add 1.0 ml to supernatant in 15 ml tube.
21. Place tubes on a rack and invert them by hand for 2 minutes to allow binding DNA.
22. Incubate the tube for 3 minutes at room temperature to allow settling of silica matrix.
23. Remove and discard 500 µl of supernatant being careful to avoid settled Binding Matrix.
24. Resuspend Binding Matrix in the remaining amount of supernatant. Transfer approximately 600 µl of the mixture to a SPIN Filter and centrifuge at 13'000 rpm for 1 minute. Empty the catch tube.
25. Do step 24 again till all the Binding Matrix have been transferred and centrifuged in the SPIN Filter.
26. Add 500 µl prepared SEWS-M (with ethanol added) and gently resuspend the pellet using the force of the liquid from the pipet tip.
27. Centrifuge at 13'000 rpm for 1 minute. Empty the catch tube and replace.
28. Without any addition of liquid, centrifuge a second time at 13'000 rpm for 2 minutes to dry the matrix of residual wash solution. Discard the catch tube and replace with a new, clean catch tube.
29. Air dry the SPIN Filter for 5 minutes at room temperature.
30. Gently resuspend Binding Matrix (above the SPIN Filter) in 100 µl of 10 mM Tris-HCl, pH 7.5.
31. Mix, close the tube, and incubated 2 minutes at room temperature.
32. Centrifuge at 13'000 rpm for 1 minute to bring eluted DNA into the clean catch tube. Discard the SPIN Filter.
33. Incubate 5 minutes at 60°C for 5 minutes.
34. Separate the DNA in two samples, in 500 µl tubes. One tube can be stored at -80°C and the other can be used or stored at -20°C. If it is used in the 2 days, it can be stored in the fridge.

A-1.1.2 JGI Illumina 16S rDNA amplicons (Itags) analysis pipeline

Document written by Julien Tremblay – 08/19/2012 Updated by Julien Tremblay on 04/25/2013

This document describes the JGI internal analysis pipeline for Itags.

1. Amplicon amplification.

Itags as a standard sequencing product will be done by using the V4 region as described by Caporaso *et al.* (2011) also used for the Earth microbiome project (earthmicrobiome.org). Hypervariable region V4 contains an average of 291 bp and spans from position 515 to 806 of an rDNA molecule. Our amplification primers are as follow:

Forward primer (black -> Illumina adapter; Red -> primer pad; Blue -> 16S V4 region forward primer)
AATGATACGGCGACCACCGAGATCTACAC TCTTTCCCTACACGACGCTCTTCCGATCT NNNNNGTCCAGCMGCCGCGGTAA
Reverse primers (black Illumina adapter; Orange -> barcode; Blue -> primer pad; Green -> 16S V4 region reverse primer). In total we have 96 different versions of this primer.
CAAGCAGAAGACGGCATACGAGAT XXXXXXXX [add 0 to 3 bases here]GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGACTACHVGGGTWTCTAAT
Forward sequencing primer ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Reverse sequencing primer GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Index sequencing primer GATCGGAAGAGCACACGTCTGAACTCCAGTCAC

For the reverse reads, we are using what we call staggered primers, meaning that from 0 to 3 nucleotides are added between the padding and V4 specific sequence. We are proceeding that way because it increases overall read quality for Itags sequencing runs (probably because it adds diversity to the pool of highly conserved 16S rRNA and/or that it helps for the initial phasing/synchronization of the Illumina base calling software). There are 96 different version of the reverse primer, each with its own unique barcode sequence (orange color).

2. Amplicon sequencing and library types.

Itags are sequenced on the Illumina MiSeq instrument with a 2x250 bp reads configuration. Each sequencing run of a MiSeq Itag run will give about 8M paired-end reads, more precisely 4M sequenced forward reads and 4M reverse reads (aka reads 1 and reads 2 respectively). Before sequencing, a PhiX spike-in shotgun library reads are added to the amplicons pool for a final concentration of about 2-25% of the ~19M paired-end reads library. Usually each sample that we receive is amplified and then pooled with 95 other samples for a total of 96 samples in a pool. This pool is sequenced on a single lane of MiSeq. We can roughly expect 50,000 – 75,000 usable QC passed reads per sample using our current 96 multiplexed design.

3. Analysis pipeline.

1- Reads are first scanned for contaminants (e.g., Illumina adapter sequences) and PhiX reads. Usually a small proportion of reads are contaminants and accordingly, 5-25% is PhiX reads.

2- From step 1, paired-end reads may be disrupted. This means that one of the read pairs might be lost due to the screening in step #1. All of these unpaired reads are discarded. This is usually a fairly small proportion of all reads.

3- All reads are then trimmed to 165 bp. This is because in the next step, read pairs will be assembled to reconstitute the original 16S rRNA amplicons and the software that we currently use is FLASH (Magoč and Salzberg, 2011) which in its first version, does not perform well when there is too much overlapping sequence in the paired reads.

4- Reads are assembled (overlapping paired assembly) with the FLASH software.

5- Primer sequences are then removed from the assembled reads. For V4 region 16S rRNA amplicons, the reverse reads will contain the reverse primer sequence, including its associated staggered bases. The forward reads do not contain a primer sequence.

6- The assembled reads from step #5 are trimmed from both 5' and 3' ends using a 20 bp sliding window quality threshold having a mean quality of 30. That is, 5' or 3' ends are cut when the mean quality score in the window is at least of 30.

7- The trimmed assembled reads from step #6 are filtered for quality. All reads having more than 5 Ns and 10 nucleotides below quality 15 are discarded. To this set of reads are added the untrimmed forward reads (after step #2) that did not assemble in step #4 (reads that were also quality controlled as described in step #5, #6 and #7). These reads will be referred to as filtered reads from now on.

8- Filtered reads are then clustered with our in-house clustering algorithm which is similar to OUT pipe (Edgar *et al.*, 2011). Briefly, reads are clustered at 100% identity and then clustered/denoised at 99% identity. Clusters having abundance lower than 3 are discarded. Remaining clusters are then scanned for chimeras with UCHIME denovo and UCHIME reference and clustered at 97% to form the final clusters/OTUs.

9- OTUs are then analyzed for taxonomic distribution using a combination of JGI in-house programs and scripts from the Qiime (qiime.org) software suite. Briefly, OTUs are classified with the RDP Classifier using an in-house training set containing the complete Greengenes database supplemented with eukaryotic sequences from the Silva databases and a customized set of mitochondria and chloroplasts 16S sequences.

10- Using taxonomic lineages obtained from step 9, a raw OTU table is generated. From that raw OTU table, an OTU table containing both bacterial and archeal organisms is generated. From this latter OTU table, two more OTU tables are generated: one rarefied to 1000 reads per samples and one rarefied to the number of reads of the sample having the lowest count/abundance.

11- A summary of reads count throughout the different steps of the pipeline is then generated.

A-1.1.3 QIIME pipeline used for alpha-diversity analysis

First, the biom files generated by the iTagger pipeline was converted to a txt files, in order to remove some samples that were not considered in this study, using this command:

```
> convert_biom.py -i otu.biom -o otu.txt -b --header_key taxonomy
```

After this, the modified OTU table was converted back to the biom format:

```
> convert_biom.py -i otu_subsample.txt -o otu_subsample.biom --biom_table_type="otu table"
```

After creating a `fasting_map.txt` file, this one was checked:

```
check_id_map.py -m fasting_map.txt -o check_id_map_output
```

Finally, alpha-diversity was computed using:

```
> echo "alpha_diversity:metrics shannon,chaol,observed_species" > alpha_params.txt
```

```
> alpha_rarefaction.py -i otu_subsample.biom -m check_id_map_output/fasting_map_corrected.txt -o  
alpha_diversity_subsample/ -p alpha_params.txt
```

A-1.1.4 DNA ethanol precipitation

- Add 2 volumes of ethanol 100% and 0.1 volume of Na-acetate 3M pH 5.2 solution to the DNA extract
- Add glycogen to a concentration of 0.05 to 1 g/L. Be also aware to glycogen concentration in DNA solution after purification should not exceed 8 g/L.
- Store the sample overnight at -20°C
- Centrifuge the sample for 20 min. at 13'000 rpm at 4°C (pre-cooled rotor)
- Empty the supernatant
- Add 1 mL of ethanol 70% (- 20°C)
- Centrifuge the sample for 15 min. at 13'000 rpm at 4°C (pre-cooled rotor)
- Empty the supernatant, remove drops, dry for max. 10 min.
- Re-suspend to pellet in TE buffer
- Incubate 15 min. at 60 °C

A-1.1.5 Sequences deposition

The following data are being deposited or have been deposited in NCBI databases:

- Metagenomics raw reads were deposited at SRA NCBI: Accession SRX672050
- V4 16S rRNA gene itag sequences:
- Raw reads were deposited at SRA under accession SRA244825.
- V4 16S rRNA gene sequences were deposited at NCBI under accession KP901405 - KP902411 and KP942780 - KP942813.
- Metagenome assembly:
- Comamonadaceae bacterium BICA1-1: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JUEA00000000. The version described in this paper is version JUEA01000000.
- Peptococcaceae bacterium BICA1-7: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JUEB00000000. The version described in this paper is version JUEB01000000.
- Pseudomonas BICA1-14: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JUEC00000000. The version described in this paper is version JUEC01000000.
- Desulfotomaculum sp. BICA1-6: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JUED00000000. The version described in this paper is version JUED01000000.
- Desulfosporosinus sp. BICA1-9: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JUEE00000000. The version described in this paper is version JUEE01000000.
- Peptococcaceae bacterium BICA1-8: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JUEF00000000. The version described in this paper is version JUEF01000000.
- All other contigs were deposited as rock porewater genome: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession JUEG00000000. The version described in this paper is version JUEG01000000.

Additionally, the IMG/MER annotation can be found with genome ID 3300002735.

A-1.2 Supplementary Tables

Table A-1.1 ICP-MS and pH measurements of borehole water samples.

Sample	Mn (ppb)	Cr (ppb)	Zn (ppb)	Si (ppb)	Al (ppb)	Fe (ppb)	Co (ppb)	Cu (ppb)	Sr (ppb)	Ni (ppb)	pH
BPC-2/2	106	0	0	0	0	4656	0	0	24871	39	7.84
BHT-1/1	125	0	0	5726	0	143	4	14	43253	43	7.82
BHT-1/2	111	0	0	4960	0	92	1	7	43454	29	6.99
BHT-1/3	72	0	0	10819	0	437	0	4	43117	22	7.75
BIC-A1/1	25	0	0	0	0	278	1	0	42858	13	8.23
BRC-3/1	282	0	0	0	0	75	65	66	47275	109	7.92
BRC-3/15	196	0	0	0	0	391	1	5	39936	37	7.82
BRC-3/17	231	0	0	0	0	508	2	0	40355	44	7.85
BRC-3/18	234	0	0	0	0	531	2	3	35288	16	7.73
BRC-3/2	402	0	61	0	0	0	15	34	47759	222	7.51
BRC-3/4	123	0	0	86	0	358	5	0	41628	17	7.56
BRC-3/13	180	0	0	0	0	486	6	11	42265	58	7.76
BRC-3/5	140	0	0	0	0	514	5	5	16049	41	7.55
BWS-A1/1	180	0	0	0	0	508	1	0	45424	19	7.45
BWS-A1/4	87	0	0	0	0	289	0	0	45422	18	7.24
BWS-A2/2	177	0	0	0	0	1901	0	0	46678	19	7.69
BWS-H2/0	386	0	0	1964	0	349	0	0	46194	95	7.89
BWS-H2/1	308	0	0	0	0	325	4	0	43076	427	6.61
BWS-H2/2	8	0	0	0	0	376	0	0	45045	37	7.58
BWS-H2/3	1	0	0	0	0	322	0	0	44977	59	7.66

Table A-1.2 IC measurements of borehole water samples.

Sample	Na ⁺ (mM)	NH ₄ ⁺ (mM)	Mg ²⁺ (mM)	K ⁺ (mM)	Ca ²⁺ (mM)	F ⁻ (mM)	Cl ⁻ (mM)	NO ₂ ⁻ (mM)	SO ₄ ²⁻ (mM)	Br ⁻ (mM)	NO ₃ ⁻ (mM)	PO ₄ ²⁻ (mM)
BPC-2/2	73.29	0.16	2.93	0.41	1.85	0.06	1.31	0.00	2.74	0.00	0.00	0.00
BHT-1/1	245.58	0.70	19.33	1.12	16.28	1.03	278.87	0.00	19.50	0.20	0.03	0.00
BHT-1/2	294.84	0.29	21.78	1.02	17.73	0.18	290.62	0.00	17.59	0.68	0.00	0.00
BHT-1/3	294.11	0.24	19.05	1.01	16.49	0.06	283.99	0.00	13.79	0.53	0.00	0.00
BIC-A1/1	313.56	0.28	18.14	1.91	18.88	0.10	290.25	0.00	15.82	0.53	0.00	0.00
BRC-3/1	269.16	0.68	26.79	1.37	21.53	0.05	320.34	0.00	25.50	0.40	0.17	0.00
BRC-3/15	303.02	0.28	23.30	1.09	18.26	0.05	301.63	0.00	21.37	2.00	0.00	0.00
BRC-3/17	325.78	0.24	26.00	1.41	20.25	0.05	303.79	0.00	22.60	0.66	0.00	0.00
BRC-3/18	324.21	0.23	24.49	1.27	19.27	0.05	304.64	0.00	20.61	9.95	0.00	0.00
BRC-3/2	298.48	0.43	25.73	1.21	20.35	0.04	312.89	0.00	22.76	0.49	0.00	0.00
BRC-3/4	285.56	0.24	24.35	0.96	16.43	0.06	301.48	0.00	24.81	0.57	0.34	0.00
BRC-3/13	310.67	0.25	19.46	0.94	14.69	0.14	256.23	0.00	18.70	1.09	0.00	0.00
BRC-3/5	295.77	0.09	19.73	2.04	15.75	0.16	305.35	0.00	18.33	4.43	0.00	0.00
BWS-A1/1	231.49	0.69	17.37	1.07	15.86	0.17	259.03	0.00	16.91	0.19	0.08	0.00
BWS-A1/4	283.36	0.26	17.43	0.90	15.18	0.04	278.61	0.00	12.62	0.55	0.00	0.00
BWS-A2/2	206.46	0.24	10.85	0.78	11.22	0.10	184.27	0.00	10.05	0.31	0.01	0.00
BWS-H2/0	236.30	0.21	12.48	4.18	10.90	0.07	225.84	0.00	11.46	0.35	0.00	0.00
BWS-H2/1	190.44	0.59	12.22	5.81	11.19	0.15	215.48	0.00	11.60	0.18	0.12	0.00
BWS-H2/2	242.54	0.26	11.93	0.81	11.29	0.09	239.32	0.00	9.94	0.57	0.00	0.00
BWS-H2/3	213.47	0.26	12.36	0.81	11.02	0.06	199.35	0.00	9.86	0.55	0.00	0.00

Table A-1.3 Number of observation of all 825 OTUs in the 23 samples, including OTU taxonomic affiliation. This table is too big for being printed in this document. It is available online here: dx.doi.org/10.5075/epfl-thesis-6727.

Table A-1.4 Genus-level microbial community composition, including oxygen dependency for each taxonomic group. Zero values are not shown (a 0.00 value indicate a positive number lower or equal to 0.005). Red entries represent aerobic microorganisms while green entries indicate anaerobic organisms. Entries in black represent facultative anaerobes. For each value, the intensity of yellow color is proportional to the contribution of the OTU to the microbial community.

BDR-TI/1	BDR-TI/2	BDR-TZ/1	BIC-A1/1	BPC-2/1	BWS-A1/1	BWS-A1/4	BWS-H2/0	BWS-H2/1	BWS-H2/2	BWS-H2/3	BRC-3/1	BRC-3/2	BRC-3/4	BRC-3/5	BRC-3/13	BRC-3/15	BRC-3/17	BRC-3/18	BWS-A2/1	BHT-1/1	BHT-1/2	BHT-1/3	Oxygen dependency*	Taxonomic affiliation	Reference					
	0.00	0.00		0.00	0.02		0.00	0.00	0.00	0.00				0.00			0.00	0.00	0.16	0.01	0.00	0.00	SAER	<i>Dietzia</i>	Goodfellow and Maldonado, 2006					
0.00	0.00		0.00	0.00							0.00	0.00	0.17	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	SAER	<i>Leucobacter</i>	Etvushenko and Takeuchi, 2006				
			0.00	0.00							0.00		0.01	0.03	0.01	0.03	0.00	0.01	0.03				0.00	SAER	<i>Rhodococcus</i>	Goodfellow and Maldonado, 2006				
	0.01	0.00		0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00		Nocardiodaceae	Dastager <i>et al.</i> , 2008				
0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.04	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.12	0.00	0.03	0.00			Other Actinobacteria					
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00			Flaviobacteriaceae	Bernadet <i>et al.</i> , 1996			
	0.00	0.00	0.00	0.00			0.00				0.16	0.12	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			Sphingobacteriaceae	Steyn <i>et al.</i> , 1998			
0.02	0.01	0.01		0.04	0.00	0.00		0.00	0.00	0.00	0.03	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	Other Bacteroidetes			
	0.00			0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00			Bacillaceae	Marquez <i>et al.</i> , 2008			
	0.00			0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			<i>Paenibacillus</i>	Shida <i>et al.</i> , 1997			
0.08	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	Other Bacilli			
	0.00	0.02	0.00	0.00	0.00	0.00	0.08	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			SANA	<i>Natronincola</i>	Zhilina <i>et al.</i> , 1998		
0.27	0.28	0.02	0.08	0.14	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.06	0.00	0.00	0.01	0.00	0.03	0.00	0.00	0.00	0.03	SANA	<i>Desulfosporosinus</i>	Spring and Rosenzweig, 2006				
0.01	0.01	0.00	0.59	0.02	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.06	0.01	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	SANA	<i>Desulfotomaculum</i>	Aüllo <i>et al.</i> , 2013	
0.05	0.06	0.00	0.02	0.13	0.01	0.00	0.00	0.01	0.00	0.01	0.01	0.01	0.00	0.10	0.01	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	SANA	Other Clostridiales	Wells and Wilkins, 1996	
	0.00	0.00	0.02	0.00	0.00	0.00		0.00	0.00	0.00	0.02	0.16	0.00	0.00	0.03	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	SANA	Other Clostridia	Wells and Wilkins, 1996	
	0.00	0.00					0.00	0.00		0.02	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				Pirellulales			
	0.00	0.00						0.00	0.02	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00				<i>Planctomyces</i>	Ward <i>et al.</i> , 2006		
			0.00	0.00				0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	Other Planctomycetes			
0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.07	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<i>Brevundimonas</i>	Vancanneyt <i>et al.</i> , 2005		
			0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.01	0.11	SAER	<i>Phenyllobacterium</i>	Eberspächer, 2005
	0.41	0.00					0.00	0.00			0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.06	0.01	0.03				<i>Afpia</i>	Weyant and Whitney, 2005		
			0.00		0.00	0.00														0.01	0.01	0.07	SAER	<i>Bosea</i>	Das, 2005					
0.00										0.00	0.01	0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.11	0.02					Cohaesibacteraceae	Hwang and Cho, 2008		
0.57	0.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<i>Pleomorphomonas</i>	Xie and Yokota, 2005	
			0.00	0.00			0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.07	0.19	SAER	<i>Xanthobacter</i>	Wiegel, 2005				
0.00	0.00	0.00		0.02	0.42	0.00	0.08	0.02	0.16	0.01	0.00	0.12	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	SAER	<i>Phaeobacter</i>	Martens <i>et al.</i> , 2006
			0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.01	0.16	0.00	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	SAER	<i>Thioclava</i>	Sorokin <i>et al.</i> , 2005
0.00	0.00	0.02	0.00	0.00	0.12	0.03	0.10	0.12	0.12	0.11	0.00	0.01	0.01	0.01	0.07	0.01	0.29	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	Other Rhodobacteraceae	Garrity <i>et al.</i> , 2005a	
0.00		0.00								0.00	0.00	0.00	0.00	0.01	0.09	0.01	0.01					0.00	0.00					Other Rhodobacterales	Garrity <i>et al.</i> , 2005b	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.02	0.05	0.04	0.07	0.04	0.03	0.06	0.20	0.07	0.04	0.00	0.01	0.12	0.00							Rhodospirillaceae	Garrity <i>et al.</i> , 2005c	
0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.02	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	SAER	<i>Sphingobium</i>	Takeuchi <i>et al.</i> , 2001
	0.00	0.03		0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.04	0.01	0.00	0.02	0.00	0.01	0.01	0.01	0.01	0.16	0.00	0.00	0.00	0.00	0.00	0.00	SAER	<i>Sphingopyxis</i>	Balkwill <i>et al.</i> , 2006
0.00	0.00	0.06	0.00	0.00	0.15	0.04	0.01	0.01	0.32	0.07	0.07	0.02	0.01	0.01	0.04	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	Sphingomonadaceae	Yabuuchi and Kosako, 2005a
0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.12	0.01	0.03	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	Other Sphingomonadales	Yabuuchi and Kosako, 2005b
0.00	0.00	0.05	0.00	0.00	0.01	0.01	0.01	0.07	0.01	0.02	0.05	0.05	0.02	0.01	0.04	0.14	0.04	0.03	0.02	0.04	0.07	0.03						Other Alphaproteobacteria		
0.00		0.00	0.00	0.00	0.00	0.00	0.00					0.04	0.00		0.15	0.00										0.01	0.01	<i>Alcaligenes</i>	Busse and Auling, 2005	
0.00	0.00	0.16																			0.01	0.01	0.00					<i>Cupriavidus</i>	Balkwill, 2005	
	0.00		0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.03	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.17	0.09	0.01					<i>Limnobacter</i>	Spring <i>et al.</i> , 2001; Lu <i>et al.</i> , 2011	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14					<i>Limnobacter</i>	Hahn <i>et al.</i> , 2010	
0.00		0.00		0.00						0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.00	SAER	<i>Variovorax</i>	Willems <i>et al.</i> , 2005				
			0.00	0.00						0.00	0.00	0.01	0.02	0.01	0.02	0.00	0.00	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	SAER	<i>Herbaspirillum</i>	Baldani <i>et al.</i> , 2005
0.00	0.00	0.00	0.00	0.01		0.00	0.00	0.00	0.01	0.11	0.0																			

Table A-1.5 Number of sequences, total number of species and number of observed species (for a same amount of sequences), Chao's 1 and Shannon's diversity indexes for all samples.

Sample	Number of sequences	Total number of observed species	Observed species*	Chao1's diversity index*	Shannon's diversity index*
BPC-2/1	56701	246	242.3	251.192	3.247
BHT-1/1	83370	129	114.4	145.861	3.957
BHT-1/2	127527	207	173.8	212.028	4.775
BHT-1/3	77240	107	91.5	134.116	3.265
BIC-A1/1	127610	87	61.6	95.867	1.951
BRC-3/1	141165	265	223.1	273.387	4.819
BRC-3/2	114810	214	184.2	219.683	4.74
BRC-3/4	113551	287	237.3	293.029	4.478
BRC-3/5	121555	286	237.5	286.787	4.208
BRC-3/13	105541	251	218.7	251.022	4.439
BRC-3/15	121535	257	222.7	271.917	5.042
BRC-3/17	119761	315	264	315.264	4.166
BRC-3/18	120201	221	174.8	231.668	2.551
BWS-A1/1	114060	134	106.2	131.457	3.349
BWS-A1/4	89699	92	73.9	103.55	2.265
BWS-A2/1	112140	169	135.1	170.738	2.932
BWS-H2/0	115941	94	67.3	102.513	2.001
BWS-H2/1	82385	94	83	104.619	3.532
BWS-H2/2	138313	83	61.2	81.857	2.726
BWS-H2/3	123971	93	67.3	99.236	2.807
BDR-T1/1	135472	65	42.5	67.03	1.81
BDR-T1/2	139145	69	45.6	74.29	1.91
BDR-T2/1	92307	123	96.1	169.134	3.063

*Calculated with 46382 sequences

Table A-1.6 Contig clustering, as outputted by CONCOCT and after a manual correction. This table is too big for being printed in this document. It is available online here: dx.doi.org/10.5075/epfl-thesis-6727.

Table A-1.7 Taxonomic affiliation of the metagenomic clusters. The final annotation was based on the 16S rRNA gene annotation (if any), followed by average nucleotide identity (ANI; for selected clusters only), by the MLTreeMap annotation (either using the default MLTM tree or the GEBA tree) and by the IMG phylogenetic distribution of identified proteins. For ‘Best BLAST hits’ column, the number in parenthesis refers to percentage of hit similarity. For ‘MLTL analysis’ and ‘IMG phylogenetic distribution’ columns, the number in parenthesis refers to the fraction of genes affiliated to the microbial group.

Cluster ID	Final taxonomic affiliation	16S rRNA gene				Taxonomic affiliation based on				IMG phylogenetic distribution	
		Contig ID	Start position	End position	Length	RDP Classifier (>0.8)	Best BLAST hits	ANI (for selected clusters only)	MLTM analysis using GEBA tree		MLTM analysis using default tree
c0	<i>Pseudomonas</i>	-	-	-	-	-	-	-	<i>Pseudomonadaceae</i> (1)	<i>Pasteurellales</i> (1)	<i>Pseudomonas</i> (0.55)
c1	<i>Comamonadaceae</i>	BICA1a_1049446	4731	6253	1523 bp	<i>Comamonadaceae</i> (97%) <i>Maiklia</i> and <i>Hydrogenophaga</i> sp. Different genera belonging to <i>Comamonadaceae</i> (<95%)	<i>Comamonadaceae</i> (0.96)	<i>Comamonadaceae</i> (1)	<i>Burkholderiales</i> (1)	<i>Acidovorax</i> (0.19)	
c2	<i>Pseudomonas</i>	-	-	-	-	-	-	-	-	-	<i>Pseudomonas</i> (0.46)
c3	<i>Desulfosporosinus</i>	BICA1a_1099571	48614	49172	559 bp	<i>Desulfosporosinus</i> sp. (97%)	<i>Desulfosporosinus</i> (0.83)	-	<i>Peptococcaceae</i> (1)	<i>Peptococcaceae</i> (1)	<i>Desulfosporosinus</i> (0.66)
c4	<i>Clostridiaceae</i>	BICA1a_1016619	144	1662	1519 bp	<i>Geosporobacter</i> (95%) <i>Salimesphibacter</i> and <i>Geosporobacter</i> sp. (95%)	-	-	<i>Clostridiales</i> (0.89)	<i>Clostridia</i> (0.83)	<i>Alkaliphilus</i> (0.17)
c5	<i>Pseudomonas</i>	BICA1a_1169816	9407	10839	1533 bp	<i>Pseudomonas chloritidismutans</i> strain AW-1 (99%) <i>Pseudomonas chloritidismutans</i> strain AW-1 (98%)	<i>Pseudomonas chloritidismutans</i> - <i>P. xanthomarina</i> (1)	<i>Pseudomonadaceae</i> (1)	<i>Pseudomonadaceae</i> (1)	<i>Pseudomonas</i> (0.92)	
c6	<i>Desulfotomaculum</i>	BICA1a_1168930	3404	4935	1532 bp	<i>Desulfotomaculum</i> sp. (94%)	<i>Peptococcaceae</i> (1)	<i>Peptococcaceae</i> (1)	<i>Peptococcaceae</i> (1)	<i>Desulfotomaculum</i> (0.70)	
c7	<i>Peptococcaceae</i>	BICA1a_1053282	1	362	362 bp	<i>Desulfotomaculum</i> and <i>Peletomaculum</i> sp. (89%) <i>Desulfotomaculum</i> , <i>Peletomaculum</i> and <i>Cryptanaerobacter</i> sp. (89%)	<i>Peptococcaceae</i> (0.87)	<i>Peptococcaceae</i> (1)	<i>Peptococcaceae</i> (1)	<i>Desulfotomaculum</i> (0.44)	
c8	<i>Peptococcaceae</i>	BICA1a_1148202	85080	85596	517 bp	<i>Peptococcaceae</i>	<i>Peptococcaceae</i> (0.83)	<i>Clostridiales</i> (0.83)	<i>Peptococcaceae</i> (1)	<i>Desulfonifera</i> (0.24)	
c9	<i>Desulfosporosinus</i>	BICA1a_1124249	1	343	343 bp	<i>Desulfosporosinus</i> facus strain STP12 (97%)	<i>Peptococcaceae</i> (1)	<i>Peptococcaceae</i> (1)	<i>Peptococcaceae</i> (1)	<i>Desulfosporosinus</i> (0.63)	
c11	<i>Bacteroides</i>	-	-	-	-	-	-	-	-	-	<i>Bacteroides</i> (0.43)
c12	<i>Peptococcaceae</i>	-	-	-	-	-	-	-	<i>Peptococcaceae</i> (1)	<i>Peptococcaceae</i> (0.9)	<i>Desulfosporosinus</i> (0.80)
c14c	<i>Streptomyces</i>	BICA1a_1192699	246	1769	1524 bp	<i>Streptomyces</i> sp. (98%)	<i>Streptomyces</i> sp. (98%)	-	<i>Streptomyces</i> (0.89)	<i>Streptomyces</i> (0.89)	<i>Streptomyces</i> (0.79)

Table A-1.8 Single-copy gene plot of each metagenomic clusters. This information highlights the purity and the completeness of clusters, which ideally should contain a single-copy of each of these genes. The table also includes for each cluster its total length (in Mbp), its number of contigs and its contribution to microbial community.

Cluster	Taxonomic affiliation	cluster length (Mbp)	Number of contigs	Single copy genes (count)																									Average coverage proportion (excluding c0, c2 and c11)									
				COG0016	COG0048	COG0049	COG0051	COG0052	COG0060	COG0072	COG0080	COG0081	COG0087	COG0088	COG0089	COG0090	COG0091	COG0092	COG0093	COG0094	COG0096	COG0097	COG0100	COG0102	COG0103	COG0130	COG0184	COG0185		COG0186	COG0197	COG0198	COG0200	COG0201	COG0244	COG0256	COG0504	COG0532
c4	<i>Geosporobacter</i>	6.129	346	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.01
c1	Comamonadaceae	2.968	128	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.03
c5	<i>Pseudomonas</i>	6.580	175	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.35	
c6	<i>Desulfotomaculum</i>	3.736	108	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.02	
c9	<i>Desulfosporosinus</i>	6.987	266	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.03	
c7	Peptococcaceae	4.745	33	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.50	
c8	Peptococcaceae	3.977	162	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.01	
c12	<i>Desulfosporosinus</i>	3.871	408	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.01	
c14c	<i>Streptomyces</i>	2.631	316	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.02	
c3	<i>Desulfosporosinus</i>	6.521	203	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.03
c0	<i>Pseudomonas</i>	1.547	96	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-
c2	<i>Pseudomonas</i>	0.021	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-
c11	Bacteroides	0.006	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-

Table A-1.9 Results of average nucleotide identity (ANI) between draft genomes of this study and known genomes that are the closest phylogenetically. For each cluster, the highest ANI value is written in bold.

	Organism	Assembly	Average Nucleotide Identity (ANI)					
			c1	c6	c7	c8	c9	c14
Comamonadaceae	<i>Acidovorax</i> sp. KKS102	GCA_000302535.1	80.10%	n. a.	n. a.	n. a.	n. a.	n. a.
	<i>Curvibacter gracilis</i> ATCC BAA-807	GCA_000518645.1	80.45%	n. a.	n. a.	n. a.	n. a.	n. a.
	<i>Hydrogenophaga intermedia</i> S1	GCA_000723405.1	81.25%	n. a.	n. a.	n. a.	n. a.	n. a.
	<i>Hydrogenophaga</i> sp. PBC	GCA_000263795.1	81.27%	n. a.	n. a.	n. a.	n. a.	n. a.
	<i>Hydrogenophaga</i> sp. T4	GCA_000576245.1	82.20%	n. a.	n. a.	n. a.	n. a.	n. a.
	<i>Ramlibacter tataouinensis</i> TTB310	GCA_000215705.1	80.50%	n. a.	n. a.	n. a.	n. a.	n. a.
	<i>Rhodoferax ferrireducens</i> T118	GCA_000013605.1	79.21%	n. a.	n. a.	n. a.	n. a.	n. a.
	<i>Candidatus Desulforudis audaxviator</i> MP104C	GCA_000018425.1	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.
	<i>Dehalobacter</i> sp. DCA	GCA_000305775.1	n. a.	i. h.	i. h.	81.88%	84.66%	n. a.
	<i>Dehalobacter restrictus</i> DSM 9455	GCA_000512895.1	n. a.	i. h.	i. h.	82.80%	i. h.	n. a.
	<i>Desulfotibacter alkalitolerans</i> DSM 16504	GCA_000620305.1	n. a.	i. h.	i. h.	91.48%	84.35%	n. a.
	<i>Desulfotibacterium hafniense</i> DCB-2	GCA_000021925.1	n. a.	i. h.	i. h.	i. h.	80.15%	n. a.
	<i>Desulfotibacterium hafniense</i> Y51	GCA_000010045.1	n. a.	i. h.	i. h.	i. h.	79.78%	n. a.
	<i>Desulfotibacterium dehalogenans</i> ATCC 51507	GCA_000243155.3	n. a.	i. h.	i. h.	i. h.	77.86%	n. a.
<i>Desulfotibacterium dichloroelimans</i> LMG P-21439	GCA_000243135.3	n. a.	i. h.	i. h.	i. h.	79.20%	n. a.	
<i>Desulfotibacterium metallireducens</i> DSM 15288	GCA_000231405.2	n. a.	i. h.	i. h.	i. h.	79.00%	n. a.	
<i>Desulfotibacterium</i> sp. PCE1	GCA_000384015.1	n. a.	i. h.	i. h.	i. h.	78.15%	n. a.	
<i>Desulfosporosinus acidiphilus</i> SJ4	GCA_000255115.3	n. a.	i. h.	i. h.	i. h.	81.00%	n. a.	
<i>Desulfosporosinus merdijei</i> DSM 13257	GCA_000231385.3	n. a.	i. h.	i. h.	i. h.	81.75%	n. a.	
<i>Desulfosporosinus orientis</i> DSM 765	GCA_000235605.1	n. a.	i. h.	i. h.	i. h.	82.01%	n. a.	
<i>Desulfosporosinus</i> sp. HMP52	GCA_000765145.1	n. a.	i. h.	81.98%	i. h.	81.98%	n. a.	
<i>Desulfosporosinus</i> sp. OT	GCA_000224515.2	n. a.	93.24%	i. h.	i. h.	84.20%	n. a.	
<i>Desulfosporosinus</i> sp. Tol-M	GCA_000770645.1	n. a.	i. h.	i. h.	93.31%	81.81%	n. a.	
<i>Desulfosporosinus youngiae</i> DSM 17734	GCA_000244895.1	n. a.	i. h.	i. h.	i. h.	82.40%	n. a.	
<i>Desulfotomaculum acetoxidans</i> DSM 771	GCA_000024205.1	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.	
<i>Desulfotomaculum alcoholivorax</i> DSM 16058	GCA_000430885.1	n. a.	79.59%	77.36%	i. h.	i. h.	n. a.	
<i>Desulfotomaculum alkaliphilum</i> DSM 12257	GCA_000711975.1	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.	
<i>Desulfotomaculum carboxydivorans</i> CO-1-SRB	GCA_000214435.1	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.	
<i>Desulfotomaculum gibsoniae</i> DSM 7213	GCA_000233715.3	n. a.	80.81%	77.90%	i. h.	82.40%	n. a.	
<i>Desulfotomaculum hydrothermale</i> Lam5 = DSM 18033	GCA_000315365.1	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.	
<i>Desulfotomaculum kuznetsovii</i> DSM 6115	GCA_000214705.1	n. a.	77.87%	76.36%	i. h.	i. h.	n. a.	
<i>Desulfotomaculum nigrificans</i> DSM 574	GCA_000189755.3	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.	
<i>Desulfotomaculum reducens</i> MI-1	GCA_000016165.1	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.	
<i>Desulfotomaculum ruminis</i> DSM 2154	GCA_000215085.1	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.	
<i>Desulfotomaculum thermocisternum</i> DSM 10259	GCA_000686645.1	n. a.	76.27%	75.76%	i. h.	i. h.	n. a.	
<i>Desulfurispora thermophila</i> DSM 16022	GCA_000376385.1	n. a.	76.65%	77.41%	i. h.	i. h.	n. a.	
<i>Pelotomaculum thermopropionicum</i> SI	GCA_000010565.1	n. a.	77.84%	76.62%	i. h.	i. h.	n. a.	
<i>Syntrophobotulus glycolicus</i> DSM 8271	GCA_000190635.1	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.	
<i>Thermincola potens</i> JR	GCA_000092945.1	n. a.	i. h.	i. h.	i. h.	i. h.	n. a.	
Pseudo-monas	<i>Pseudomonas chloritidismutans</i> AW-1	GCA_000495915.1	n. a.	n. a.	n. a.	n. a.	n. a.	97.02%
	<i>Pseudomonas knackmussii</i> B13	GCA_000689415.1	n. a.	n. a.	n. a.	n. a.	n. a.	79.75%
	<i>Pseudomonas stutzeri</i> ATCC 17588	GCA_000219605.1	n. a.	n. a.	n. a.	n. a.	n. a.	86.60%
	<i>Pseudomonas stutzeri</i> A1501	GCA_000013785.1	n. a.	n. a.	n. a.	n. a.	n. a.	86.60%
	<i>Pseudomonas xanthomarina</i> S11	GCA_000825645.1	n. a.	n. a.	n. a.	n. a.	n. a.	97.81%

n. a. stands for not analyzed

i. h. stands for insufficient hits

Table A-1.10 Pathway annotation of the metagenomic clusters, for anaerobic respiration, central carbon metabolism, ATP synthesis, carbon fixation, nitrogen fixation and carbon oxidation in sulfate-reducing bacteria. The acetyl-CoA pathway can also be used in the other direction, for oxidizing acetate to carbon dioxide by sulfate-reducing bacteria. The table also includes the number of electron transferring proteins identified in each cluster. The completeness of pathways is labeled by a green scale indicated below the table.

Cluster	Taxonomic affiliation	Anaerobic respiration			Central carbon metabolism							Electron transfer				Carbon fixation													
		Disimilatory sulfate reduction	Disimilatory nitrate reduction	Denitrification	Key enz.	Glycolysis (EMP pathway)	Pentose phosphate pathway (Pentose phosphate cycle)	Entner-Doudoroff pathway	Pyruvate to acetyl-coA	Citrate cycle (TCA cycle, Krebs cycle)	Anaplerotic sequence (requires enzymes form glucose to oxalacetate, via TCA cycle)	Glyoxylate cycle	Incomplete TCA fork	modified TCA cycle in the incomplete (acetate) oxidizing SRB	electron transferring flavoproteins	electron transferring quinones	electron transferring [Fe-S] proteins (excluding hydrogenases)	electron transferring cytochromes	ATP synthase	Key enz.	Reductive acetyl-CoA pathway (Wood-Ljungdahl pathway)	Key enz.	Reductive pentose phosphate cycle (Calvin cycle)	Hydroxypropionate-hydroxybutyrate cycle	3-Hydroxypropionate bi-cycle	Reductive citrate cycle (Arnon-Buchanan cycle)	Dicarboxylate-hydroxybutyrate cycle	Nitrogen fixation	
c4	<i>Geosporobacter</i>	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c1	Comamonadaceae	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c12	<i>Desulfosporosinus</i>	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c5	<i>Pseudomonas</i>	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c6	<i>Desulfotomaculum</i>	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c9	<i>Desulfosporosinus</i>	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c7	Peptococcaceae	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c8	Peptococcaceae	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c14c	<i>Streptomyces</i>	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c3	<i>Desifosporosinus</i>	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c11	Bacteroides	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

0-49% of genes present
 50-74% of genes present
 75-99% of genes present (or 1 gene missing)
 100% of genes present

Table A-1.11 Number of hydrogenases observed in each cluster, and pathway annotation of fermentation identified in each cluster. The completeness of pathways is labeled by a green scale indicated below the table.

Cluster	Taxonomic affiliation	Hydrogenases							Fermentation																				
		uncharacterized [NiFe]	[NiFe] group 1	[NiFe] group 2b	[NiFe] group 3a	[NiFe] group 3d	[NiFe] group 3b	[NiFe] group 4	FeFe-only	succinate production	formate production	ethanol production	butanol production	butyrate production	acetate production	lactate production	propionate production	acetoin production	butanediol production	acetone production	isopropanol production								
c4	<i>Geosporobacter</i>	2	1	0	4	0	0	0	13	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c1	Comamonadaceae	1	0	1	1	0	2	0	2	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c12	<i>Desulfosporosinus</i>	3	2	0	1	0	0	0	14	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c5	<i>Pseudomonas</i>	0	0	0	1	0	0	0	2	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c6	<i>Desulfotomaculum</i>	2	1	0	8	0	0	0	7	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c9	<i>Desulfosporosinus</i>	11	4	0	6	0	0	3	33	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c7	Peptococcaceae	1	1	0	9	0	0	1	7	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c8	Peptococcaceae	4	3	0	5	0	0	0	17	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c14c	<i>Streptomyces</i>	0	0	0	0	0	0	0	2	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c3	<i>Desifosporosinus</i>	1	2	0	4	0	0	2	31	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
c11	Bacteroides	0	0	0	0	0	0	0	0	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

all enzymes found
 1 enzyme missing

Table A-1.12 Number of genes encoding for phosphatases, lipases, proteases, nucleases, and involved in other extracellular organic matter degradation, and in thiosulfate oxidation.

Cluster	Taxonomic affiliation	S ₂ O ₃ ²⁻ oxidation		Phosphatases					Lipases			Proteases															Nucleases					Other extracellular enzymes degrading organic matter												
		Sox complex (thiosulfate oxidation)	APS reductase (thiosulfate oxidation)	Phytases	Nonspecific phosphatases	Other specific organic phosphatases	Glucose dehydrogenase (for gluconic acid synthesis)	Polyphosphatase	Inorganic pyrophosphatase	ABC transport system	Chemotaxis	esterase/lipase	membrane maturation	secretory	SOS	ABC	signal	SpoIVB	Germination	Secreted	Pilus	periplasmic	chaperone	other	total	restriction	RNA, ribonuclease	repair	polymerase	others	total	Laminarinase	Invertase	Asparaginase	Glutaminase	Urease	Alpha-amylase (starch hydrolysis)	Amorphous cellulose hydrolysis	Xylanase	Chitinase				
c4	<i>Geosporobacter</i>	0	0	0	0	1	0	2	3	0	1	23	20	0	3	2	1	7	1	2	0	4	7	6	97	150	11	7	3	3	31	55	0	0	1	3	0	0						
c1	Comamonadaceae	0	0	0	3	1	0	1	1	3	1	16	11	0	2	2	1	4	0	0	0	0	4	9	32	65	6	12	0	2	21	41	0	0	1	1	1	0						
c12	<i>Desulfosporosinus</i>	0	0	0	1	0	0	0	2	1	1	22	16	0	1	0	5	5	1	1	0	1	6	1	64	101	3	2	1	2	23	31	0	0	1	0	0	0						
c5	<i>Pseudomonas</i>	0	1	1	1	2	1	1	1	2	3	34	16	0	2	4	2	6	0	0	3	1	10	7	63	114	6	18	3	4	42	73	0	0	1	1	2	4	yes					
c6	<i>Desulfotomaculum</i>	0	0	0	0	0	0	1	2	0	1	10	9	0	2	1	2	5	1	1	0	3	6	4	40	74	4	6	3	3	18	34	0	0	0	1	0	0						
c9	<i>Desulfosporosinus</i>	0	1	0	0	1	0	0	2	0	1	31	16	0	1	3	3	5	2	1	0	2	7	3	67	110	5	4	1	1	23	34	0	0	3	1	0	0						
c7	Peptococcaceae	0	0	0	0	0	0	1	3	0	3	15	11	1	3	2	2	4	1	1	0	5	4	6	46	86	6	5	2	1	23	37	0	0	0	1	0	0	yes					
c8	Peptococcaceae	0	0	0	0	2	0	1	3	0	2	11	16	0	1	1	0	6	0	1	0	2	3	2	46	78	4	6	3	2	19	34	0	0	2	1	0	0						
c14c	<i>Streptomyces</i>	0	1	0	2	0	0	0	0	0	0	14	4	1	0	1	0	0	0	0	0	0	2	6	39	53	7	0	1	2	13	23	0	0	0	0	0	2	yes	yes	yes			
c3	<i>Desulfosporosinus</i>	0	0	0	0	1	0	0	0	2	0	5	34	22	2	3	2	1	8	1	1	0	2	7	3	68	120	8	4	1	6	27	46	0	0	0	2	0	0					
c11	Bacteroides	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0						

Table A-1.13 Gene list from *Pseudomonas* c5 that might be involved in processes described in Figure 3.4.A, including gene annotation, COG, KO and EC number. The number of the first row corresponds to the reaction number in Figure 3.4.A.

Reaction no.	Protein coding gene ID	Gene product	COG number	KO number	EC number
1	BICA1a_110947557	Glucokinase	COG0837	K00845	EC:2.7.1.2
1	BICA1a_112635730	Glucokinase	COG0837	K00845	EC:2.7.1.2
2	BICA1a_106034735	Glucose-6-phosphate isomerase	COG0166	K01810	EC:5.3.1.9
3	BICA1a_108164131	Fructose-1-phosphate kinase and related fructose-6-phosphate kinase (PfkB)	COG1105	K16370	EC:2.7.1.11
4	BICA1a_113178852	Fructose/tagatose biphosphate aldolase	COG0191	K01624	EC:4.1.2.13
5	BICA1a_106034752	Triosephosphate isomerase	COG0149	K01803	EC:5.3.1.1
6	BICA1a_10205496	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	COG0057	K00134	EC:1.2.1.12
6	BICA1a_110947545	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	COG0057	K00134	EC:1.2.1.12
6	BICA1a_113741333	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	COG0057	K00134	EC:1.2.1.12
6	BICA1a_117442050	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	COG0057	K00134	EC:1.2.1.12
6	BICA1a_113178850	3-phosphoglycerate kinase	COG0126	K00927	EC:2.7.2.3
7	BICA1a_113958220	Phosphoglyceromutase	COG0696	K15633	EC:5.4.2.1
8	BICA1a_110194248	Enolase	COG0148	K01689	EC:4.2.1.11
8	BICA1a_10830801	Enolase	COG0148	K01689	EC:4.2.1.11
8	BICA1a_110194248	Enolase	COG0148	K01689	EC:4.2.1.11
8	BICA1a_11370831	Enolase	COG0148	K01689	EC:4.2.1.11
9	BICA1a_10121372	Pyruvate kinase	COG0469	K00873	EC:2.7.1.40
9	BICA1a_11279464	Pyruvate kinase	COG0469	K00873	EC:2.7.1.40
9	BICA1a_1184721114	Pyruvate kinase	COG0469	K00873	EC:2.7.1.40
9	BICA1a_118472330	Pyruvate kinase	COG0469	K00873	EC:2.7.1.40
10	BICA1a_112359925	Na+-transporting methylmalonyl-CoA/oxaloacetate decarboxylase, gamma subunit	COG3630	K01573	EC:4.1.1.3
10	BICA1a_112359927	Na+-transporting methylmalonyl-CoA/oxaloacetate decarboxylase, beta subunit	COG1883	K01572	EC:4.1.1.3
10	BICA1a_112359926	Pyruvate/oxaloacetate carboxyltransferase	COG5016	K01571	EC:4.1.1.3
10	BICA1a_1108478125	Pyruvate/oxaloacetate carboxyltransferase	COG5016	K01571	EC:4.1.1.3
10	BICA1a_1108478124	Na+-transporting methylmalonyl-CoA/oxaloacetate decarboxylase, gamma subunit	COG3630	K01573	EC:4.1.1.3
10	BICA1a_1108478126	Na+-transporting methylmalonyl-CoA/oxaloacetate decarboxylase, beta subunit	COG1883	K01572	EC:4.1.1.3
11	BICA1a_118472167	Predicted dehydrogenase - COG0579	COG0579	K00116	EC:1.1.5.4
12	BICA1a_11698166	Fumarase	COG0114	K01679	EC:4.2.1.2
12	BICA1a_1184721111	Tartrate dehydratase alpha subunit/Fumarate hydratase class I, N-terminal domain	COG1951	K01676	EC:4.2.1.2
12	BICA1a_118472333	Tartrate dehydratase alpha subunit/Fumarate hydratase class I, N-terminal domain	COG1951	K01676	EC:4.2.1.2
12	BICA1a_10080083	Fumarase	COG0114	K01679	EC:4.2.1.2
13	BICA1a_102158991	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG1053	K00239	EC:1.3.99.1
13	BICA1a_102158992	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	COG0479	K00240	EC:1.3.99.1
13	BICA1a_102158989	Succinate dehydrogenase/fumarate reductase, cytochrome b subunit	COG2009	K00241	
13	BICA1a_102158990	Succinate dehydrogenase, hydrophobic anchor subunit	COG2142	K00242	
14	BICA1a_10551331	Lactate dehydrogenase and related dehydrogenases	COG1052	K03778	EC:1.1.1.28
14	BICA1a_10430031	D-isomer specific 2-hydroxyacid dehydrogenase, catalytic domain	COG1052	K03778	EC:1.1.1.28
14	BICA1a_102841990	Lactate dehydrogenase and related dehydrogenases	COG1052	K03778	EC:1.1.1.28
15	BICA1a_100346189	Pyruvate dehydrogenase complex, dehydrogenase (E1) component	COG2609	K00163	EC:1.2.4.1
15	BICA1a_10215231	Pyruvate/2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, eukaryotic type, alpha subunit	COG1071	K00161	EC:1.2.4.1
15	BICA1a_111304567	Pyruvate/2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, eukaryotic type, alpha subunit	COG1071	K00161	EC:1.2.4.1
15	BICA1a_111304569	Pyruvate/2-oxoglutarate dehydrogenase complex, dehydrogenase (E1) component, eukaryotic type, beta subunit	COG0022	K00162	EC:1.2.4.1
15	BICA1a_100346190	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide acyltransferase (E2) component, and related enzymes	COG0508	K00627	EC:2.3.1.12
15	BICA1a_111304571	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide acyltransferase (E2) component, and related enzymes	COG0508	K00627	EC:2.3.1.12
15	BICA1a_102158995	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component, and related enzymes	COG1249	K00382	EC:1.8.1.4
16	BICA1a_104479765	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
16	BICA1a_101191451	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
16	BICA1a_112635790	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
16	BICA1a_11235992	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
16	BICA1a_10424799	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
16	BICA1a_1110704107	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
16	BICA1a_118008915	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
17	BICA1a_10010216	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
17	BICA1a_119535985	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
17	BICA1a_11693573	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
17	BICA1a_100800832	3-hydroxyacyl-CoA dehydrogenase	COG1250	K01825	EC:1.1.1.35
17	BICA1a_119495027	3-hydroxyacyl-CoA dehydrogenase	COG1250	K07516	EC:1.1.1.35
17	BICA1a_100800832	3-hydroxyacyl-CoA dehydrogenase	COG1250	K01825	EC:1.1.1.35
18	BICA1a_112041614	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
18	BICA1a_112041615	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
18	BICA1a_116981637	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
18	BICA1a_1173957115	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
18	BICA1a_118008923	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
18	BICA1a_118008921	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
18	BICA1a_119535992	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
19	BICA1a_1003461114	Acyl-CoA dehydrogenases	COG1960	K00248	EC:1.3.8.1
19	BICA1a_117395763	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
19	BICA1a_10675241	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
19	BICA1a_114050710	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
19	BICA1a_113178513	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
19	BICA1a_119535993	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
19	BICA1a_119535991	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
19	BICA1a_112359915	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
19	BICA1a_10531224	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
19	BICA1a_116981644	Acyl-CoA dehydrogenases	COG1960	K06445	EC:1.3.99.-
20-22	BICA1a_110703741	Acetaldehyde dehydrogenase (acetylating)	COG4569	K04073	EC:1.2.1.10
20-22	BICA1a_110703741	Acetaldehyde dehydrogenase (acetylating)	COG4569	K04073	EC:1.2.1.10
21	BICA1a_1126357110	Predicted dehydrogenases and related proteins - COG0673	COG0673	K00100	
21	BICA1a_112359910	Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	COG1028	K00100	
23	BICA1a_110194242	Zn-dependent alcohol dehydrogenases, class III	COG1062	K00121	EC:1.1.1.1
23	BICA1a_10668026	Zn-dependent alcohol dehydrogenases	COG1064	K13953	EC:1.1.1.1
23	BICA1a_114451235	Zn-dependent alcohol dehydrogenases, class III	COG1062	K00121	EC:1.1.1.1
23	BICA1a_11609786	Zn-dependent alcohol dehydrogenases	COG1064	K13953	EC:1.1.1.1
23	BICA1a_102115929	Alcohol dehydrogenase, class IV	COG1454	K00001	EC:1.1.1.1
23	BICA1a_1173957105	NADPH:quinone reductase and related Zn-dependent oxidoreductases	COG0604	K00001	EC:1.1.1.1
23	BICA1a_120250481	Alcohol dehydrogenase, class IV	COG1454	K13954	EC:1.1.1.1
23	BICA1a_102115929	Alcohol dehydrogenase, class IV	COG1454	K00001	EC:1.1.1.1
24	BICA1a_118472189	Phosphotransacetylase	COG0280	K13788	EC:2.3.1.8
25	BICA1a_118472190	Acetate kinase	COG0282	K00925	EC:2.7.2.1
26	BICA1a_112635785	Coenzyme F420-reducing hydrogenase, beta subunit	COG1035		

Table A-1.14 Gene list from Peptococcaceae c7 that might be involved in processes described in Figure 3.4.B, including gene annotation, COG, KO and EC number. The number of the first row corresponds to the reaction number in Figure 3.4.B.

Reaction no.	Protein coding gene ID	Gene product	COG number	KO number	EC number
27	BICA1a_118430117	Transcriptional regulator/sugar kinase	COG1940	K00845	EC:2.7.1.2
28	BICA1a_1094883194	Glucose-6-phosphate isomerase	COG0166	K15916	EC:5.3.1.8
28	BICA1a_118430118	Glucose-6-phosphate isomerase	COG0166	K15916	EC:5.3.1.8
29	BICA1a_107544837	6-phosphofructokinase	COG0205	K00850	EC:2.7.1.11
29	BICA1a_1172615311	6-phosphofructokinase	COG0205	K00850	EC:2.7.1.11
30	BICA1a_102506412	Fructose/tagatose biphosphate aldolase	COG0191	K01624	EC:4.1.2.13
30	BICA1a_1202110278	DhnA-type fructose-1,6-bisphosphate aldolase and related enzymes	COG1830	K01623	EC:4.1.2.13
31	BICA1a_118430128	Triosephosphate isomerase	COG0149	K01803	EC:5.3.1.1
32	BICA1a_1172615257	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	COG0057	K00134	EC:1.2.1.12
32	BICA1a_117261542	Predicted phosphoglycerate mutase, AP superfamily - COG3635	COG3635	K01834	EC:5.4.2.1
32	BICA1a_118430126	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	COG0057	K00134	EC:1.2.1.12
32	BICA1a_118430127	3-phosphoglycerate kinase	COG0126	K00927	EC:2.7.2.3
33	BICA1a_113638828	Fructose-2,6-bisphosphatase	COG0406	K01834	EC:3.1.3.73
33	BICA1a_118430129	Phosphoglyceromutase	COG0696	K01834	EC:5.4.2.1
34	BICA1a_118430130	Enolase	COG0148	K01689	EC:4.2.1.11
35	BICA1a_1172615310	Pyruvate kinase	COG0469	K00873	EC:2.7.1.40
36	BICA1a_1172615679	Malate/lactate dehydrogenases	COG0039	K00016	EC:1.1.1.27
37	BICA1a_114820258	Na ⁺ -transporting methylmalonyl-CoA/oxaloacetate decarboxylase, beta subunit	COG1883	K01572	EC:4.1.1.3
37	BICA1a_117261513	Pyruvate/oxaloacetate carboxyltransferase	COG5016	K01571	EC:4.1.1.3
37	BICA1a_1172615162	Pyruvate/oxaloacetate carboxyltransferase	COG5016	K01571	EC:4.1.1.3
38	BICA1a_117261514	Malate/lactate dehydrogenases	COG0039	K00024	EC:1.1.1.37
39	BICA1a_117261522	Tartrate dehydratase alpha subunit/Fumarate hydratase class I, N-terminal domain	COG1951	K01677	EC:4.2.1.2
39	BICA1a_117261521	Tartrate dehydratase beta subunit/Fumarate hydratase class I, C-terminal domain	COG1838	K01678	EC:4.2.1.2
39	BICA1a_112103665	Tartrate dehydratase alpha subunit/Fumarate hydratase class I, N-terminal domain	COG1951	K01677	EC:4.2.1.2
39	BICA1a_112103664	Tartrate dehydratase beta subunit/Fumarate hydratase class I, C-terminal domain	COG1838	K01678	EC:4.2.1.2
39	BICA1a_112103675	Tartrate dehydratase alpha subunit/Fumarate hydratase class I, N-terminal domain	COG1951	K01677	EC:4.2.1.2
40	BICA1a_112103678	Succinate dehydrogenase, hydrophobic anchor subunit	COG2142	K00242	
40	BICA1a_112103677	Succinate dehydrogenase/fumarate reductase, cytochrome b subunit	COG2009	K00241	
40	BICA1a_112103679	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	COG0479	K00240	EC:1.3.99.1
40	BICA1a_112103680	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG1053	K00244	EC:1.3.99.1
40	BICA1a_115715170	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	COG0479	K00245	EC:1.3.99.1
40	BICA1a_115715168	Succinate dehydrogenase/fumarate reductase, cytochrome b subunit	COG2009	K00246	
40	BICA1a_115715169	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG1053	K00244	EC:1.3.99.1
41		Could not be identified in <i>Desulfotomaculum</i> cluster c7 genome. Corresponds to EC:2.8.3.-			
42	BICA1a_1089845200	Methylmalonyl-CoA mutase, N-terminal domain/subunit	COG1884	K01848	EC:5.4.99.2
42	BICA1a_1089845199	Methylmalonyl-CoA mutase, C-terminal domain/subunit (cobalamin-binding)	COG2185	K01849	EC:5.4.99.2
42	BICA1a_1075448220	Methylmalonyl-CoA mutase, N-terminal domain/subunit	COG1884	K01848	EC:5.4.99.2
42	BICA1a_1075448221	Methylmalonyl-CoA mutase, C-terminal domain/subunit (cobalamin-binding)	COG2185	K01849	EC:5.4.99.2
42	BICA1a_117261517	Methylmalonyl-CoA mutase, C-terminal domain/subunit (cobalamin-binding)	COG2185	K01849	EC:5.4.99.2
42	BICA1a_117261518	Methylmalonyl-CoA mutase, N-terminal domain/subunit	COG1884	K01848	EC:5.4.99.2
43	BICA1a_117261516	Methylmalonyl-CoA mutase, N-terminal domain/subunit	COG3185	K05606	EC:5.1.99.1
44	BICA1a_117261515	Acetyl-CoA carboxylase, carboxyltransferase component (subunits alpha and beta)	COG4799	K01966	EC:6.4.1.3
45	BICA1a_1094883111	Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, alpha subunit	COG0674	K03737	EC:1.2.7.-
46	BICA1a_115715179	Alcohol dehydrogenase, class IV	COG1454	K04072	EC:1.1.1.1
47	BICA1a_115715179	Alcohol dehydrogenase, class IV	COG1454	K04072	EC:1.1.1.1
47	BICA1a_1172615560	Threonine dehydrogenase and related Zn-dependent dehydrogenases	COG1063	K00121	EC:1.1.1.1
47	BICA1a_1202110382	Alcohol dehydrogenase, class IV	COG1454	K00001	EC:1.1.1.1
47	BICA1a_1202110482	Alcohol dehydrogenase, class IV	COG1454	K13954	EC:1.1.1.1
48	BICA1a_120211067		COG1042	K01905	EC:6.2.1.13
49	BICA1a_108984534	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
49	BICA1a_1126667204	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
49	BICA1a_1075448202	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
49	BICA1a_1075448216	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
49	BICA1a_1172615519	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
49	BICA1a_1202110266	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
49	BICA1a_1202110268	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
49	BICA1a_1202110251	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
50	BICA1a_112103691	3-hydroxyacyl-CoA dehydrogenase	COG1250	K07516	EC:1.1.1.35
50	BICA1a_108984596	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
50	BICA1a_108984535	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
50	BICA1a_1011693290	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
50	BICA1a_1011693297	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
50	BICA1a_1126667206	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
50	BICA1a_112666750	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
50	BICA1a_112666748	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
50	BICA1a_1075448201	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
50	BICA1a_1172615518	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
50	BICA1a_1172615686	3-hydroxyacyl-CoA dehydrogenase	COG1250	K00074	EC:1.1.1.157
51	BICA1a_1075448200	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01715	EC:4.2.1.55
51	BICA1a_1202110250	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01715	EC:4.2.1.55
51	BICA1a_10384662	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
51	BICA1a_108984595	Enoyl-CoA hydratase/carnithine racemase	COG1024	K13767	EC:4.2.1.17
51	BICA1a_1011693300	Enoyl-CoA hydratase/carnithine racemase	COG1024	K13767	EC:4.2.1.17
51	BICA1a_112680448	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
51	BICA1a_1126667208	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
51	BICA1a_112103639	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
51	BICA1a_1094883180	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
51	BICA1a_113876928	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
51	BICA1a_1184301225	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
51	BICA1a_1202110293	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
51	BICA1a_1202110263	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17
51	BICA1a_1202110220	Enoyl-CoA hydratase/carnithine racemase	COG1024	K01692	EC:4.2.1.17

Table A-1.14 (continued)

Reaction no.	Protein coding gene ID	Gene product	COG number	KO number	EC number
52	BICA1a_1075448217	Acyl-CoA dehydrogenases	COG1960	K00248	EC:1.3.8.1
52	BICA1a_112608451	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
52	BICA1a_1126667207	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
52	BICA1a_1094883181	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
52	BICA1a_1157151127	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
52	BICA1a_1172615606	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
52	BICA1a_120211044	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
52	BICA1a_120211042	Acyl-CoA dehydrogenases	COG1960	K00249	EC:1.3.8.7
52	BICA1a_1075448206	Acyl-CoA dehydrogenases	COG1960	K00248	EC:1.3.99.2
53	BICA1a_1172615303	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	COG0365	K01896	EC:6.2.1.2
54	BICA1a_115715179	Alcohol dehydrogenase, class IV	COG1454	K04072	EC:1.1.1.1
55		Could not be identified in Desulfotomaculum cluster c7 genome. Corresponds to EC:1.1.1.-			
56	BICA1a_1172615278	CO dehydrogenase/acetyl-CoA synthase delta subunit (corrinoid Fe-S protein)	COG2069	K00194	EC:2.1.1.245
56	BICA1a_1172615280	CO dehydrogenase/acetyl-CoA synthase gamma subunit (corrinoid Fe-S protein)	COG1456	K00197	EC:2.1.1.245
56	BICA1a_1172615282	6Fe-6S prismatic cluster-containing protein	COG1151	K00190	EC:1.2.7.4
56-57	BICA1a_1172615281	CO dehydrogenase/acetyl-CoA synthase beta subunit	COG1614	K14138	EC:1.2.7.4
56-57	BICA1a_1202110393	CO dehydrogenase/acetyl-CoA synthase beta subunit	COG1614	K14138	EC:1.2.7.4
56-57	BICA1a_1202110394	6Fe-6S prismatic cluster-containing protein	COG1151	K00190	EC:1.2.7.4
57	BICA1a_1172615277	Pterin binding enzyme	COG1410	K15023	
57	BICA1a_117826647	Aerobic-type carbon monoxide dehydrogenase, small subunit CoxS/CutS homologs	COG2080	K03518	EC:1.2.99.2
58	BICA1a_1172615266	5,10-methylenetetrahydrofolate reductase	COG0685	K00297	EC:1.5.1.20
59-60	BICA1a_1172615153	5,10-methylene-tetrahydrofolate dehydrogenase/Methenyl tetrahydrofolate cyclohydrolase	COG0190	K01491	EC:1.5.1.5
61	BICA1a_109488329	Formyltetrahydrofolate synthetase	COG2759	K01938	EC:6.3.4.3
62	BICA1a_1089845305	Uncharacterized anaerobic dehydrogenase	COG3383	K00123	EC:1.2.1.2
62	BICA1a_1089845306	Uncharacterized anaerobic dehydrogenase	COG3383	K00336	EC:1.6.5.3
63	BICA1a_1089845296	Uncharacterized anaerobic dehydrogenase	COG3383	K00123	EC:1.2.1.2
63	BICA1a_1089845305	Uncharacterized anaerobic dehydrogenase	COG3383	K00123	EC:1.2.1.2
64	BICA1a_120211075	Formate hydrogenlyase subunit 3/Multisubunit Na ⁺ /H ⁺ antiporter, MnhD subunit	COG0651	K12137	EC:1.-
64	BICA1a_120211074	Formate hydrogenlyase subunit 4	COG0650		
64	BICA1a_120211073	Hydrogenase 4 membrane component (E)	COG4237	K12140	EC:1.-
64	BICA1a_120211072	Formate hydrogenlyase subunit 3/Multisubunit Na ⁺ /H ⁺ antiporter, MnhD subunit	COG0651	K12141	EC:1.-
64	BICA1a_120211071	Ni,Fe-hydrogenase III large subunit	COG3261		
64	BICA1a_120211070	Ni,Fe-hydrogenase III small subunit	COG3260		
65	BICA1a_1094883206	Ni,Fe-hydrogenase I large subunit	COG0374	K06281	EC:1.12.99.6
65	BICA1a_1094883207	Ni,Fe-hydrogenase I small subunit	COG1740	K06282	EC:1.12.99.6
66	BICA1a_1184301203	NADH:ubiquinone oxidoreductase subunit 3 (chain A)	COG0838	K00330	EC:1.6.5.3
66	BICA1a_1184301204	NADH:ubiquinone oxidoreductase 20 kD subunit and related Fe-S oxidoreductases	COG0377	K00331	EC:1.6.5.3
66	BICA1a_1184301205	NADH:ubiquinone oxidoreductase 27 kD subunit	COG0852	K00332	EC:1.6.5.3
66	BICA1a_1184301206	NADH:ubiquinone oxidoreductase 49 kD subunit 7	COG0649	K00333	EC:1.6.5.3
66	BICA1a_1184301207	NADH:ubiquinone oxidoreductase subunit 1 (chain H)	COG1005	K00337	EC:1.6.5.3
66	BICA1a_1184301208	Formate hydrogenlyase subunit 6/NADH:ubiquinone oxidoreductase 23 kD subunit (chain I)	COG1143	K00338	EC:1.6.5.3
66	BICA1a_1184301209	NADH:ubiquinone oxidoreductase subunit 6 (chain J)	COG0839	K00339	EC:1.6.5.3
66	BICA1a_1184301210	NADH:ubiquinone oxidoreductase subunit 11 or 4L (chain K)	COG0713	K00340	EC:1.6.5.3
66	BICA1a_1184301211	NADH:ubiquinone oxidoreductase subunit 5 (chain L)/Multisubunit Na ⁺ /H ⁺ antiporter, MnhA subunit	COG1009	K00341	EC:1.6.5.3
66	BICA1a_1184301212	NADH:ubiquinone oxidoreductase subunit 4 (chain M)	COG1008	K00342	EC:1.6.5.3
66	BICA1a_1184301213	NADH:ubiquinone oxidoreductase subunit 2 (chain N)	COG1007	K00343	EC:1.6.5.3
67	BICA1a_118430176	ATP sulfurylase (sulfate adenyllyltransferase)	COG2046	K00958	EC:2.7.7.4
68	BICA1a_118430132	Inorganic pyrophosphatase	COG3808	K01507	EC:3.6.1.1
69	BICA1a_1172615263		COG1150	K03390	EC:1.8.98.1
69	BICA1a_1172615264	Heterodisulfide reductase, subunit B	COG2048	K03389	EC:1.8.98.1
69	BICA1a_1172615265	Pyridine nucleotide-disulphide oxidoreductase	COG1148	K03388	EC:1.8.98.1
69	BICA1a_1172615268	Coenzyme F420-reducing hydrogenase, delta subunit	COG1908		
69	BICA1a_1172615269	NADPH-dependent glutamate synthase beta chain and related oxidoreductases	COG0493		
69	BICA1a_1172615273	Coenzyme F420-reducing hydrogenase, beta subunit	COG1035		
69	BICA1a_1172615275	Coenzyme F420-reducing hydrogenase, delta subunit	COG1908		
70	BICA1a_1075448204	Electron transfer flavoprotein, alpha subunit	COG2025	K03522	
70	BICA1a_1075448205	Electron transfer flavoprotein, beta subunit	COG2086	K03521	
71	BICA1a_118430179		COG1148	K03388	EC:1.8.98.1
71	BICA1a_118430180	Methyl-viologen-reducing hydrogenase, delta subunit	COG1148	K03388	EC:1.8.98.1
72	BICA1a_118430177	Adenosine-5'-phosphosulfate reductase beta subunit			
72	BICA1a_118430178	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG1053	K00394	EC:1.8.99.2
73	BICA1a_102506429	Dissimilatory sulfite reductase (desulfovirlidin), gamma subunit	COG2920	K11179	EC:2.8.1.-
73	BICA1a_102506430	Fe-S oxidoreductase	COG0247		
73	BICA1a_102506431	Nitrate reductase gamma subunit	COG2181		
73	BICA1a_102506439	Dissimilatory sulfite reductase (desulfovirlidin), alpha and beta subunits	COG2221	K11181	EC:1.8.99.3
73	BICA1a_102506440	Dissimilatory sulfite reductase (desulfovirlidin), alpha and beta subunits	COG2221	K11180	EC:1.8.99.3
74	BICA1a_108207114	FOF1-type ATP synthase, subunit a	COG0356	K02108	EC:3.6.3.14
74	BICA1a_108207115	FOF1-type ATP synthase, subunit c/Archaeal/vacuolar-type H ⁺ -ATPase, subunit K	COG0636	K02110	EC:3.6.3.14
74	BICA1a_108207118	FOF1-type ATP synthase, alpha subunit	COG0056	K02111	EC:3.6.3.14
74	BICA1a_108207119	FOF1-type ATP synthase, gamma subunit	COG0224	K02115	EC:3.6.3.14
74	BICA1a_108207116	FOF1-type ATP synthase, subunit b	COG0711	K02109	EC:3.6.3.14
74	BICA1a_108207117	FOF1-type ATP synthase, delta subunit (mitochondrial oligomycin sensitivity protein)	COG0712	K02113	EC:3.6.3.14
74	BICA1a_108207120	FOF1-type ATP synthase, beta subunit	COG0055	K02112	EC:3.6.3.14
74	BICA1a_108207121	FOF1-type ATP synthase, epsilon subunit (mitochondrial delta subunit)	COG0355	K02114	EC:3.6.3.14
75	BICA1a_1157151196	Coenzyme F420-reducing hydrogenase, delta subunit	COG1908	K14127	EC:1.12.99.-
75	BICA1a_1157151197	Coenzyme F420-reducing hydrogenase, gamma subunit	COG1941	K14128	EC:1.12.99.-
75	BICA1a_1157151198	Coenzyme F420-reducing hydrogenase, alpha subunit	COG3259	K14126	EC:1.12.99.-
76	BICA1a_1172615156	2-polyprenylphenol hydroxylase and related flavodoxin oxidoreductases	COG0543	K00528	EC:1.18.1.2

A-1.3 Supplementary Figures

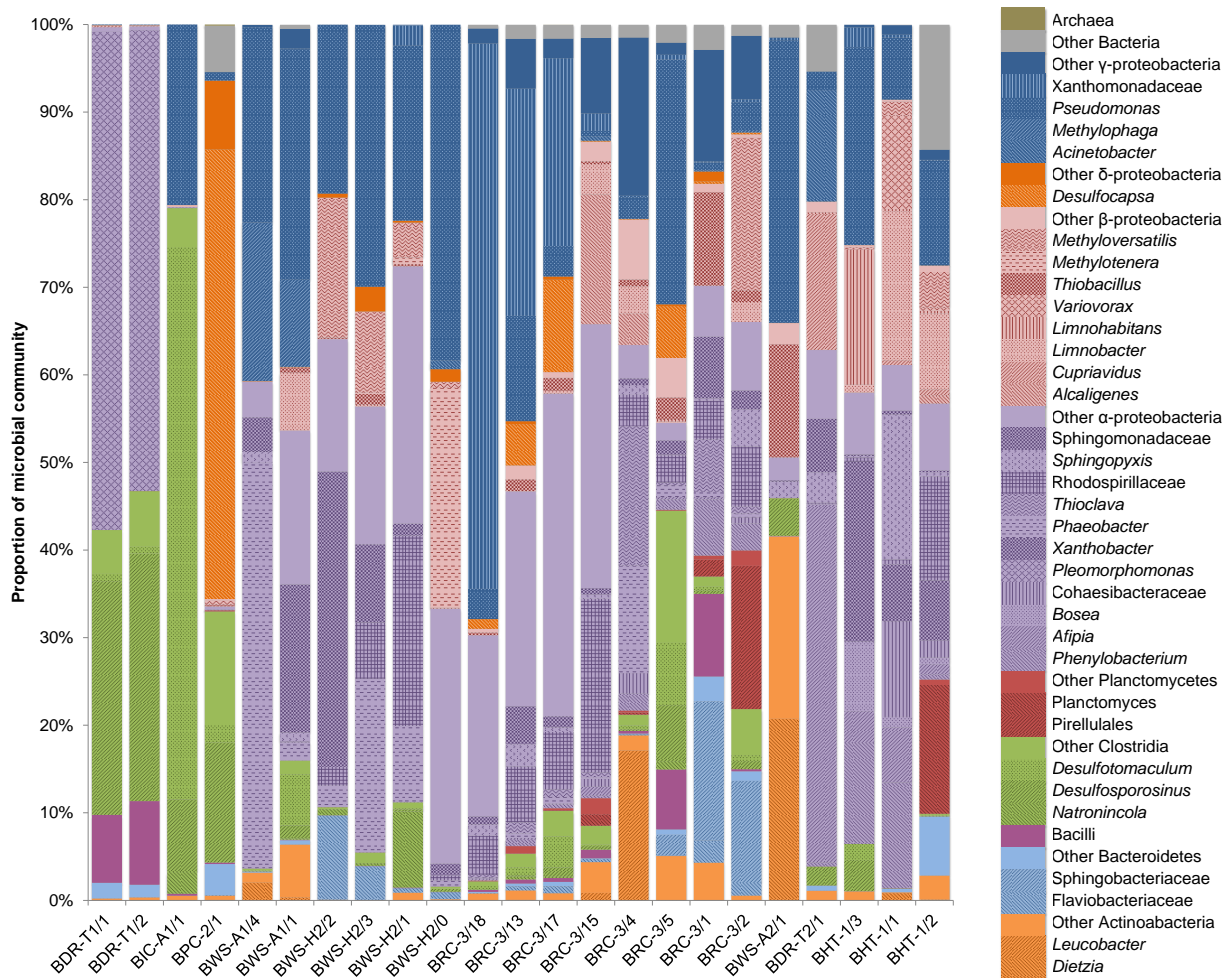


Figure A-1.1 Microbial community composition of all samples, based on the region V4 of the 16S rRNA gene. Microbial groups are color-coded as follow: light orange for Actinobacteria, light blue for Bacteroidetes, purple for Bacilli, green for Clostridia, red for Planctomycetes, light purple for Alphaproteobacteria, light mauve for Betaproteobacteria, orange for Deltaproteobacteria, blue for Gammaproteobacteria, grey for other Bacteria and light brown for Archaea.

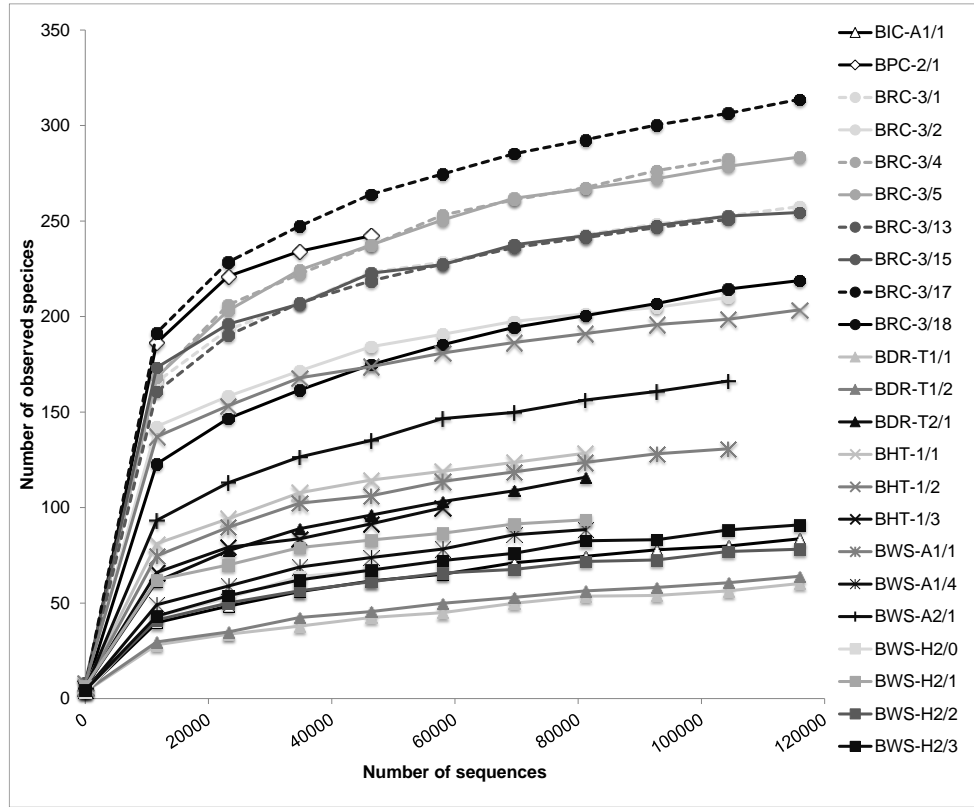


Figure A-1.2 Rarefaction curve of the number of observed species, based on the V4 region of the 16S rRNA gene sequencing.

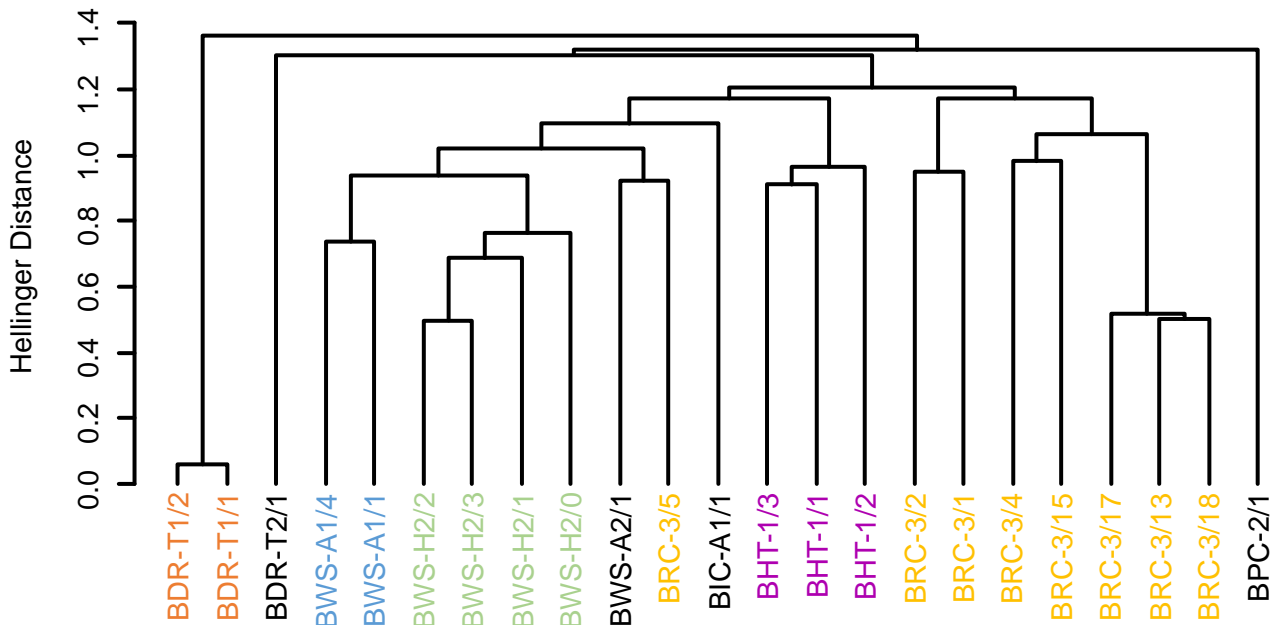


Figure A-1.3 Sample clustering based on their microbial community, using Hellinger distances and average linkage method. The samples are color-coded according to the borehole origin.

Annex 2 Supplementary information to
Chapter 4

A-2.1 Supplementary Materials and Methods

A-2.1.1 Calculation of dissolved gas concentration

In order to back-calculate the dissolved gas concentration, first the volume of sampled water was determined.

$$\begin{aligned} & \text{volume of sampled water [mL]} \\ &= \frac{\text{bottle weight after sampling [g]} - \text{bottle weight before sampling [g]}}{\text{water sample density [g} \cdot \text{mL}^{-1}]} \end{aligned}$$

Then, headspace volume can be calculated.

$$\text{headspace volume [mL]} = \text{total bottle volume [mL]} - \text{volume of sampled water [mL]}$$

Then, the final pressure of the sampling bottle was calculated.

$$\begin{aligned} & \text{bottle pressure after sampling [abs bar]} \\ &= \text{bottle pressure before sampling [abs bar]} \times \frac{\text{total bottle volume [mL]}}{\text{headspace volume [mL]}} \end{aligned}$$

The following calculations were performed for each analyzed gas. The gas partial pressure was calculated.

$$\begin{aligned} & \text{gas partial pressure [Pa]} \\ &= \text{gas GC measurement [\%]} \times \text{bottle pressure after sampling [abs bar]} \\ &\quad \times 10^5 \end{aligned}$$

Then, the gas concentration in headspace was calculated.

$$\begin{aligned} & \text{gas concentration in headspace [M]} \\ &= \frac{\text{gas partial pressure [Pa]}}{8.314 [\text{J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}] \times \text{incubation temperature [K]}} \times \frac{1}{1000} \end{aligned}$$

Then, the number of moles of gas in the headspace was calculated.

$$\begin{aligned} & \text{gas amount in headspace [mol]} \\ &= \text{gas concentration in headspace [M]} \times \text{headspace volume [mL]}/1000 \end{aligned}$$

Then, Henry's law constant $k_{H,cc}$, which gives the ratio between gas concentration in water and gas concentration in headspace at equilibrium and for a given temperature, was calculated for each gas and for incubation temperature (Sander, 1999).

$$\begin{aligned}
 & \text{Henry's law constant } k_{H,cp} \text{ [at a temperature } T] \\
 & = k_{H,cp} \text{ [at a temperature of } 298.15 \text{ K]} \\
 & \times e^{(\text{constant } C \text{ of van't Hoff equation [K]} \times (1/\text{incubation temperature } T \text{ [K]} - 1/298.15 \text{ [K])})}
 \end{aligned}$$

where:

$$k_{H,cp} = 3.4 \times 10^{-2} \text{ for CO}_2$$

$$k_{H,cp} = 7.8 \times 10^{-4} \text{ for H}_2$$

$$k_{H,cp} = 1.4 \times 10^{-3} \text{ for CH}_4$$

and where:

$$\text{constant } C \text{ of van't Hoff equation} = 2.4 \times 10^3 \text{ for CO}_2$$

$$\text{constant } C \text{ of van't Hoff equation} = 5.0 \times 10^2 \text{ for H}_2$$

$$\text{constant } C \text{ of van't Hoff equation} = 1.6 \times 10^3 \text{ for CH}_4$$

$$\begin{aligned}
 & \text{Henry's law constant } k_{H,cc} \text{ at a temperature } T \\
 & = k_{H,cp} \text{ at a temperature } T \cdot \text{incubation temperature } T \text{ [K]} / 12.2
 \end{aligned}$$

Then, gas concentration in water and after equilibrium can be calculated.

$$\begin{aligned}
 & \text{gas concentration in water after equilibrium [M]} \\
 & = \text{gas concentration in headspace [M]} \\
 & \times \text{Henry's law constant } k_{H,cc} \text{ at a temperature } T
 \end{aligned}$$

Then, gas amount in water and after equilibrium can be calculated.

$$\begin{aligned}
 & \text{gas amount in water after equilibrium [mol]} \\
 & = \text{gas concentration in water after equilibrium [M]} \\
 & \times \text{volume of sampled water [mL]} / 1000
 \end{aligned}$$

Then, the total amount of gas before equilibrium can be calculated.

$$\begin{aligned}
 & \text{total gas amount before equilibrium [mol]} \\
 & = \text{gas amount in water after equilibrium [mol]} \\
 & + \text{gas amount in headspace [mol]}
 \end{aligned}$$

Finally, sample gas concentration can be calculated.

$$\text{gas concentration in sample [M]} = \frac{\text{total gas amount before equilibrium [mol]}}{\text{volume of sampled water [mL]} / 1000}$$

A-2.1.2 Phenol-chloroform DNA extraction method

1. Resuspend biomass in TE buffer pH 7.5-8.0.
2. Add lysozyme to 150 mg/L.
3. Incubate 2 hours at 37 °C.
4. Add SDS to 0.1%.
5. Freeze/thaw the sample 3 times.
6. Treat the lysate with Proteinase K (100 mg/L) for 1 hour at 55 °C.
7. Add one volume of phenol/chloroform/isoamyl alcohol (pH 7.5-8.0, TE-saturated).
8. Mix by inverting rapidly the tubes.
9. Centrifuge 1 minute at 14'000 xg at room temperature.
10. Pipet the aqueous phase to a fresh tube.
11. Repeat steps 6 to 9 with the aqueous phase until no protein is visible at the interface of the phases.
12. Add to the aqueous phase one volume of chloroform/isoamyl alcohol.
13. Repeat steps 7 to 9.
14. Add 0.1 volume of Na-acetate 3M pH 5.2.
15. Add glycogen 1 µL of a solution of 20 g/L of glycogen.
16. Add 2 volume of EtOH 100%.
17. Mix gently.
18. Mix and incubate over night at 4 °C or -20 °C, or 1 hour at -20 °C or -80 °C.
19. Centrifuge 20 minutes at 16'000 xg at (4 °C).
20. Discard carefully the supernatant.
21. Add some 70% EtOH (at -20 °C if possible).

22. Centrifuge 15 minutes at 16'000 xg at 4 °C.
23. Remove carefully the supernatant and remove all drops around the pellet.
24. Let it dry for 10 minutes.
25. Resuspend the pellet in TE buffer, pH 7.5-8.0.
26. Incubate at 60 °C for 15 minutes.
27. Freeze the sample at -20 °C if needed to be store for more than 2 days (otherwise at 4 °C).

A-2.1.3 Sequences deposition

The raw 16S itag reads from JGI were deposited to NCBI SRA under accession number SRA244825.

The raw 16S reads from Research & Testing Laboratory were deposited to NCBI SRA under accession numbers SRX957365 and SRX957366.

16S rRNA gene sequences were deposited to NCBI GenBank under accession numbers KP901405 - KP902411 and KP942780 - KP942813.

The raw metagenomic reads from Lausanne Genomic Technologies Facility were deposited to NCBI SRA under accession numbers SRX951251, SRX955797, SRX955810, SRX955847, SRX957166, SRX957167, SRX957183, SRX957184, SRX957186 and SRX957187.

The raw metagenomic reads from JGI were deposited to NCBI SRA under accession numbers SRA246967, SRA246971, SRA246972, SRA246976, SRA246982, SRA246984.

The 15 selected draft genomes were deposited in this way:

- Gammaproteobacteria bacterium BRH_c0: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADM00000000. The version described in this paper is version LADM01000000.
- Peptococcaceae bacterium BRH_c4a: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADN00000000. The version described in this paper is version LADN01000000.
- Peptococcaceae bacterium BRH_c4b: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADO00000000. The version described in this paper is version LADO01000000.

- Peptococcaceae bacterium BRH_c8a: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADP00000000. The version described in this paper is version LADP01000000.
- *Hoeflea* sp. BRH_c9: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADQ00000000. The version described in this paper is version LADQ01000000.
- *Desulfatitalea* sp. BRH_c12: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADR00000000. The version described in this paper is version LADR01000000.
- Desulfobulbaceae bacterium BRH_c16a: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADS00000000. The version described in this paper is version LADS01000000.
- Clostridiaceae bacterium BRH_c20a: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADT00000000. The version described in this paper is version LADT01000000.
- *Hyphomonas* sp. BRH_c22: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADU00000000. The version described in this paper is version LADU01000000.
- Peptococcaceae bacterium BRH_c23: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADV00000000. The version described in this paper is version LADV01000000.
- Hyphomonadaceae bacterium BRH_c29: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADW00000000. The version described in this paper is version LADW01000000.
- *Pseudomonas* sp. BRH_c35: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADX00000000. The version described in this paper is version LADX01000000.
- *Roseovarius* sp. BRH_c41: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADY00000000. The version described in this paper is version LADY01000000.

- Flavobacteriales bacterium BRH_c54: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADZ00000000. The version described in this paper is version LADZ01000000.
- Rhodospirillaceae bacterium BRH_c57: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LAEA00000000. The version described in this paper is version LAEA01000000.

All the other contigs (unclassified) were deposited in this way:

- Rock porewater metagenome: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LADL00000000. The version described in this paper is version LADL01000000.

Gene annotation is available at IMG JGI portal, under Taxon ID 3300002468.

A-2.2 Supplementary Tables

Table A-2.1 Number of OTUs observation for every sample. The last column gives a taxonomic affiliation for each OTU. The sequences of the two last samples (day_482 and day_505) were obtained from a different sequencing facility, but were mapped on OTUs defined for the 31 other samples. OTUs 1065 to 1098 corresponds to OTUs only present in these two last samples (day_482 and day_505). More information is available in the section 4.2.5. OTU labels are the same that the ones from Table A-1.3. This table is too big for being printed in this document. It is available online here: dx.doi.org/10.5075/epfl-thesis-6727.

Table A-2.2 Clustering table indicating the raw CONCOCT output and the same output that was manually corrected. This table is too big for being printed in this document. It is available online here: dx.doi.org/10.5075/epfl-thesis-6727.

shaded in grey. For those, ANI analysis was carried out. For each 16S rRNA gene, its contig ID, its position on this contig, its length and its completeness are indicated. IMG phylogenetic distribution is also indicated for each cluster, even though this information was not used for assessing the taxonomic affiliation of clusters.

16S rRNA gene		MLtreeMap taxonomic annotation					
Cluster	RDP taxonomic affiliation	NCBI BLASTN best hit	Average Nucleotide Identity	using default tree	using geBa tree	IMG phylogenetic distribution	Final annotation
0			<i>Halilea salixgens</i> (77%)	<i>Saccharophagus</i> (71%)	Gamma proteobacteria (100%)	Gamma proteobacteria (89%)	Gamma proteobacteria
1						<i>Pseudomonas</i> (55%)	
2	<i>Roseovarius</i>	<i>Roseovarius tolerans</i> (95%)		Rhodobacteraceae (100%)	Rhodobacteraceae (100%)	Rhodobacteraceae (85%)	<i>Roseovarius</i>
3				<i>Erythrobacter</i> (94%)	Sphingomonadales (100%)	Sphingomonadales (85%)	Sphingomonadales
4a	Peptococaceae	<i>Desulfotomaculum gibsoniae</i> (92%)	<i>Desulfotomaculum gibsoniae</i> (79%)	Peptococaceae (100%)	Clostridia (100%)	Peptococaceae (71%)	Peptococaceae
4b	Peptococaceae	<i>Cryptanaerobacter phenolicus</i> (90%)					
4c	Peptococaceae	<i>Pelotomaculum terephthalicum</i> (88%)	<i>Desulfotomaculum gibsoniae</i> (82%)	Peptococaceae (100%)	Peptococaceae (100%)	Peptococaceae (74%)	Peptococaceae
5	Alphaproteobacteria	<i>Pannonibacter phragmitetus</i> (89%)		Peptococaceae (100%)	Peptococaceae (89%)	Peptococaceae (62%)	Peptococaceae
6				Rhodospirillales (100%)	Alphaproteobacteria (100%)	Alphaproteobacteria (74%)	Alphaproteobacteria
7a	<i>Gracilibacter</i>	<i>Gracilibacter thermotolerans</i> (97%)		<i>Desulfifibacterium</i> (83%)	<i>Desulfifibacterium</i> (83%)	Peptococaceae (69%)	<i>Gracilibacter</i>
7b						Clostridiales (45%)	
8a	Peptococaceae	<i>Desulfotomaculum gibsoniae</i> (92%)	<i>Desulfotomaculum alcoholivorax</i> (78%)	Peptococaceae (100%)	Peptococaceae (100%)	<i>Desulfotomaculum</i> (88%)	Peptococaceae
8b	Peptococaceae	<i>Desulfotomaculum gibsoniae</i> (95%)					
8c	Peptococaceae	<i>Desulfotomaculum thermosporovans</i> (93%)					
9	<i>Hoeflea</i>	<i>Hoeflea marina</i> (89%)	<i>Hoeflea</i> sp. (83%)	Rhizobiales (100%)	Rhizobiales (92%)	<i>Desulfosporosinus</i> (82%)	<i>Hoeflea</i>
10				Clostridia (67%)	Peptococaceae (67%)	<i>Hoeflea</i> (52%)	<i>Hoeflea</i>
11	<i>Rhizobium</i>	<i>Rhizobium selenitireducens</i> (89%)		Rhizobaceae (100%)	Rhizobium (100%)	Clostridia (51%)	Clostridia
12	<i>Desulfitalea</i>	<i>Desulfitalea lepidiphila</i> (97%)	<i>Desulfococcus multivorans</i> (77%)	Deltaproteobacteria (79%)	<i>Desulfococcus</i> (93%)	Rhizobaceae (80%)	Rhizobium
13	Xanthomonadaceae	<i>Rehabacterium terrae</i> (95%)		Xanthomonadaceae (100%)	Xanthomonadaceae (100%)	<i>Desulfitalea</i>	<i>Desulfitalea</i>
14						Alphaproteobacteria (78%)	Xanthomonadaceae
15	<i>Brevundimonas</i>	<i>Brevundimonas intermedia</i> (99%)					
16a	<i>Desulfobulbaceae</i>	<i>Desulfobacterium thiosulfatigenes</i> (93%)					
16b	<i>Desulfotomaculum</i>	<i>Desulfotomaculum vacuolans</i> (93%)	<i>Desulfocapsa sulfexigens</i> (76%)	<i>Desulfitalea</i> (79%)	<i>Desulfitalea</i> (79%)	Desulfobulbaceae (42%)	Desulfobulbaceae
17							
18	<i>Desulfifibacter</i>	<i>Desulfifibacter alkalicolerans</i> (95%)		Enterobacteriaceae (100%)	<i>Shigella</i> (100%)	Alphaproteobacteria (77%)	Enterobacteriaceae
19	<i>Desulfifibacter</i>	<i>Desulfifibacter alkalicolerans</i> (95%)		Clostridiales (80%)	Clostridia (80%)	<i>Pseudomonas</i> (86%)	<i>Desulfifibacter</i>
20a	Firmicutes	<i>Proteinibacter ethanologenes</i> (83%)	No genome match	Clostridia (100%)	Clostridia (70%)	<i>Escherichia coli</i> (87%)	
20b				<i>Moorella</i> (100%)	Clostridiales (64%)	Peptococaceae (34%)	
21	<i>Hyphomonas</i>	<i>Hyphomonas neptunium</i> (97%)	<i>Hyphomonas oceanitis</i> (83%)	<i>Thiobacillus</i> (100%)	<i>Thiobacillus</i> (100%)	Clostridia (72%)	Firmicutes
22				Hyphomonadaceae (100%)	Hyphomonadaceae (100%)	<i>Thiobacillus</i> (71%)	Clostridia
23			<i>Desulfotomaculum gibsoniae</i> (81%)	<i>Desulfifibacterium</i> (100%)	<i>Desulfifibacterium</i> (100%)	<i>Hyphomonas</i> (54%)	<i>Hyphomonas</i>
						<i>Desulfosporosinus</i> (63%)	Peptococaceae

Table A-2.4 (continued)

Cluster	RDP taxonomic affiliation	16S rRNA gene	MLtreeMap taxonomic annotation					
			NCBI BLASTN best hit	Average Nucleotide Identity	using default tree	using gaba tree	IMG phylogenetic distribution	Final annotation
24	<i>Coralliomargarita</i>	<i>Coralliomargarita akajimensis</i> (86%)			Chlamydiales (100%) Clostridia (94%)	<i>Opitulus</i> (100%) Clostridiaceae (88%)	Alphaproteobacteria (43%) Clostridiales (63%) Alphaproteobacteria (67%)	<i>Coralliomargarita</i> Clostridia
25								
26	<i>Anaerolinea</i>	<i>Anaerolinea thermophila</i> (87%)			Chloroflexi (78%)	Chloroflexi (78%)	<i>Anaerolinea</i> (46%)	<i>Anaerolinea</i>
27	<i>Escherichia/Shigella</i>	<i>Escherichia fergusoni</i> (99%)						<i>Escherichia/Shigella</i>
28	<i>Hyphomonas</i>	<i>Hyphomonas adhaerens</i> (99%)	<i>Hyphomonas</i> sp. (83%)		<i>Hyphomonas</i> (100%) <i>Gramella</i> (100%) Bacteria (100%) Bacteroidetes/Chlorobi group (100%)	<i>Hyphomonas</i> (100%) Flavobacteriaceae (100%) Bacteria (100%) Bacteroidetes/Chlorobi group (100%)	Hyphomonadaceae (73%) Flavobacteriaceae (73%) Deltaproteobacteria (53%) Ignavibacteriaceae (67%) Bacteria (95%)	Hyphomonadaceae Flavobacteriaceae Bacteria Bacteria
29								
30								
31								
32	Bacteria	<i>Ignavibacterium album</i> (86%)			Alteromonadaceae (100%) <i>Pseudomonas</i> (100%) Alphaproteobacteria (100%)	Gammaproteobacteria (100%) <i>Pseudomonas</i> (100%) Alphaproteobacteria (100%)	Gammaproteobacteria (73%) <i>Pseudomonas</i> (93%) Rhizobiales (71%)	Gammaproteobacteria <i>Pseudomonas</i> Alphaproteobacteria
33								
34								
35	<i>Pseudomonas</i>	<i>Pseudomonas chloritidis/mutans</i> (99%)	<i>Pseudomonas xanthomarina</i> (98%)		<i>Desulfotribacterium</i> (100%)	<i>Desulfotribacterium</i> (100%)	<i>Desulfosporosinus</i> (61%)	<i>Desulfosporosinus</i>
36	<i>Desulfosporosinus</i>	<i>Desulfosporosinus lacus</i> (96%)			<i>Rhodopirellula</i> (83%)	<i>Rhodopirellula</i> (100%)		Planctomycetes
37	<i>Desulfosporosinus</i>	<i>Desulfosporosinus lacus</i> (97%)						
38								
39								
40								
40*								
41	<i>Roseovarius</i>	<i>Roseovarius mucosus</i> (99%)	<i>Roseovarius</i> sp. (98%)		Alphaproteobacteria (95%) Sphingomonadales (100%) Rhodobacteraceae (100%) Rhodobacteraceae (100%) Sphingomonadales (100%) Flavobacteriaceae (100%)	Alphaproteobacteria (100%) Sphingomonadales (100%) Flavobacteriaceae (100%) Aquificales (50%)	Planctomycetes (82%) Alphaproteobacteria (67%) Alphaproteobacteria (77%) Sphingomonadales (71%) <i>Roseovarius</i> (84%)	Planctomycetes Alphaproteobacteria Alphaproteobacteria Sphingomonadales <i>Roseovarius</i>
42								
43								
44								
45								
46	<i>Parvularcula</i>	<i>Parvularcula oceani</i> (92%)			Alphaproteobacteria (98%) <i>Biridobacterium</i> (100%)	Alphaproteobacteria (100%) <i>Beutenbergia</i> (100%)	Alphaproteobacteria (73%) Actinomycetales (76%) Alphaproteobacteria (64%)	<i>Parvularcula</i> Bacteria
47								
48								
49	Bacteria	<i>Caloramator indicus</i> (83%)			Clostridia (70%)	Firmicutes (90%)	Clostridia (56%)	Bacteria
50	Bacteria	<i>Caloramator indicus</i> (81%)						
51								
52	<i>Lutibacter</i>	<i>Lutibacter maritimus</i> (95%)			<i>Erythrobacter</i> (95%)	<i>Erythrobacter</i> (100%)	Erythrobacteraceae (75%) Rhodobacteraceae (56%)	Erythrobacteraceae
53*	<i>Sphingobium</i>	<i>Sphingobium xenophagum</i> (99%)			Flavobacteriaceae (100%) Sphingomonadales (100%) Bacteroidetes (100%)	Flavobacteriaceae (100%) Sphingomonadaceae (100%) Bacteroidetes (100%)	Flavobacteriaceae (80%) Sphingomonadales (70%) Bacteroidetes (89%)	<i>Lutibacter</i> <i>Sphingobium</i> Owenweeksia
54	<i>Owenweeksia</i>	<i>Owenweeksia hongkongensis</i> (89%)	<i>Crocinitomix catalasitica</i> (73%)		Gammaproteobacteria (67%)	Gammaproteobacteria (67%)	Gammaproteobacteria (51%)	Gammaproteobacteria
55								
56								
57	Rhodospirillaceae	<i>Marispirillum indicum</i> (94%)	<i>Caenispirillum salinarum</i> (77%)		Rhodospirillaceae (93%) Sphingomonadales (100%)	Rhodospirillaceae (93%) Sphingomonadaceae (100%)	Rhodospirillaceae (65%) Sphingomonadales (70%)	Rhodospirillaceae <i>Sphingobium</i>
58*	<i>Sphingobium</i>	<i>Sphingobium xenophagum</i> (99%)						

* Clusters 53 and 40x have been merged to a newly defined cluster 58

Table A-2.5 Analysis of 16S rRNA genes recovered from metagenomic binning. The table includes 16S RNA gene localization and length, its RDP classification, its NCBI BLASTN best hit, and its corresponding OTU number from the 16S rRNA sequencing project encompassing 23 samples coming from 8 boreholes (Table A-1.5).

NCBI BLASTN										Comparison with itag sequencing (region V4 of 16S rRNA gene)				
Cluster	Contig ID	Gene start position	Gene stop position	Gene length (bp)	Is 16S rRNA gene complete?	RDP taxonomic affiliation	Best hit	Reference	Identity	OTU #	Alignment length	Alignment identity	number of OTU observations	RDP taxonomic affiliation
2	BRHa_1000274	531	1996	1466	yes	Roseovarius	<i>Roseovarius tolerans</i>	NR_026405.1	96%		no match found	no match found		
4a	BRHa_1002702	2	494	493	no	Peptococcaceae	<i>Desulfotomaculum gibsoniae</i>	NR_114759.1	92%	29	253	100	13/23	<i>Desulfotomaculum</i>
4b	BRHa_1004082	38170	39701	1531	yes	Peptococcaceae	<i>Cryptanaerobacter phenolicus</i>	NR_025757.1	90%		no match found	no match found		
5	BRHa_1000636	1	324	324	no	Peptococcaceae	<i>Peblotomaculum terethylalicum</i>	NR_040948.1	88%		no match found	no match found		
7a	BRHa_1000961	18676	20170	1495	yes	Alphaproteobacteria	<i>Pannonibacter pragmatitus</i>	NR_028009.1	89%	11	253	100	20/23	Rhodospirillaceae
	BRHa_1002652	3979	5488	1520	yes	Gracilbacter	<i>Gracilbacter thermotolerans</i>	NR_115692.1	97%	70	253	100	10/23	Clostridiales
	BRHa_1005448	227403	227813	411	no	Peptococcaceae	<i>Desulfotomaculum gibsoniae</i>	NR_103939.1	92%		no match found	no match found		
8a	BRHa_1000843	1	559	559	no	Peptococcaceae	<i>Desulfotomaculum gibsoniae</i>	NR_103939.1	93%		no match found	no match found		
	BRHa_1004053	339	1868	1530	yes	Peptococcaceae	<i>Desulfotomaculum thermosapovorans</i>	NR_044948.1	93%	7	253	100	19/23	<i>Desulfotomaculum</i>
	BRHa_1004053	5834	7322	1489	no	Peptococcaceae	<i>Desulfotomaculum gibsoniae</i>	NR_114761.1	93%	7	253	100	19/23	<i>Desulfotomaculum</i>
9	BRHa_1003556	1	1000	1000	no	Hoeflea	<i>Hoeflea marina</i>	NR_043007.1	99%		no match found	no match found		
11	BRHa_1003273	4602	6083	1482	yes	Rhizobium	<i>Rhizobium selenitireducens</i>	NR_044216.1	99%	65	253	100	10/23	Rhizobiales
12	BRHa_1002295	424014	425575	1562	yes	Desulfatitalea	<i>Desulfatitalea tepidiphila</i>	NR_113315.1	97%	46	253	100	13/23	Desulfobacteraceae
13	BRHa_1004451	10345	11881	1537	yes	Xanthomonadaceae	<i>Rehabacterium terrae</i>	NR_118587.1	95%	4	253	100	22/23	Xanthomonadaceae
15	BRHa_1001180	2116	3573	1458	yes	Brevundimonas	<i>Brevundimonas intermedia</i>	NR_041966.1	99%	48	253	100	21/23	<i>Brevundimonas</i>
	BRHa_1000106	2	1385	1384	no	Desulfonispota	<i>Desulfonispota thiosulfatigenes</i>	NR_026497.1	93%	24	253	100	19/23	Peptococcaceae
16a	BRHa_1003205	4192	5733	1542	yes	Desulfobulbaceae	<i>Desulfobacterium catecholicum</i>	NR_028895.1	95%	0	253	100	23/23	<i>Desulfocapsa</i>
	BRHa_1002043	112079	112441	363	no	Desulforhopalus	<i>Desulforhopalus vacuolatus</i>	NR_044653.1	93%		no match found	no match found		
	BRHa_1000294	1	348	348	no	Desulforhopalus	<i>Desulforhopalus vacuolatus</i>	NR_044653.1	93%		no match found	no match found		
19	BRHa_1004424	73008	73383	376	no	Desulfifitibacter	<i>Desulfifitibacter alkalicolerans</i>	NR_042962.1	95%		no match found	no match found		
	BRHa_1005141	47219	47602	384	no	Desulfifitibacter	<i>Desulfifitibacter alkalicolerans</i>	NR_042962.1	95%		no match found	no match found		
20a	BRHa_1004490	145	1673	1529	yes	Firmicutes	<i>Proteiniborus ethanoligenes</i>	NR_044093.1	85%	26	253	100	11/23	Clostridiales
22	BRHa_1004732	3751	5202	1452	yes	Hyphomonas	<i>Hyphomonas neptunium</i>	NR_074092.1	97%	34	253	100	18/23	Hyphomonas
24	BRHa_1000784	317	1848	1532	yes	Coralliomargarita	<i>Coralliomargarita akeimensis</i>	NR_074901.1	86%	49	253	100	11/23	Punicetocaceae
27	BRHa_1007068	9652	10007	356	no	Anaerolinea	<i>Anaerolinea thermophila</i>	NR_074383.1	87%		no match found	no match found		
28	BRHa_1004480	3294	4831	1538	yes	Escherichia/Shigella	<i>Escherichia fergusonii</i>	NR_074902.1	99%	590	253	100	3/23	<i>Escherichia</i>
29	BRHa_1007353	12	780	769	no	Hyphomonas	<i>Hyphomonas adhaerens</i>	NR_024937.1	99%	12	253	100	21/23	Hyphomonas
32	BRHa_1006978	36856	37389	534	no	Bacteria	<i>Ignavibacterium album</i>	NR_074698.1	86%		no match found	no match found		
35	BRHa_1000795	3721	5253	1533	yes	Pseudomonas	<i>Pseudomonas chloritidismutans</i>	NR_115115.1	99%	2	253	100	23/23	<i>Pseudomonas</i>
	BRHa_1002890	2	377	376	no	Desulfosporosinus	<i>Desulfosporosinus facus</i>	NR_042202.1	96%		no match found	no match found		
37	BRHa_1004950	1	333	333	no	Desulfosporosinus	<i>Desulfosporosinus facus</i>	NR_042202.1	96%		no match found	no match found		
	BRHa_1004477	1	365	365	no	Desulfosporosinus	<i>Desulfosporosinus facus</i>	NR_042202.1	97%		no match found	no match found		
41	BRHa_1003404	4349	5815	1467	yes	Roseovarius	<i>Roseovarius mucosus</i>	NR_042159.1	99%	3	253	100	23/23	Rhodobacteraceae
46	BRHa_1000455	2895	4356	1462	yes	Parvularcula	<i>Parvularcula oceanii</i>	NR_125641.1	92%	60	253	100	8/23	Rhodobacterales
49	BRHa_1003600	36114	36476	363	no	Bacteria	<i>Caloramator indicus</i>	NR_026134.1	83%		no match found	no match found		
	BRHa_1002025	2	324	323	no	Bacteria	<i>Caloramator indicus</i>	NR_026134.1	81%		no match found	no match found		
52	BRHa_1007118	120	1631	1512	yes	Lutibacter	<i>Lutibacter maritimus</i>	NR_116738.1	95%	42	253	99.6	16/23	Flavobacteriaceae
54	BRHa_1002744	1	1328	1328	no	Owenweekia	<i>Owenweekia hongkongensis</i>	NR_074100.1	89%	188	253	100	6/23	Flavobacteria
57	BRHa_1007414	585	2069	1485	yes	Rhodospirillaceae	<i>Marispirillum indicum</i>	NR_044545.1	94%	1	253	100	23/23	<i>Novispirillum</i>
58	BRHa_1003529	3616	5101	1486	yes	Springobium	<i>Springobium xenophagum</i>	NR_026304.1	99%	89	253	99.21	18/23	<i>Springobium</i>

Table A-2.7 ANI analysis of the 15 selected draft genomes.

Cluster	Most similar genome	ANI	Assembly	Genomes tested
c0	<i>Haliera salexigens</i> DSM 19537	77.38%	GCA_000423125.1	Each Alteromonadaceae genera
c4a	<i>Desulfotomaculum gibsoniae</i> DSM 7213	78.77%	GCA_000233715.3	All <i>Desulfotomaculum</i> and <i>Pelotomaculum</i>
c4b	<i>Desulfotomaculum gibsoniae</i> DSM 7213	82.44%	GCA_000233715.3	All <i>Desulfotomaculum</i> and <i>Pelotomaculum</i>
c8a	<i>Desulfotomaculum alcoholivorax</i> DSM 16058	78.00%	GCA_000430885.1	All <i>Desulfotomaculum</i> and <i>Pelotomaculum</i>
c9	<i>Hoeflea</i> sp. BAL378	83.47%	GCA_000759435.1	All <i>Hoeflea</i>
c12	<i>Desulfococcus multivorans</i> DSM 2059	76.95%	GCA_000422185.1	15 top NCBI BLASTN hits of 16S rRNA gene(s)
c16a	<i>Desulfocapsa sulfexigens</i> DSM 10523	76.30%	GCA_000341395.1	15 top NCBI BLASTN hits of 16S rRNA gene(s)
c20a	<i>No genome match</i>			
c22	<i>Hyphomonas oceanitis</i> SCH89	83.31%	GCA_000685295.1	All <i>Hyphomonas</i>
c23	<i>Desulfotomaculum gibsoniae</i> DSM 7213	80.88%	GCA_000233715.3	All Peptococcaceae
c29	<i>Hyphomonas</i> sp. BH-BN04-4	83.06%	GCA_000682695.1	All Hyphomonadaceae
c35	<i>Pseudomonas xanthomarina</i> S11	98.21%	GCA_000825645.1	15 top NCBI BLASTN hits of 16S rRNA gene(s)
c41	<i>Roseovarius</i> sp. 217	98.07%	GCA_000152845.1	All <i>Roseovarius</i>
c54	<i>Crocinitomix catalasitica</i> ATCC 23190	73.33%	GCA_000621625.1	15 top NCBI BLASTN hits of 16S rRNA gene(s)
c57	<i>Caenispirillum salinarum</i> AK4	77.15%	GCA_000315795.1	15 top NCBI BLASTN hits of 16S rRNA gene(s)

Table A-2.8 Protein profile of sample recovered at day 483. Average adjusted protein NSAF was calculated by averaging the NSAF of run 1 (Table A-2.9) and run 2 (Table A-2.10) and then by multiplying this value by 10^5 . This table is too big for being printed in this document. It is available online here: dx.doi.org/10.5075/epfl-thesis-6727.

Table A-2.9 Protein profile of the first run of the sample recovered at day 483. Adjusted NSAF were calculated by multiplying NSAF by 10^5 . This table is too big for being printed in this document. It is available online here: dx.doi.org/10.5075/epfl-thesis-6727.

Table A-2.10 Protein profile of the second run of sample recovered at day 483. Adjusted NSAF were calculated by multiplying NSAF by 10^5 . This table is too big for being printed in this document. It is available online here: dx.doi.org/10.5075/epfl-thesis-6727.

Table A-2.11 List of proteins belonging to Desulfobulbaceae c16a and involved in processes described in Figure 4.10. In the second column, P stands for proteome, while G for genome.

Reaction no.	Protein coding gene ID	In proteome (P) or genome (G) only?	Subcellular localization	Gene product	COG number	KO number	EC number
1	BRHa_100687050	P	Cytoplasmic membrane	Ni ₂ Fe-hydrogenase I small subunit	COG1740	K06282	EC:1.12.99.6
1	BRHa_100687051	P	Cytoplasmic membrane	Ni ₂ Fe-hydrogenase I large subunit	COG0374	K06281	EC:1.12.99.6
2	BRHa_100267644	P	Unknown	Heterodisulfide reductase, subunit A and related polyferredoxins	COG1148	K03388	EC:1.8.98.1
2	BRHa_100267645	P	Cytoplasm	Heterodisulfide reductase, subunit A and related polyferredoxins	COG1148	K03388	EC:1.8.98.1
2	BRHa_100267646	P	Cytoplasmic membrane	Heterodisulfide reductase, subunit C	COG1150	-	-
3	BRHa_1004263109	G	Cytoplasmic membrane	Polysulfide reductase	COG5557	K00185	EC:1.2.7.-
3	BRHa_1004263110	G	Unknown	Fe-S-cluster-containing hydrogenase components 1	COG0437	-	-
3	BRHa_1004263112	P	Cytoplasm	Fe-S oxidoreductase	COG0247	-	-
3	BRHa_1004263113	P	Cytoplasmic membrane	Nitrate reductase gamma subunit	COG2181	-	-
4	BRHa_100252619	P	Cytoplasm	ATP sulfurylase (sulfate adenyltransferase)	COG2046	K00958	EC:2.7.7.4
5	BRHa_100267642	P	Unknown	Adenosine-5'-phosphosulfate reductase beta subunit;	COG4231	K00395	EC:1.8.99.2
5	BRHa_100267643	P	Cytoplasm	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG1053	K00394	EC:1.8.99.2
6	BRHa_100622935	P	Cytoplasm	Dissimilatory sulfite reductase (desulfoviridin), alpha and beta subunits	COG2221	K11180	EC:1.8.99.3
6	BRHa_100622936	P	Cytoplasm	Dissimilatory sulfite reductase (desulfoviridin), alpha and beta subunits	COG2221	K11181	EC:1.8.99.3
6	BRHa_100622937	G	Unknown	Dissimilatory sulfite reductase D (Dsr/D)	-	-	-
7	BRHa_100622972	P	Unknown	F0F1-type ATP synthase, subunit b	COG0711	K02109	EC:3.6.3.14
7	BRHa_100622973	P	Unknown	F0F1-type ATP synthase, subunit b	COG0711	K02109	EC:3.6.3.14
7	BRHa_100622974	P	Cytoplasm	F0F1-type ATP synthase, delta subunit (mitochondrial oligomycin sensitivity protein)	COG0712	K02113	EC:3.6.3.14
7	BRHa_100622975	P	Cytoplasm	F0F1-type ATP synthase, alpha subunit	COG0056	K02111	EC:3.6.3.14
7	BRHa_100622976	G	Cytoplasm	F0F1-type ATP synthase, gamma subunit	COG0224	K02115	EC:3.6.3.14
7	BRHa_100622977	P	Cytoplasm	F0F1-type ATP synthase, beta subunit	COG0055	K02112	EC:3.6.3.14
7	BRHa_100622978	P	Unknown	F0F1-type ATP synthase, epsilon subunit (mitochondrial delta subunit)	COG0355	K02114	EC:3.6.3.14
7	BRHa_100622955	G	Cytoplasmic membrane	F0F1-type ATP synthase, subunit c/Archaeal/vacuolar-type H ⁺ -ATPase, subunit K	COG0636	K02110	EC:3.6.3.14
7	BRHa_100622954	G	Cytoplasmic membrane	F0F1-type ATP synthase, subunit a	COG0356	K02108	EC:3.6.3.14
8	BRHa_1001515120	P	Cytoplasm	Coenzyme F420-reducing hydrogenase, gamma subunit	COG1941	K14128	EC:1.12.99.-
8	BRHa_1001515121	P	Cytoplasm	Coenzyme F420-reducing hydrogenase, alpha subunit	COG3259	K14126	EC:1.12.99.-
9	BRHa_1007437112	P	Cytoplasm	2-polyphenylphenol hydroxylase and related flavodoxin oxidoreductases	COG0543	K00528	EC:1.18.1.2
10	BRHa_100617011	P	Cytoplasm	NADPH:quinone reductase and related Zn-dependent oxidoreductases	COG0604	K00344	EC:1.6.5.5
11	BRHa_10068707	P	Cytoplasmic membrane	Predicted NADH:ubiquinone oxidoreductase, subunit RnfB - COG2878	COG2878	K03616	-
11	BRHa_10068708	P	Cytoplasmic membrane	Predicted NADH:ubiquinone oxidoreductase, subunit RnfA - COG4657	COG4657	K03617	-
11	BRHa_10068709	G	Cytoplasmic membrane	Predicted NADH:ubiquinone oxidoreductase, subunit RnfE - COG4660	COG4660	K03613	-
11	BRHa_100687010	G	Cytoplasm	Predicted NADH:ubiquinone oxidoreductase, subunit RnfG - COG4659	COG4659	K03612	-
11	BRHa_100687011	G	Cytoplasmic membrane	Predicted NADH:ubiquinone oxidoreductase, subunit RnfD - COG4658	COG4658	K03614	-
11	BRHa_100687012	P	Cytoplasm	Predicted NADH:ubiquinone oxidoreductase, subunit RnfC - COG4656	COG4656	K03615	-
12	BRHa_1007437101	P	Cytoplasm	Carbonic anhydrases/acetyltransferases, isoleucine patch superfamily	COG0663	-	-
13	BRHa_100354542	G	Cytoplasm	Uncharacterized anaerobic hydrogenase	COG3383	K00123	EC:1.2.1.2
13	BRHa_100354543	P	Cytoplasm	Uncharacterized anaerobic dehydrogenase	COG3383	K00123	EC:1.2.1.2
14	BRHa_100354519	P	Cytoplasm	Formyltetrahydrofolate synthetase	COG2759	K01938	EC:3.6.3.4.3
15-16	BRHa_100354520	P	Cytoplasm	5,10-methylene-tetrahydrofolate dehydrogenase/Methenyl tetrahydrofolate cyclohydrolase	COG0190	K01491	EC:1.5.1.5
17	BRHa_10037249	P	Cytoplasm	5,10-methylenetetrahydrofolate reductase	COG0685	K00297	EC:1.5.1.20
18	BRHa_100354562	P	Cytoplasm	Pterin binding enzyme	COG1410	-	-
19	BRHa_100354559	P	Cytoplasm	6Fe-6S prismatic cluster-containing protein	COG1151	K00198	EC:1.2.99.2
19	BRHa_100354565	P	Cytoplasm	6Fe-6S prismatic cluster-containing protein	COG1151	K00190	EC:1.2.7.4
20	BRHa_100354566	P	Cytoplasm	CO dehydrogenase/acetyl-CoA synthase delta subunit (corrinoid Fe-S protein)	COG2069	K00194	EC:2.1.1.245
20	BRHa_100354563	P	Unknown	CO dehydrogenase/acetyl-CoA synthase gamma subunit (corrinoid Fe-S protein)	COG1456	K00197	EC:2.1.1.245
20	BRHa_100354564	P	Cytoplasm	CO dehydrogenase/acetyl-CoA synthase beta subunit	COG1614	K14138	EC:1.2.7.4
21	BRHa_1001804275	P	Unknown	Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, alpha subunit	COG0674	K03737	EC:1.2.7.-
21	BRHa_1000308122	P	Unknown	Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, alpha subunit	COG0674	K00169	EC:1.2.7.-
22	BRHa_1004263141	P	Cytoplasm	Pyruvate carboxylase	COG1038	K01958	EC:6.4.1.1
22	BRHa_100115516	P	Periplasm	Pyruvate/oxaloacetate carboxyltransferase	COG5016	K01960	EC:6.4.1.1
22	BRHa_100227747	P	Cytoplasm	Pyruvate/oxaloacetate carboxyltransferase	COG5016	K01960	EC:6.4.1.1
23	BRHa_100332486	P	Cytoplasm	Citrate synthase	COG0372	K01647	EC:2.3.3.1
24	BRHa_100042063	G	Cytoplasm	Aconitase B	COG1049	K01682	EC:4.2.1.3
24	BRHa_100332494	G	Cytoplasm	Aconitase B	COG1049	K01682	EC:4.2.1.3
24	BRHa_100598884	G	Cytoplasm	Aconitase A	COG1048	K01681	EC:4.2.1.-
25	BRHa_100255622	G	Cytoplasm	Isocitrate dehydrogenases	COG0538	K00031	EC:1.1.1.42
25	BRHa_100042058	G	Cytoplasm	Isocitrate dehydrogenases	COG0538	K00031	EC:1.1.1.42
26	BRHa_100598858	P	Cytoplasm	Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, alpha subunit	COG0674	K00174	EC:1.2.7.3
26	BRHa_100598857	G	Cytoplasm	Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, beta subunit	COG1013	K00175	EC:1.2.7.3
26	BRHa_100598856	G	Unknown	Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, gamma subunit	COG1014	K00177	EC:1.2.7.3
26	BRHa_100598859	G	Unknown	-	COG1152	K00176	EC:1.2.7.3
27	BRHa_100227751	P	Cytoplasm	Succinyl-CoA synthetase, alpha subunit	COG0074	K01902	EC:6.2.1.5
27	BRHa_100227752	P	Cytoplasm	Succinyl-CoA synthetase, beta subunit	COG0045	K01903	EC:6.2.1.5
28	BRHa_100705860	G	Cytoplasmic membrane	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	COG1053	K00239	EC:1.3.99.1
28	BRHa_100705861	P	Cytoplasmic membrane	Succinate dehydrogenase/fumarate reductase, Fe-S protein subunit	COG0479	K00240	EC:1.3.99.1
28	BRHa_100705859	G	Cytoplasmic membrane	-	-	K00241	-
29	BRHa_100029439	G	Cytoplasm	Tartrate dehydratase alpha subunit/Fumarate hydratase class I, N-terminal domain	COG1951	K01676	EC:4.2.1.2
29	BRHa_100029420	G	Cytoplasm	Fumarase	COG0114	K01679	EC:4.2.1.2
30	BRHa_1001804178	P	Unknown	Malate/lactate dehydrogenases	COG0039	K00024	EC:1.1.1.37
30	BRHa_100372494	P	Unknown	Malate/lactate dehydrogenases	COG0039	K00024	EC:1.1.1.37
31	BRHa_100267631	P	Cytoplasm	Phosphoenolpyruvate carboxykinase (GTP)	COG1274	K01596	EC:4.1.1.32
32	BRHa_1000308196	P	Cytoplasm	Enolase	COG0148	K01689	EC:4.2.1.11
33	BRHa_100151541	G	Cytoplasm	Predicted phosphoglycerate mutase, AP superfamily - COG3635	COG3635	K15635	EC:5.4.2.1
33	BRHa_100204380	G	Cytoplasm	Phosphoglyceromutase	COG0696	K01834	EC:5.4.2.1
33	BRHa_100348113	G	Cytoplasm	Fructose-2,6-bisphosphatase	COG0406	K01834	EC:5.4.2.1
34	BRHa_1003749104	G	Cytoplasm	3-phosphoglycerate kinase	COG0126	K00927	EC:2.7.2.3
35	BRHa_100622964	P	Cytoplasm	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	COG0057	K00134	EC:1.2.1.12
35	BRHa_1003749106	P	Cytoplasm	Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate dehydrogenase	COG0057	K00134	EC:1.2.1.12
36	BRHa_1003749103	G	Cytoplasm	Triosephosphate isomerase	COG0149	K01803	EC:5.3.1.1
37	BRHa_1001515284	P	Cytoplasm	Fructose/tagatose bisphosphate aldolase	COG0191	K01624	EC:4.1.2.13
38	BRHa_1001804297	G	Cytoplasm	Archaeal fructose 1,6-bisphosphatase	COG1980	K01622	EC:3.1.3.11

Table A-2.11 (continued)

Reaction no.	Protein coding gene ID	In proteome (P) or genome (G) only?	Subcellular localization	Gene product	COG number	KO number	EC number
39	BRHa_100227759	P	Cytoplasm	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	COG0365	K01895	EC:6.2.1.1
39	BRHa_100622969	P	Cytoplasm	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	COG0365	K01895	EC:6.2.1.1
40	BRHa_100227744	P	Cytoplasm	Methylmalonyl-CoA mutase, N-terminal domain/subunit	COG1884	K01847	EC:5.4.99.2
41	BRHa_1003324154	P	Cytoplasm	Acetyl-CoA carboxylase, carboxyltransferase component (subunits alpha and beta)	COG4799	K01966	EC:6.4.1.3
41	BRHa_100227750	P	Cytoplasm	Acetyl-CoA carboxylase, carboxyltransferase component (subunits alpha and beta)	COG4799	K01966	EC:6.4.1.3
41	BRHa_1003324155	P	Cytoplasm	Acetyl/propionyl-CoA carboxylase, alpha subunit	COG4770	K01965	EC:6.4.1.3
42	BRHa_100227753	P	Cytoplasm	Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	COG0365	K01908	EC:6.2.1.17
43	BRHa_1004263142	P	Cytoplasm	Pyruvate carboxylase	COG1038	K01961	EC:6.3.4.14
43	BRHa_100082530	G	Cytoplasm	Acetyl-CoA carboxylase alpha subunit	COG0825	K01962	EC:6.4.1.2
44	BRHa_100227733	P	Cytoplasm	Malonyl-CoA decarboxylase (MCD)	-	K01578	EC:4.1.1.9
45	BRHa_100252659	P	Cytoplasm	(acyl-carrier-protein) S-malonyltransferase	COG0331	K00645	EC:2.3.1.39
45	BRHa_100227720	G	Cytoplasm	3-oxoacyl-[acyl-carrier-protein] synthase III	COG0332	K00648	EC:2.3.1.180
45	BRHa_100111556	G	Cytoplasm	3-oxoacyl-[acyl-carrier-protein] synthase III	COG0332	K00648	EC:2.3.1.180
45	BRHa_1003324146	G	Cytoplasm	3-oxoacyl-[acyl-carrier-protein] synthase III	COG0332	K00648	EC:2.3.1.180
46	BRHa_1000308102	P	Cytoplasm	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
46	BRHa_1003749190	P	Cytoplasm	Acetyl-CoA acetyltransferase	COG0183	K00626	EC:2.3.1.9
47	BRHa_1001804165	P	Cytoplasmic membrane	Long-chain acyl-CoA synthetases (AMP-forming)	COG1022	K01897	EC:6.2.1.3
47	BRHa_100372481	P	Cytoplasmic membrane	Long-chain acyl-CoA synthetases (AMP-forming)	COG1022	K01897	EC:6.2.1.3
48	BRHa_100613146	P	Cytoplasm	Acetyl-CoA acetyltransferase	COG0183	K00632	EC:2.3.1.16
49	BRHa_100598844	P	Outer membrane	Long-chain fatty acid transport protein	COG2067	K06076	-
50	BRHa_100598838	P	Cytoplasm	Acyl-CoA dehydrogenases	COG1960	K00248	EC:1.3.99.2
50	BRHa_100598839	P	Cytoplasmic membrane	Fe-S oxidoreductase	COG0247	-	-
51	BRHa_100227716	P	Cytoplasm	Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	COG1028	K00059	EC:1.1.1.100
52	BRHa_100111212	P	Cytoplasm	Ketol-acid reductoisomerase	COG0059	K00053	EC:1.1.1.86
53	BRHa_100151542	P	Cytoplasm	Homoserine dehydrogenase	COG0460	K00003	EC:1.1.1.3
54	BRHa_100297220	P	Cytoplasm	Isocitrate/isopropylmalate dehydrogenase	COG0473	K00052	EC:1.1.1.85
55	BRHa_1001515241	P	Cytoplasm	Phosphoglycerate dehydrogenase and related dehydrogenases	COG0111	K00058	EC:1.1.1.95
55	BRHa_100722337	P	Cytoplasm	Phosphoglycerate dehydrogenase and related dehydrogenases	COG0111	K00058	EC:1.1.1.95
56	BRHa_100725848	P	Cytoplasm	Dihydropicolinate reductase	COG0289	K00215	EC:1.3.1.26
57	BRHa_1007437113	P	Cytoplasm	NADPH-dependent glutamate synthase beta chain and related oxidoreductases	COG0493	K00266	EC:1.4.1.13
58	BRHa_100725854	P	Cytoplasm	Ornithine carbamoyltransferase	COG0078	K00611	EC:2.1.3.3
59	BRHa_1004263104	P	Cytoplasm	5-enolpyruvylshikimate-3-phosphate synthase	COG0128	K00800	EC:2.5.1.19
60	BRHa_100725855	P	Cytoplasm	Ornithine/acetylornithine aminotransferase	COG4992	K00818	EC:2.6.1.11
61	BRHa_100541474	P	Cytoplasm	Glucosamine 6-phosphate synthetase	COG0449	K00820	EC:2.6.1.16
62	BRHa_100705853	P	Cytoplasm	Phosphoserine aminotransferase	COG1932	K00831	EC:2.6.1.52
63	BRHa_1007437171	P	Cytoplasm	Aspartate/tyrosine/aromatic aminotransferase	COG1448	K00832	EC:2.6.1.57
64	BRHa_1007437174	P	Cytoplasm	Aspartokinases	COG0527	K00928	EC:2.7.2.4
65	BRHa_100297219	P	Cytoplasm	Phosphoribosylpyrophosphate synthetase	COG0462	K00948	EC:2.7.6.1
66	BRHa_1007437130	P	Cytoplasm	Thiamine pyrophosphate-requiring enzymes	COG0028	K01652	EC:2.2.1.6
67	BRHa_100725849	P	Cytoplasm	Dihydropicolinate synthase/N-acetylneuraminatase lyase	COG0329	K01714	EC:4.3.3.7
68	BRHa_100722334	P	Cytoplasm	Peptidyl-prolyl cis-trans isomerase (rotamase) - cyclophilin family	COG0652	K01802	EC:5.2.1.8
69	BRHa_1006229124	P	Cytoplasm	Glutamine synthetase	COG0174	K01915	EC:6.3.1.2
70	BRHa_100725853	P	Cytoplasm	Argininosuccinate synthase	COG0137	K01940	EC:6.3.4.5
71	BRHa_100705865	P	Cytoplasm	Aspartate/tyrosine/aromatic aminotransferase	COG0436	K01206	EC:2.6.1.83
72	BRHa_100372414	P	Cytoplasm	NADPH-dependent glutamate synthase beta chain and related oxidoreductases	COG0493	-	-
73	BRHa_100354510	P	Cytoplasm	Aspartate-semialdehyde dehydrogenase	COG0136	K00133	EC:1.2.1.11
74	BRHa_100252635	P	Cytoplasm	Saccharopine dehydrogenase and related proteins	COG1748	K00290	EC:1.5.1.7
75	BRHa_100725845	P	Cytoplasm	Transaldolase	COG0176	K00616	EC:2.2.1.2
76	BRHa_100722359	P	Cytoplasm	Threonine synthase	COG0498	K01733	EC:4.2.3.1
77	BRHa_100687040	P	Cytoplasm	Predicted ornithine cyclodeaminase, mu-crystallin homolog - COG2423	COG2423	K01750	EC:4.3.1.12
78	BRHa_100426372	P	Cytoplasm	O-acetylhomoserine sulfhydrylase	COG2873	K01470	EC:2.5.1.49
79	BRHa_100146147	P	Extracellular	Flagellin and related hook-associated proteins	COG1344	K02406	-
79	BRHa_100146156	P	Extracellular	Flagellin and related hook-associated proteins	COG1344	K02406	-
80	BRHa_10025264	P	Extracellular	-	COG4968	-	-
81	BRHa_100297236	P	Unknown	Tip pilus assembly protein, tip-associated adhesin PiliY1	COG3419	-	-
81	BRHa_100252616	P	Unknown	-	COG3419	K02674	-
82	BRHa_1003749111	P	Cytoplasm	Tip pilus assembly protein, ATPase PilM	COG4972	K02662	-
83	BRHa_1003749115	P	Outer membrane	-	COG4796	K02666	-
84	BRHa_100722394	P	Cytoplasmic membrane	Sulfate permease and related transporters (MFS superfamily)	COG0659	-	-
85	BRHa_100676489	P	Periplasm	ABC-type branched-chain amino acid transport systems, periplasmic component	COG0683	K01999	-
85	BRHa_100332442	P	Periplasm	ABC-type branched-chain amino acid transport systems, periplasmic component	COG0683	K01999	-
85	BRHa_100705819	P	Periplasm	ABC-type branched-chain amino acid transport systems, periplasmic component	COG0683	K01999	-
85	BRHa_100657099	P	Periplasm	ABC-type branched-chain amino acid transport systems, periplasmic component	COG0683	K01999	-
85	BRHa_100594516	P	Periplasm	ABC-type branched-chain amino acid transport systems, periplasmic component	COG0683	K01999	-
85	BRHa_100426345	P	Unknown	ABC-type branched-chain amino acid transport systems, periplasmic component	COG0683	K01999	-
86	BRHa_100297226	P	Periplasm	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	COG0834	K09969	-
87	BRHa_100030891	P	Periplasm	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain	COG0834	K02030	-
88	BRHa_1003724106	P	Cytoplasmic membrane	ABC-type polar amino acid transport system, ATPase component	COG1126	K02028	EC:3.6.3.21
92	BRHa_100692244	P	Cytoplasm	Inorganic pyrophosphatase/exopolyphosphatase	COG1227	K01507	EC:3.6.1.1
93	BRHa_1006081109	P	Cytoplasm	Phosphate uptake regulator	COG0704	K02039	-
93	BRHa_10074317	P	Cytoplasm	Phosphate uptake regulator	COG0704	K02039	-
90	BRHa_1006081105	P	Cytoplasmic membrane	ABC-type phosphate transport system, ATPase component	COG1117	K02036	EC:3.6.3.27
90	BRHa_100743111	P	Cytoplasmic membrane	ABC-type phosphate transport system, ATPase component	COG1117	K02036	EC:3.6.3.27
91	BRHa_100743110	P	Cytoplasmic membrane	ABC-type phosphate transport system, periplasmic component	COG0226	K02040	-
91	BRHa_1006081106	P	Cytoplasmic membrane	ABC-type phosphate transport system, periplasmic component	COG0226	K02040	-
89	BRHa_100393024	P	Periplasm	ABC-type phosphate/phosphonate transport system, periplasmic component	COG3221	K02044	-
94	BRHa_100151530	P	Periplasm	ABC-type Fe3+ transport system, periplasmic component	COG1840	K02012	-
95	BRHa_1003749212	P	Cytoplasm	Ferritin-like protein	COG1528	K02217	EC:1.16.3.1
96	BRHa_1003724142	P	Unknown	ABC-type sugar transport system, periplasmic component	COG1653	K02027	-
96	BRHa_1001804226	P	Unknown	ABC-type_sugar_transport_system_periplasmic_component	COG1653	K02027	-

Table A-2.12 List of proteins identified in metaproteomic data that are involved in processes described in Figure 4.10, Figure 4.11 or mentioned in the main text. For an exhaustive list of proteins measured, refer to Table A-2.8. This table is too big for being printed in this document. It is available online here: dx.doi.org/10.5075/epfl-thesis-6727.

A-2.3 Supplementary Figure

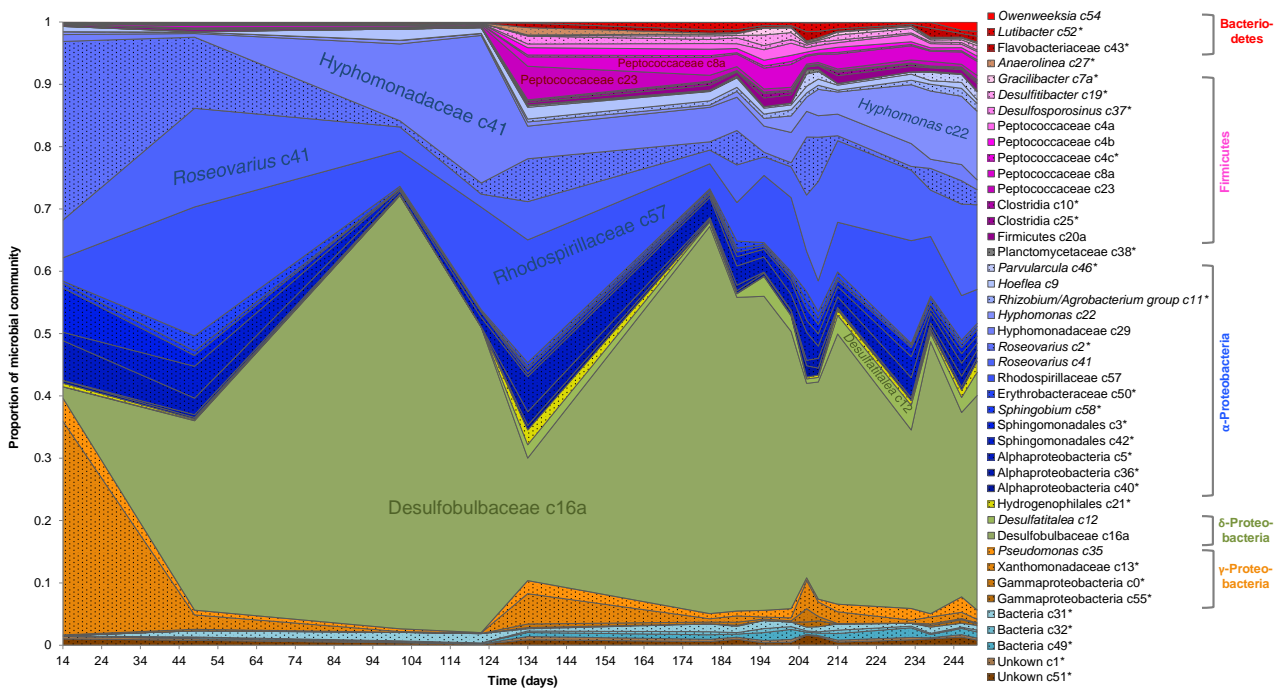


Figure A-2.1 Microbial community composition over time, based on the taxonomic affiliation of genomes obtained with metagenomic analysis. The dashed area and clusters labeled with * correspond to clusters that are not pure or not complete enough to be considered as draft genomes. See Table A-2.3 for more information.

References

- Abicht HK, Mancini S, Karnachuk OV, Solioz M. (2011). Genome sequence of *Desulfosporosinus* sp. OT, an acidophilic sulfate-reducing bacterium from copper mining waste in Norilsk, northern Siberia. *J Bacteriol* **193**:6104–6105.
- Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, *et al.* (2014). Binning metagenomic contigs by coverage and composition. *Nat Methods* **11**:1144–1146.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990). Basic local alignment search tool. *J Mol Biol* **215**:403–410.
- Amend JP, Teske A. (2005). Expanding frontiers in deep subsurface microbiology. *Palaeogeogr Palaeoclimatol Palaeoecol* **219**:131–155.
- Amy PS. (1997). Microbial dormancy and survival in the subsurface. In: *Microbiology of the Terrestrial Deep Subsurface*, Amy, PS & Haldeman, DL (eds), CRC Press: Boca Ration, pp. 185–203.
- Auelo T, Ranchou-Peyruse A, Ollivier B, Magot M. (2013). *Desulfotomaculum* spp. and related gram-positive sulfate-reducing bacteria in deep subsurface environments. *Front Microbiol* **4**:362.
- Aüilo T, Ranchou-Peyruse A, Ollivier B, Magot M. (2013). *Desulfotomaculum* spp. and related gram-positive sulfate-reducing bacteria in deep subsurface environments. *Front Microbiol* **4**. doi:10.3389/fmicb.2013.00362.
- Bagramyan K, Trchounian A. (2003). Structural and functional features of formate hydrogen lyase, an enzyme of mixed-acid fermentation from *Escherichia coli*. *Biochem Mosc* **68**:1159–1170.
- Baldani JJ, Baldani VLDi, Dobereiner J. (2005). *Herbaspirillum*. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, G (ed) Vol. 2 Part C, pp. 629–636.
- Balkwill DL. (2005). *Cupriavidus*. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, pp. 600–604.
- Barr NF, Allen AO. (1959). Hydrogen Atoms in the Radiolysis of Water. *J Phys Chem* **63**:928–931.
- Battaglia F, Gaucher E. (2003). Mont Terri project pore water chemistry experiment. Microbial and particle characterisation in the borehole water. Mont Terri Project: St-Ursanne, Switzerland.
- Bernardet J-F, Segers P, Vancanneyt M, Berthe F, Kersters K, Vandamme P. (1996). Cutting a gordian knot: emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium hydatis* nom. nov. (basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *Int J Syst Bacteriol* **46**:128–148.
- Boisvert S, Raymond F, Godzaridis É, Laviolette F, Corbeil J. (2012). Ray Meta: scalable de novo metagenome assembly and profiling. *Genome Biol* **13**:R122.
- Bonin B, Colin M, Dutfoy A. (2000). Pressure building during the early stages of gas production in a radioactive waste repository. *J Nucl Mater* **281**:1–14.
- Borcard D, Gillet F, Legendre P. (2011). *Numerical ecology with R*. Springer: New-York <http://www.springer.com/statistics/life+sciences%2C+medicine+%26+health/book/978-1-4419-7975-9> (Accessed May 1, 2014).
- Bossart P, Meier PM, Moeri A, Trick T, Mayor J-C. (2002). Geological and hydraulic characterisation of the excavation disturbed zone in the Opalinus Clay of the Mont Terri rock laboratory. *Eng Geol* **66**:19–38.
- Bossart P, Thury M. (2008). Mont Terri rock laboratory. project, programme 1996 to 2007 and results. Federal office of topography swisstopo: Bern.
- Bowman JP. (2005). *Methylophaga*. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part B, Springer: New-York, pp. 190–192.
- Bowman JP, McMeekin TA. (2005). *Alteromonadaceae*. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part B, Springer: New-York, pp. 443–444.

- Brown AD. (1990). Microbial water stress physiology. Principles and perspectives. John Wiley & Sons: Chichester.
- Brown SD, Thompson MR, Verberkmoes NC, Chourey K, Shah M, Zhou J, *et al.* (2006). Molecular dynamics of the *Shewanella oneidensis* response to chromate stress. *Mol Cell Proteomics MCP* **5**:1054–1071.
- Le Caër S. (2011). Water Radiolysis: Influence of Oxide Surfaces on H₂ Production under Ionizing Radiation. *Water* **3**:235–253.
- Calow P. (1977). Conversion efficiencies in heterotrophic organisms. *Biol Rev* **52**:385–409.
- Canfield DE. (2001). Isotope fractionation by natural populations of sulfate-reducing bacteria. *Geochim Cosmochim Acta* **65**:1117–1124.
- De Cannière P, Schwarzbauer J, Höhener P, Lorenz G, Salah S, Leupin OX, *et al.* (2011). Biogeochemical processes in a clay formation in situ experiment: Part C – Organic contamination and leaching data. *Appl Geochem* **26**:967–979.
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**:335–336.
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, *et al.* (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* **108**:4516–4522.
- Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, *et al.* (2014). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* **42**:D459–D471.
- Chen L, Hu M, Huang L, Hua Z, Kuang J, Li S, *et al.* (2014). Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage. *ISME J*. doi:10.1038/ismej.2014.245.
- Chivian D, Brodie EL, Alm EJ, Culley DE, Dehal PS, DeSantis TZ, *et al.* (2008). Environmental Genomics Reveals a Single-Species Ecosystem Deep Within Earth. *Science* **322**:275–278.
- Chourey K, Jansson J, Verberkmoes N, Shah M, Chavarria KL, Tom LM, *et al.* (2010). Direct Cellular Lysis/Protein Extraction Protocol for Soil Metaproteomics. *J Proteome Res* **9**:6615–6622.
- Chourey K, Nissen S, Vishnivetskaya T, Shah M, Pfiffner S, Hettich RL, *et al.* (2013). Environmental proteomics reveals early microbial community responses to biostimulation at a uranium- and nitrate-contaminated site. *Proteomics* **13**:2921–2930.
- Cline JD. (1969). Spectrophotometric determination of hydrogen sulfide in natural waters. *Limnol Oceanogr* **14**:454–458.
- Colwell FS, Smith RP. (2004). Unifying principles of the deep terrestrial and deep marine biospheres. In: *The Subseafloor Biosphere at Mid-Ocean Ridges*, Wilcock, WSD, Delong, EF, Kelley, DS, Baross, JA, & Cary, SC (eds), American Geophysical Union, pp. 355–367.
- Courdouan A, Christl I, Meylan S, Wersin P, Kretzschmar R. (2007). Characterization of dissolved organic matter in anoxic rock extracts and in situ pore water of the Opalinus Clay. *Appl Geochem* **22**:2926–2939.
- Cragg BA, Law KM, Cramp A, Parkes RJ. (1998). The response of bacterial populations to sapropels in deep sediments of the Eastern Mediterranean (Site 969). *Proc Ocean Drill Program Sci Results* **160**:303–307.
- Darling AE, Mau B, Perna NT. (2009). Progressive Mauve: Multiple alignment of genomes with gene flux and rearrangement. *ArXiv09105780 Q-Bio*. <http://arxiv.org/abs/0910.5780> (Accessed November 19, 2014).
- Das SK. (2005). Borea. In: *Bergey's Systematic Manual of Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, pp. 459–461.
- Dastager SG, Lee J-C, Ju Y-J, Park D-J, Kim C-J. (2008). *Marmoricola bigeumensis* sp nov., a member of the family Nocardioideaceae. *Int J Syst Evol Microbiol* **58**:1060–1063.
- Daumas S, Labat M, Lasne E. (2014). HT (hydrogen transfer) experiment: Molecular study and classical microbiological analyses. Technical Note 2014-18. St-Ursanne, Switzerland.
- Daumas S, Necib S. (2014). Microbiological and molecular analysis of the water circulating in the innercircuit of the BIC-1 borehole. Technical Note 2013-81. St-Ursanne, Switzerland.

- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, *et al.* (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**:5069–5072.
- Dong Y, Kumar CG, Chia N, Kim P-J, Miller PA, Price ND, *et al.* (2014). Halomonas sulfidaeris-dominated microbial community inhabits a 1.8 km-deep subsurface Cambrian Sandstone reservoir. *Environ Microbiol* **16**:1695–1708.
- Eberspaecher J. (2005). Phenyllobacterium. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 part C, Springer: New-York, pp. 316–324.
- Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. doi:10.1093/bioinformatics/btq461.
- Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**:2194–2200.
- Edwards KJ, Becker K, Colwell F. (2012). The deep, dark energy biosphere: intraterrestrial life on earth. *Annu Rev Earth Planet Sci* **40**:551–568.
- Enning D, Garrelfs J. (2014). Corrosion of Iron by Sulfate-Reducing Bacteria: New Views of an Old Problem. *Appl Environ Microbiol* **80**:1226–1236.
- Eschbach M, Schreiber K, Trunk K, Buer J, Jahn D, Schobert M. (2004). Long-term anaerobic survival of the opportunistic pathogen *Pseudomonas aeruginosa* via pyruvate fermentation. *J Bacteriol* **186**:4596–4604.
- Evtushenko LI, Takeuchi M. (2006). The family Microbacteriaceae. In: *The Prokaryotes*, Dworkin, M, Falkow, S, Rosenberg, E, Schleifer, K-H, & Stackebrandt, E (eds) Vol. 3, Springer: New-York, pp. 1020–1098.
- Fredrickson JK, Balkwill DL. (2006). Geomicrobial Processes and Biodiversity in the Deep Terrestrial Subsurface. *Geomicrobiol J* **23**:345–356.
- Fredrickson JK, Fletcher M. (2001). *Subsurface Microbiology and Biogeochemistry*. Wiley.
- Fredrickson JK, Garland TR, Hicks RJ, Thomas JM, Li SW, McFadden KM. (1989). Lithotrophic and Heterotrophic Bacteria in Deep Subsurface Sediments and Their Relation to Sediment Properties. *Geomicrobiol J* **7**:53–66.
- Fredrickson JK, McKinley JP, Bjornstad BN, Long PE, Ringelberg DB, White DC, *et al.* (1997). Pore-size constraints on the activity and survival of subsurface bacteria in a late cretaceous shale-sandstone sequence, northwestern New Mexico. *Geomicrobiol J* **14**:183–202.
- Fruttschi M, Bernier-Latmani R. (2010a). DR Experiment: evaluation of the role of glycerol in microbial growth at the Mt Terri rock laboratory. Technical Note 2009-35. Mont Terri Project: St-Ursanne, Switzerland.
- Fruttschi M, Bernier-Latmani R. (2010b). PC (porewater chemistry) Experiment: porewater microbial community stimulated by glycerol at the Mt Terri rock laboratory - Final report. Technical Note 2009-36. Mont Terri Project: St-Ursanne, Switzerland.
- Furnes H, Staudigel H. (1999). Biological mediation in ocean crust alteration: how deep is the deep biosphere? *Earth Planet Sci Lett* **166**:97–103.
- Garrity GM, Bell JA, Lilburn T. (2005a). Rhodobacteraceae fam. nov. In: *Bergey's Manual of Systematic Bacteriology*, Vol. 2, Springer, p. 161.
- Garrity GM, Bell JA, Lilburn T. (2005b). Rhodobacterales. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, p. 161.
- Garrity GM, Bell JA, Lilburn T. (2005c). Rhodospirillaceae. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 part C, Springer: New-York, p. 1.
- Gautschi A. (2001). Hydrogeology of a fractured shale (Opalinus Clay): Implications for deep geological disposal of radioactive wastes. *Hydrogeol J* **9**:97–107.
- Ghiorse WC, Wilson JT. (1988). Microbial ecology of the terrestrial subsurface. *Adv Appl Microbiol* **33**:107–172.

- Gimmi T, Leupin OX, Eikenberg J, Glaus MA, Van Loon LR, Waber HN, *et al.* (2014). Anisotropic diffusion at the field scale in a 4-year multi-tracer diffusion and retention experiment – I: Insights from the experimental data. *Geochim Cosmochim Acta* **125**:373–393.
- Gobat J-M, Aragno M, Matthey W. (2004). *The living soil: fundamentals of soil science and soil biology*. Science Publishers, U.S.: Enfield, NH.
- Gold T. (1992). The deep, hot biosphere. *Proc Natl Acad Sci* **89**:6045–6049.
- Golyshin PN, Harayama S, Timms KN, Yakimov MM, Brenner DJ, Krieg NR, *et al.* (2005). Alcanivorax. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM (ed) Vol. 2 Part B, Springer: New-York, pp. 295–299.
- Goodfellow M, Maldonado LA. (2006). The families Dietziaceae, Gordoniaceae, Nocardiaceae and Tsukamurellaceae. In: *The Prokaryotes*, Dworking, M, Falkow, S, Rosenberg, E, Schleifer, K-H, & Stackebrandt, E (eds) Vol. 3, Springer: New-York, pp. 843–888.
- Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**:81–91.
- De Haan H. (1974). Effect of a fulvic acid fraction on the growth of a *Pseudomonas* from Tjeukemeer (The Netherlands). *Freshw Biol* **4**:301–310.
- Hahn MW, Kasalický V, Jezbera J, Brandt U, Jezberová J, Šimek K. (2010). *Limnohabitans curvus* gen. nov., sp. nov., a planktonic bacterium isolated from a freshwater lake. *Int J Syst Evol Microbiol* **60**:1358–1365.
- Hansen TA. (1993). Carbon Metabolism of Sulfate-Reducing Bacteria. In: *The Sulfate-Reducing Bacteria: Contemporary Perspectives*, Odom, JM & Jr, RS (eds) Brock/Springer Series in Contemporary Bioscience, Springer: New-York, pp. 21–40.
- Holmes DE, Bond DR, Lovley DR. (2004). Electron Transfer by *Desulfobulbus propionicus* to Fe(III) and Graphite Electrodes. *Appl Environ Microbiol* **70**:1234–1237.
- Hwang CY, Cho BC. (2008). *Cohaesibacter gelatinilyticus* gen. nov., sp. nov., a marine bacterium that forms a distinct branch in the order Rhizobiales, and proposal of *Cohaesibacteraceae* fam. nov. *Int J Syst Evol Microbiol* **58**:267–277.
- Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**:119.
- Jähne B, Heinz G, Dietrich W. (1987). Measurement of the diffusion coefficients of sparingly soluble gases in water. *J Geophys Res Oceans* **92**:10767–10776.
- Jannasch HW, Taylor CD. (1984). Deep-Sea Microbiology. *Annu Rev Microbiol* **38**:487–487.
- Johnson SS, Hebsgaard MB, Christensen TR, Mastepanov M, Nielsen R, Munch K, *et al.* (2007). Ancient bacteria show evidence of DNA repair. *Proc Natl Acad Sci* **104**:14401–14405.
- Jones DM, Tipping E. (1998). Organic geochemical analysis of Opalinus clay samples from Mont Terri. Mont Terri Project: St-Ursanne, Switzerland.
- Junier P, Junier T, Podell S, Sims DR, Detter JC, Lykidis A, *et al.* (2010). The genome of the Gram-positive metal- and sulfate-reducing bacterium *Desulfotomaculum reducens* strain MI-1. *Environ Microbiol* **12**:2738–2754.
- Kalyuzhnaya MG, Bowerman S, Lara JC, Lidstrom ME, Chistoserdova L. (2006). *Methylotenera mobilis* gen. nov., sp. nov., an obligately methylamine-utilizing bacterium within the family Methylophilaceae. *Int J Syst Evol Microbiol* **56**:2819–2823.
- Kalyuzhnaya MG, Marco PD, Bowerman S, Pacheco CC, Lara JC, Lidstrom ME, *et al.* (2006). *Methyloversatilis universalis* gen. nov., sp. nov., a novel taxon within the Betaproteobacteria represented by three methylotrophic isolates. *Int J Syst Evol Microbiol* **56**:2517–2522.
- Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* **42**:D199–D205.
- Kashefi K, Lovley DR. (2000). Reduction of Fe(III), Mn(IV), and Toxic Metals at 100°C by *Pyrobaculum islandicum*. *Appl Environ Microbiol* **66**:1050–1056.

- Keller LM, Holzer L, Gasser P, Erni R, Rossell MD. (2015). Intergranular pore space evolution in MX80 bentonite during a long-term experiment. *Appl Clay Sci* **104**:150–159.
- Keller LM, Holzer L, Wepf R, Gasser P. (2011). 3D geometry and topology of pore pathways in Opalinus clay: Implications for mass transport. *Appl Clay Sci* **52**:85–95.
- Kelly DP, Wood AP, Stackebrandt E, Brenner DJ, Krieg NR, Staley JT. (2005). Thiobacillus. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, G (ed) Vol. 2 Part C, Springer: New-York, pp. 764–769.
- Kim BH, Gadd GM. (2008). *Bacterial Physiology and Metabolism*. 1st edition. Cambridge University Press: Cambridge New York.
- Kim M, Oh H-S, Park S-C, Chun J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* **64**:346–351.
- Kuever J, Rainey FA, Widdel F. (2005a). Desulfobulbaceae. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, p. 988.
- Kuever J, Rainey FA, Widdel F. (2005b). Desulfocapsa. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, pp. 992–994.
- Laban NA, Tan B, Dao A, Foght J. (2015). Draft genome sequence of uncultivated Desulfosporosinus sp. strain Tol-M, obtained by stable isotope probing using [¹³C₆]toluene. *Genome Announc* **3**:e01422–14.
- Langmead B, Salzberg SL. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**:357–359.
- Lehman RM, Roberto FF, Earley D, Bruhn DF, Brink SE, O'Connell SP, *et al.* (2001). Attached and unattached bacterial communities in a 120-meter corehole in an acidic, crystalline rock aquifer. *Appl Environ Microbiol* **67**:2095–2106.
- Lerouge C, Grangeon S, Gaucher EC, Tournassat C, Agrinier P, Guerrot C, *et al.* (2011). Mineralogical and isotopic record of biotic and abiotic diagenesis of the Callovian–Oxfordian clayey formation of Bure (France). *Geochim Cosmochim Acta* **75**:2633–2663.
- Leschine SB. (1995). Cellulose degradation in anaerobic environments. *Annu Rev Microbiol* **49**:399–426.
- Leys N, Valcke E. (in prep.). Technical Note 2014-50. St-Ursanne, Switzerland.
- Libert M, Bildstein O, Esnault L, Jullien M, Sellier R. (2011). Molecular hydrogen: An abundant energy source for bacterial activity in nuclear waste repositories. *Phys Chem Earth Parts ABC* **36**:1616–1623.
- Li M. (2011). DUK - a fast and efficient kmer based sequence mMatching tool. *Lawrence Berkeley Natl Lab*. <http://escholarship.org/uc/item/5cb5810k> (Accessed May 2, 2014).
- Lovley D. (1991). Dissimilatory Fe(III) and Mn(IV) reduction. *Microbiol Rev* **55**:259–287.
- Lovley DR, Chapelle FH. (1995). Deep subsurface microbial processes. *Rev Geophys* **33**:365–381.
- Lu H, Sato Y, Fujimura R, Nishizawa T, Kamijo T, Ohta H. (2011). *Limnobacter litoralis* sp. nov., a thiosulfate-oxidizing, heterotrophic bacterium isolated from a volcanic deposit, and emended description of the genus *Limnobacter*. *Int J Syst Evol Microbiol* **61**:404–407.
- Lunau M, Lemke A, Walther K, Martens-Habbena W, Simon M. (2005). An improved method for counting bacteria from sediments and turbid environments by epifluorescence microscopy. *Environ Microbiol* **7**:961–968.
- Lütke L, Moll H, Bachvarova V, Selenska-Pobell S, Bernhard G. (2013). The U(VI) speciation influenced by a novel *Paenibacillus* isolate from Mont Terri Opalinus clay. *Dalton Trans Camb Engl* **2003** **42**:6979–6988.
- Madigan MT, Martinko JM, Stahl D, Clark DP. (2010). *Brock Biology of Microorganisms*. 13th edition. Benjamin Cummings: San Francisco.
- Magoč T, Salzberg SL. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**:2957–2963.
- Mais A. (1998). Mercury porosimetry and H₂O determination. Mont Terri Project: St-Ursanne, Switzerland.

- Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Grechkin Y, *et al.* (2012). IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res* **40**:D115–D122.
- Márquez MC, Carrasco IJ, Xue Y, Ma Y, Cowan DA, Jones BE, *et al.* (2008). *Aquisalibacillus elongatus* gen. nov., sp. nov., a moderately halophilic bacterium of the family Bacillaceae isolated from a saline lake. *Int J Syst Evol Microbiol* **58**:1922–1926.
- Marsh GP, Taylor KJ. (1988). An assessment of carbon steel containers for radioactive waste disposal. *Corros Sci* **28**:289–320.
- Martens T, Heidorn T, Pukall R, Simon M, Tindall BJ, Brinkhoff T. (2006). Reclassification of *Roseobacter gallaeciensis* Ruiz-Ponte *et al.* 1998 as *Phaeobacter gallaeciensis* gen. nov., comb. nov., description of *Phaeobacter inhibens* sp. nov., reclassification of *Ruegeria algicola* (Lafay *et al.* 1995) Uchino *et al.* 1999 as *Marinovum algicola* gen. nov., comb. nov., and emended descriptions of the genera *Roseobacter*, *Ruegeria* and *Leisingera*. *Int J Syst Evol Microbiol* **56**:1293–1304.
- Mauclaire L, McKenzie JA, Schwyn B, Bossart P. (2007). Detection and cultivation of indigenous microorganisms in Mesozoic claystone core samples from the Opalinus Clay Formation (Mont Terri Rock Laboratory). *Phys Chem Earth Parts ABC* **32**:232–240.
- Mazurek M. (1999). Mineralogy of the Opalinus clay. In: *Mont Terri rock laboratory: results of the hydrogeological, geochemical and geotechnical experiments performed in 1996 and 1997*, Thury, M & Bossart, P (eds) Geological Report Vol. 23, Federal office of topography swisstopo: Bern, pp. 15–18.
- Moll H, Lütke L, Bachvarova V, Cherkouk A, Selenska-Pobell S, Bernhard G. (2014). Interactions of the Mont Terri Opalinus Clay Isolate Sporomusa sp. MT-2.99 with Curium(III) and Europium(III). *Geomicrobiol J* **31**:682–696.
- Moore ERB, Tindall BJ, Santos VAPM Dos, Pieper DH, Ramos J-L, Palleroni NJ. (2006). Nonmedical: *Pseudomonas*. In: *The Prokaryotes*, Dworkin, M, Falkow, S, Rosenberg, E, Schleifer, K-H, & Stackebrandt, E (eds) Vol. 6, Springer: New-York, pp. 646–703.
- Moore RL, Weiner RM, Gebers R. (1984). Notes: Genus *Hyphomonas* Pongratz 1957 nom. rev. emend., *Hyphomonas polymorpha* Pongratz 1957 nom. rev. emend., and *Hyphomonas neptunium* (Leifson 1964) comb. nov. emend. (*Hyphomicrobium neptunium*). *Int J Syst Bacteriol* **34**:71–73.
- Mortimer RJG, Galsworthy AMJ, Bottrell SH, Wilmot LE, Newton RJ. (2011). Experimental evidence for rapid biotic and abiotic reduction of Fe (III) at low temperatures in salt marsh sediments: a possible mechanism for formation of modern sedimentary siderite concretions. *Sedimentology* **58**:1514–1529.
- Muyzer G, Stams AJM. (2008). The ecology and biotechnology of sulphate-reducing bacteria. *Nat Rev Microbiol* **6**:441–454.
- Nagra. (2004). Annual report 2004. National Cooperative for the Disposal of Radioactive Waste: Wettingen.
- Nagra. (2008). Le bon moment pour agir. National Cooperative for the Disposal of Radioactive Waste: Wettingen [http://www.nagra.ch/data/documents/database/dokumente/\\$default/Default%20Folder/Publikationen/Broschueren%20Themenhefte/f_zeit_zum_handeln.pdf](http://www.nagra.ch/data/documents/database/dokumente/$default/Default%20Folder/Publikationen/Broschueren%20Themenhefte/f_zeit_zum_handeln.pdf).
- Nagra. (2015). Siting regions for deep geological repositories. National Cooperative for the Disposal of Radioactive Waste: Wettingen [http://www.nagra.ch/data/documents/database/dokumente/\\$default/Default%20Folder/Publikationen/Broschueren%20Themenhefte/e_Faltblatt%20Einengung%202015_klein.pdf](http://www.nagra.ch/data/documents/database/dokumente/$default/Default%20Folder/Publikationen/Broschueren%20Themenhefte/e_Faltblatt%20Einengung%202015_klein.pdf).
- Nagra. (2013). Wussten Sie,... – Erstaunliches zu Radioaktivität und Entsorgung. National Cooperative for the Disposal of Radioactive Waste: Wettingen [http://www.nagra.ch/data/documents/database/dokumente/\\$default/Default%20Folder/Publikationen/Broschueren%20Themenhefte/d_th6%20wussten%20sie%202013.pdf](http://www.nagra.ch/data/documents/database/dokumente/$default/Default%20Folder/Publikationen/Broschueren%20Themenhefte/d_th6%20wussten%20sie%202013.pdf).
- Nealson KH, Inagaki F, Takai K. (2005). Hydrogen-driven subsurface lithoautotrophic microbial ecosystems (SLiMEs): do they exist and why should we care? *Trends Microbiol* **13**:405–410.
- Onstott TC, Colwell FS, Kieft TL, Murdoch L, Phelps TJ. (2009). New Horizons for Deep Subsurface Microbiology. *Microbe* **4**:499–505.
- Onstott TC, Phelps TJ, Colwell FS, Ringelberg D, White DC, Boone DR, *et al.* (1998). Observations pertaining to the origin and ecology of microorganisms recovered from the deep subsurface of Taylorsville Basin, Virginia. *Geomicrobiol J* **15**:353–385.
- Paoletti AC, Parmely TJ, Tomomori-Sato C, Sato S, Zhu D, Conaway RC, *et al.* (2006). Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc Natl Acad Sci U S A* **103**:18928–18933.

- Parker RJ, Wellsbury P. (2004). Deep Biospheres. In: *Microbial Diversity and Bioprospecting*, Bull, AT (ed), ASM press: Wahsington, pp. 120–129.
- Parkes RJ, Cragg BA, Wellsbury P. (2000). Recent studies on bacterial populations and processes in subseafloor sediments: A review. *Hydrogeol J* **8**:11–28.
- Pearson FJ. (2002). PC Experiment: Recipe for Artificial Pore Water.
- Pearson FJ, Acros D, Boisson J-Y, Fernandez AM, Gäbler HE, Gaucher E, *et al.* (2003). Mont Terri project: geochemistry of water in the Opalinus clay formation at the Mont Terri rock laboratory. Federal office for water and geology FOWG: Bern <http://www.bafu.admin.ch/publikationen/publikation/00633/index.html?lang=en>.
- Pedersen K. (2010). Analysis of copper corrosion in compacted bentonite clay as a function of clay density and growth conditions for sulfate-reducing bacteria. *J Appl Microbiol* **108**:1094–1104.
- Pedersen K. (2000). Exploration of deep intraterrestrial microbial life: current perspectives. *FEMS Microbiol Lett* **185**:9–16.
- Pedersen K. (1993). The deep subterranean biosphere. *Earth-Sci Rev* **34**:243–260.
- Pedersen K, Hallbeck L, Arlinger J, Erlandson A-C, Jahromi N. (1997). Investigation of the potential for microbial contamination of deep granitic aquifers during drilling using 16S rRNA gene sequencing and culturing methods. *J Microbiol Methods* **30**:179–192.
- Poulain S, Sergeant C, Simonoff M, Le Marrec C, Altmann S. (2008). Microbial investigations in Opalinus clay, an argillaceous formation under evaluation as a potential host rock for a radioactive waste repository. *Geomicrobiol J* **25**:240–249.
- Ramamoorthy S, Sass H, Langner H, Schumann P, Kroppenstedt RM, Spring S, *et al.* (2006a). Desulfosporosinus Lacus Sp. Nov., a Sulfate-Reducing Bacterium Isolated from Pristine Freshwater Lake Sediments. *Int J Syst Evol Microbiol* **56**:2729–2736.
- Ramamoorthy S, Sass H, Langner H, Schumann P, Kroppenstedt RM, Spring S, *et al.* (2006b). Desulfosporosinus lacus sp. nov., a sulfate-reducing bacterium isolated from pristine freshwater lake sediments. *Int J Syst Evol Microbiol* **56**:2729–2736.
- R Development Core Team. (2008). R: a language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria <http://www.R-project.org>.
- Rodríguez H, Fraga R. (1999). Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnol Adv* **17**:319–339.
- Rogosa M. (1971). Peptococcaceae, a New Family To Include the Gram-Positive, Anaerobic Cocci of the Genera Peptococcus, Peptostreptococcus, and Ruminococcus. *Int J Syst Bacteriol* **21**:234–237.
- Rothschild LJ, Mancinelli RL. (2001). Life in extreme environments. *Nature* **409**:1092–1101.
- Russel C. (1997). The collection of subsurface samples by mining. In: *The Microbiology of the Terrestrial Deep Subsurface*, Amy, PS & Halderman, DL (eds), CRC Press: Boca Raton, pp. 45–49.
- Saddler GS, Bradbury JF. (2005). Xanthomonadaceae. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part B, Springer: New-York, p. 63.
- Sander R. (1999). Compilation of Henry's law constants for inorganic and organic species of potential importance in environmental chemistry. Max-Planck Institute of Chemistry, Air Chemistry Department Mainz, Germany <http://www.ceset.unicamp.br/~mariaacm/ST405/Lei%20de%20Henry.pdf> (Accessed January 13, 2015).
- Santelli CM, Banerjee N, Bach W, Edwards KJ. (2010). Tapping the Subsurface Ocean Crust Biosphere: Low Biomass and Drilling-Related Contamination Calls for Improved Quality Controls. *Geomicrobiol J* **27**:158–169.
- Schink B. (1997). Energetics of syntrophic cooperation in methanogenic degradation. *Microbiol Mol Biol Rev* **61**:262–&.
- Schippers A, Neretin LN. (2006). Quantification of microbial communities in near-surface and deeply buried marine sediments on the Peru continental margin using real-time PCR. *Environ Microbiol* **8**:1251–1260.
- Schreiber K, Boes N, Eschbach M, Jaensch L, Wehland J, Bjarnsholt T, *et al.* (2006). Anaerobic survival of *Pseudomonas aeruginosa* by pyruvate fermentation requires an usp-type stress protein. *J Bacteriol* **188**:659–668.

- Selenska S, Klingmüller W. (1991). Direct detection of nif-gene sequences of Enterobacter agglomerans in soil. *FEMS Microbiol Lett* **80**:243–245.
- Sharma R, Dill BD, Chourey K, Shah M, VerBerkmoes NC, Hettich RL. (2012). Coupling a detergent lysis/cleanup methodology with intact protein fractionation for enhanced proteome characterization. *J Proteome Res* **11**:6008–6018.
- Shida O, Takagi H, Kadowaki K, Nakamura LK, Komagata K. (1997). Transfer of Bacillus alginolyticus, Bacillus chondroitinus, Bacillus curdlanolyticus, Bacillus gluconolyticus, Bacillus kobensis, and Bacillus thiaminolyticus to the genus Paenibacillus and emended description of the genus Paenibacillus. *Int J Syst Bacteriol* **47**:289–298.
- Skinner BJ, Porter SC, Parker J. (2004). Dynamic earth. An introduction to physical geology. Wiley: Hoboken.
- Sorokin DY, Tourova TP, Spiridonova EM, Rainey FA, Muyzer G. (2005). Thioclava pacifica gen. nov., sp. nov., a novel facultatively autotrophic, marine, sulfur-oxidizing bacterium from a near-shore sulfidic hydrothermal area. *Int J Syst Evol Microbiol* **55**:1069–1075.
- Spring S, Kämpfer P, Schleifer KH. (2001). Limnobacter thiooxidans gen. nov., sp. nov., a novel thiosulfate-oxidizing bacterium isolated from freshwater lake sediment. *Int J Syst Evol Microbiol* **51**:1463–1470.
- Spring S, Lapidus A, Schroder M, Gleim D, Sims D, Meincke L, et al. (2009). Complete genome sequence of Desulfotomaculum acetoxidans type strain (5575T). *Stand Genomic Sci* **1**:242–253.
- Spring S, Rosenzweig F. (2006). The genera Desulfitobacterium and Desulfosporosinus: taxonomy. In: *The Prokaryotes*, Dworkin, M, Falkow, S, Rosenberg, E, Schleifer, K-H, & Stackebrandt, E (eds) Vol. 4, pp. 771–786.
- Stams AJM, Kremer DR, Nicolay K, Weenk GH, Hansen TA. (1984). Pathway of propionate formation in Desulfobulbus propionicus. *Arch Microbiol* **139**:167–173.
- Stark M, Berger SA, Stamatakis A, Mering C von. (2010). MLTreeMap - accurate maximum likelihood placement of environmental DNA sequences into taxonomic and functional reference phylogenies. *BMC Genomics* **11**:461.
- Stevens TO, McKinley JP. (1995). Lithoautotrophic Microbial Ecosystems in Deep Basalt Aquifers. *Science* **270**:450–455.
- Steyn PL, Segers P, Vancanneyt M, Sandra P, Kersters K, Joubert JJ. (1998). Classification of heparinolytic bacteria into a new genus, Pedobacter, comprising four species: Pedobacter heparinus comb. nov., Pedobacter piscium comb. nov., Pedobacter africanus sp. nov. and Pedobacter saltans sp. nov. Proposal of the family Sphingobacteriaceae fam. nov. *Int J Syst Bacteriol* **48**:165–177.
- Stookey LL. (1970). Ferrozine - a new spectrophotometric reagent for iron. *Anal Chem* **42**:779–781.
- Stroes-Gascoyne S, Schippers A, Schwyn B, Poulain S, Sergeant C, Simonoff M, et al. (2007). Microbial community analysis of Opalinus Clay drill core samples from the Mont Terri underground research laboratory, Switzerland. *Geomicrobiol J* **24**:1–17.
- Stroes-Gascoyne S, Sergeant C, Schippers A, Hamon CJ, Nèble S, Vesvres M-H, et al. (2011). Biogeochemical processes in a clay formation in situ experiment: Part D – Microbial analyses – Synthesis of results. *Appl Geochem* **26**:980–989.
- Stumm W, Morgan JJ. (1996). Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters. 3rd Edition. Wiley-Blackwell: New York.
- Suess E. (1888). Das Antlitz der Erde. Tempsky: Vienna.
- Suzina NE, Muliukin AL, Kozlova AN, Shorokhova AP, Dmitriev VV, Barinova ES, et al. (2004). Ultrastructure of resting cells of some non-spore-forming bacteria. *Mikrobiologiya* **73**:516–529.
- Tabb DL, Fernando CG, Chambers MC. (2007). MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J Proteome Res* **6**:654–661.
- Takeuchi M, Hamana K, Hiraishi A. (2001). Proposal of the genus Sphingomonas sensu stricto and three new genera, Sphingobium, Novosphingobium and Sphingopyxis, on the basis of phylogenetic and chemotaxonomic analyses. *Int J Syst Evol Microbiol* **51**:1405–1417.
- The Uniprot Consortium. (2014). Activities at the universal protein resource (UniProt). *Nucleic Acids Res* **42**:D191–D198.

- Thompson MR, VerBerkmoes NC, Chourey K, Shah M, Thompson DK, Hettich RL. (2007). Dosage-dependent proteome response of *Shewanella oneidensis* MR-1 to acute chromate challenge. *J Proteome Res* **6**:1745–1757.
- Thury M, Bossart P. (1999). The Mont Terri rock laboratory, a new international research project in a Mesozoic shale formation, in Switzerland. *Eng Geol* **52**:347–359.
- Tikhonov VV, Yakushev AV, Zavgorodnyaya YA, Byzov BA, Demin VV. (2010). Effects of humic acids on the growth of bacteria. *Eurasian Soil Sci* **43**:305–313.
- Towner M. (2006). The Genus *Acinetobacter*. In: *The Prokaryotes*, Dworkin, M, Falkow, S, Rosenberg, E, Schleifer, K-H, & Stackebrandt, E (eds) Vol. 6, Springer: New-York, pp. 746–758.
- Uroz S, Calvaruso C, Turpault M-P, Frey-Klett P. (2009). Mineral weathering by bacteria: ecology, actors and mechanisms. *Trends Microbiol* **17**:378–387.
- Vancanneyt M, Segers P, Wolf-Rainer A, De Vos P. (2005). *Brevundimonas*. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, pp. 308–316.
- Vignais PM, Billoud B. (2007). Occurrence, Classification, and Biological Function of Hydrogenases: An Overview. *Chem Rev* **107**:4206–4272.
- Vinsot A, Appelo CAJ, Cailteau C, Wechner S, Pironon J, De Donato P, *et al.* (2008). CO₂ data on gas and pore water sampled in situ in the Opalinus Clay at the Mont Terri rock laboratory. *Phys Chem Earth Parts ABC* **33**, **Supplement 1**:S54–S60.
- Vinsot A, Appelo C a. J, Lundy M, Wechner S, Lettry Y, Lerouge C, *et al.* (2014). In situ diffusion test of hydrogen gas in the Opalinus Clay. *Geol Soc Lond Spec Publ* **400**:563–578.
- Vreeland RH, Rosenzweig WD, Powers DW. (2000). Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. *Nature* **407**:897–900.
- Wall JD, Krumholz LR. (2006). Uranium Reduction. *Annu Rev Microbiol* **60**:149–166.
- Wang Q, Garrity GM, Tiedje JM, Cole JR. (2007). Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**:5261–5267.
- Wang W, Sun J, Hartlep M, Deckwer W-D, Zeng A-P. (2003). Combined use of proteomic analysis and enzyme activity assays for metabolic pathway analysis of glycerol fermentation by *Klebsiella pneumoniae*. *Biotechnol Bioeng* **83**:525–536.
- Ward N, Staley JT, Fuerst JA, Giovanni S, Schlesner H, Stackebrandt E. (2006). The order Planctomycetales, including the genera *Planctomyces*, *Pirellula*, *Gemmata* and *Isosphaera* and the candidatus genera *Brocadia*, *Kueneenia* and *Scalindua*. In: *The Prokaryotes*, Dworkin, M (ed) Vol. 7, pp. 757–793.
- Wells CL, Wilkins TD. (1996). Clostridia: sporeforming anaerobic Bacilli. In: *Medical Microbiology*, Baron, S (ed), University of Texas Medical Branch at Galveston: Galveston (TX).
- Wersin P, Leupin OX, Mettler S, Gaucher EC, Mäder U, De Cannière P, *et al.* (2011). Biogeochemical processes in a clay formation in situ experiment: Part A – Overview, experimental design and water data of an experiment in the Opalinus Clay at the Mont Terri Underground Research Laboratory, Switzerland. *Appl Geochem* **26**:931–953.
- Weyant RS, Whitney AM. (2005). *Afipia*. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, pp. 443–448.
- Whitman WB, Coleman DC, Wiebe WJ. (1998). Prokaryotes: The unseen majority. *Proc Natl Acad Sci* **95**:6578–6583.
- Widdel F. (2006). The Genus *Desulfotomaculum*. In: *The Prokaryotes: A Handbook on the Biology of Bacteria*, Dworkin, M, Falkow, S, Rosenberg, E, Schleifer, K-H, & Stackebrandt, E (eds) Vol. 4, Springer: New-York, pp. 787–794.
- Wiegel JKW. (2005). *Xanthobacter*. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, pp. 555–566.
- Wiegel J, Tanner R, Rainey FA. (2006). An Introduction to the Family Clostridiaceae. In: *The Prokaryotes: A Handbook on the Biology of Bacteria*, Dworkin, M, Falkow, S, Rosenberg, E, Schleifer, K-H, & Stackebrandt (eds) Vol. 4, Springer: New-York, pp. 654–678.

- Wildi W, Appel D, Buser M, Dermange F, Eckhardt A, Hufschmied P, *et al.* (2000). Disposal Concepts for Radioactive Waste. Federal Office of Energy: Bern, Switzerland.
- Willems A, Mergaert J, Swings J. (2005). *Variovorax*. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, G, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, pp. 732–735.
- Willerslev E, Hansen AJ, Poinar HN. (2004). Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends Ecol Evol* **19**:141–147.
- Willerslev E, Hansen AJ, Rønn R, Brand TB, Barnes I, Wiuf C, *et al.* (2004). Long-term persistence of bacterial DNA. *Curr Biol CB* **14**:R9–10.
- Wu S, Zhu Z, Fu L, Niu B, Li W. (2011). WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* **12**:444.
- Xie C-H, Yokota A. (2005). *Pleomorphomonas oryzae* gen. nov., sp. nov., a nitrogen-fixing bacterium isolated from paddy soil of *Oryza sativa*. *Int J Syst Evol Microbiol* **55**:1233–1237.
- Xiong W, Giannone RJ, Morowitz MJ, Banfield JF, Hettich RL. (2015). Development of an enhanced metaproteomic approach for deepening the microbiome characterization of the human infant gut. *J Proteome Res* **14**:133–141.
- Yabuuchi E, Kosako Y. (2005). Sphingomonadaceae. In: *Bergey's Manual of Systematic Bacteriology*, Garrity, GM, Brenner, DJ, Krieg, NR, & Staley, JT (eds) Vol. 2 Part C, Springer: New-York, p. 233.
- Yi H, Schumann P, Chun J. (2007). *Demequina aestuarii* gen. nov., sp. nov., a novel actinomycete of the suborder Micrococccineae, and reclassification of *Cellulomonas fermentans* Bagnara *et al.* 1985 as *Actinotalea fermentans* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **57**:151–156.
- Yoon J-H, Kang S-J, Park S, Lee S-Y, Oh T-K. (2007). Reclassification of *Aquaspirillum itersonii* and *Aquaspirillum peregrinum* as *Novispirillum itersonii* gen. nov., comb. nov. and *Insolitspirillum peregrinum* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **57**:2830–2835.
- Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, *et al.* (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**:1608–1615.
- Zhilina TN, Detkova EN, Rainey FA, Osipov GA, Lysenko AM, Kostrikina NA, *et al.* (1998). *Natronoincola histidinovorans* gen. nov., sp. nov., a new alkaliphilic acetogenic anaerobe. *Curr Microbiol* **37**:177–185.
- Zopfi J, Ferdelman TG, Fossing H. (2004). Distribution and fate of sulfur intermediates—sulfite, tetrathionate, thiosulfate, and elemental sulfur—in marine sediments. *Geol Soc Am Spec Pap* **379**:97–116.

Curriculum Vitae

Alexandre Bagnoud

born in May 17th 1984

Avenue de la Gare 45

Swiss nationality

1022 Chavannes-près-Renens

Switzerland

phone: +41 79/766'36'41

email: alexandre.bagnoud@gmail.ch



EDUCATION

- Coursera MOOC - Finding Hidden Messages in DNA (Bioinformatics I) 2015
- GTPB Training program in Bioinformatics – Analysing Whole-Genome Shotgun Metagenomic Data, Oeiras (Portugal) 2013
- Satellite symposium of SME 2013 – Metagenomic data analysis (Morat) 2013
- PhD thesis at EPFL (Environmental Microbiology Laboratory), Lausanne since 2010
- Master of Sciences in Biogeosciences, University of Neuchâtel, grade « excellent » 2006-2008
- Bachelor of Sciences in Biology, University of Neuchâtel, grade « very good », Landry price laureate 2003-2006
- High School Diploma in Lycée-Collège des Creusets, Sion, in Biology-Chemistry, specific option Physic 1998-2003

SCIENTIFIC EXPERIENCES

- PhD thesis, “Assessment of the impact of the Opalinus clay microbial community on nuclear waste storage” at EPFL, Lausanne since 2010
- Bioinformatics internship in Anders Andersson’s laboratory, at SciLifeLab, Stockholm 2014 (2 weeks)
- Participation to a field sampling campaign in a uranium contaminated peatbog, within the project of Dr. Yuheng Wang. 2012 (2 weeks)
- Training course of molecular biology (DNA extraction from rock samples) Institute of Radiochemistry, Forschungszentrum Dresden-Rossendorf, Germany. 2010 (1 week)

- In charge of research for the European project CO2 SOL STOCK, 2009-2010
Institute of Geology and Paleontology, University of Lausanne
- Master thesis, “Ecology of anammox in soils”, grade 6 2007-2008
- Training course of limnology and lacustrine microbiology, 2008 (2 weeks)
University of Valencia, Spain,
within the framework of the Ph.D. studies of Walter Xavier Alexis
- Training course of molecular tools in microbial ecology, 2008 (1 week)
Helmoltz Zentrum München, Germany

PARTICIPATION TO TEACHINGS

- In charge of the practicals in microbiology for the bachelor of Sciences since 2013
in Environmental Engineering, EPFL, Lausanne
- Teaching assistant in the “Microbiology for engineers” course for the bachelor of since 2010
Sciences and Environmental Engineering, EPFL, Lausanne
- In charge of the practicals in microbiology for the first year of Medicine, 2008
University of Neuchâtel
- In charge of the practicals in microbiology for the bachelor of Biology, 2008
- Assistantship for different field trips of the Biogeosciences master, 2008
University of Neuchâtel

SCIENTIFIC EXPERTISE

- Molecular biology tools:
DNA extraction from various types of samples, DNA purification and concentration, PCR, qPCR, DGGE, FISH, T-RFLP, cloning-sequencing
- Bioinformatics:
454 and Illumina data trimming (QIIME, Uparse), microbial community 16S rRNA gene data analysis (QIIME), statistical tools for microbial community analysis (R), familiar with UNIX, basic Bash and Python programming, NGS metagenomic contigs binning (CONCOCT), genes and metabolic pathways annotation
- Culture methods:
Classic culture methods, aerobic and anaerobic incubations, measurement of activity rates, SYBR Green microbial counting, running and maintenance of long-term recirculating cultural experiments in *in-situ* conditions
- Geochemical analysis:
Determination of nitrogen, iron and sulfur species concentrations and other physico-chemical parameters in water and soil, Ion chromatography, ICP-MS, determination of dissolved gases

LANGUAGES

- French: mother tongue
- English: good speaking, writing and reading skills
- German: school knowledge

REFERENCES

- Dr. Jakob Zopfi, Research Associate at University of Basel:
 - e-mail: jakob.zopfi@unibas.ch
 - phone: +41 (0)61 267 36 15
- Prof. Rizlan Bernier-Latmani, Associate Professor at EPFL, Lausanne:
 - e-mail: rizlan.bernier-latmani@epfl.ch
 - phone: +41 (0)21 693 50 01

PUBLICATIONS

- Wang Y., Bagnoud A., Suvorova E., McGivney E., Chesaux L., Phrommavanh V., Descostes M. and Bernier-Latmani R. (2014) Geochemical control on uranium(IV) speciation in a mining-impacted wetland. *Environmental Science & Technology*.
- Bagnoud A., Humbert S., Hai B., Schloter M., Conen F., Seth B. and Zopfi J. Anammox in soil: distribution, diversity and activity along a soil transect in the Jura Mountains (Bellefontaine, F). In prep.
- Bagnoud A., deBruijn I., Andersson A., Diomidis N., Leupin O., Schwyn B. and Bernier-Latmani R. A minimalistic microbial food web in an excavated deep subsurface clay rock. In prep.
- Bagnoud A., Chourey K., Hettich R., deBruijn I., Andersson A., Leupin O., Schwyn B. and Bernier-Latmani R. Hydrogen-driven carbon cycling in the deep subsurface. In prep.
- Bagnoud A., Leupin O., Schwyn B. and Bernier-Latmani R. Microbial hydrogen oxidation and sulfate reduction rate in Opalinus Clay. In prep.

ORAL PRESENTATIONS

- Bagnoud A., Schwyn B., Leupin O., de Bruijn I., Andersson A. and Bernier-Latmani R. (2014) Microbial hydrogen oxidation and carbon cycling in the deep subsurface. Presented at the ninth international symposium on subsurface microbiology, Pacific Grove, CA.
- Bagnoud A., Andersson A., de Bruijn I., Leupin O., Schwyn B., and Bernier-Latmani R. (2014) Microbial Oxidation of Hydrogen in the Deep Subsurface. Presented at the 15th International Symposium on Microbial Ecology, Seoul, South Korea.
- Bagnoud A., Schwyn B., and Bernier-Latmani R. (2012) Assessing the Metabolic Potential of Microbes in Opalinus Clay Rock. Presented at the Goldschmidt, Montréal, Canada.
- Bagnoud A., Schwyn B., Leupin X., and Bernier-Latmani R. (2013) In Situ Microbial Oxidation of H₂ in the Deep Subsurface. Presented at the 11th wwiss geoscience meeting, Lausanne, Switzerland.

Zopfi J., Humbert S., Bagnoud A., Tarnawski S., Conen F. and Seth B., (2009) Anammox in terrestrial ecosystems: Distribution, diversity and activity. *Geochim. Cosmochim. Acta* 73, A1538-A1538. Presented at the 19th Annual VM Goldschmidt Conference, 2009, Davos, Switzerland.

Humbert S., Tarnawski S., Bagnoud A., Conen F., Seth B., Mallet M.P., Fromin N., Aragno M. and Zopfi J., (2008) Terrestrial anammox bacteria: diversity and activity. Presented at the 12th International Symposium in Microbial Ecology, Cairns, Australia.

POSTER PRESENTATIONS

Bagnoud A., Mueller H, Schwyn B, Bernier-Latmani R, (2011). Characterization of the metabolic potential of microorganisms in Opalinus Clay at Mont Terri underground rock laboratory. Federation of European Microbiological Societies 2011, Geneva, Switzerland.

Bagnoud A., Humbert S, Conen F, Seth B and Zopfi J, (2009). Distribution, diversity and activity of anammox bacteria along a soil transect in the Jura Mountains in Bellefontaine (F). *Swiss Microbial Ecology* 2009, Einsideln, Switzerland.