Hydrogenases from a hydrothermal deep-sea vent metagenome and the development of a novel activity-based screen to identify H_2 uptake active enzymes from metagenomic libraries

Dissertation with the aim of achieving a doctoral degree at the Faculty of Mathematics, Informatics and Natural Sciences

Department of Biology

of University Hamburg

submitted by

Nicolas Rychlik

2015 in Hamburg

Approved by the Department of Biology, University of Hamburg at request of Jun.-Prof. Dr. Mirjam Perner

Second evaluator oft he dissertation: Prof. Dr. Wolfgang Streit

Date of defense (Disputation): 10. Juni 2016

Table of contents

Figures

Tables

Abbreviations

aa	amino acid(s)
Acc. No.	accession number
ad	up to
amp	ampicillin
ATP	adenosine triphosphate
bidest	bidistilled water
BLAST	Basic Local Alignment Search Tool
bp	base pairs
BSA	bovine serum albumin
°C	degree Celsius
Da	dalton
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DMF	dimethylformamide
DMSO	dimethyl-sulfoxide
dNTP	deoxyribonucleotide triphosphate
DSMZ	German collection of Microorganisms and Cell Cultures (“Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH”)
EC number	enzyme commission number
EDTA	ethylene-diamine-tetraacetid-acid
et al.	et alii (Latin: and others)
EtOH	ethanol
Fig.	figure
g	gram(s)
g	radial centrifugation force
GC %	percentage of G and C in DNA sequences
h	hour(s)
IPTG	isopropyl thio-β-D-galactoside
kb	kilobases
l	litre(s)
LB	Luria Bertani
µ	micro- (1×10^{-6})
m	milli- (1×10^{-3})

M	molar
mA	milliampere
MCS	multiple cloning site
min	minute(s)
mRNA	messenger RNA
MV	methyl viologen
n	nano- (1×10^{-9})
NCBI	National Center for Biotechnology Information
OD	optical density
ORF	open reading frame
PCR	polymerase chain reaction
pH	negative logarithm of the molar concentration of dissolved hydronium ions
RNA	ribonucleic acid
RNase	ribonuclease
rpm	rotations per minute
rRNA	ribosomal RNA
RT	room temperature
SDS	Sodium dodecyl sulphate
sec	second(s)
t	time
T_{ann}	annealing temperature
Taq	thermostable DNA polymerase of *Thermus aquaticus*
TE	Tris-EDTA
T_m	melting temperature
UV	ultraviolet
V	Volt
wt	wild type
w/vol	weight per volume
vol.	volume
vol/vol	volume per volume
vol/w	volume per weight
X-Gal	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Aminoacids

Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
Cys	Cysteine
Gln	Glutamine
Glu	Glutamic acid
Gyl	Glycine
His	Histindine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

Nucleobases

A	Adenine
C	Cytosine
G	Guanine
T	Thymine

1 Introduction

1.1 Hydrogenases

Hydrogen (H_2) is one of the most abundant chemical substances and is usually found bound in H_2O and organic hydrocarbons (Schlapbach and Zuttel, 2001). In most prokaryotic and eukaryotic metabolic pathways, H_2 plays a central role and is actively involved in the carbon, nitrogen and sulfur cycle (Dilling and Cypionka, 1990; Bothe *et al.*, 2010; Tang *et al.*, 2011). H_2, even in low concentrations, is an ideal electron carrier between organisms because it can freely diffuse through cytoplasmic membranes (Thauer *et al.*, 2010). In several fermentative processes H_2 can be produced to regenerate reducing equivalents (Hallenbeck, 2009). The ability to use H_2 as an energy source is widely distributed among chemolithotrophic organisms (Vignais and Billoud, 2007) and its oxidation provides up to -237 kJ mol^{-1} energy (Amend and Shock, 2001).

The interconversion between molecular H_2 and protons and electrons is catalyzed by hydrogenases ($H_2 <--> 2\ H^+ + 2\ e^-$), which can be found in Bacteria, Archaea and Eukarya (Takai *et al.*, 2005; Vignais and Billoud, 2007). The direction of the reaction depends on the redox potential of the components able to interact with the enzyme (Cord-Ruwisch *et al.*, 1988). Hydrogenases play a central role in energy metabolism of many microorganisms originating from various habitats (Takai *et al.*, 2005; Guiral *et al.*, 2006; Vignais and Billoud, 2007). The activation of H_2 is mechanistically challenging, and the catalytic mechanism is of considerable interest (Thauer *et al.*, 2010). Hydrogenases can be categorized according to the metal contents in the active site into three distinct phylogenetic classes: [NiFe]-, [FeFe]- and [Fe]-hydrogenases (Vignais and Billoud, 2007) (Figure 1).

The [NiFe]-hydrogenases are the best studied class of hydrogenases and are widespread present enzymes in Eubacteria and in Archaea (Lojou, 2011). Their biosynthesis, maturation and processing follows a complex pathway, which involves at least seven auxiliary proteins (Vignais and Billoud, 2007). This set of proteins directs for example the synthesis and incorporation of the metal center and their non-protein ligands into the active site of the large subunit or the removal of a C-terminal extension from the precursor of the large subunit after the insertion of the metallocentre (Vignais and Billoud, 2007). In contrast proteins necessary for the biosynthesis of [FeFe]-hydrogenases are not so numerous, just two radical S-adenosyl-methionine (SAM) proteins are required for the assembly of the active site (H-cluster biosynthesis) (Vignais and Billoud, 2007).

Hydrogenases are usually considered as highly O_2 sensitive, where [FeFe]-hydrogenases are known to react irreversibly with O_2, but [NiFe]-hydrogenases can react with O_2 to yield products which can be reactivated upon reduction (Lee *et al.*, 1973). Although exceptions exist, [NiFe]-hydrogenases are commonly associated with H_2 consumption, [FeFe]-hydrogenases are usually involved in H_2 production (Vignais and Billoud, 2007) and [Fe]-hydrogenases are restricted to methanogenesis (Thauer, 1998).

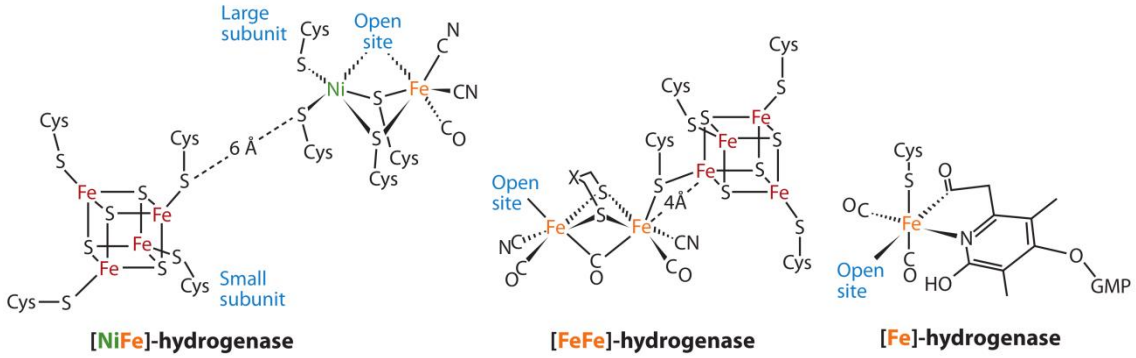


Figure 1: The metal contents in the active sites of the three types of hydrogenases: [NiFe]-hydrogenases, [FeFe]-hydrogenases and [Fe]-hydrogenases. These phylogenetically distinct hydrogenases have unusual structural features in common, such as CO- and CN ligands (Thauer *et al.*, 2010)

1.2 Hydrothermal deep-sea vents

Deep-sea hydrothermal venting biotopes can provide oases of life in the otherwise mostly vast deep-sea. The biomass in deep-sea hydrothermal vent environments is mostly of autochthonous nature, since hardly any primary produced organic matter from the surface of the oceans arrives at the deeper seafloor (Karl *et al.*, 1980; Jannasch and Taylor, 1984; Prieur, 1997). The local primary production at deep-sea vents does not base on sunlight because sunlight cannot reach down to the deep-sea (Felbeck and Somero, 1982; Jannasch and Taylor, 1984; Prieur, 1997).

However, chemical energy (e.g. hydrogen- or sulfide oxidation) fuels primary biomass production instead, where reduced inorganic electron donors (e.g. hydrogen, sulfide) are supplied by the emanating hot hydrothermal fluids (Figure 2). Depending on the type of host rock, different inorganic compounds can be enriched in the fluids, resulting from water-rock reactions or magma degassing and are transported from inner earth to habitable biotopes (Karl *et al.*, 1980; Felbeck and Somero, 1982; Jannasch and Taylor, 1984; Prieur, 1997; Kelley *et al.*, 2002; Fouquet *et al.*, 2010). Ultramafic rock has an igneous origin, it is assumed to have formed by slow cooling and has a relatively coarse-grained (> 3 mm) texture (Le Maitre *et al.*, 2002). This type of rock has a low silica

content (< 45 %), is rich in magnesium and iron and is classified according to their content of mafic minerals (Le Maitre *et al.*, 2002). In ultramafic-hosted venting systems hydrogen (H_2) is considerably enriched and its oxidation is proposed to be the major energy source for biomass synthesis (McCollom, 2007). In contrast, basalt are igneous rocks with a higher silica content between 45 – 52 % (Le Maitre *et al.*, 2002). As a result in basalt-hosted hydrothermal systems sulfide is considerably higher and sulfide oxidation is estimated to provide the energy for most primary production (McCollom and Shock, 1997). As hot, highly reduced hydrothermal emissions rise from the inner earth and come in contact with cold, oxygenated ambient seawater, mixing processes constitute multiple microhabitats with steep physico-chemical gradients. Besides electron donor (e.g. hydrogen or sulfide) availability, temperature and oxygen concentrations can also vary considerably in hydrothermal habitats (Mullineaux *et al.*, 2003; Shock and Canovas, 2010; Kim *et al.*, 2013; Perner *et al.*, 2013). This leads to the fact, that hydrothermal biotopes can facilitate a range of niches with different thermal and chemical conditions resulting in highly diverse deep-sea vent microbial communities (Reysenbach *et al.*, 2000; Luther *et al.*, 2001). Due to this metabolic, physiologic and taxonomic richness of hydrothermal deep-sea vent niches, theses habitats have a great potential to identify e.g. thermostable enzymes or enzymes with other attractive abilities for biotechnological applications (Prieur, 1997; Deming, 1998; Guezennec, 2002).

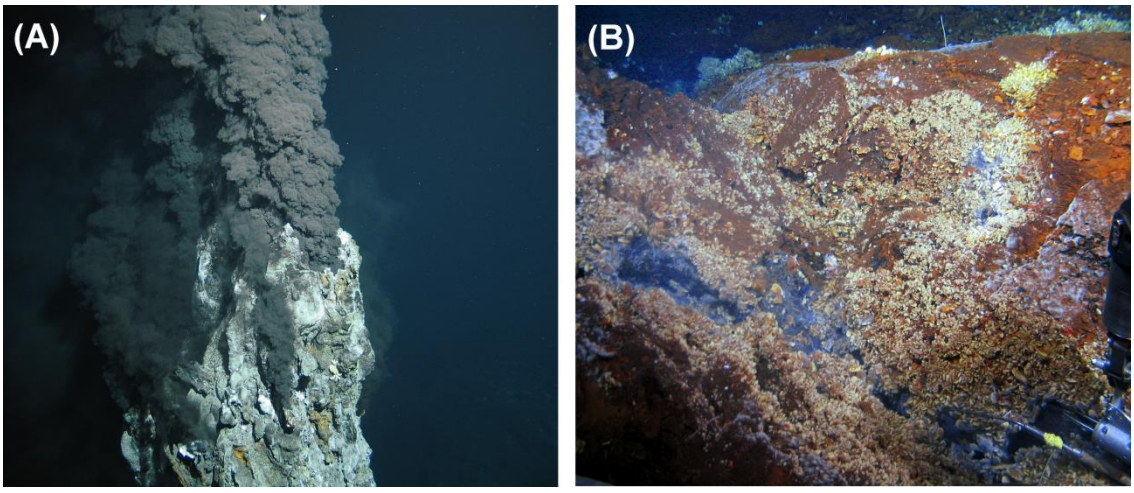


Figure 2: Examples for the uniqueness of hydrothermal deep-sea vent biotops. (A) A black smoker emitting hot hydrothermal fluids. (B) An exemplary warm water vent emitting diffuse hydrothermal fluids with clamps, shrimps and bacterial mats (© IFM GEOMAR)

Since H_2 can be enriched in the hydrothermal emissions, it is not surprising that a large number of abundant and phylogenetically diverse H_2 oxidizers can be recognized in these niches (Campbell *et al.*, 2006; Perner *et al.*, 2010; Flores *et al.*, 2011). Given that different H_2 consuming microbes can colonize abiotically distinct hydrothermal habitats (H_2-rich versus H_2-poor, oxygen-rich versus oxygen-poor, high versus low temperature),

these habitats provide a large potential for recovering novel hydrogenases with specific biochemical properties (temperature or oxygen tolerant). With respect to these abiotic conditions, the rich energy source provided by H_2 oxidation and the large numbers of diverse H_2-oxidizing microorganisms, hydrothermal vents facilitate ideal conditions for seeking oxygen tolerant hydrogenases.

However, working directly with hydrogenase enzymes from hydrothermal deep-sea vent habitats is difficult due to several reasons: (A) Sample sizes are normally limited because of the logistical challenges to gain the required deep-sea samples. (B) Additionally, during the time-consuming salvage of the collected deep-sea samples hydrogenase proteins can be irreversible inhibited in consequence of oxygen exposure. On the contrary, working with DNA originated from deep-sea samples is easier because DNA is more resistant towards degradation and in addition larger amounts of DNA can be relatively easy collected. Furthermore, the metagenomic potential of recovered hydrothermal deep-sea vent samples is accessible by storing the received DNA in genomic libraries.

1.3 Metagenomics and metagenomic hydrogenases

Novel hydrogenases have been recovered traditionally through cultivation dependent approaches (cf. Adams and Mortenson, 1984; Brugna-Guiral *et al.*, 2003; Guiral *et al.*, 2006; Ludwig *et al.*, 2009; Vargas *et al.*, 2011). However, since less than 1 % of all microbes are cultivatable (cf. Amann *et al.*, 1995; Handelsman, 2004), capturing the genetic information of all organisms in a sample by culture-independent methods becomes crucial to tap the tremendous resources in the environment. A metagenome comprises the entire genomes of all microorganisms in a specific habitat and metagenomics are the culture-independent investigations of these genomes. The DNA information of a specific habitat can be stored in metagenomic large insert libraries by using the total isolated DNA, which is ligated into a vector (e.g. fosmid) and transferred into a well-cultivable and suitable host, most often *E. coli* (Handelsman *et al.*, 1998). Metagenomic libraries are powerful tools of metagenomics to gain access to the physiology and genetics of uncultured organisms and to identify new enzymes, antibiotics and other reagents in libraries from diverse environments (Handelsman, 2004). Traditionally, novel biocatalysts have been obtained through cultivation dependent approaches (Ogawa and Shimizu, 1999). In the last years, metagenomics could be applied successfully for discovering novel biocatalysts, specific genes or other valuable enzymes by using either sequence-based technologies, especially since next

generation sequencing methods (NGS) are available, or by function-based approaches (Streit and Schmitz, 2004; Ferrer *et al.*, 2009; Uchiyama and Miyazaki, 2009; Iqbal *et al.*, 2012).

A handicap in discovering genes by sequence-based tools is that comparative sequence analysis only taps the potential resources but not the actual activity and functionality and that identification of genes is based on conserved regions of already known genes (Daniel, 2004; Yun and Ryu, 2005; Perner *et al.*, 2011b; Rabausch *et al.*, 2013). Nevertheless this is also the main advantage of sequence-based screening methods, because of the independency of correct expression, folding, maturation, activation and secretion of heterologous proteins in the host strains. Additionally, the development of NGS technology and improvements in bioinformatics have significantly improved this methodology (Simon and Daniel, 2011) and might be useful for obtaining insight into the theoretical enzymatic potential of a specific habitat.

However, truly novel enzymes possessing new motifs can only be discovered by function-based screening approaches (Rees *et al.*, 2003; Handelsman, 2004; Ferrer *et al.*, 2005; Yun and Ryu, 2005; Heath *et al.*, 2009; Perner *et al.*, 2011b; Rabausch *et al.*, 2013). Thus, there is a considerable interest in the use, development and improvement of function-based metagenome screening technologies (Tuffin *et al.*, 2009). Inherent limitations also exist when using a function-based screening because of the dependence on a foreign host for expressing environmental DNA from an unknown source organism, which can appear to be difficult (Yun and Ryu, 2005; Perner *et al.*, 2011b). Generally if *E. coli* is used as heterologous host it is assumed that only 40 % of foreign genes are correctly expressed to an active enzyme and thereby detectable by function-based screening processes (Rondon *et al.*, 1999; Gabor *et al.*, 2004). A different codon usage could be one problem for the functional expression of metagenomic proteins in heterologous hosts (Streit *et al.*, 2004; Warren *et al.*, 2008; Uchiyama and Miyazaki, 2009). However, the usage of alternative hosts could resolve these problems and several alternative hosts were successfully used for metagenomic function-basted screening processes, for example *Pseudomonas putida* (Martinez *et al.*, 2004), *Streptomyces lividans* (Courtois *et al.*, 2003; McMahon *et al.*, 2012), *Sinorhizobium meliloti* (Wang *et al.*, 2006; Schallmey *et al.*, 2011), *Rhizobium leguminosarum* (Li *et al.*, 2005; Wexler *et al.*, 2005) and *Desulfovibiro* sp. (Rousset *et al.*, 1998). Furthermore, to exploit the potential involved in using an alternative host for expression of metagenomic genes, a broad-host-range vector should be used for the construction of metagenomic libraries, which is efficiently transferable from *E. coli* to the alternative host and should be stable maintained in the host (Aakvik *et al.*, 2009).

For the development of a function-based screening for [NiFe]- H_2 uptake hydrogenases or similar enzymes, the recombinant expression of a corresponding enzyme in a heterologous host can indeed encounter several difficulties (Blokesch *et al.*, 2001; Casalot and Rousset, 2001). Several proteins are involved in the complex pathway of assembly, maturation and activation, because of the difficult architecture of the enzyme (Casalot and Rousset, 2001; Bock *et al.*, 2006; Vignais and Billoud, 2007). However, several studies have demonstrated that a heterologous expression of hydrogenases can be successful (Friedrich *et al.*, 1984; Casalot and Rousset, 2001; Bock *et al.*, 2006; Maroti *et al.*, 2009) if proteins for assembly and maturation of the heterologous host work for the heterologous hydrogenase as well or if the required corresponding maturation genes are also heterologously expressed. The observation that the structural genes encoding [NiFe]-hydrogenases are mostly clustered together with accessory genes for maturation and activation might be advantageous for heterologous expression of metagenomic hydrogenases located on metagenomic large insert libraries.

1.4 Biotechnological relevance of hydrogenases

Hydrogenases are extremely interesting for biotechnological applications. In consequence of the limitation of fossil fuels and the impact of their usage on global climate change an alternative, clean, renewable energy fuel is required. H_2 is being regarded as an important renewable fuel for future renewable energy technologies (Karyakin *et al.*, 2005; Armstrong *et al.*, 2009). H_2 converting enzymes have become interesting because of their ability to produce H_2 from H_2O (biocatalyzed electrolysis). Additionally, H_2 delivers the highest energy output relative to molecular weight, and its combustion only produces water and heat. H_2 / O_2 fuel cells can directly generate electricity by using H_2 as chemical energy source and oxidize H_2 into water. The commercially available H_2 / O_2 fuel cells contain platinized anode and cathode compartments for this electrocatalysis. The electrocatalysis by platinum doesn't satisfy all requirements of fuel cells, due to the inactivation by CO and H_2S (Karyakin *et al.*, 2005; Tye *et al.*, 2005), which is a serious problem because of impurities within the used H_2 fuel. One biotechnological application for hydrogenases is in enzymatic H_2 / O_2 fuel cells. Here hydrogenases seem to be good candidates to replace expensive chemicals and rare elements as redox mediators like platinum. On the one hand Armstrong and Hirst (2011) could show that enzymes are the most efficient electrocatalysts and can set new standards for catalysts of future energy technologies. Additionally, hydrogenase enzymes have the further advantages of high specify and low over-potential for H_2

oxidation, biodegradability and bioavailability (Lamle *et al.*, 2003; Lojou, 2011). On the other hand for the development of such an enzymatic fuel cell many challenges have to be overcome like the high sensitivity to oxygen of most hydrogenases, the immobilization of the enzyme on the electrode and the stability of both hydrogenase and electrode are only a few examples (Lojou, 2011). One of the most crucial challenges in these biotechnological applications is to resolve the problem associated with the oxygen sensitivity of hydrogenases (Ludwig *et al.*, 2009; Lojou, 2011). Summarizing, enzyme-based fuel cells using hydrogenases as catalyst could be realized, but the oxygen inactivation of the enzyme and the associated limitation of the power output is continuously a serious problem (Lojou *et al.*, 2008; Wait *et al.*, 2010). Further developments of platinum free H_2 / O_2 fuel cells could come from discovery of hydrogenases having even greater O_2-tolerance or by genetic modification of hydrogenase systems known to date (Wait *et al.*, 2010).

1.5 A functional screening for [NiFe]-hydrogenases originated from metagenomic libraries

For the identification and investigation of novel [NiFe]-hydrogenases or similar enzymes originated from hydrothermal deep-sea vent habitats, the construction of large insert metagenomic libraries is a promising means to capture the genetic potential for later expression in a heterologous host. As described before, a functional metagenomic approach using a metagenomic library constructed with recovered DNA originated from a hydrothermal deep-sea vent sample is a good alternative to working with hydrogenase enzymes harvested directly from the same sample due to logistical reasons. The sampling procedure can result in irreversibly inhibited proteins owing to oxygen exposure in contrast to working with DNA, which is more resistant toward degradation.

Additionally, the time-consuming and labor-intensive cultivating of hydrothermal deep-sea vent microbes with potential to exhibit novel [NiFe]-hydrogenases could be avoided due to using the culture independent metagenomic methods. Furthermore, metagenomics enable the access to enzymes originated from uncultivable microbes and allow investigating these enzymes.

For the development of a functional screening for seeking novel [NiFe]-hydrogenases, it must be taken into consideration that several proteins are involved in assembly, maturation and activation of [NiFe]-hydrogenases and that these accessory proteins are usually highly specific. As mentioned above, that is one reason why a heterologous

expression of metagenomic hydrogenases in a foreign host can be difficult, although respective genes are mostly clustered together with the structural genes encoding the [NiFe]-hydrogenases (Friedrich *et al.*, 1984; Casalot and Rousset, 2001; Bock *et al.*, 2006; Maroti *et al.*, 2009). Therefore, an appropriate host with maturation proteins capable of assembling the [NiFe]-hydrogenases targeted in the metagenome needs to be chosen to enable a functional screening.

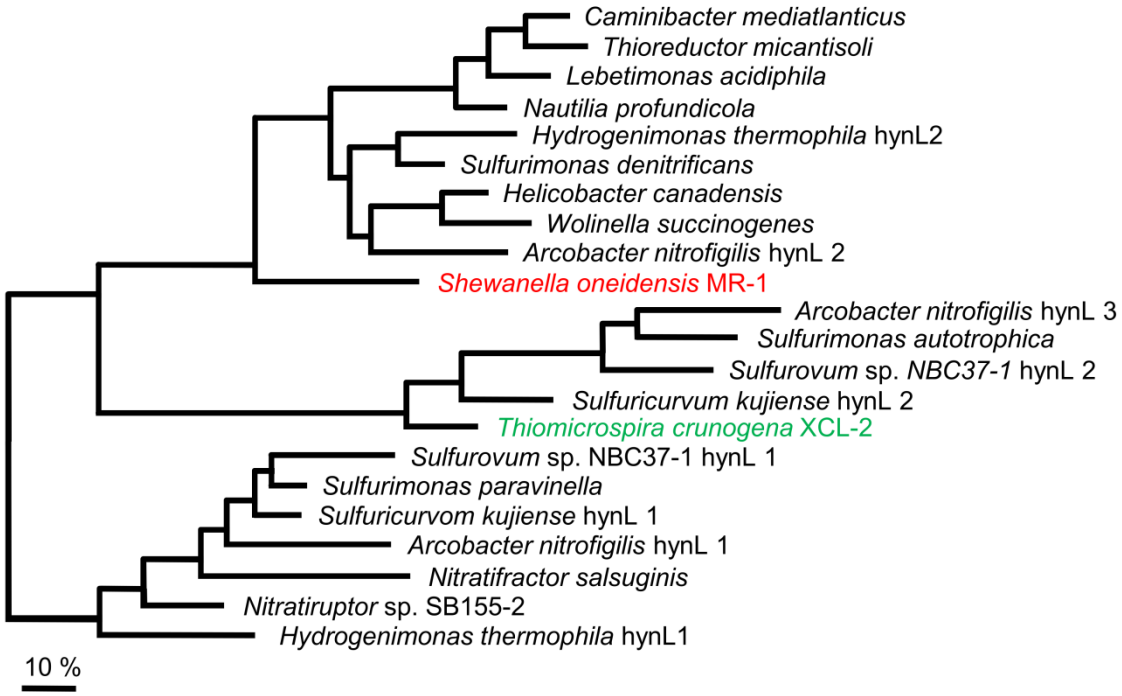


Figure 3: Phylogenetic tree of [NiFe]-hydrogenases originated from *Epsilonproteobacteria* and the closest relatives originated from other phyla. Besides several *Epsilonproteobacteria* (black), the closest relatives are the *Gammaproteobacteria S. oneidensis* MR-1 (red) and *Thiomicrospira crunogena* XCL-2 (green). The phylogenetic tree is based on amino acid sequence similarities and is computed with a maximum likelihood algorithm using the GTR substitution model.

Epsilonproteobacteria seem to be predominant hydrogen consumers in hydrothermal deep-sea vent habitats (Campbell *et al.*, 2006; Perner *et al.*, 2010; Flores *et al.*, 2011) and are associated with the production of primary biomass in such habitats if hydrothermal fluids are enriched with hydrogen. Therefore, one aim of the proposed functional screening was to allow the detection of metagenomic hydrogenases originated from *Epsilonproteobacteria*, but the prokaryotic model organism and versatile host *E. coli* can unlikely assemble epsilonproteobacterial hydrogenases (Blokesch *et al.*, 2001; Casalot and Rousset, 2001; Böck *et al.*, 2006; Perner *et al*. 2011). With respect to hydrogenase phylogeny, the hydrogenase from *Shewanella oneidensis* MR-1 shows similarities to those from *Epsilonproteobacteria* (Figure 3 written in red) and therefore it can be assumed that the essential assembly and maturation apparatus may be compatible for maturing hydrogenases from certain *Epsilonproteobacteria*.

Shewanella oneidensis MR-1 is a facultative aerobic, non-sporulating Gram-negative *Gammaproteobacteria* isolated from anaerobic sediments (Venkateswaran *et al.*, 1999). It is a mesophilic polar flagellated organism with an optimal growth temperature of 30°C. *S. oneidensis* MR-1 can metabolize a broad range of typical fermentative end products (e.g. acetate, fumarate, lactate, malate, pyruvate and succinate), but cannot use glucose as a carbon and energy source. Given its evident preference for C-3 and C-1 carbon sources, *S. oneidensis* MR-1 will depend on anaplerotic reactions, the glyoxylate pathway, gluconeogenesis and the non-oxidative pentose pathway reactions, to generate the intermediates required for biosynthetic metabolism. Additionally, *S. oneidensis* can utilize amino acids as carbon source and fatty acids are also used as carbon and energy source.

S. oneidensis MR-1 belongs to a class of bacteria known as “Dissimilatory Metal Reducing Bacteria” (DMRB). Under O_2-deficient conditions it can reduce oxidized metals like Fe(III) (Lovley *et al.*, 1989). It can couple the oxidation of H_2 - catalyzed by a sole [NiFe]-hydrogenase - to the reduction of oxidized metals (Lovley *et al.*, 1989; Petrovskis *et al.*, 1994; Dawood and Brozel, 1998; Heidelberg *et al.*, 2002; Meshulam-Simon *et al.*, 2007). In defined culture conditions and if in *S. oneidensis* MR-1 H_2 uptake hydrogenases are active, H_2 is used as electron donor and electrons are transferred to Fe(III)citrate which is then reduced to Fe(II)citrate.

*Reaction: 2 Fe(III)citrate + H_2 --> 2 Fe(II)citrate$^-$ + 2 H^+ (ΔG_r = -65.87 kJ*mol^{-1})*

Additionally to one [NiFe]-hydrogenase, a second putative hydrogenase, a [FeFe]-hydrogense, was identified in the genome of *S. oneidensis* MR-1 (Heidelberg *et al.*, 2002). However, the [FeFe]-hydrogenase of *S. oneidensis* MR-1 is described as hydrogen forming hydrogenase without the capability to consume and utilize H_2 as electron source (Meshulam-Simon *et al.*, 2007). Therefore, an influence on the metal reducing capability of *S. oneidensis* MR-1 is not described for the putative [FeFe]-hydrogenase in literature (Lovley *et al.*, 1989; Petrovskis *et al.*, 1994; Dawood and Brozel, 1998; Heidelberg *et al.*, 2002; Meshulam-Simon *et al.*, 2007) and seems to be highly unlikely.

An in-frame deletion introduced in the large subunit of the structural [NiFe]-hydrogenase gene (*hyaB*) of *S. oneidensis* MR-1 should result into a deletion mutant which can no longer oxidize H_2 and transfer electrons from this reaction to oxidized metals like Fe(III). The missing ability of the deletion mutant should be used in a function based screening procedure to recover H_2 uptake activity from metagenomic large insert library clones. Recombinant expression of hydrogenase genes encoded on corresponding metagenomic fosmids should complement the *S. oneidensis* MR-1 [NiFe]-hydrogenase

deletion mutant and result in active hydrogenases which restore detectable metal reducing capacities.

1.6 Intention of this work

The aim of this study was the identification and investigation of novel, preferably oxygen tolerant hydrogenase enzymes encoded on metagenomic material from hydrothermal deep-sea vent habitats. Therefore, a function-based screening approach for metagenomic libraries was developed to investigate metagenomic material conserved in broad host range fosmid libraries to identify the searched hydrogenase enzymes.

Since heterologous expression of functional hydrogenases in the standard host *Escherichia coli* is difficult because complex interactions of maturation and assembly proteins are often needed, an alternative host was chosen to optimize the capability of maturation and assembly of heterologous hydrogenases and to circumvent the tap of the tremendous resources of uncultivated hydrogen oxidizing microbes in hydrothermal deep-sea vent habitats. Therefore, a function-based screening method with the alternative heterologous host *Shewanella oneidensis* Δ*hyaB,* a [NiFe]-hydrogenase deletion mutant of the *S. oneidensis* MR-1 strain, was established.

This would be the first function-based screen for seeking H_2 uptake active hydrogenases from metagenomic fosmid libraries, based on restored metal reducing capabilities of the constructed *S. oneidensis* Δ*hyaB* mutant if hydrogenases are heterologously expressed in the surrogate host.

2 Material & Methods

2.1 Bacterial strains, vectors, primers and constructs

In the following table 1 bacterial strains, vectors and vector-constructs used in this study are listed. Furthermore, in the following table 2 primers used in this study are listed.

Table 1: Bacterial strains and vectors used in this study

Strain or plasmid	Relevant genotype or descriptions	Reference or source
Bacterial strains		
Escherichia coli		
WM3064	donor strain for conjugation *(thr*B1004 *pro thi rps*L *hsd*S *lacZ*ΔM15 RP4-1360 Δ(*ara*BAD)*567* Δ*dap*A1341::[*erm pir*(wt)])	(Lassak *et al.*, 2010)
EPI300-T1^{R}	*E. coli* host for fosmid library; Phage T1-resistant and lacZ^{-} strain with L-arabinose-induced chromosomally expressed TrfA, (F^{-} *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) (StrR) φ80d*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ^{-} r*ps*L *nup*G *trf*A *ton*A *dhfr)*	Epicentre
DH5α	host for *E. coli* cloning (F^{-} Φ80*lac*ZM15 Δ(l*ac*ZYA-argF)U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A, *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^{-})	Gibco BRL
S17-1	donor strain for conjugation (*RP4-2-Tc::Mu-Km::Tn7, pro, res*$^{-}$ *mod*$^{+}$, *Tp*^{r} *Sm*^{r})	
HBH-101 (pRK2013)	mobilizing strain for triparental conjugation (F^{-} *mcr*B *mrr hsd*S20(r_B^- m_B^-) *rec*A13 *leu*B6 *ara*-14 *pro*A2 *lac*Y1 *gal*K2 *xyl*-5 *mtl*-1 *rps*L20(*Sm*^{R}) *gln*V44 λ^{-} *tra*$^{+}$ *Km*^{r})	(Figurski and Helinski, 1979)
Shewanella oneidensis		
MR-1	wild type strain	(Venkateswaran *et al.*, 1999)
Δ*hyaB*	deletion mutant of the large subunit of the sole [NiFe]-hydrogenase	this study

Aciduliprofundum boonei T469	wild type strain	(Reysenbach *et al.*, 2006)
Aciduliprofundum sp. MAR08-339	wild type strain	Material was kindly provided by Anna-Louise Reysenbach
Vectors / vector-constructs		
pNPTS138-R6KT	suicide plasmid for in-frame deletions; *mobRP4*$^{+}$*ori*-R6K *sacB*; *Km*^{r}	(Lassak *et al.*, 2010)
pNPTS138-R6KT::ΔhyaB-gen^{R}	*hyaB* deletion fragment separated with gentamycin resistance cassette in pNPTS138-R6KT; *Gm*^{r}	this study
pBBR1MCS-5	broad-host-range cloning vector; *Gm*^{r}*; P*$_{lac}$	(Kovach *et al.*, 1995)
pGEM-t	cloning vector; *Ap*^{r}*; P*$_{T7}$*; P*$_{SP6}$*; orif1*	Promega
pGEM-t::hyaBup	pGEM-t with an upstream fragment of the *hyaB* region	this study
pGEM-t::hyaBdw	pGEM-t with an downstream fragment of the *hyaB* region	this study
pGEM-t::gen^{R}	pGEM-t with gentamycin resistance cassette from pBBR1MCS-5	this study
pGEM-t::hyaBup-gen^{R}	pGEM-t with gentamycin resistance cassette from pBBR1MCS-5 and an upstream fragment of the *hyaB* region	this study
pGEM-t::ΔhyaB-gen^{R}	*hyaB* deletion fragment separated with gentamycin resistance cassette in pGEM-t	this study
pBSK+	pBluescript SK+ phagemid vector; Ap^{r}; P$_{T3}$; P$_{T7}$; orif1; pUCori	Stratagene
pBSK+::P18-F1	pBSK+ with subcloned segment F1 amplified by PCR using pRS44::SP-P18F11 as template	this study
pBSK+::P18-F1+2	pBSK+ with subcloned segments F1 and F2 amplified by PCR using pRS44::SP-P18F11 as template	this study
pRS44	cloning vector for fosmid libraries; Cm^{r}; Km^{r}; *oriT; ori2, oriV; parABCDE; cos; loxP*	(Aakvik *et al.*, 2009)
pRS44:: hyaB	a *S. oneidensis* MR-1 genome fragment of ~35,000 bp including the whole [NiFe]-hydrogenase operon in pRS44 for complementation	this study

pRS44::SP-P05H01	fosmid clone with hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P18F11	fosmid clone with hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::P18-BamHI-SC1	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-BamHI-SC2	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-BamHI-SC3a	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-BamHI-SC3bc	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-BamHI-SC4ab	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-BamHI-SC5	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-StuI-SC1	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-StuI-SC2ab	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-StuI-SC3	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-StuI-SC4	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-StuI-SC5	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-BsaAI-SC1	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-SacII-SC2	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-SacII-SC3	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::P18-PCR-F1+2	subclone of the pRS44::SP-P18F11 fosmid clone with hydrogenase activity	this study
pRS44::SP-P11C02	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study

pRS44::SP-P11H02	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P11H04	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P30B05	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P30F07	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P31A02	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P31A05	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P31B08	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P35F04	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P37B01	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P37B03	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P44E07	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study
pRS44::SP-P44H11	metagenomic library fosmid with false positive hydrogenase activity from a hydrothermal deep-sea vent metagenomic library	this study

Table 2: Primers used in this study. Restriction sites in the primer sequences are underlined and were used for site directed cloning.

Primer name	Sequence
gen_EcoRI_F	CGC GAA TTC GAC GCA CAC CGT GGA AA
gen_KpnI_RC	GGC GGT ACC GGC GGC GTT GTG ACA ATT T
hyaBup_PstI_F	GGT CCT GCA GTG GCC GTT TTG ATG CAG GTG
hyaBup_EcoRI_RC	GCG GAA TTC TTT CAG TAT GAC TTC AAT
hyaBdw_KpnI_F	GAC GGT ACC TCG ATG CTG TCA ATG CCC TG
hyaBdw_SalI_RC	GAC GTC GAC CGA ATG CGG GTT TCA GAA TGG
hyaB_cont_RC	CGG CCA AGC AAT ATA GAA TCC G
hyaB_cont_F	GCC GCT GCG ATT ATC TCT GTT G
modT7_prom	CGA CGG CCA GTG AAT TGT AAT ACG
pCC1FOS_rev	CTC GTA TGT TGT GTG GAA TTG TGA GC
Sub2_T7_PCR	GGG GCA GAG AAT TGC TAT AGG
hyp454_P2_F	CCT CTG AGA TCG GTC ATT GGT AT
hyp454_P2_RC	AGG CCT TGC TGA CGA TGT TGA TCT G

2.2 Media and supplements

2.2.1 Antibiotics and other supplements

If required, after autoclaving used medium was supplemented with different antibiotics and supplementary substances accordingly Table 3.

Table 3: Overview of required antibiotics and supplementary substances

substance	solvent	stock concentration	end concentration
gentamycin	H_2O_{dest}	10 mg/ml	10 µg/ml
chloramphenicol	Ethanol (100 %)	12.5 mg/ml	12.5 µg/ml
kanamycin	H_2O_{dest}	20 mg/ml	20 µg/ml
ampicillin	H_2O_{dest}	100 mg/ml	100 µg/ml
IPTG	H_2O_{dest}	0.1 mol	0.2 mmol
X-Gal	DMF	50 mg/ml	50 µg/ml
diaminopimelic acid	H_2O_{dest}	30 mM	0.3 mM
sucrose	H_2O	50 % (w/v)	10 % (w/v)
Copy Control Induction Solution	H_2O	1000×	1×
Copy Control Autoinduction Solution	H_2O	500×	1×

All substances dissolved in H_2O are sterilized by filtration through a cellulose syringe with a pore size of 0.22 µm or PTFE filter with a pore size of 0.2 µm.

2.2.2 LB medium

LB medium contained:

tryptone	10	g/l
yeast extract	5	g/l
NaCl	10	g/l
H_2O	ad 1000	ml

For LB agar plates, medium was solidified by adding 1.4 % (w/v) agar before medium was autoclaved at 121 °C for 20 min.

2.2.3 **Modified FW medium (Lovley et al., 1989)**

FW medium contained:

$NaHCO_3$	2.50	g/l
KCl	0.10	g/l
NH_4Cl	1.50	g/l
$NaH_2PO_4 \times H_2O$	0.60	g/l
$CaCl \times 2H_2O$	0.10	g/l
Fe(III)citrate	3.00	g/l
L-arginine-hydrochloride	0.02	g/l
L-glutamine	0.02	g/l
D / L-serine	0.04	g/l
trace elements solution	10	ml/l
H_2O	ad 900	ml

trace elements solution (Sako *et al.*, 1996):

Nitrilotriacetic acid (NTA)	1.50	g/l
$MgSO_4 \times 7H_2O$	3.00	g/l
$MnSO_4 \times 2H_2O$	0.50	g/l
NaCl	1.00	g/l
$FeSO_4$ x $7H_2O$	0.10	g/l
$CoSO_4 \times 7H_2O$	0.18	g/l
$CaCl_2 \times 7H_2O$	0.10	g/l
$ZnSO_4 \times 7H_2O$	0.18	g/l
$CuSO_4 \times 5H_2O$	0.01	g/l
$KAl(SO_4)_2 \times 12H_2O$	0.02	g/l
H_3BO_3	0.01	g/l
$Na_2MoO_4 \times 2H_2O$	0.01	g/l
$NiCl_2 \times 6H_2O$	0.025	g/l
$Na_2SeO_3 \times 5H_2O$	0.30	g/l
H_2O	900	ml

First dissolve nitrilotriacetic acid and adjust pH to 6.5 with KOH, then add minerals. Final pH 7.0 (with KOH). Ad 1000 ml with H_2O.

vitamin solution (Balch *et al.*, 1979):

Biotin	2.0	mg/l
Folic acid	2.0	mg/l
Vitamin B6	10.0	mg/l
Vitamin B1	5.0	mg/l
Vitamin B2	5.0	mg/l
Niacin	5.0	mg/l
Vitamin B5	5.0	mg/l
Vitamin B12	0.1	mg/l
4-Aminobenzoic acid	5.0	mg/l
Lipoic acid	5.0	mg/l

The anaerobic cultivation of *S. oneidensis* in FW medium was conducted in serum bottles with rubber stoppers. Oxygen was removed from the FW medium by flushing hot medium with N_2 (5.0) for 1 hour. The pH of the medium was adjusted to 7.0 with 1 M NaOH and again flushed with N_2 (5.0) for 15 minutes. Serum bottles with a volume of 120 ml were

purged with N_2, filled with 50 ml FW medium, sealed with butyl-rubber stoppers and autoclaved at 121 °C for 20 min. Afterwards, the headspace was replaced with H_2/CO_2 gas (80 % / 20 % v/v) and 0.5 ml vitamin solution was added by passing a cellulose syringe filter to finish the preparation of the medium.

2.2.4 **Aciduliprofundum medium (DSMZ medium 1083)**

Aciduliprofundum medium contained:

NaCl	30.00	g
$MgSO_4 \times 7\ H_2O$	3.50	g
$MgCl_2 \times 6\ H_2O$	2.75	g
$CaCl_2 \times 2\ H_2O$	0.38	g
KCl	0.33	g
NaBr	0.05	g
$(NH_4)_2SO_4$	0.10	g
KH_2PO_4	0.28	g
Na_3-citrate	2.94	g
yeast extract	1.00	g
tryptone	1.00	g
sulfur, powdered	10.00	g
$Na_2S \times 9\ H_2O$	0.50	g
Wolfe's mineral elixier	1.00	ml
resazurin	0.50	mg
H_2O	ad 1000	ml

Wolfe's mineral elexier:(Sako *et al.*, 1996)

$MgSO_4 \times 7\ H_2O$	30.00	g
$MnSO_4 \times H_2O$	5.00	g
NaCl	10.00	g
$FeSO_4$ x $7\ H_2O$	1.00	g
$CoCl_2 \times 6\ H_2O$	1.80	g
$CaCl_2 \times 7\ H_2O$	1.00	g
$ZnSO_4 \times 7\ H_2O$	1.80	g
$CuSO_4 \times 5\ H_2O$	0.10	g
$KAl(SO_4)_2 \times 12\ H_2O$	0.18	g
H_3BO_3	0.10	g
$Na_2MoO_4 \times 2\ H_2O$	0.10	g
$(NH4)2Ni(SO4)_2 \times 6\ H_2O$	2.80	g
$Na_2WO_4 \times 2\ H_2O$	0.10	g
Na_2SeO_4	0.10	g
H_2O	1000	ml

First adjust pH to 1.0 with diluted H_2SO_4, then add and dissolve the salts.

Dissolve ingredients (except sulfide), boil medium for 1 min, then cool to room temperature under 80 % N_2 and 20 % CO_2 gas mixture. Dispense under same gas atmosphere in culture vessels containing the appropriate amount of powdered sulfur and

autoclave for 60 min at 105 °C. Prior to inoculation add sulfide from a sterile, anoxic stock solution prepared under N_2 and adjust pH of completed medium to 4.5.

2.3 Culture conditions

2.3.1 Cultivation of bacterial and archaeal strains

Shewanella oneidensis, Aciduliprofundum boonei T469 and the *Aciduliprofundum* sp. MAR08-339 and *Escherichia coli* strains used in this study are listed in Table 1.

S. oneidensis strains were cultivated at 28°C liquid LB medium, on LB agar plates, or fresh water (FW) medium.

E. coli strains were cultivated at 37°C in liquid LB medium or on LB agar plates. Medium for *E. coli* WM3064 (Table 1) were supplemented with DAP (Lassak *et al.*, 2010).

The *Aciduliprofundum boonei* T469 strain and the *Aciduliprofundum* sp. MAR08-339 strain were cultivated in Aciduliprofundum medium (2.2.4) at 70 °C as previously described (Reysenbach *et al.*, 2006; Reysenbach and Flores, 2008; Schouten *et al.*, 2008). Cells were harvested as late log phase of growth.

2.3.2 Determination of cell density

For the determination of the cell density of liquid cultures, the optical density (OD) was measured at a wavelength of 600 nm (OD_{600}) with a BioPhotometer (Eppendorf, Hamburg, Germany). For *E. coli*, an OD_{600} of 1.0 corresponds to a cell number of 8×10^8 *E. coli* cells per ml.

2.3.3 Strain maintenance

For long-term storage of *E. coli*, sterile glycerol was added to overnight cultures to a final volume of 33 % (vol / vol). For long-term storage of *S. oneidensis* strains sterile DMSO was added to overnight cultures to a final volume of 10 % (vol / vol). *E. coli* glycerol stocks or *S. oneidensis* DMSO stocks were stored at −70 °C in a screw-cap tube.

2.4 Molecular cloning

2.4.1 Storage of genomic DNA, plasmids, primer and other nucleic acids

All nucleic acids were stored at -20 °C for long-term DNA preservation, as well as for short-term retention.

2.4.2 Plasmid minipreparation

Plasmids and cloning constructs in *E. coli* were isolated according to the protocol “Preparation of Plasmid DNA by Alkaline Lysis with SDS: Minipreparation” described in Molecular Cloning: A Laboratory Manual by Sambrook and Russsell (2001) with minor modifications. The described optional step of the phenol : chloroform extraction was not conducted, where practicable. DNA quantity and integrity was determined by a spectrophotometer (2.4.6) and agarose gel electrophoresis (2.4.7).

2.4.3 Plasmid isolation with a plasmid mini kit

Pure plasmid DNA was desperately needed for sequencing and other applications. In this case, a “HighSpeed Plasmid mini kit” (Avegene life science, Taipei, Taiwan) was used according to the manufacturer's instructions with minor modifications to obtain pure plasmid DNA. The elution step was conducted with 70 °C warm elution buffer, which incubated for more than 1h on the silica membrane. DNA quantity and integrity was determined by spectrophotometry (2.4.6) and agarose gel electrophoresis (2.4.7).

2.4.4 Gel extraction of DNA

For extraction and purification of single DNA bands from agarose gels, the “MinElute Gel Extraction Kit” (Qiagen, Venlo, Netherlands) or the “Gel/PCR DNA Fragments Extraction Kit” (Avegene Life Science, Taipei, Taiwan) was used according to the manufacturer's instructions with minor modifications. The elution step was conducted with 70 °C warm elution buffer, which incubated for more than 1h on the silica membrane. DNA quantity and integrity was determined by spectrophotometry (2.4.6) and agarose gel electrophoresis (2.4.7).

2.4.5 Purification of PCR products and DNA of other applications

When a purification of PCR products and DNA of other applications (e.g. restriction, dephosphorylation etc.) was necessary, the “Gel/PCR DNA Fragments Extraction Kit” (Avegene Life Science, Taipei, Taiwan) was used according to the manufacturer's instructions with minor modifications. The elution step was conducted with 70 °C warm elution buffer, which incubated for more than 1h on the silica membrane. DNA quantity and integrity was determined by spectrophotometry (2.4.6) and agarose gel electrophoresis (2.4.7).

2.4.6 Spectrophotometric determination of DNA concentration and purity

For quantifying the amount of DNA, the absorbance was measured at 260 nm and 280 nm. For the determination of DNA concentration and purity, a Synergy™ HT microplate reader with a Take 3 plate (BioTek, Winooski, VT, USA) was used. A solution with an A_{260} of 1 corresponds to ~50 µg/ml double-stranded DNA (Sambrook and Russell, 2001). The ratio of absorbance at 260 nm and 280 nm (OD_{260} : OD_{280}) is used as a test for contamination of a preparation of DNA with proteins. Pure preparations of DNA have OD_{260} : OD_{280} values of 1.8 and 2.0, respectively (Sambrook and Russell, 2001).

2.4.7 Agarose gel electrophoresis

Agarose gel electrophoresis of DNA was used to separate, identify and purify DNA fragments. This method was conducted according to the protocols of the chapter “Gel Electrophoresis of DNA and Pulsed-field Agarose Gel Electrophoresis” described in “Molecular Cloning: A Laboratory Manual” by Sambrook and Russsell (2001). Agarose gel electrophoresis was applied in a Horizon© 58 gel chamber (Life Technologies, Carlsbad, CA, USA) with the power supply PowerPac Basic (Bio-Rad, Hercules, CA, USA). The location of DNA within the gel was determined directly by staining with ethidium bromide and was detected by direct examination of the gel in UV light with the Molecular Imager GelDoc XR+ (Bio-Rad, Hercules, CA, USA) supported by the Quantity One 1-D-Analysis software(Bio-Rad, Hercules, CA, USA).

2.4.8 Polymerase chain reaction (PCR)

The PCR was used to amplify a specific region of a DNA strand (the DNA target) across several orders of magnitude, generating millions of copies of a particular DNA sequence. PCR products were analyzed using agarose gel electrophoresis (2.4.7) and were purified by gel extraction (2.4.4) or directly purified by silica adsorption (2.4.5).

2.4.8.1 PCR primers

Primers were ordered at Eurofins Genomics (Ebersberg, Germany). If the PCR was performed with Taq polymerase, the melting temperature T_m was calculated with the following equation (Sambrook and Russell, 2001; Kibbe, 2007) and used as annealing temperature:

$T_m = 64.9 + 41 \times (yG + zC - 16.4) / (wA + xT + yG + zC)$

If the PCR was performed with Phusion polymerase, T_m was calculated in consideration to the modified Breslauer's thermodynamics (Breslauer *et al.*, 1986) with the T_m calculator webtool of Thermo Fisher Scientific (http://www.thermoscientificbio.com/webtools/tmc/).

The lower melting temperature T_m of both Primers was used to perform the PCR.

2.4.8.2 PCR conditions

PCRs were performed in different PCR cyclers with heated lid (e.g. C1000 Thermal Cycler, Bio-Rad, Hercules, CA, USA).

The PCR reaction conditions, performed with commercial DNA polymerases, were applied according to the described protocol for GoTaq DNA Polymerase (Promega Corporation, Madison, USA) or Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA).

Table 4: PCR reaction conditions

Initial denaturation	95 °C	2 min	
Denaturation	95 °C	45 sec	32 ×
Annealing	T_m (2.4.8.1)	45 sec	
Elongation	72 °C	1 min/kb	
Final elongation	72 °C	5 min	

Innumerable PCR reactions were performed with a homemade Taq polymerase from *Thermus aquaticus* which was heterologous expressed in *E. coli* and purified by the

working group of Prof. W. Streit. The PCR reaction conditions for this Taq polymerase were described in the protocols of “Molecular Cloning: A Laboratory Manual” by Sambrook and Russsell (2001) and were conducted with minor modifications like illustrated in Table 4.

The PCR reactions contained the following ingredients:

10× amplification buffer	2.5	µl
dNTPs [2 mM each]	2.0	µl
forward primer [10 µM]	1.0	µl
reverse primer [10 µM]	1.0	µl
homemade Taq	1.0	µl
template DNA	1 - 100	ng
H_2O	ad 25	µl

The 10× amplification buffer was prepared as described in “Molecular Cloning: A Laboratory Manual” by Sambrook and Russsell (2001) and the dNTPs were supplied by Thermo Fisher Scientific (Waltham, MA, USA).

2.4.8.3 Direct colony PCR

Direct colony PCR was applied to verify putative positive clones after transformation (2.4.11) or transduction (2.7.6), especially during a sequence based screening of (meta)genomic libraries (2.7.9) or during difficult cloning procedures for seeking the required correct clone. Colony PCRs were mostly conducted in a 96 scale for a high throughput. Therefore, the described standard reaction conditions (2.4.8.2) were used, but the respective colonies were suspended directly in the PCR reaction as template.

2.4.9 Enzymatic modification of DNA

2.4.9.1 Enzymatic digestion of DNA

For site specific digestion of DNA restriction endonucleases of Thermo Fisher Scientific (Waltham, MA, USA) were applied. Two categories of enzymatic digestion were conducted (A) analytical restriction studies and (B) preparative digestions.

(A) Analytical restriction

10fold buffer	1 µl
DNA	0.2 – 1 µg
restriction enzymes	5 U per enzyme
H_2O	ad 10 µl

(B) Preparative digestion

10fold buffer	2.5 µl
DNA	1 – 5 µg
restriction enzymes	20 U per enzyme
H_2O	ad 25 µl

The reaction mixture of both digestion categories varied in the volume, the amount of DNA and the amount of the endonucleases. Analytical restriction studies were incubated for 1 h and preparative digestions were incubated for 2 h at the respective optimal temperature. Both digestions were analyzed by gel electrophoresis (2.4.7) and required DNA bands of the preparative digestion were extracted from the agarose gel (2.4.4).

2.4.9.2 Dephosphorylation of cloning vector DNA

To prevent recircularization during ligation digested cloning vector DNA was dephosphorylated by “FastAP Thermosensitive Alkaline Phosphatase” (Thermo Fisher Scientific, Waltham, MA, USA). Due to the fact that this phosphatase is active in all restriction enzyme buffers, dephosphorylation was conducted simultaneous to the digestion of plasmids. For each 10 U restriction endonuclease in the reaction mixture 2 U phosphatase was added.

2.4.9.3 End-Repair and phosphorylation of DNA fragments for cloning

End-it™ DNA End-Repair Kit (Epicentre Biotechnologies, Madison, WI, USA) was used to fill up restriction enzyme-digested DNA fragments or to phosphorylate PCR products, amplified with DNA polymerases with proof reading. This kit was applied to convert DNA fragments to 5’-phosphorylated, blunt-ended DNA for an efficient blunt-end ligation (2.4.9.4).

2.4.9.4 Ligation of DNA with digested ends or PCR products generated by polymerases with proof-reading

Restriction enzyme generated DNA fragments and PCR products generated by polymerases with proof-reading were purified (2.4.4 or 2.4.5) and ligated into prepared plasmids by using T4 DNA ligase (Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions with minor modifications. The ligation was incubated at 22 °C for 1 h or at 4 °C overnight, heat inactivated at 70 °C for 10 min and transformated into competent *E. coli* cells (2.4.11).

2.4.9.5 Ligation of PCR products (generated by Taq polymerases)

After being purified (2.4.5), PCR products generated by Taq polymerases were ligated into the pGEM-T vector system (Promega Corporation, Madison, USA) by TA cloning according to the manufacturer's instructions.

2.4.10 Preparation of chemically competent *E. coli* cells for heat shock transformation

Chemically competent *E. coli* cells were prepared as described in "The QIAexpressionist™ - A handbook for high-level expression and purification of 6xHis-tagged proteins" (Fifth edition, 2003).

2.4.11 Heat shock transformation of *E. coli*

Competent *E. coli* cells (2.4.10) were transformated by heat shock with recombinant plasmids or fosmids according to the protocol "The Hanahan Method for Preparation and Transformation of Competent E. coli: High-efficiency Transformation" described in "Molecular Cloning: A Laboratory Manual" by Sambrook and Russsell (Sambrook and Russell, 2001) with minor modification. Instead of adding SOB medium to the cells and plating on SOC agar plates with corresponding antibiotic, LB medium and LB agar plates were used.

2.4.12 Conjugation

Conjugation, also called mating, is the transfer of genetic material between bacterial cells by direct cell-to-cell contact or by a bridge-like connection between two cells. Two types of conjugation were applied: Diparental mating and triparental mating.

2.4.12.1 Diparental conjugation

Vector constructs with an origin of transfer (*oriT*) were transferred directly from the donor strain to the acceptor strain, if the donor strains had integrated a *tra* region into the genome (e.g. *E. coli* S17-1 or *E. coli* WM3064). Therefore both strains were cultivated overnight according their optimal growth conditions. 1 ml of overnight cultures, donor and acceptor strain, were centrifugated (5,000 g, 5 min) in the same tube and the supernatant were discarded. The cell pellet was washed with 1 ml LB medium without any antibiotics and was finally resuspended in 50 µl LB medium. The cell suspension was dropped on a LB agar plate and was incubated overnight at the lower optimal growth temperature of both strains, if their growth conditions differ from each other. Cell material was scratched out on selection LB agar plates and incubated over night at optimal growth temperature of the acceptor strain.

2.4.12.2 Triparental conjugation

A triparental conjugation was applied, if the donor strain hadn't integrated a *tra* region into the genome. To conduct the mating, the additional *E. coli* HBH-101 helper strain with the plasmid pRK2013 carrying a *tra* locus necessary to perform the conjugation was required. The triparental mating was performed like the diparental mating (2.4.12.1), but the helper strain *E. coli* HBH-101 (pRK2013) was added to the cell suspension of the acceptor and donor strain.

2.5 Genome mutation

2.5.1 Construction of the *S. oneidensis* Δ*hyaB* deletion mutant

The development of site-directed deletion mutant was possible by allelic exchange via a double cross over event using the suicide vector pNPTS138-R6KT leaving only short N- and C-terminal sections of the target *hyaB* gene, the structural large subunit of the [NiFe]-hydrogenase. The definition of a suicide vector is a plasmid that cannot replicate in a particular host. In the consequence after transferring the vector into the host, this vector has to be integrated into the host chromosomal DNA by homologous recombination to retain e.g. the antibiotic resistance encoded on the suicide vector.

For this purpose, a 518 bp fragment upstream (hyaBup) and a 577 bp fragment downstream (hyaBdw) of the *S. oneidensis* MR-1 large structural subunit of the [NiFe]-hydrogenase (*hyaB*) and a gentamycin resistant cassette (gen^R) from the broad host range plasmid pBBR1MCS-5 (Kovach *et al.*, 1995) were amplified by PCR using the corresponding primer pairs (Table 2) and GoTaq DNA polymerase (Promega Corporation, Madison, USA). The DNA fragments were purified by agarose gel electrophoresis (2.4.4), were ligated into the pGEM-T vector system (2.4.9.5) and were transformated into the host *E. coli* DH5α (2.4.11). The vector constructs pGEM-T::hyaBup, pGEM-T::hyaBdw, pGEM-T::gen^R (Table 1) were isolated by a Plasmid Mini Kit (2.4.3). An insert consisting of the upstream fragment (hyaBup) and downstream fragment (hyaBdw) separated by gen^R was cloned into the multiple cloning site (MCS) of the pGEM-T vector system, resulting in the pGEM-t::ΔhyaB-gen^R vector construct (cloning strategy see Figure 4).

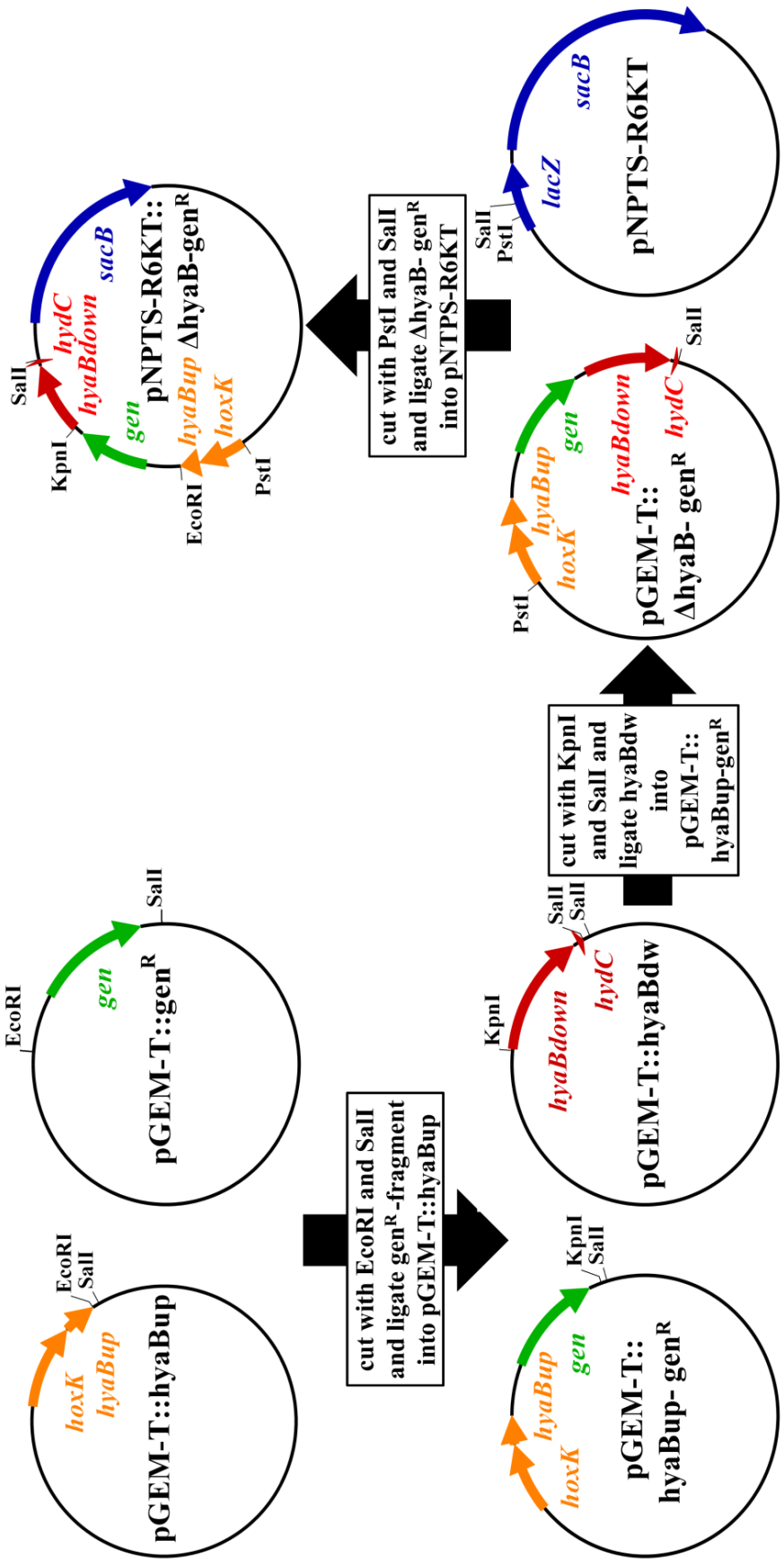


Figure 4: Cloning strategy for the construction of required suicide vector pNTPS-R6KT::ΔhyaB-gen^R. The particular fragments were amplified by PCR (2.4.8) and cloned into the pGEM-T vector (2.4.13). The fragments hyaBup, gen^R, hyaBup-gen^R, hyaBdw, ΔhyaB-gen^R and the vector pNTPS-R6KT were excised by enzymatic digestion (2.4.9), purified by agarose gel extraction (2.4.4) and ligated as described (2.4.9.4).

An *E. coli* DH5α clone with the vector construct pGEM-t::ΔhyaB-gen^R was grown overnight on 5 ml LB containing ampicillin and gentamycin (Table 3), the vector was isolated by a plasmid mini kit (2.4.3) and was digested with KpnI and SalI (2.4.9, Figure 4). The excised ΔhyaB-gen^R fragment was purified by gel extraction (2.4.4) and ligated into the KpnI and SalI restricted (2.4.9; Figure 4) and dephosphorylated (2.4.9.2) suicide vector pNPTS138-R6KT (Lassak *et al.*, 2010).

Chemically competent *E. coli* WM3064 cells (2.4.10) were transformed with the ligation reaction product pNPTS138-R6K::ΔhyaB-gen^R (2.4.11) and was used as donor strain for diparental conjugative mating (2.4.12.1) to introduce the suicide vector construct into the *S. oneidensis* MR-1 genome (Figure 5). The donor strain *E. coli* WM3064 was excluded due to their DAP auxotrophy and *S. oneidensis* crossover integration clones were selected on LB plates containing gentamycin and kanamycin. Both antibiotic resistances were encoded on the transferred and in the genome integrated plasmid pNPTS138-R6K::ΔhyaB-gen^R (Figure 5), whereby the kanamycin resistance was a selection marker needed for a successful first crossover and the gentamycin resistance was an additional marker bracketing by homologous *hyaB* regions which should replace the *hyaB* wild type gene in the aspired mutant. Single kanamycin- and gentamycin resistant colonies were grown overnight in liquid LB medium without any antibiotics and then plated on LB plates containing 10 % sucrose to counter select for double-crossover recombinants (Figure 5). Because the crossover segregation events are rare, as counter-selectable marker a gene for a levansucrase (*sacB*) from *Bacillus subtilis* is encoded on pNPTS138-R6KT, which expression is toxic for gram-negative bacteria when grown in presence of sucrose and therefore provides a direct selection for loss of the plasmid. Under ideal conditions the probability is about 50 % that the second cross over event result in a restored wild type phenotype (loss of the plasmid and the mutant gene) or in the intended deletion mutant (loss of the plasmid and the wild type gene). To eliminate clones with the restored wild type phenotype (Figure 5), obtained single colonies were checked for kanamycin sensitivity and gentamycin resistance. The corresponding colonies were screened for the *hyaB* gene deletion (Δ*hyaB*) by colony PCR (2.4.8.3) using primer bracketing the deleted region (hyaB_cont_F / hyaB_cont_RC; Table 2) and primer targeting the gentamycin resistance cassette (gen_EcoRI_F / gen_KpnI_RC; Table 2). PCR products were checked by agarose gel electrophoresis (2.4.7), whereby products of the right size have indicated an in-frame hydrogenase deletion mutant. Additionally to confirm the respective successful *hyaB* deletion, PCR products were purified (2.4.5) and sequenced (2.6.1).

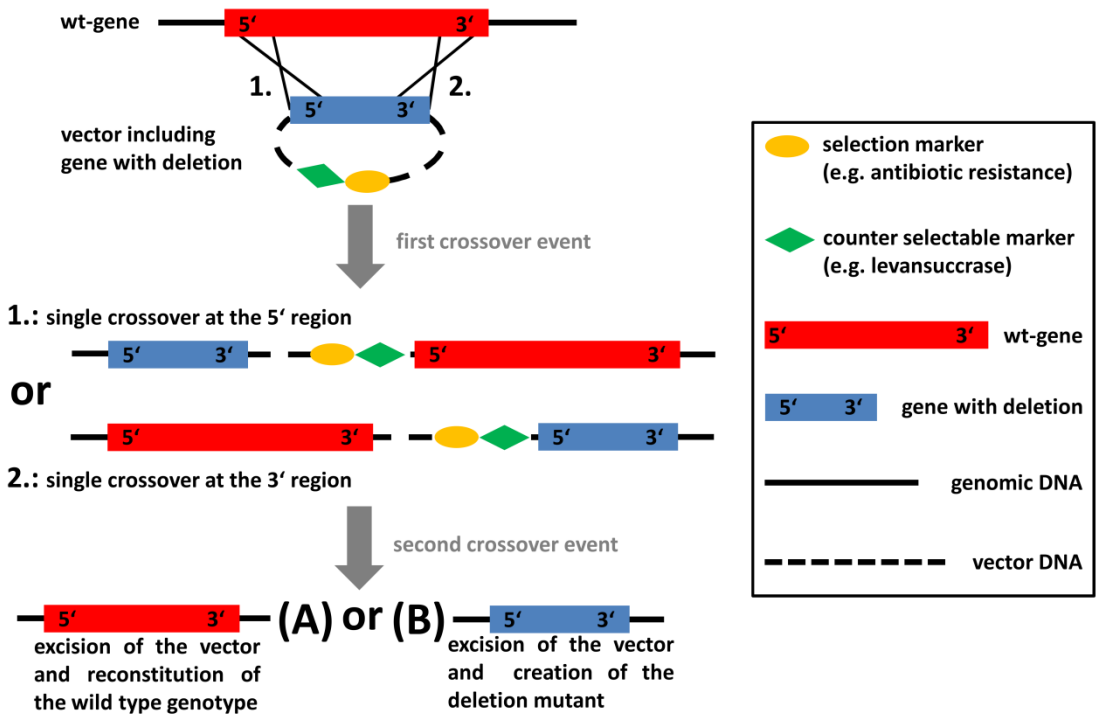


Figure 5: General depiction of the used method to obtain site directed mutant strains by a double crossover event. The suicide plasmid was integrated via a first crossover event into the genome closely to the 5' region (1.) or 3' region (2.) of the target gene. Single crossover transformants were selected by a selection marker (e.g. antibiotic resistance). Because crossover segregation events are rare, a counter selectable marker (e.g. a levansucrase) enabled the selection for a second single crossover event. The second crossover event resulted with a probability of 50 % either into a restored wild type strain (A) or a desired deletion mutant (B). (Figure modified from Navas *et al.*, 2001)

2.6 Sequencing of DNA

2.6.1 ABI sequencing

ABI sequencing was carried out by Eurofins MWG Operon (Ebersberg, Germany) or at the IKMB (Christian Albrechts University, Kiel, Germany). DNA concentrations of purified samples (2.4.5, 2.4.4 or 2.4.3) were adjusted and primers were added according to instruction of the corresponding company or institute.

2.6.2 Next generation sequencing (454 pyrosequencing and Illumina sequencing)

Sequencing of selected fosmids was accomplished with a 454 FLX sequencer system (Roche 454, Branford, USA) and with the Illumina HiSeq2000 technology.

Pyrosequencing was conducted by the group of Prof. S. Schreiber at the IKMB and Illumina sequencing at the GATC Biotech AG (Konstanz, Germany).

2.7 Construction of fosmid libraries

Two (meta)genomic fosmid libraries were constructed. One of *S. oneidensis* MR-1, to recover a fosmid containing *hyaB*, which was used to complement the *S. oneidensis* Δ*hyaB* in order to restore the wild type phenotype. The other was constructed from a hydrothermal vent sample and used for screening to identify H_2 uptake active hydrogenases.

2.7.1 Isolation and preparation of pRS44 for the construction of (meta)genomic libraries

Fosmid libraries were created using the broad-host-range plasmid vector pRS44 (Aakvik *et al.*, 2009). An *E. coli* EPI300-T1 clone with the pRSS44 plasmid was cultivated in 5 ml LB containing kanamycin and chloramphenicol (kanamycin- and chloramphenicol resistance encoded on the pRS44 plasmid) over night. From this preculture, 2.5 ml was used to inoculate 25 ml LB containing kanamycin, chloramphenicol and a corresponding amount Copy Control Induction Solution (Epicentre Biotechnologies, Madison, WI, USA) (Table 3), which switched the copy number of the plasmid from single copy to high-copy by inducing the chromosomally *trfA* gene of the *E. coli* EPI300-T1 host. After incubation for 5 h, the pRS44 vector was isolated from the induced culture using the HiSpeed Plasmid Midi Kit (Qiagen, Venlo, Netherlands) according to the manufacturer's protocol. The pRS44 plasmid was preparative digested with Eco72I (2.4.9) and dephosphorylated (2.4.9.2). The pRS44 fosmid library was constructed according to the manufacturer's protocol for pCC1FOS libraries (Epicentre Biotechnologies, Madison, WI, USA) with minor modifications described in Simon *et al.*, 2010 and Weiland *et al.*, 2010.

2.7.2 Cell material for a genomic library of *Shewanella oneidensis* MR-1 and isolation of genomic DNA

For the genomic *S. oneidensis* MR-1 fosmid library chromosomal DNA of a 200 ml overnight culture was isolated by enzymatic lysis and phenol/chloroform extraction according to an established protocol (Streit *et al.*, 1993). Isolated genomic DNA was

analyzed for purity, concentration and integrity by spectrophotometry and agarose gel electrophoresis (2.4.7 and 2.4.6).

2.7.3 Environmental sample of a hydrothermal vent massive sulfide chimney sample

For the metagenomic library from a hydrothermal vent sample, a massive sulfide chimney of the black smoker Sisters Peak located on the southern Mid-Atlantic Ridge (MAR) (4.80°S 12.37°W) was collected. The chimney structure was sampled by a remote operated vehicle (ROV 6000, GEOMAR, Kiel) during the MAR-SUED V cruise (March/April 2009) with the RV Meteor. After collection the chimney sample (274 ROV 1B) was stored at -80°C until further processing. The chromosomal DNA was extracted from 246 g of the sulfide chimney according to a described protocol by enzymatic lysis and phenol/chloroform extraction (Streit *et al.*, 1993) with the following modifications: The massive sulfide sample of the chimney structure was homogenized using mortar and pestle and suspended in 1 M NaCl, the enzyme concentration was increased for lysozyme (10 mg/ml) and proteinase K (10 mg/ml) and the sample was incubated in buffer containing proteinase K for 16 h at 55°C. To have an adequate amount of material for the construction of a metagenomic library, multiple displacement amplification (MDA) with phi29 DNA polymerase (REPLI-g Kit, Qiagen, Venlo, Netherlands) was conducted according to the manufacturer's instructions. MDA was employed in three parallels with each 2.5 µl (50-100 ng/µl) bulk DNA vent chimney starting material which resulted in 2.5 µg/µl pooled, purified DNA, analyzed by spectrophotometry and agarose gel electrophoresis (2.4.7 and 2.4.6).

2.7.4 End-Repair

Damaged or incompatible ends of purified (meta)genomic DNA was converted by a DNA end-repair kit (Epicentre, Madison, WI, USA) into blunt-ended, 5'-phosohorylated DNA, which was required for a efficient ligation with the prepared pRS44 vector. End-Repair reactions were conducted according to the manufacturer's instructions (End-it™ DNA End-Repair Kit, Epicentre, Madison, WI, USA) and were cleaned-up by dialysis. Therefore, a membrane filter with a pore size of 0.025 µm (EDM Millipore, Billerica, MA, USA) was placed on the surface of a culture dish filled with sterile H_2O. The end-repair reactions were transferred onto the membrane filter and were incubated for 30 minutes at room temperature. The cleaned-up end-repair reactions were stored at -20°C.

2.7.5 Ligation and packaging

The cleaned-up blunt-ended DNA (2.7.4) was ligated into the pRS44 fosmid (Aakvik *et al.*, 2009) using the Fast-Link DNA Ligase and all corresponding reagents of Epicentre (Madison, WI, USA) according to the manufacturer's protocol with minor modifications. The optimal desired molar ratio of isolated insert DNA to pRS44 vector DNA was determined with 1:5. For packaging of pRS44 fosmids with metagenomic insert DNA, MaxPlax Lambda Packaging Extracts (Epicentre, Madison, WI, USA) were used and the packaging procedure was conducted according to manufacturer's protocol.

2.7.6 Preparation of phage competent cells and transduction

A (meta)genomic library, packaged into λ phages, was transduced into *E. coli* EPI300-T^R cells (Table 1). For preparation of phage competent *E. coli* EPI300-T^R cells, LB medium containing additional 10 mM $MgCl_2$ and 1 % maltose was inoculated with 1 % (v/v) of an *E. coli* EPI300-T^R overnight culture. The culture was incubated at 37°C until an OD_{600} 0.8 - 1.0 is reached (2.3.2). The cells were stored at 4°C and could be used for 72 h.

Before spreading the metagenomic library, the titer of the phage particles was determined by transduction of competent *E. coli* EPI300-T^R cells with different volumes of phage packaging reactions. A defined volume of phage particles (2.5 µl, 5 µl, 10 µl, 15 µl and 20 µl) was added in triplicates to 100 µl phage competent *E. coli* EPI300-T^R cells, mixed and the mixture was incubated at 37°C for 1 h. Phage infected *E. coli* EPI300-T^R cells were spread on LB agar plates with 12.5 µg/ml chloramphenicol and were incubated at 37°C overnight. On the bases of counted colonies per plate, the titer of different volumes of the packaged phage particles was calculated and a titer leading to 100-300 colonies per plate was chosen for the remained phage packaging reactions. Further transductions with the corresponding volume of packaging reaction for 100-300 colonies was conducted as described above. The resulting LB chloramphenicol plates with infected *E. coli* EPI300-T^R colonies were stored at 4 °C.

2.7.7 Storage of (meta)genomic libraries

The infected *E. coli* EPI300-T^R colonies were picked into microtiter plates, containing 150 µl LB medium with 12.5 µg/ml chloramphenicol and incubated at 37°C overnight. Each well will be mixed with glycerol to a final concentration of 35 %. The plates were copied into new prepared microtiter plates, were also incubated overnight at 37°C and

glycerol to a final concentration of 35 % was added into each well. All fosmid libraries and their copies were stored at -70°C.

2.7.8 Checking insert size and quality

The quality of all (meta)genomic libraries were checked by restriction analysis and by sequencing. Therefore, random clones were cultivated in 5 ml LB containing 12.5 µg/ml chloramphenicol and a corresponding amount autoinduction solution (Epicentre Biotechnologies, Madison, WI, USA) (Table 3) under shacking overnight at 37 °C. Cells were harvested by centrifugation (16.000 g, 1 min, 4 °C) and fosmids were isolated with a kit (2.4.3). Isolated fosmids were analytical restricted (2.4.9) with three different restriction enzymes: BamHI, EcoRI and HindIII. The restriction patterns were analyzed by agarose gel electrophoresis (2.4.7) to calculate the insert size of the fosmids. Finally, the isolated fosmids were sequenced by ABI sequencing (2.6.1) using vector primers to confirm the correct origin of the insert by high sequence similarities to estimated organisms.

2.7.9 PCR based screening of the genomic *S. oneidensis* MR-1 library for the *hyaB* genes

A genomic fosmid library with 960 clones of the *S. oneidensis* MR-1 wt genome was screened via PCR for the *hyaB* gene, which was used for complementation experiments with the constructed *S. oneidensis* Δ*hyaB* deletion mutant. For this purpose the fosmid library was copied into new microtiter plates (2.7.7) and crude extracts of fosmid pools consisting of 96 clones were prepared. Therefore, 100 µl of each clone was transferred into a falcon tube, cells were harvested by centrifugation (5,000 g, 10 min), washed with 5 ml Tris (10 mM, pH 8.0) and harvested again (5,000 g, 10 min) to elute the cells finally in 1 ml Tris (10 mM, pH 8.0). The pools were boiled for 10 min at 97 °C, cell debris was removed by centrifugation (16,000 g, 5 min) and the supernatant was used as template for PCR (2.4.8.3) using the primers hyaB_cont_F and hyaB_cont_RC (Table 2). To circumscribe putative fosmids carrying the *hyaB* gene from the positive tested 96 clone pools, the described PCRs were repeated for pooled material from rows and from columns of the microtiter plates to locate positive hits among 96 clone pools. The putative fosmid clones with the *hyaB* gene were isolated from the *E coli* EPI300-T1 host (2.4.3) and were investigated by sequencing the fosmid ends (2.6.1) to verify that the identified fosmids contained the *hyaB* gene using the primers modT7_prom and pCC1FOS_rev.

2.7.10 Transfer of metagenomic libraries from *E. coli* to *S. oneidensis* Δ*hyaB*

The fosmid library derived from the black smoker Sisters Peak was transferred by triparental conjugation from *E. coli* EPI300-T^R to the alternative host *S. oneidensis* Δ*hyaB*, like described (2.4.12.2) with minor modifications resulting from the simultaneously conjugation of 96 plasmids. To conduct the triparental conjugation, the donor-, acceptor- and helper-strains were cultivated and prepared as follows:

(1) The *E. coli* EPI300-T^R donor working cultures of the fosmid library were cultivated in 96-deep-well plates with 1.5 ml LB medium containing 12.5 µg/ml chloramphenicol and incubated at 37°C under shaking until an OD_{600} of 0.8-1.0 (2.3.2) was reached. The donor cells were harvested directly in the 96-deep-well plates (centrifugation: 2,250 g, 30 min, 4° C) and stored on ice until needed.

(2) For each 96-deep-well donor plate one 50 ml culture of the acceptor *S. oneidensis* Δ*hyaB* was prepared. The day before performing conjugation, *S. oneidensis* Δ*hyaB* cultures were inoculated in 50 ml LB medium containing 10 µg/ml gentamycin. The acceptor cultures were incubated at 28°C under shaking overnight and stored on ice until needed.

(3) For the triparental conjugation, the helper strain *E. coli* HBH101 (pRK2013) was essential and for each 96-deep-well donor plate one flask with 50 ml LB medium containing 30 µg/ml kanamycin was inoculated. The helper strain cultures were incubated at 37°C under shaking until an OD_{600} of 0.8-1.0 (2.3.2) was reached and stored on ice until needed.

To perform the triparental mating, 0.5 ml of the stored acceptor *S. oneidensis* Δ*hyaB* and the conjugation helper strain *E. coli* HBH101 (pRK2013) were added to each well of the 96-deep-well containing the *E. coli* EPI300-T^R donor strain. The cells were harvested directly in the 96-deep-well plates (2,250 g, 30 min, 4 °C) and pellets were resuspended in 20 µl LB medium. Cells were transferred by an 8 channel pipette onto LB agar plates as single, distinct dots and incubated overnight at 28°C. New microtiter plates filled with 150 µl LB medium containing 12.5 µg/ml chloramphenicol and 10 µg/ml gentamycin were prepared and inoculated with cells from the triparental conjugation to select for pRS44 fosmids in the new host *S. oneidensis* Δ*hyaB*. After incubation at 28°C for one day, the transferred fosmid library in the host *S. oneidensis ΔhyaB* was used to conduct the functional screening for [NiFe]-hydrogenases. Each well of the microtiter plates was mixed with dimethyl sulfoxide (DMSO) to a final concentration of 10 % for storage at -70°C.

2.8 Function-based screening for seeking H_2 uptake active enzymes

2.8.1 Complementation experiments

The hydrogenase deletion mutant *S. oneidensis* Δ*hyaB* was complemented with the fosmid pRS44::hyaB. The fosmid was transferred into *S. oneidensis* Δ*hyaB* by triparental mating (2.4.12.2). The *S. oneidensis* Δ*hya*B carrying the plasmid pRS44::hyaB was tested for restored H_2 uptake activity. Additionally, the *S. oneidensis* MR-1 wt was tested as positive control and the *S. oneidensis* Δ*hyaB* deletion mutant was tested as negative control for H_2 uptake activity. Therefore, pre-cultures of all three strains grown in 5 ml LB containing 12.5 µg/ml chloramphenicol and 10 µg/ml gentamycin were harvested (16,000 g, 5 min) and washed with FW medium. FW medium supplemented with Fe(III)citrate and H_2 as sole energy source in the headspace was inoculated with the washed cells of the *S. oneidensis* MR-1 wt and *S. oneidensis* Δ*hyaB* complemented with pRS44::hyaB and cultivated under anaerobic conditions at 28 °C. For the cultivation of the *S. oneidensis* Δ*hyaB* deletion mutant, the FW medium was additional supplemented with 10 mM pyruvate and 10 mM fumarate, because this strain can't utilize H_2 as energy source.

Additionally, the *S. oneidensis* Δ*hyaB* clone complemented with pRS44::hyaB was cultivated together with 95 other *S. oneidensis* Δ*hya*B clones (not exhibiting H_2 uptake activity) in a corresponding 96-deep-well plate with 1.5 ml LB with the corresponding antibiotics, pooled harvested by centrifugation and in one culture bottle to illustrate that on a 1:96 scale the complemented *S. oneidensis* Δ*hyaB* clone was detectable by a color change in the medium.

2.8.2 Detection of H_2 uptake active enzymes in metagenomic libraries by complementation of *S. oneidensis* Δ*hyaB* and H_2 uptake activity in the *S. oneidensis* wt strain

The function-based screening was performed with the metagenomic pRS44 library from the massive sulfide chimney sample of the black smoker Sisters Peak in *S. oneidensis* Δ*hyaB*. Therefore, serum bottles with rubber stoppers filled with 50 ml FW medium containing Fe(III)citrate and H_2/CO_2 (80 % / 20 % v/v) in the headspace were prepared. Furthermore, the 96 well microtiter plates of the metagenomic pRS44 library in *S. oneidensis* Δ*hyaB* were stamped on LB plates containing gentamycin (10 µg/ml) and

chloramphenicol (12.5 µg/ml) and were incubated overnight at 28°C, whereby a half microtiter plate with 48 clones were stamped on one LB agar plate. The grown colonies of one microtiter plate were swept off the plates with FW medium to obtain a clone pool of 96 *S. oneidensis* Δ*hyaB* clones containing fosmids with metagenomic DNA. The clone pools were harvested by centrifugation (16,000 g, 5 min), washed with FW medium and afterwards used to inoculate the prepared serum bottles (100 µl inoculum). After an incubation time of 6-21 days at 28°C, recombinant hydrogenase activity should be detectable by color change of the medium from yellow (Fe(III)citrate) to colorless (Fe(II)citrate).

To locate the *S. oneidensis* Δ*hyaB* clones with an active hydrogenase from positive tested 96-clone-pools, the described function-based screening was repeated for *S. oneidensis* Δ*hyaB* fosmids pools of 8 clones. Therefore, a clone pool for each microtiter column was prepared and inoculated in serum bottles (100 µl inoculum) like described previously, to detect recombinant hydrogenase activity. If clones of a column tested positive, the described function-based screen was repeated for each clone of these columns.

2.8.3 Analysis of *S. oneidensis*- and *Acidulipofundum* strains for existing H_2 uptake activity capabilities

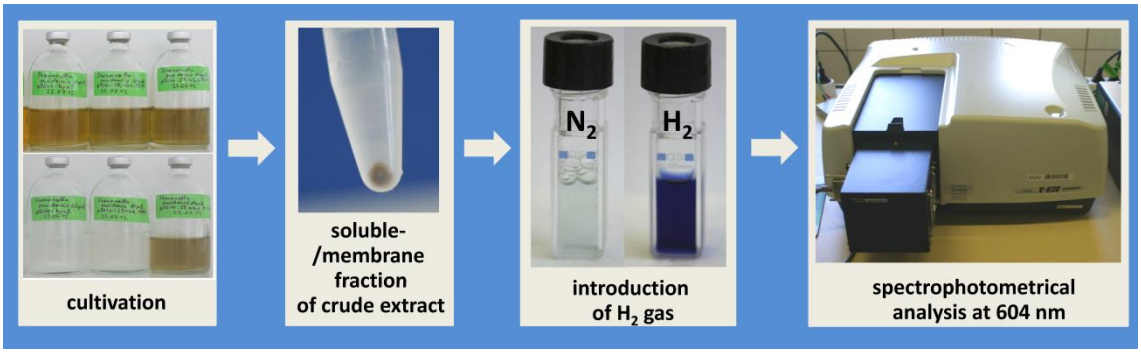


Figure 6: Schematic workflow illustration for implementation of a partial purification of a metagenomic hydrogenase followed by spectrophotometric hydrogenase activity assay. *S oneidensis* Δ*hyaB* mutants complemented with metagenomic fosmids were cultivated in FW medium with H_2 as sole energy source (2.8.2). Hydrogenases of the recovered H_2 uptake active clones were purified partially by preparing soluble and membrane fractions of crude extracts (2.8.3.1). The hydrogenase activity assay based on the reduction of the artificial electron donor methyl viologen (from colorless to blue) via hydrogenase activity and can be detected spectrophotometrically at 604 nm after introduction of H_2 (2.8.4).

Detected metagenomic H_2 uptake active enzymes were analyzed directly in the host *S. oneidensis* Δ*hyaB*. Therefore, a partial purification of the metagenomic recombinant hydrogenases was performed (for details see below), spectrophotometric hydrogenase

activity assays were conducted to determine activity and biochemical properties (Figure 6). Additionally, other *S. oneidensis*- and *Aciduliprofundum* strains were also investigated.

2.8.3.1 Preparation of crude cell extracts and partial purification of recombinant H_2 uptake active enzymes

S. oneidensis MR-1 wt, *S. oneidensis* Δ*hyaB*, *S. oneidensis* Δ*hyaB* complemented with pRS44::hyaB, and *S. oneidensis* Δ*hyaB* clones containing a metagenomic gene encoding enzyme, which can couple the H_2 oxidation to the reduction of Fe(III)citrate (e.g. pRS44::SP-P05F01 and pRS44::SP-P18F11), were cultivated anaerobically on a large scale. A total culture volume of 400 ml to 1,000 ml were grown in several serum bottles containing 50 ml FW medium under a H_2/CO_2 atmosphere (80 % / 20 % v/v) at 28°C for up to 10 days. Cells were harvested by centrifugation (16,000 g, 4°C, 30 min) and washed with 20 mM sodium phosphate buffer (pH 7.0) containing 1 mM dithithreitol (DTT).

The partial purification of the hydrogenases of all H_2 uptake active clones were performed as previously described (Maroti *et al.*, 2009) with some modifications. Cells were disrupted by sonication with the Sonicator UP50H using sonotrode MS2 (Hieschler, Germany) under anaerobic conditions on ice (five times for 30 seconds, 70 % amplitude, cycle 0.5). Cell debris was removed by centrifugation (10,000 g, 5 min) and the supernatant was ultra-centrifugated (100,000 g, 4°C, 1h) to separate the membrane- from the soluble fraction. The pellet, containing the membrane fraction, was resuspended in 20 mM sodium phosphate buffer (pH 7.0) containing 1 mM DTT. The supernatant of the ultracentrifugation step was saved as the soluble fraction. Both fractions were tested for H_2 uptake activity. The protein concentrations of all samples were estimated according to Bradford (Bradford, 1976).

2.8.3.2 Preparation of crude cell extracts and partial purification of Aciduliprofundum

A. boonei T469 and *Aciduliprofundum* sp. MAR08-339 were cultivated anaerobically in a total culture volume of 400 ml in several serum bottles containing 100 ml Aciduliprofundum medium under a H_2/CO_2 atmosphere (80 % / 20 % v/v) at 70 °C for 3 days. Cells were harvested under anaerobic conditions by centrifugation (16,000 g, 4°C, 30 min) and washed with 20 mM sodium phosphate buffer (pH 7.0) containing 1 mM dithithreitol (DTT).

The partial purification of the hydrogenases of all H_2 uptake active clones were performed as described (2.8.3.1). The generated membrane- and soluble fractions were tested for H_2 uptake activity. The protein concentrations of all samples were estimated according to Bradford (Bradford, 1976).

2.8.3.3 Determination of the extinction coefficient of methyl viologen

For the calculation of the H_2 uptake activity, the extinction coefficient of methyl viologen was estimated for the applied temperatures 25 °C, 55 °C and 70 °C. Therefore, the extinction at 604 nm for different concentrations of fully reduced metyl viologen was measured spectrophotometrically at least threefold. Anoxic reaction mixtures with a final volume of 950 µl were prepared in quartz cuvettes containing 20 mM sodium phosphate buffer (pH 7.0), 1 mM DTT and different concentrations of methyl viologen and were sealed with rubber stoppers. The cuvettes were flushed with N_2 for 5 minutes, the methyl viologen were reduced completely by adding 50 µl sodium dithionite solution (50 mg/ml) and the extinction at 602 nm was measured. The methyl viologen extinction coefficient for the applied temperatures 25 °C, 55 °C and 70 °C was calculated via the Beer-Lambert law ($A=\varepsilon \cdot l \cdot c$). The spectrophotometric results were evaluated by linear regression to determine the linear correlation of extinction $_{604nm}$ and methyl viologen concentration.

2.8.3.4 Hydrogenase uptake activity assay

The hydrogenase activity was determined by a H_2 consumption assay, using oxidized methyl viologen as an artificial electron acceptor and H_2 as the electron donor. This assay was performed as previously described (Takai *et al.*, 2005; Guiral *et al.*, 2006; Maroti *et al.*, 2009) with some modifications under strictly anoxic conditions. Anoxic reaction mixtures containing 20 mM sodium phosphate buffer (pH 7.0), 1 mM DTT and 20 mM oxidized methyl viologen (MV) were placed into quartz cuvettes and sealed with rubber stoppers. The cuvettes were flushed with N_2 for 5 minutes and then traces of 5 mM sodium dithionite solution were added until a stable bluish color could be obtained that indicated anoxic conditions. The prepared membrane and soluble fractions were added to the reaction mixtures. The reactions were started by flushing the cuvettes for 20 seconds with 100 % H_2. The reduction of the colorless form of oxidized MV^{2+} to the reduced deep blue form MV^{+} by an active H_2 uptake active enzymes were followed spectrophotometrically at 604 nm. The H_2 uptake activity of the samples was determined at three different incubation temperatures (25 °C, 55 °C and 70 °C).

2.8.4 H_2 consumption measurements

For the identified *S. oneidensis* strains with H_2 uptake active enzymes encoded on the fosmids and for the *S. oneidensis* Δ*hyaB* hydrogenase deletion mutant, the H_2 consumption rates were determined in triplicates with a gas chromatograph (Thermo Scientific, Trace GC Ultra, Waltham, MA, USA). All *S. oneidensis* strains were cultivated in serum bottles with anoxic FW medium supplemented with Fe(III)citrate and approximately 10 % (v/v) H_2 as sole energy source in the headspace. The *S. oneidensis* cultures were incubated at 28°C and the consumption of H_2 were monitored against the time by sampling and analyzing the headspace. After a nearly complete H_2 consumption, approximately 10 % (v/v) H_2 was added to the headspace of the concerned *S. oneidensis* cultures and the monitoring were continued to confirm the H_2 observed consumption.

2.8.5 H_2 evolution assay

For the H_2 evolution assays prepared membrane- or soluble fractions of *S. oneidensis* stains or other organisms were needed and prepared as described above (2.8.3.1 and 2.8.3.2). The assays were performed using reduced methyl viologen (MV) as artificial electron carrier and sodium dithionite as electron donor as previously described (Maroti *et al.*, 2009) with minor modifications. In Hungate tubes sealed with rubber stoppers, anoxic reaction mixtures with a final volume of 2 ml were prepared containing 20 mM sodium phosphate buffer (pH 7.0), 1 mM DTT, MV (20 mM) and 150 µl of the prepared membrane- or soluble fractions of *S. oneidensis* stains. After flushing the Hungate tubes with N_2 for 5 minutes, the reactions were started by adding 100 µl sodium dithionite (50 mg/ml). The reaction mixtures were incubated at 28°C for 1h and were stopped by adding 200 µl trichloroacetic acid (40 % w/v). Gas samples of the headspace were analyzed by a gas chromatograph (Thermo Scientific, Trace GC Ultra, Waltham, MA, USA) to determine the amount of evolved H_2.

3 Results

3.1 Development of a function-based screening method for seeking H_2 uptake active enzymes

A function-based screening approach for metagenomic libraries was developed to investigate metagenomic material from hydrothermal deep-sea vent habitats conserved in broad-host range fosmid libraries to identify and study novel, preferably oxygen tolerant hydrogenase enzymes. The screening was realized by using the deletion mutant *S. oneidensis* Δ*hyaB* as surrogate host for constructed metagenomic libraries due to the fact that und anaerobic conditions the activity of the [NiFe]-hydrogenase (Hya) of *S. oneidensis* MR-1 enables the utilization of oxidized metals as electron acceptor and thus *S. oneidensis* MR-1 is categorized as “Dissimilatory Metal Reducing Bacteria” (DMRB). In the *S. oneidensis* Δ*hyaB* deletion mutant, heterologous expressed and functional H_2 uptake hydrogenases can restore the lost ability to reduce metals like Fe(III) under anaerobic conditions via coupling to H_2 oxidation.

Therefore, a site-directed deletion mutant of *S. oneidensis* MR-1, targeting the large structural subunit of the sole [NiFe]-hydrogenase, was constructed via a double crossover event using the suicide vector pNPTS138-R6KT (3.1.1). Subsequently, metal reducing capabilities coupled the H_2 oxidation of the *S. oneidensis* MR-1 wt and the absence thereof in the constructed *S. oneidensis* Δ*hyaB* mutant were confirmed (3.1.3). Afterward, the capability of complementing the *S. oneidensis* Δ*hyaB* mutant with the fosmid encoded hydrogenase operon of *S. oneidensis* MR-1 wt were investigated to verify the potential to use the deletion mutant in the function-based screening procedure to detect heterologous expressed hydrogenases (3.1.5).

3.1.1 Construction of the *S. oneisensis* Δ*hyaB* deletion mutant

The [NiFe]-hydrogenase deletion mutant of *S. oneidensis* MR-1 (*S. oneidensis* Δ*hyaB*) was constructed by site directed mutagenesis via a double crossover event using the suicide vector pNPTS138-R6KT leaving only short N- and C-terminal sections of the target *hyaB* gene (2.5.1) resulting in an expected *S. oneidensis* phenotype without H_2 uptake activity. *S. oneidensis* clones with a putative successful site directed allele exchange were checked for the *hyaB* gene deletion (Δ*hyaB*) by colony PCR (2.4.8.3).

An amplicon with the predicted product size of 2,466 bp (line 2; Figure 7) using the primer hyaB_cont_F / hyaB_cont_RC (Table 2) could be observed for the *S. oneidensis* MR-1 wt template as result of the PCR bracketing the hyaB gene. In contrast the PCR using the primer hyaB_cont_F / hyaB_cont_RC (Table 2) and bracketing the hyaB gene of putative S. oneidensis mutant clone resulted in the expected product size of 2,283 bp (line 1; Figure 7). Additionally, for the PCR targeting the inserted gentamycin resistance cassette using the primer gen_EcoRI_F / gen_KpnI_RC (Table 2), no product could be observed by using a *S. oneidensis* MR-1 wt clone (line 5; Figure 7) in contrast to the putative *S. oneidensis* mutant clone where a product with the expected size of 874 bp could be observed (line 4 Figure 7).

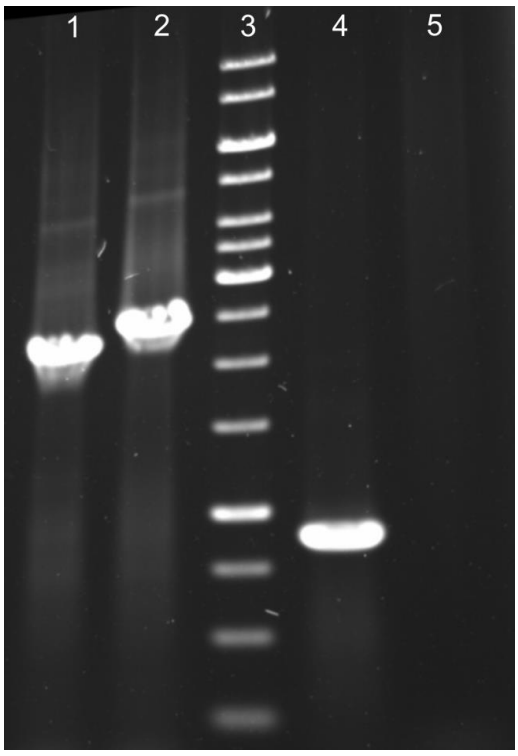


Figure 7: Colony PCR with cell material of the *S. oneidensis* MR-1 wt and the putative *S. oneidensis* Δ*hyaB* using primer bracketing the target *hyaB* region and targeting the gentamycin resistance cassette. (1) *hyaB* PCR product using a putative *S. oneidensis* Δ*hyaB* mutant as template (2) *hyaB* PCR product using *S. oneidensis* MR-1 wt as template (4) PCR product targeting the inserted gentamycin resistance cassette using a putative *S. oneidensis* Δ*hyaB* mutant as template (5) PCR product targeting the gentamycin resistance cassette using *S. oneidensis* MR-1 wt as template (3) GeneRuler™ 1 kb DNA Ladder (Thermo Fisher Scientific, Waltham, MA, USA)

Furthermore, the PCR amplicon bracketing the *hyaB* gene of the putative *S. oneidensis* Δ*hyaB* mutant was sequenced (2.6.1) to confirm the correct replacement of an indispensable part of the hydrogenase with the gentamycin resistance cassette. As expected the sequence analysis confirmed, that the putative *S. oneidensis* Δ*hyaB* mutant contained a gentamycin resistance cassette (gen^R) of 867 bp, which has replaced an essential 1,050 bp section of the *hyaB* gene (see Figure 21 in appendix 9.1).

3.1.2 Seeking a fosmid with the whole *S. oneidensis* [NiFe]-hydrogenase operon

A genomic fosmid library with 960 clones of the *S. oneidensis* MR-1 wt genome was generated (2.7) and the resulting clones were screened by PCR for the *hyaB* gene (2.7.9). Three fosmid clones with an insert carrying the *hyaB* gene were found and the fosmid ends were sequenced (2.6.1) using vector primer (pCC1FOS_rev / modT7_prom; Table2). Thereby, the pRS44 fosmid clone pRS44::hyaB (pRS44::Sowt-P05H02) containing a genomic fragment of ~35 kbp which included the whole *S. oneidensis* MR-1 [NiFe]-hydrogenase operon encoding structural-, maturation- and assembly proteins was found (Figure 8) and was used for complementation experiments. The fosmid pRS44::hyaB was used for complementation experiments to prove the capability to restore the wild type phenotype of the constructed *S. oneidensis* Δ*hyaB* deletion mutant.

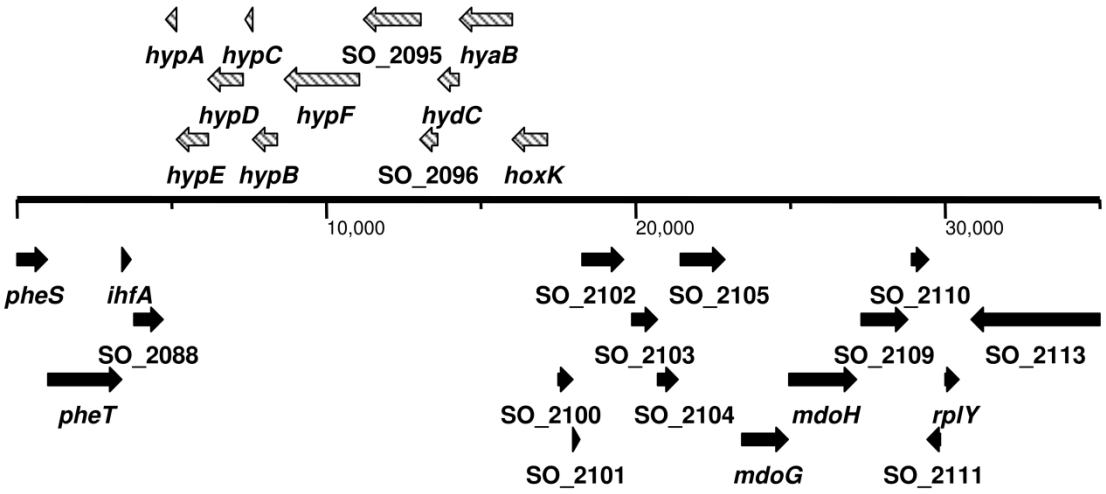


Figure 8: ORFs arrangement on the insert of the fosmid pRS44::hyaB. The fosmid contains a genome fragment of ~35 kbp which includes the whole [NiFe]-hydrogenase operon of *S. oneidensis* MR-1. All ORFs of the [NiFe]-hydrogenase operon are illustrated as hatched white arrows in the upper part of the figure, in contrast to the other ORFs which are exemplified as black arrows in the lower part of the figure. All ORF designations correspond to the annotation of the *S. oneidensis* MR-1 genome (NCBI Reference Sequence: NC_004347).

3.1.3 Validation of the function-based screening method for seeking metagenomic H_2 uptake active enzymes by complementation experiments

The developed functional screen for detecting H_2 uptake activity based on the ability of *S. oneidensis* MR-1 to couple H_2 oxidation to Fe(III) reduction and was implemented by

a colorimetric detection of complemented *S. oneidensis* Δ*hyaB* [NiFe]-hydrogenase deletion mutant with H_2 uptake active enzymes encoded on metagenomic fosmids.

For validation of a function-based screening method, the *S. oneidensis* MR-1 wt strain (So wt) and the *S. oneidensis* Δ*hyaB* deletion mutant (So Δ*hyaB*) were cultivated under anaerobic conditions in FW medium supplemented with Fe(III)citrate and H_2 as sole energy source (2.8.1). The So wt with an active H_2 uptake hydrogenase could oxidize and use H_2 as electron donor and electrons were transferred to Fe(III)citrate which then was reduced to Fe(II)citrate. Activity was detected by a color change of the medium from yellow (Fe(III)citrate) to colorless (Fe(II)citrate) (Figure 9 A, B). In contrast, So Δ*hyaB* could not oxidize H_2 because of the deficient [NiFe]-hydrogenase and thereby no electrons for the Fe(III)citrate reduction were available. Therefore, the So Δ*hyaB* cultures did not show a color change and remained yellow (Figure 9 C, D). Additionally, this result suggested that the second putative hydrogenase of *S. oneidensis* MR-1, a [FeFe]-hydrogenase, did not have any influences on the metal reducing capabilities and therefore the [FeFe]-hydrogenase did not seem to have any effects on the developed function-based screening method.

In theory, complementation of *S. oneidensis* Δ*hyaB* with a H_2 uptake activity encoding fosmid should restore the original phenotype and the color of the medium should change to colorless. To ensure that a complemented *S. oneidensis* Δ*hyaB* deletion mutant showed this rescued phenotype, the *S. oneidensis* Δ*hyaB* deletion mutant was complemented with the fosmid pRS44::hyaB (So pRS44::hyaB), which contained the whole [NiFe]-hydrogenase operon with structural-, maturation- and assembly proteins of *S. oneidensis* MR-1 (3.1.2) by triparental mating (2.8.1)

This So pRS44::hyaB was tested for restored H_2 uptake activity by cultivating under anaerobic conditions in FW medium supplemented with Fe(III)citrate and H_2 as sole energy source in the headspace (2.8.1). As expected, a restored H_2 uptake activity could be observed by the culture medium's color change indicating the use of H_2 for the reduction of Fe(III)citrate to Fe(II)citrate (Figure 9 E, F).

Additionally, a single So pRS44::hyaB was cultivated together with 96 *S. oneidensis* Δ*hya*B clones (not exhibiting H_2 uptake activity) in one culture bottle to illustrate that on a 1:96 scale the complemented *S. oneidensis* Δ*hyaB* clone was detectable by a color change in the medium (2.8.1). Here a H_2 uptake activity was also detectable by a color change of the medium.

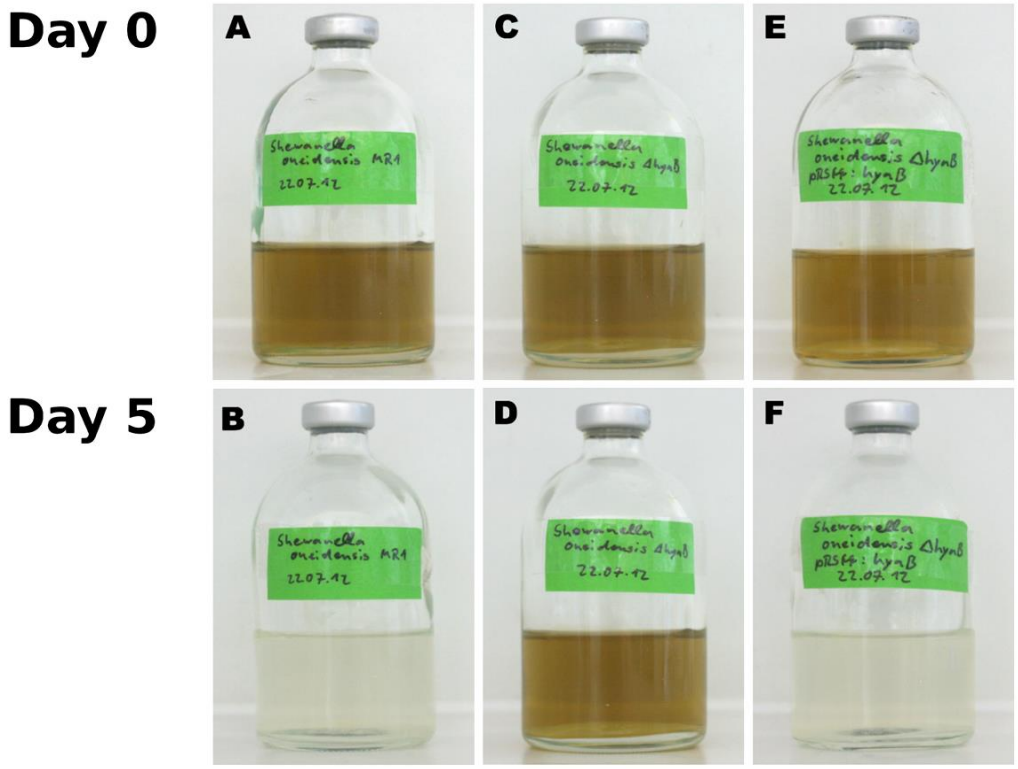


Figure 9: Complementation experiments with the *S. oneidensis* ΔhyaB strain complemented with the fosmid pRS44::hyaB.: *S. oneidensis* MR-1 wt (So wt) immediately after inoculation (A) and after 5 days incubation (B), *S. oneidensis* ΔhyaB (So Δ*hyaB*) immediately after inoculation (C) and after 5 days incubation (D), and *S. oneidensis* Δ*hyaB* complemented with pRS44::hyaB (So pRS44::hyaB) immediately after inoculation (E) and after 5 days incubation (F). All samples were cultivated in FW medium supplemented with Fe(III)citrate and H_2/CO_2 (80 %/20 %, v/v) and incubated at 28°C for several days. A color change from yellow to colorless could be observed if active H_2 uptake active enzymes were present and catalyzed H_2 oxidation whereby the yielded electrons were transferred to Fe(III)citrate (yellow) which were reduced to Fe(II)citrate (colorless).

3.1.4 Estimation of the extinction coefficient of methyl viologen for the calculation of H_2 uptake activities

Besides the color change in the medium indicating H_2 uptake activity, the H_2 uptake enzyme activity of the *S. oneidensis* MR-1 wt (So wt), the *S. oneidensis* Δ*hyaB* mutant (So Δ*hyaB*) and the *S. oneidensis* ΔhyaB mutant complemented with the fosmid pRS44::hyaB (So pRS44::hyaB) was measured (2.8.3.4). The measurements were based on the reduction of methyl viologen and were conducted as described before (Takai *et al.*, 2005; Guiral *et al.*, 2006; Maroti *et al.*, 2009).

For the calculation of the H_2 uptake activity, it was necessary to estimate the extinction coefficient of methyl viologen. As different measurement conditions (wavelength, temperature, pH, solvent) result into different extinction coefficients, the extinction of different methyl viologen concentrations was estimated at constant conditions in 20 mM

sodium phosphate buffer (pH 7.0) with 1 mM DTT at a wavelength of 602 nm with the exception of the temperature (2.8.3.4). The linear correlation of extinction at 602 nm and methyl viologen concentration was determined for different temperatures and the extinction coefficient was estimated by linear regression with 5,199 L mol^{-1} cm^{-1} for 25 °C, 9,358 L mol^{-1} cm^{-1} for 55 °C and 10,848 L mol^{-1} cm^{-1} for 70 °C (see Figure 22 in appendix 9.2).

3.1.5 Evaluation of the function-based screening method by H_2 uptake activity-, H_2 consumption- and H_2 evolution measurements of accomplished complementation experiments

For H_2 uptake activity *in vitro* measurements, cells of the *S. oneidensis* MR-1 wt (So wt), the *S. oneidensis* Δ*hyaB* mutant (So Δ*hyaB*) and the *S. oneidensis* ΔhyaB mutant complemented with the fosmid pRS44::hyaB (So pRS44::hyaB) were cultivated (2.8.1). Crude cell extracts were prepared to gain their soluble- and membrane protein fractions as described (2.8.3.1). In the membrane fraction of So wt a H_2 uptake enzyme activity of 238.0 ± 44.6 nmol * min^{-1} * mg^{-1} total protein was measured at 25 °C (Figure 10). The H_2 uptake activity in the membrane fraction of So Δ*hyaB* was negligible (25°C: 2.7 ± 2.2 nmol * min^{-1} * mg^{-1} total protein; 55°C: 9.6 ± 5.6 nmol * min^{-1} * mg^{-1} total protein; 70°C: 16.0 ± 2.3 nmol * min^{-1} * mg^{-1} total protein).). Additionally, this result suggested that the second putative hydrogenase of *S. oneidensis* MR-1, a [FeFe]-hydrogenase, did not seem to have a H_2 uptake activity. In contrast, for So pRS44::hyaB a H_2 uptake activities of 259.7 ± 44.5 nmol * min^{-1} * mg^{-1} total protein at 25°C, 1362.6 ± 352.7 nmol * min^{-1} * mg^{-1} total protein at 55°C and 3239.1 ± 1770.0 nmol * min^{-1} * mg^{-1} total protein at 70°C were measured in the membrane fraction. For the 25 °C measurements the activity was in the same range as for the So wt activity (Figure 10). In all cases, no significant hydrogenase activity (P value <0.001) was detected in the tested soluble fractions in comparison to the corresponding membrane fractions (data not shown).

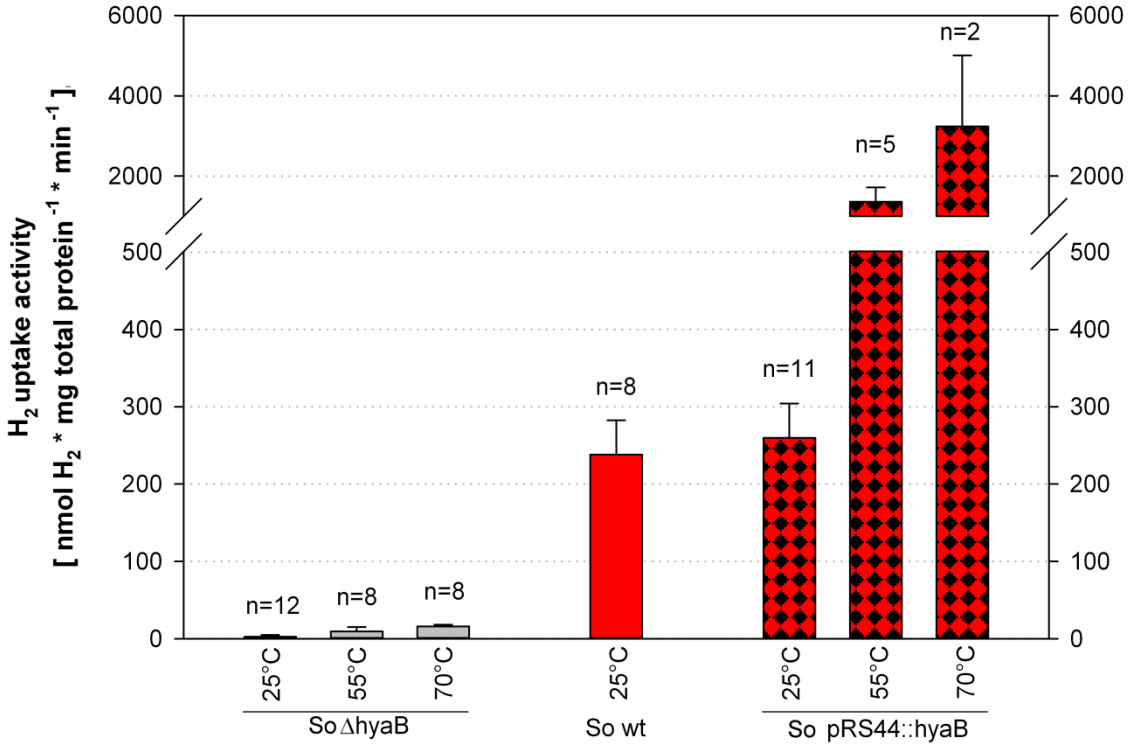


Figure 10: H_2 uptake activity measurements for the *S. oneidensis* MR-1 wt (So wt), the *S. oneidensis* Δ*hyaB* [NiFe]-hydrogenase deletion mutant (So ΔhyaB) and the *S. oneidensis* Δ*hyaB* mutant complemented with the [NiFe]-hydrogenase gene operon of the *S. oneidensis* MR-1 wt encoded on the fosmid pRS44::hyaB (So pRS44::hyaB). The hydrogenase uptake activities were measured spectrophotometrically by the reduction of the artificial electron donor methyl viologen (from colorless to blue) for partially purified crude cell extracts at different temperatures. The numbers of performed measurements are denoted above the single bars (n).

The H_2 consumption of So wt, So Δ*hyaB* and So pRS44::hyaB were determined *in vivo* with a gas chromatograph by measuring the concentration of H_2 in the headspace of the cultures against time (2.8.4). Only a slight decrease of H_2 concentration could be observed for So Δ*hyaB* (Figure 11), indicating the missing H_2 consumption ability of this strain. Consequently, also this result suggested a missing H_2 uptake activity for the mentioned second hydrogenase of *S. oneidensis* MR-1, a putative [FeFe]-hydrogenase. In contrast a strong decrease of H_2 concentration over the time could be observed for So wt and So pRS44::hyaB (Figure 11) indicating a consumption of H_2 by these strains. For So pRS44::hyaB these measurements accordingly confirmed, that the H_2 consumption activity was mediated by genes encoded on the fosmid pRS44::hyaB.

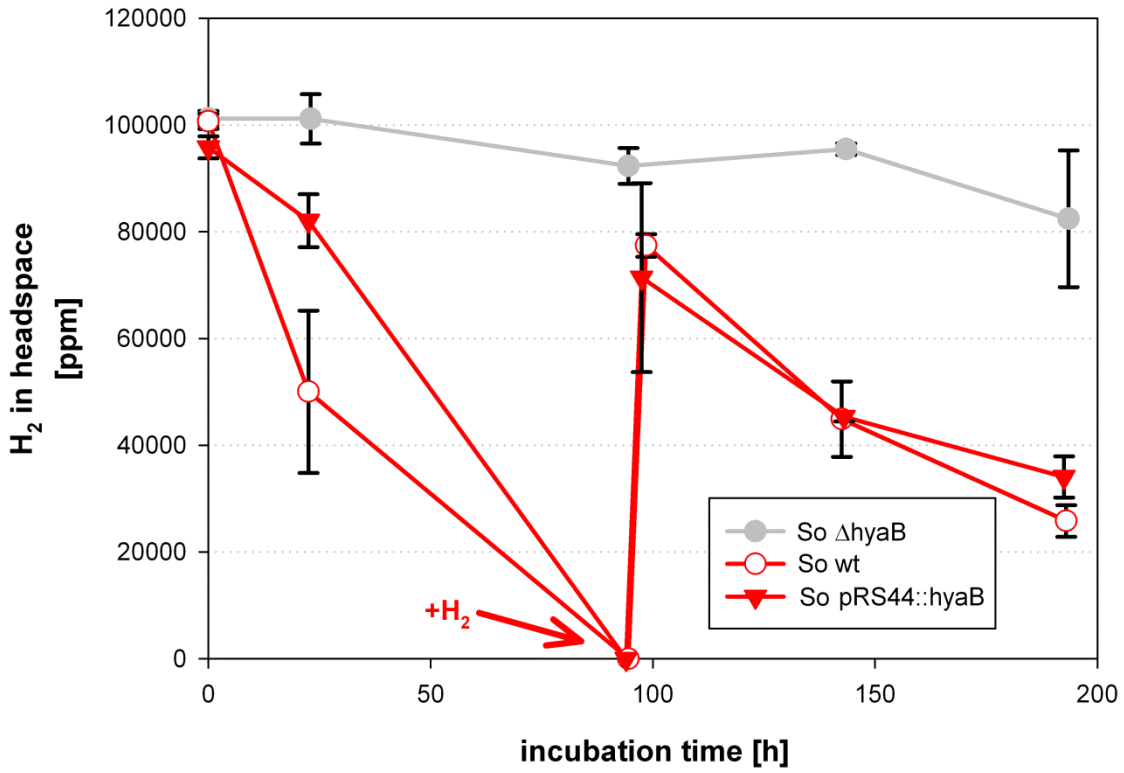


Figure 11: H_2 consumption measurements for the *S. oneidensis* MR-1 wt strain (So wt), the *S. oneidensis* Δ*hyaB* [NiFe]-hydrogenase deletion mutant (So ΔhyaB) and the *S. oneidensis* Δ*hyaB* mutant complemented with the [NiFe]-hydrogenase gene operon of the *S. oneidensis* MR-1 wt encoded on the fosmid pRS44::hyaB (So pRS44::hyaB). The H_2 consumption rates were measured in triplicates with a gas chromatograph. For So wt and So pRS44::hyaB, all hydrogen in the headspace was utilized after 96 hours and additional hydrogen was injected to the cultures (red arrow with +H_2).

The H_2 evolution rates of the So wt, So ΔhyaB and So pRS44::hyaB strain were analyzed *in vitro* by measuring the H_2 produced by the oxidation of completely reduced methyl viologen catalyzed by partially purified crude cell extracts (2.8.5). For So wt, a H_2 evolution rate of 1.31 ±0.20 µmol H_2 * min^{-1} * mg^{-1} total protein was measured in the membrane fraction (Figure 12).The H_2 evolution rate in the membrane fraction of So pRS44::hyaB was determined with 1.15±0.24 µmol H_2 * min^{-1} * mg^{-1} total protein (Figure 12), which was comparable to the *S. oneidensis* MR-1 wt evolution rate. For So ΔhyaB no H_2 evolution was detectable in the membrane fraction (Figure 12). This result for the membrane fraction of So ΔhyaB suggested that the second putative hydrogenase of *S. oneidensis* MR-1, a [FeFe]-hydrogenase, did not seem to have a H_2 evolution activity for the *in vitro* method applied here.

In summary, these measurements demonstrated that the newly developed activity-based screening method to identify H_2 uptake activity enzymes worked.

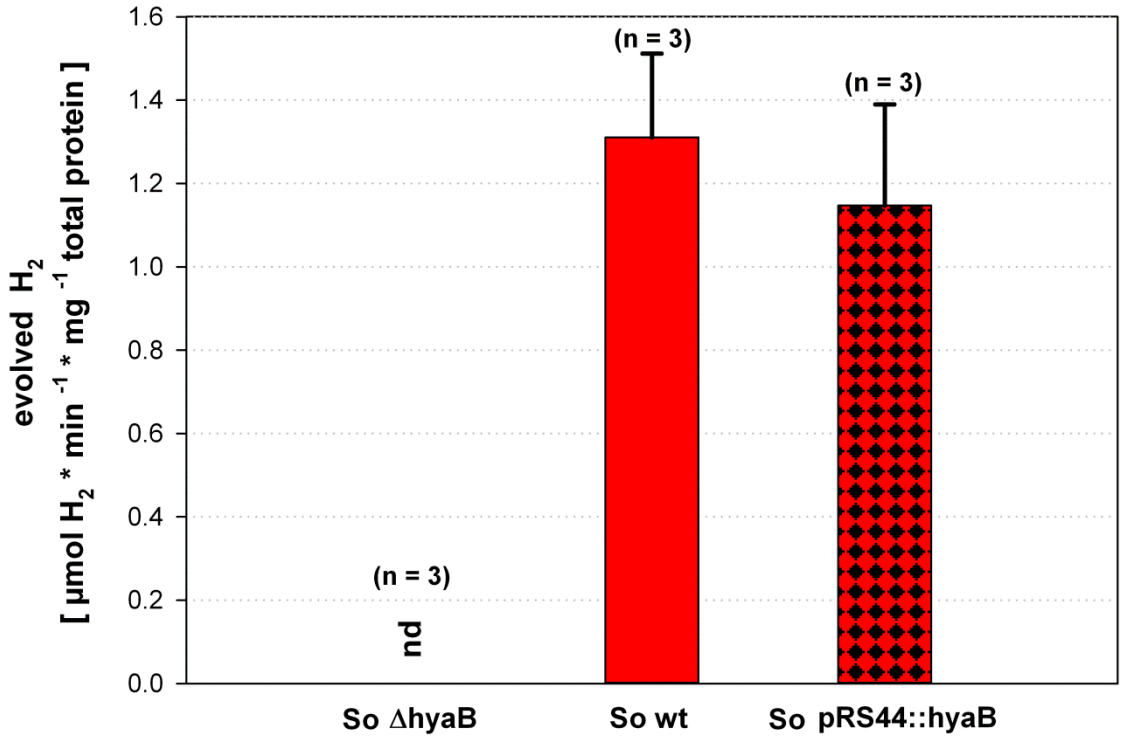


Figure 12: H_2 evolving measurements for the *S. oneidensis* MR-1 wt strain (So wt), the *S. oneidensis* Δ*hyaB* [NiFe]-hydrogenase deletion mutant (So ΔhyaB) and the *S. oneidensis* Δ*hyaB* mutant complemented with the [NiFe]-hydrogenase gene operon of the *S. oneidensis* MR-1 wt encoded on the fosmid pRS44::hyaB (So pRS44::hyaB). The H_2 evolving activities were measured in triplicates with a gas chromatograph by the oxidation of the completely reduced artificial electron donor methyl viologen catalyzed by partially purified crude cell extracts in sealed Hungate tubes after incubation for one hour at 28 °C.

3.2 Construction and quality control of a metagenomic fosmid library from a hydrothermal deep-sea vent sample

A metagenomic hydrothermal vent fosmid library with 7,776 clones was generated (2.7). Fosmids were randomly checked for insert size via fosmid isolation and subsequent analytical digestion (2.4.3, 2.4.9). Their average insert size was estimated with 22 kbp, and the percentage of clones carrying inserts was close to 92 % (exemplary restriction analysis see Figure 13).

Additionally, fosmid insert T7 ends of the Sisters Peak fosmid library were sequenced (Eurofins MWG Operon, Ebersberg, Germany) to confirm vent origin of the fosmid inserts. Sequences exhibited similarities to microbes typically found at hydrothermal deep-sea vent habitats like *Aquificales* (e.g. *Hydrogenivirga* sp.) *Epsilonproteobacteria*

(e.g. *Sulfurovum* sp.) and *Euryarchaeota* (e.g. *Aciduliprofundum boonei*) and showed that the metagenomic fosmid library has a vent origin of the fosmid inserts (Table 5).

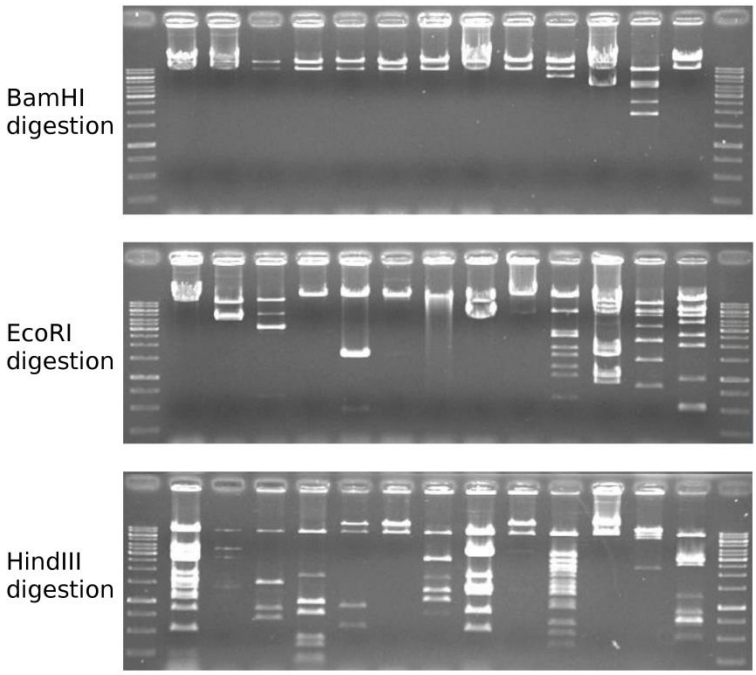


Figure 13: Restriction analysis of 13 random fosmids from the constructed metagenomic broad-host-range large insert library of the Sisters Peak chimney massive sulfide sample. Each fosmid was digested severally with the restriction enzymes BamHI, EcoRI and HindIII and analyzed in parallel by agarose gel electrophoresis to estimate their average insert size.

Table 5: Results of blastx analysis of random fosmids from the constructed metagenomic broad-host-range large insert library of the Sisters Peak chimney massive sulfide sample. Analyzed sequences of the fosmid insert T7 ends can be revised in the appendix 9.4.

Sequenced clone	Description	Max score	Total score	Query cover	E value	Identity
SP_B01_T7	uridine phosphorylase *[Aciduliprofundum boonei* T469]	36.6	36.6	46 %	4.9	22 %
SP_C01_T7	protein of unknown function DUF820 [*Thiorhodovibrio* sp. 970]	35.8	35.8	23 %	7.1	35 %
SP_D01_T7	hypothetical protein Arcve_1673 [*Archaeoglobus veneficus* SNP6]	144	144	68 %	1E-39	41 %
SP_F01_T7	hypothetical protein HG1285_08961 [*Hydrogenivirga* sp. 128-5-R1-1]	135	135	60 %	3E-35	43 %
SP_G01_T7	leucyl aminopeptidase [Ralstonia eutropha H16]	244	244	82 %	5E-75	74 %
SP_H01_T7	RmlA [Geobacillus tepidamans]	288	288	68 %	7E-95	74 %
SP_A02_T7	hypothetical protein Theam_1360 [Thermovibrio ammonificans HB-1]	88.2	88.2	71 %	6E-18	32 %
SP_B02_T7	ABC transporter, permease protein [Aciduliprofundum boonei T469]	191	191	49 %	6E-57	73 %
SP_D02_T7	IS200-like transposase [Thermococcus kodakarensis KOD1]	157	157	41 %	4E-46	70 %

3.3 Detection of H_2 uptake activities by the developed functional screening method

All 7,776 clones of the constructed metagenomic hydrothermal vent fosmid library from the massive sulfide chimney of the black smoker Sisters Peak were screened for H_2 uptake hydrogenase activity by the developed functional screening method.

Initially, the first 2,592 clones were screened for H_2 uptake hydrogenase activity in a first round. With ca. one and a half year distance, the remaining 5,184 clones were analyzed in a second round with support of technical assistants.

The hydrothermal vent metagenomic library was successfully transferred by triparental conjugation from the host *E. coli* EPI300-T^R to the alternative host *S. oneidensis* Δ*hyaB* and stored at -70 °C (2.7.10). The function-based screening was performed by cultivating pools of 96 clones in one serum bottle with FW medium (2.8.2). Putative H_2 uptake activities in the clone pools were detectable by a color change of the inoculated FW medium from yellow to colorless. To identify the clones with recombinant hydrogenase activity in the positive tested 96 clone pools, the described function-based screening was repeated several times with decreasing clone numbers (2.8.2).

During the first round of the screening procedure, which included 2,592 examined fosmid clones, six potential 96 clone pools were identified by the expected color change in the medium. These identified putative positive 96 clone pools are listed in Table 6. The two clones) with hydrogenase activity pRS44::SP-P05F01 (P05F01) and pRS44::SP-P18F11 (P18F11) could be finally identified from the six clone pools (Table 6).

From the residual 5,184 clones, five 96 clone pools with putative H_2 uptake activity were identified in the second round of the screening process (Table 6). From the five clone pools 14 single clones with putative recombinant hydrogenase activity were identified during the second round of the functional H_2 uptake activity screening (Table 6).

For the 96 clone pools of the plates 20, 23 and 24, single H_2 uptake active clones couldn't be identified successfully, due to ambiguous results of the screening process. Instead the focus was directed on the identified single fosmid clones and their investigation.

Two of the identified fosmid clones were further analyzed in detail on a sequence and activity level, due to their early identification date: the fosmid clones P05F01 and P18F11 (3.4). The other detected 12 fosmid clones with potential H_2 uptake activity were not investigated with the same range, but the screening procedure was repeated several times and the 12 fosmid clones were also analyzed on a sequence and activity level. In

contrast to the fosmid clones P05F01 and P18F11, a H_2 uptake hydrogenase activity couldn't be confirmed for any of these 12 fosmid clones (3.5).

Table 6: Putative positive tested 96 clone pools and putative single active fosmid clones with recombinant hydrogenase activity identified by the H_2 uptake hydrogenase activity screening.

96 clone pools	Putative active clones		
plate 05[a]	P05F01 [a]		
plate 11 [a]	P11C02	P11H04	
plate 18 [a]	P18F11 [a]		
plate 20 [a]			
plate 23 [a]			
plate 24 [a]			
plate 30	P30B05	P30F07	
plate 31	P31A02	P31A05	P31B08
plate 35	P35F04		
plate 37	P37B01	P37B03	
plate 44	P44E07	P44H11	

[a] in underlined: 96 clone pools and single active fosmid clones identified in the first round of the screening process

3.4 Analysis of the putative H_2 uptake active fosmid clones pRS44::SP-P05F01 and pRS44::SP-P18F11

3.4.1 Sequence analysis of the fosmid clones pRS44::SP-P05F01 containing putative metagenomic H_2 uptake active genes

The fosmid insert size of the pRS44::SP-P05F01 (P05F01) fosmid clone was approximately 10.7 kbp and the predicted ORFs resembled genes of the hydrogenase operon from *N. salsuginis* DSM 16511 (Table 7): The ORFs of the fosmid P05F01 resembled two identical structural genes for a large subunit of a uptake [NiFe]-hydrogenase (*hyaB*) (maximal identity 83 %) and three structural genes for a small subunit of the [NiFe]-hydrogenase (*hyaA*) (maximal identity of 88 %) (Figure 14, Table 7). Additionally, two ORFs with similarities to the [NiFe]-hydrogenase cytochrome b subunit (*hyaC*) (max. ident. 82 %) were identified, but 258 nucleotides (83 amino acids) are missing in the ORF at the end of the fosmid insert (Figure 14, Table 7). Furthermore, one ORF resembled a hydrogenase maturation protease (*hydD*) with a maximal identity of 86 % (Figure 14, Table 7).

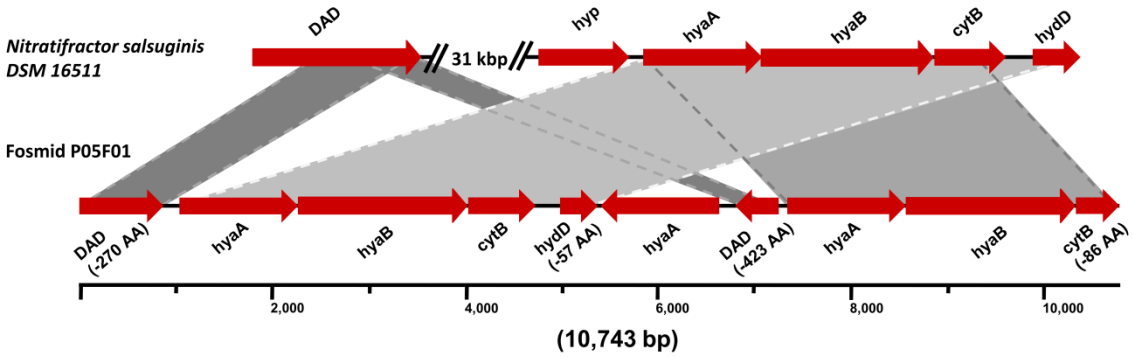


Figure 14: Schematic representation of the compared ORF arrangement of the sequenced metagenomic fosmid pRS44::SP-P05F01 (P05F01) and the [NiFe]-hydrogenase operon of *N. salsuginis* DSM16511. The arrows represent the predicted ORFs and show the direction of the supposed gene transcription. The dashed lines embracing the gray areas show the allocation of ORFs on the fosmid to corresponding similar ORFs on the genome of *N. salsuginis* DSM 16511. DAD= dihydroxy-acid dehydratase; *hyaA* = hydrogenase (nife) small subunit; *hyaB* = nickel-dependent hydrogenase large subunit; *cytB* = Ni-Fe hydrogenase cytochrome b subunit; *hydD* = hydrogenase maturation protease

Despite the similarity of the ORFs on the fosmid P05F01 to the genes of *N. salsuginis*, the gene arrangement of the ORFs on the fosmid differed from the arrangement of the genes found in the hydrogenase operon of *N. salsuginis* (Figure 14) (Anderson *et al.*,

2011). While on the *N. salsuginis* genome only one copy of the uptake hydrogenase gene cluster exists, on the fosmid P05F01 parts of hydrogenase gene clusters could be found in duplicates or triplicates (Figure 14, Table 7). In addition to one gene operon consisting of a small structural [NiFe]-hydrogenase subunit, large structural [NiFe]-hydrogenase subunit, [NiFe]-hydrogenase cytochrome B structural subunit and a hydrogenase maturation protease, a second similar gene cluster with a small, large and cytochrome B structural [NiFe]-hydrogenase subunit could be located on the fosmid P05F01 (both clusters light gray shaded in Figure 14). Additionally, a third copy of a small structural [NiFe]-hydrogenase subunit was found. Furthermore, two ORFs were located on the fosmid P05F01 with partly similarities to a dihydroxy-acid dehydratase (DAD) originated from *Sulfurovum* sp. NBC37-1 (dark gray shaded in Figure 14).

Table 7: Blastx results of the detected ORFs of the metagenomic fosmid pRS44::SP-P05F01 [a].

ORFs	Accession No.	Description	Query coverage	E-value	Max identity[1]
DAD	YP_003303114	dihydroxy-acid dehydratase [*Sulfurovum* sp. NBC37-1]	100 %	1e-152	81 %
hyaA	YP_004167096	hydrogenase (nife) small subunit [*Nitratifractor salsuginis* DSM 16511]	99 %	0.0	88 %
hyaB	YP_004167097	nickel-dependent hydrogenase large subunit [*Nitratifractor salsuginis* DSM 16511]	99 %	0.0	83 %
cytB	YP_004167098	Ni-Fe hydrogenase, cytochrome b subunit [*Nitratifractor salsuginis* DSM 16511]	100 %	1e-140	82 %
hydD	YP_004167099	hydrogenase maturation protease [*Nitratifractor salsuginis* DSM 16511]	96 %	2e-70	86 %
hyaA	YP_004167096	hydrogenase (nife) small subunit [*Nitratifractor salsuginis* DSM 16511]	99 %	0.0	88 %
DAD	YP_003303114	dihydroxy-acid dehydratase [*Sulfurovum* sp. NBC37-1]	93 %	1e-70	82 %
hyaA	YP_004167096	hydrogenase (nife) small subunit [*Nitratifractor salsuginis* DSM 16511]	99 %	0.0	88 %
hyaB	YP_004167097	nickel-dependent hydrogenase large subunit [*Nitratifractor salsuginis* DSM 16511]	99 %	0.0	83 %
cytB	YP_004167098	Ni-Fe hydrogenase, cytochrome b subunit [*Nitratifractor salsuginis* DSM 16511]	99 %	3e-83	82 %

[a] Listed are the best matches to next cultivated relative after a blastx search against the NCBI nr database.

[1] based on amino acids.

3.4.2 Sequence analysis of the fosmid clone pRS44::SP-P18F11 containing putative metagenomic H_2 uptake active genes

The fosmid insert size of the clone pRS44::SP-P18F11 (P18F11) had approximately 20 kbp. All ORFs showed high similarities to genes of the Euryarchaeota *Aciduliprofundum* sp. MAR08-339. This species is an unnamed isolate from a *Aciduliprofundum* population, which seems to be clearly distinct from known *Aciduliprofundum* species, but has not been characterized by traditional methods until now. Although H_2 uptake activity was identified for *S. oneidensis* Δ*hyaB* complemented with this fosmid, none of the identified ORFs showed similarities to a known hydrogenase. The fosmid sequence is characterized by a repetition of identical ORFs with similarities to a formate dehydrogenase (FDH), NADH:ubiquinone oxidoreductase NADH-binding 51 kDa subunits, NADH:ubiquinone oxidoreductase 24 kDa subunits, molybdopterin-guanine dinucleotide biosynthesis protein A and cytosine deaminase-like metal-dependent hydrolase (Figure 15).

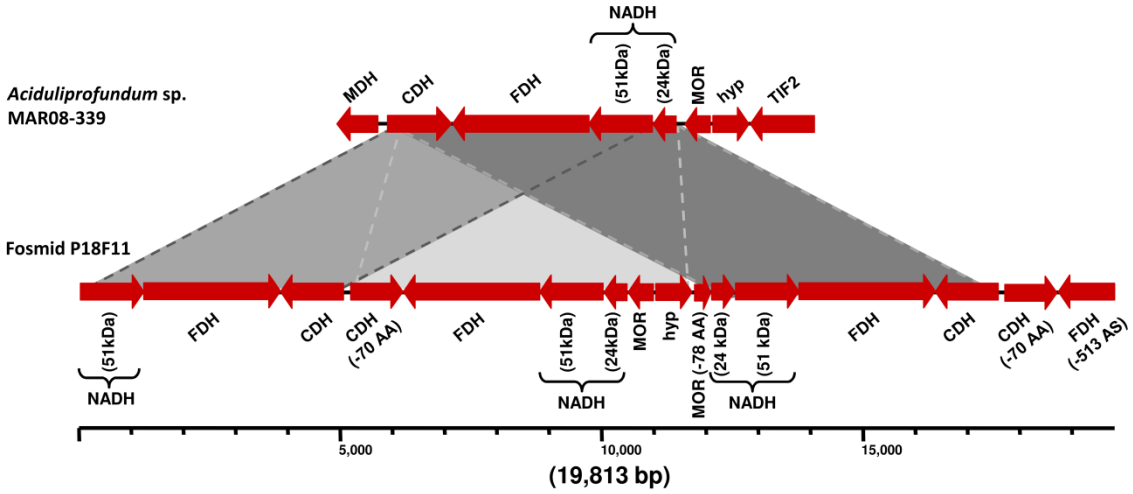


Figure 15: Schematic representation of the compared ORF arrangement of the sequenced metagenomic fosmid pRS44::SP-P18F11 (P18F11) and a homologous region on the *Aciduliprofundum* sp. MAR08-339 genome. The arrows represent the predicted ORFs and show the direction of the supposed gene transcription. The dashed lines embracing the gray areas show the allocation of ORFs on the fosmid to corresponding similar ORFs on the genome of *Aciduliprofundum* sp. MAR08-339. MDH = metal dependent hydrolase; CDH = cytosine deaminase-like metal dependent hydrolase; FDH = formate dehydrogenase, alpha subunit archaeal-type; NADH (51kDa) = NADH:ubiquinone oxidoreductase, NADH-binding 51 kDa subunit; NADH (24kDa) = NADH:ubiquinone oxidoreductase, 24 kDa subunit; MOR = molybdopterin-guanine dinucleotide biosynthesis protein A; *hyp* = hypothetical protein AciM339_0272; TIF2 = translation initiation factor 2 subunit gamma

Table 8: Blastx results of the detected ORFs of the metagenomic fosmid pRS44::SP-P18F11 [a].

ORFs	Accession No.	Description	Query coverage	E value	Max identity[1]
NADH (51 kDa)	ZP_04873436	NADH:ubiquinone oxidoreductase, NADH-binding 51 kDa subunit [*Aciduliprofundum* sp. MAR08-339]	100 %	0.0	82 %
FDH	ZP_04874240	formate dehydrogenase, alpha subunit, archaeal-type [*Aciduliprofundum* sp. MAR08-339]	100 %	0.0	85 %
CDH	YP_007246330	cytosine deaminase-like metal-dependent hydrolase [*Aciduliprofundum* sp. MAR08-339]	99 %	0.0	90 %
CDH (-70 AA)	YP_007246330	cytosine deaminase-like metal-dependent hydrolase [*Aciduliprofundum* sp. MAR08-339]	99 %	0.0	89 %
FDH	ZP_04874240	formate dehydrogenase, alpha subunit, archaeal-type [*Aciduliprofundum* sp. MAR08-339]	100 %	0.0	85 %
NADH (51 kDa)	ZP_04873436	NADH:ubiquinone oxidoreductase, NADH-binding 51 kDa subunit [*Aciduliprofundum* sp. MAR08-339]	100 %	0.0	82 %
NADH (24 kDa)	WP_015282791	NADH:ubiquinone oxidoreductase, 24 kDa subunit [*Aciduliprofundum* sp. MAR08-339]	100 %	6e-49	74 %
MOR	YP_007246334	molybdopterin-guanine dinucleotide biosynthesis protein A [*Aciduliprofundum* sp. MAR08-339]	100 %	2e-105	84 %
hyp	YP_007246335	hypothetical protein AciM339_0272 [*Aciduliprofundum* sp. MAR08-339]	100 %	5e-132	96 %
MOR (-78 AA)	YP_007246334	molybdopterin-guanine dinucleotide biosynthesis protein A [*Aciduliprofundum* sp. MAR08-339]	99 %	6e-46	79 %
NADH (24 kDa)	WP_015282791	NADH:ubiquinone oxidoreductase, 24 kDa subunit [*Aciduliprofundum* sp. MAR08-339]	100 %	6e-49	74 %
NADH (51 kDa)	ZP_04873436	NADH:ubiquinone oxidoreductase, NADH-binding 51 kDa subunit [*Aciduliprofundum* sp. MAR08-339]	100 %	0.0	82 %
FDH	ZP_04874240	formate dehydrogenase, alpha subunit [*Aciduliprofundum* sp. MAR08-339]	100 %	0.0	85 %
CDH	YP_007246330	cytosine deaminase-like metal-dependent hydrolase [*Aciduliprofundum* sp. MAR08-339]	99 %	0.0	90 %
CDH (-70 AA)	YP_007246330	cytosine deaminase-like metal-dependent hydrolase [*Aciduliprofundum* sp. MAR08-339]	99 %	0.0	89 %
FDH (-513 AA)	ZP_04874240	formate dehydrogenase, alpha subunit [*Aciduliprofundum* sp. MAR08-339]	99 %	0.0	83 %

[a] Listed are the best matches to next cultivated relative after a blastx search against the NCBI nr database.

[1] based on amino acids.

Four ORFs resembled genes of the α-subunit of a formate dehydrogenase (FDH) with a maximal identity of 85 %, three identical ORFs for a NADH:ubiquinone oxidoreductase NADH-binding 51 kDa subunit with a maximal identity of 82 %, two identical ORFs for a NADH:ubiquinone oxidoreductase 24 kDa subunit with a maximal identity of 74 %, two ORFs resembled genes of a molybdopterin-guanine dinucleotide biosynthesis protein A with a maximal identity of maximal identity 83 % and four ORFs resembled genes of a cytosine deaminase-like metal-dependent hydrolase (CDH) with a maximal identity of 90 % (Figure 15, Table 8). Additionally, there is one non-repeating ORFs with a similarity to the hypothetical protein AciM339_0272 (maximal identity 96 %) (Figure 15, Table 8). The ORF arrangement of P18F11 is similar to the arrangement found on the genome of *Aciduliprofundum* sp. MAR08-339. But in contrast to the *A.* sp. MAR08-339 genome with only one copy of each ORF (CDH, FDH, NADH (24 kDa), NADH (51kDa), MOR, hyp), three slightly different copies of the ORF arrangement are found on the fosmid P18F11 (gray shaded in Figure 15).

3.4.3 Hydrogenase capacity analysis of the fosmid clones pRS44::SP-P05F01 and pRS44::SP-P18F11 containing putative metagenomic H_2 uptake active enzymes in *S. oneidensis* Δ*hyaB*

3.4.3.1 H_2 consumption of *S. oneidensis* Δ*hyaB* strains complemented with pRS44::SP-P05F01 or pRS44::SP-P18F11

The H_2 consumption of the *S. oneidensis* Δ*hyaB* strain, *S. oneidensis* Δ*hyaB* with the fosmid pRS44::SP-P05F01 (P05F01) and *S. oneidensis* Δ*hyaB* with the fosmid pRS44::SP-P19F11 (P18F11) were determined *in vivo* with a gas chromatograph by measuring the concentration of H_2 in the headspace of the cultures against time (2.8.4). As before in these measurements only a slight decrease of H_2 concentration could be observed for the *S. oneidensis* Δ*hyaB* deletion mutant (Figure 16) against the time, indicating the lost H_2 consumption abilities of this strain. In contrast a strong decrease of H_2 concentration could be observed for the *S. oneidensis* Δ*hyaB* with the fosmids P05F01 or P18F11 (Figure 16) against the time, indicating a consumption of H_2 by these strains and accordingly confirm the H_2 consumption activity mediated by genes encoded on the fosmids.

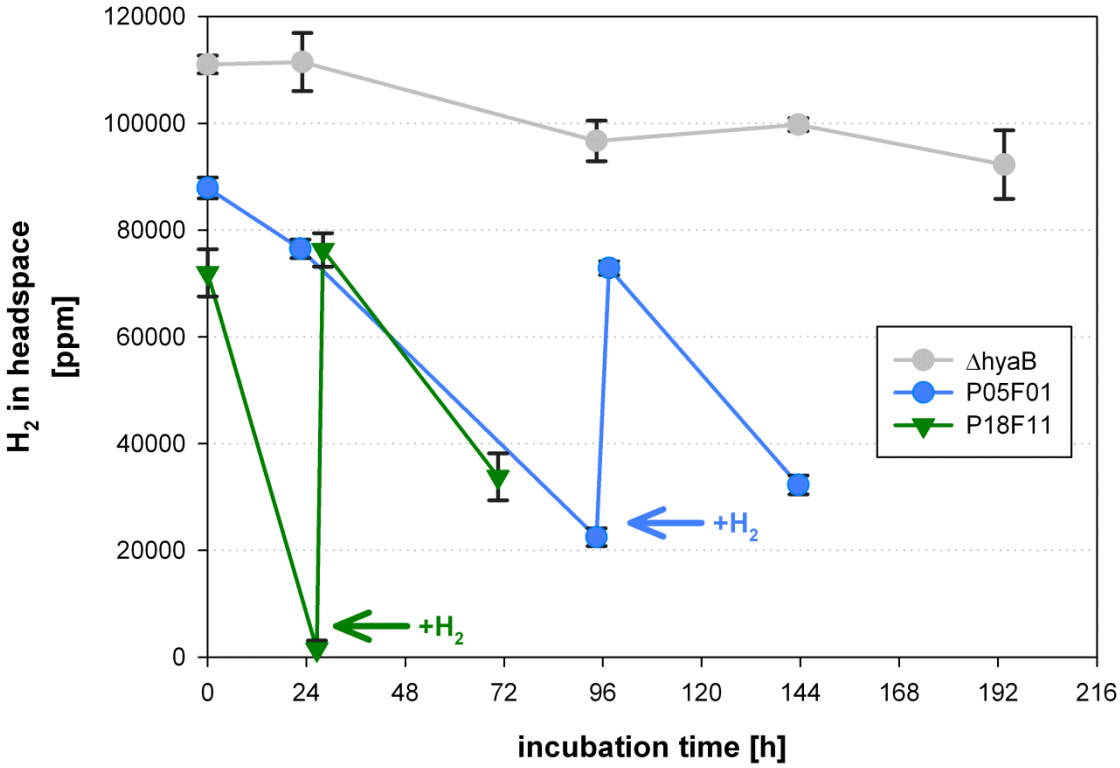


Figure 16: H_2 consumption measurements for the the *S. oneidensis* Δ*hyaB* [NiFe]-hydrogenase deletion mutant (ΔhyaB) and the *S. oneidensis* Δ*hyaB* mutant strain which was complemented with the two H_2 uptake active fosmids pRS44::SP-P05F01 (P05F01) and pRS44::SP-P18F11 (P18F11). The H_2 consumption rates were measured in triplicates with a gas chromatograph. For P05F01 more than 75 % of the initial hydrogen in the headspace was utilized after 96 hours and for P18F11 all hydrogen in the headspace was utilized after 26 hours. To all cultures additional hydrogen was injected after its detected utilization (blue arrow with +H_2 for P05F01 and green arrow with +H_2 for P18F11).

3.4.3.2 H_2 uptake activity of *S. oneidensis* Δ*hyaB* strains complemented with pRS44::SP-P05F01 or pRS44::SP-P18F11

The turnover rates of the potentially H_2 uptake active enzymes encoded on the fosmid pRS44::SP-P05F01 (P05F01) and pRS44::SP-P18F11 (P18F11), and additionally the H_2 uptake activity of the *S. oneidensis* Δ*hyaB* deletion mutant (Δ*hyaB)*, *Aciduliprofundum boonei* T469 and *Aciduliprofundum* sp. MAR08-339 were analyzed spectrophotometrically and *in vitro* by hydrogenase activity assays at the different temperatures of 25°C, 55°C and 70°C with protein crude extracts of membrane fractions originated from corresponding bacterial and archaeal cultures (2.3.1, 2.8.3.1 and 2.8.3.4).

For the fosmid P05F01, H_2 uptake activities of 219.5 ± 54.4 nmol * min^{-1} * mg^{-1} total protein at 25°C, 2677.0 ± 941.4 nmol * min^{-1} * mg^{-1} total protein at 55°C and 12001.5 ± 1035.1 nmol * min^{-1} * mg^{-1} total protein at 70°C were measured in the membrane fraction (Figure 17).

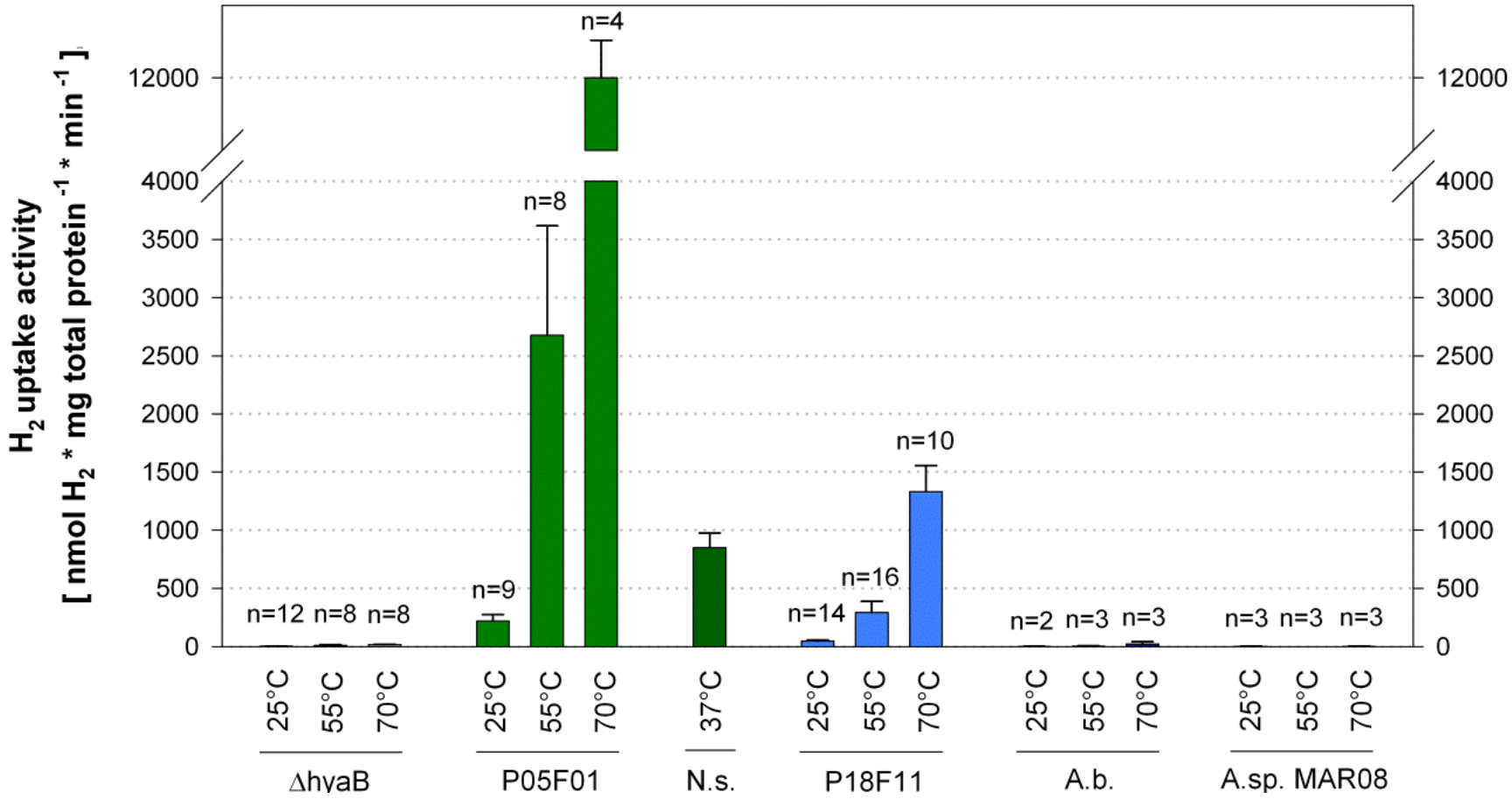


Figure 17: H_2 uptake activity measurements for the membrane fractions of *S. oneidensis* Δ*hyaB* (ΔhyaB), for the *S. oneidensis* Δ*hyaB* deletion mutants which were complemented with the two H_2 uptake active fosmids pRS44::SP-P05F01 (P05F01) and pRS44::SP-P18F11 (P18F11), for *Aciduliprofundum boonei* T469 (A.b.), for *Aciduliprofundum* sp. MAR08-339 (A.sp. MAR08) and for *Nitratifractor salsuginis* DSM16511 (N.s.). The hydrogenase uptake activities were measured spectrophotometrically by the reduction of the artificial electron donor methyl viologen (from colorless to blue) for partially purified crude cell extracts at different temperatures, expect for the H_2 uptake activity of *N. salsuginis* DSM 16511 which was abstracted by Takai et al. (2005). The numbers of performed measurements are denoted above the single bars (n).

The H_2 uptake activities in the membrane fraction of the second fosmid P18F11 were 45.8 ± 9.1 nmol * min ^{1} * mg^{-1} total protein at 25°C, 292.6 ± 96.7 nmol * min^{-1} * mg^{-1} total protein at 55°C and 1331.5 ± 223.5 nmol * min^{-1} * mg^{-1} total protein at 70°C (Figure 17).

For Δ*hyaB* minor H_2 uptake activities were determined (25°C: 2.7 ± 2.2 nmol * min^{-1} * mg^{-1} total protein; 55°C: 9.6 ± 5.6 nmol * min^{-1} * mg^{-1} total protein; 70°C: 16.0 ± 2.3 nmol * min^{-1} * mg^{-1} total protein; Figure 17). Also for *A. boonei* T469 negligible H_2 uptake activities could be measured (25°C: 1.2 ± 1.3 nmol * min^{-1} * mg^{-1} total protein; 55°C: 3.4 ± 4.1 nmol * min^{-1} * mg^{-1} total protein; 70°C: 21.6 ± 19.5 nmol * min^{-1} * mg^{-1} total protein; Figure 17) as well as for *Aciduliprofundum* sp. MAR08-339 (25°C: 0.4 ± 3.3 nmol * min^{-1} * mg^{-1} total protein; 55°C: -2.4 ± 1.7 nmol * min^{-1} * mg^{-1} total protein; 70°C: 1.5 ± 2.6 nmol * min^{-1} * mg^{-1} total protein; Figure 17).

In all cases, no significant hydrogenase activities were detectable in the tested soluble fractions (data not shown).

3.4.3.3 H_2 evolution of *S. oneidensis* Δ*hyaB* strains complemented with pRS44::SP-P05F01 or pRS44::SP-P18F11

The H_2 evolution rates of the potentially H_2 uptake active enzymes in *S. oneidensis* membrane fractions encoded on the fosmid pRS44::SP-P05F01 (P05F01) or pRS44::SP-P18F11 (P18F11) were analyzed by determining the produced H_2 with a gas chromatograph (2.8.5). For the fosmid P05F01, a H_2 evolution rate of 1.15 ± 0.20 µmol H_2 * min^{-1} * mg^{-1} total protein was measured in the membrane fraction at 28 °C (Figure 18). The H_2 evolution rate in the membrane fraction of the fosmid P18F11 was determined with 0.29 ± 0.02 µmol H_2 * min^{-1} * mg^{-1} total protein at 28 °C (Figure 18). For the negative control, H_2 evolution rate for the *S. oneidensis* Δ*hyaB* strain membrane fraction was measured and no H_2 evolution was detectable at 28 °C.

The H_2 evolution rates of *Pyrococcus furiosus* and *Thermococcus kodakaraensis* were gathered from foreign publications. The H_2 evolution rates of *Pyrococcus furiosus* with 4.00 µmol H_2 * min^{-1} * mg^{-1} total protein was abstracted by Sapra *et al.* (2000). For *Thermococcus kodakaraensis* a with H_2 evolution rate of 1.09 µmol H_2 * min-1 * mg^{-1} total protein was abstracted by Kanai *et al.* (2003). Both H_2 evolution rates were determined with a gas chromatograph at 80 °C.

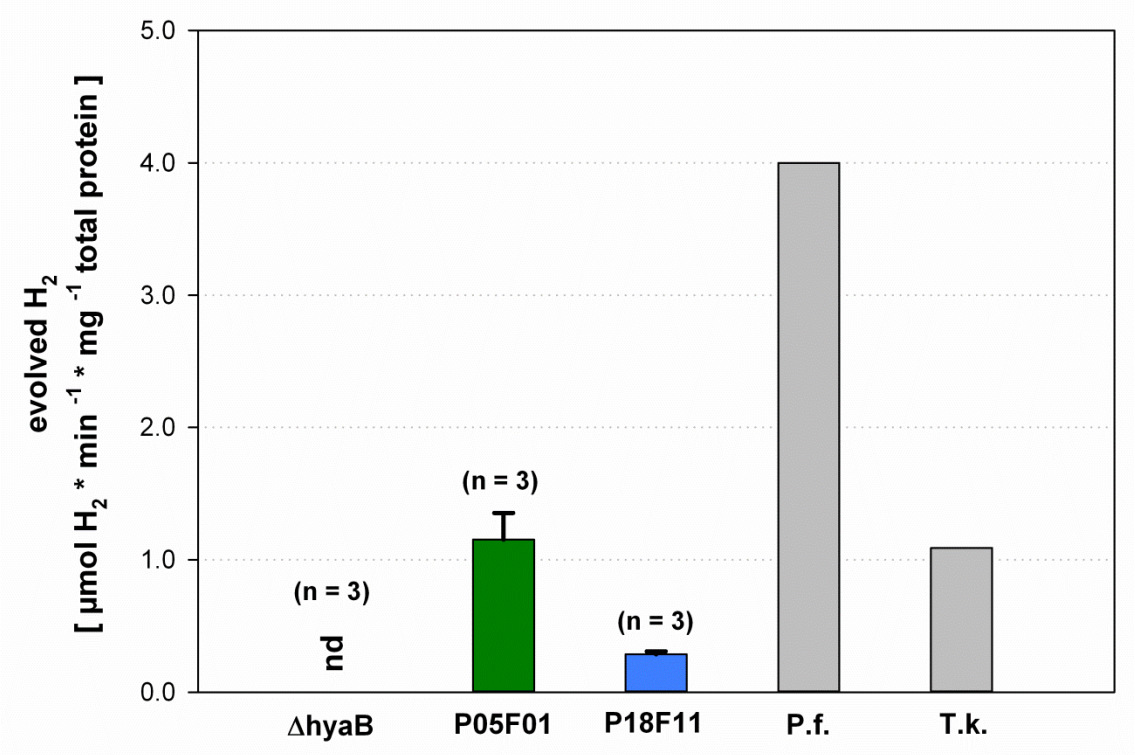


Figure 18: H_2 evolving measurements for the *S. oneidensis* Δ*hyaB* [NiFe]-hydrogenase deletion mutant (ΔhyaB), for the *S. oneidensis* Δ*hyaB* deletion mutants which were complemented with the two H_2 uptake active fosmids pRS44::SP-P05F01 (P05F01) and pRS44::SP-P18F11 (P18F11), for *Pyrococcus furiosus* (P.f.) and for *Thermococcus kodakaraensis* (T.k.). The H_2 evolving activities were measured in triplicates with a gas chromatograph by the oxidation of the completely reduced artificial electron donor methyl viologen catalyzed by partially purified crude cell extracts in sealed Hungate tubes after incubation for one hour at 28 °C. The H_2 evolution rates of *Pyrococcus furiosus* and *Thermococcus kodakaraensis* were abstracted by Sapra et al. (2000) for *Pyrococcus* and by Kanai et al. (2003) for *Thermococcus*.

3.4.4 Identification of the ORFs responsible for the H_2 uptake activity of the fosmid pRS44::SP-P18F11

In contrast to the H_2 uptake active fosmid pRS44::SP-P05F01 (P05F01), no ORF associated with hydrogenase activity could be found on the sequenced fosmid pRS44::SP-P18F11 (P18F11) (3.4.1). To identify the ORFs which mediated the H_2 uptake activity of this fosmid, P18F11 was subcloned by generating fragments via restriction with StuI, BamHI, BasAI and SacII. The restricted P18F11 fragments (2.4.9) were length separated by agarose gel electrophoresis (2.4.7), all generated fragments were extracted from the agarose gel (2.4.4) and cloned into pRS44 (2.4.9.4) resulting in 14 pRS44 constructs.

Table 9: List of subclones originated from the restriction of fosmid pRS44::SP-P18F11 (P18F11), generated by restriction enzymes, and the intact ORFs present on the corresponding subclone.

used restriction enzyme	subclone no	construct name	tested	length [bp]	complete ORFs on construct [a]
BamHI	SC 1	pRS44::P18-BamHI-SC1	5 times	4,095	hyp; MOR (-78AA); NADH (24 kDa); NADH (51 kDa)
	SC 2	pRS44::P18-BamHI-SC2	5 times	3,215	NADH (24 kDa); NADH (51 kDa)
	SC 3a	pRS44::P18-BamHI-SC3a	3 times	2,422	NADH (51 kDa)
	SC 3b, 3c	pRS44::P18-BamHI-SC3bc	3 times	2,463	CDH; CDH (-70AA)
	SC 4a, 4b, 4c	pRS44::P18-BamHI-SC4ab	3 times	1,377	--
	SC 5	pRS44::P18-BamHI-SC5	3 times	1,022	--
StuI	SC 1	pRS44::P18-StuI-SC1	3 times	4,850	NADH (51 kDa); FDH
	SC 2a, 2b	pRS44::P18-StuI-SC2ab	4 times	5,431	CDH (-70AA); FDH; NADH (51 kDa)
	SC 3	pRS44::P18-StuI-SC3	2 times	2,075	--
	SC 4	pRS44::P18-StuI-SC4	2 times	1,185	MOR
	SC 5	pRS44::P18-StuI-SC5	2 times	841	MOR (-78AA)
BsaAI	SC 1	pRS44::P18-BsaAI-SC1	3 times	6,151	hyp; MOR -78AA; NADH (24 kDa); NADH (51 kDa); FDH
SacII	SC 2	pRS44::P18-SacII-SC2	2 times	7,817	NADH (24 kDa); NADH (51 kDa); FDH; CDH; CDH -70AA
	SC 3	pRS44::P18-SacII-SC3	2 times	1,402	hyp

[a] CDH = cytosine deaminase-like metal dependent hydrolase; FDH = formate dehydrogenase, alpha subunit archaeal-type; NADH (51kDa) = NADH:ubiquinone oxidoreductase, NADH-binding 51 kDa subunit; NADH (24kDa) = NADH:ubiquinone oxidoreductase, 24 kDa subunit; MOR = molybdopterin-guanine dinucleotide biosynthesis protein A; hyp = hypothetical protein AciM339_0272

All these 14 created pRS44 constructs (Figure 19, Table 9) were transferred to *S. oneidensis* Δ*hyaB* (2.4.12) and were tested for restored H_2 uptake activity (3.1.3). None of these pRS44 constructs (Figure 19, Table 9) could restore the ability of *S. oneidensis* Δ*hyaB* to oxidize H_2 and to couple this reaction to the reduction of Fe(III) to Fe(II), which would be detectable by a color change from yellow to colorless.

In addition to the subcloning of P18F11, a second approach with a PCR strategy was attempted to identify the ORFs on the fosmid P18F11, which restored the H_2 uptake ability in *S. oneidensis* Δ*hyaB*. Therefore, two P18F11 segments were amplified (2.4.8) by using unique primer sites: modT7prom + hyp454_P2_RC for amplification of segment F1 and Sub2_T7_PCR + hyp454_P2_F for amplification of segment F2 (Table 2; Figure 19). Segment F1 was cloned into blunt end digested (EcorV) pBSK+ (pBSK+::P18-F1) and subsequently digested with StuI targeting the restriction site within the primer hyp454_P2_RC (2.4.9; Table 2). Segment F2 was ligated into the linearized pBSK+::P18-F1 to create a scarless ligated fragment in pBSK+ (pBSK+::P18-F1+2) consisting both segments (2.4.9.4). The ligated fragment was cut out by restriction with SalI /NotI, the fragment F1+2 was polished by an endrepair (2.4.9.3) and was ligated into pRS44 resulting in the fosmid pRS44::P18-PCR-F1+2 (Figure 19; Table 10). The successful creation of pRS44::P18-PCR-F1+2 was verified by sequencing (2.6.1). The fosmid pRS44::P18-PCR-F1+2 contained the following ORFs: two formate dehydrogenase, two NADH:ubiquinone oxidoreductase NADH-binding 51 kDa subunits, one NADH:ubiquinone oxidoreductase 24 kDa subunit, one molybdopterin-guanine dinucleotide biosynthesis protein A, one hypothetical protein and two cytosine deaminase-like metal-dependent hydrolases, whereby only one of these CDH ORFs was intact (Figure 19). Their orientation and assembly is illustrated in Figure 19. Nevertheless, the pRS44::P18F11-PCR-F1+2 construct couldn't restore the ability of *S. oneidensis* Δ*hyaB* to reduce Fe(III) by utilization of H_2 either (2.4.12).

Table 10: Subclone from the fosmid pRS44::SP-P18F11 generated by a PCR strategy and the intact ORFs present on the constructed subclone.

PCR strategy	subclone no	construct name	tested	length [bp]	complete ORFs on construct [a]
PCR-F1+2	F1+2	pRS44::P18-PCR-F1+2	3 times	11,803	NADH (51 kDa); FDH; CDH; CDH (-70AA); FDH; NADH (51 kDa); NADH (24 kDa); MOR; hyp

[a] CDH = cytosine deaminase-like metal dependent hydrolase; FDH = formate dehydrogenase; NADH (51kDa) = NADH:ubiquinone oxidoreductase, NADH-binding 51 kDa subunit; NADH (24kDa) = NADH:ubiquinone oxidoreductase, 24 kDa subunit; MOR = molybdopterin-guanine dinucleotide biosynthesis protein A; hyp = hypothetical protein

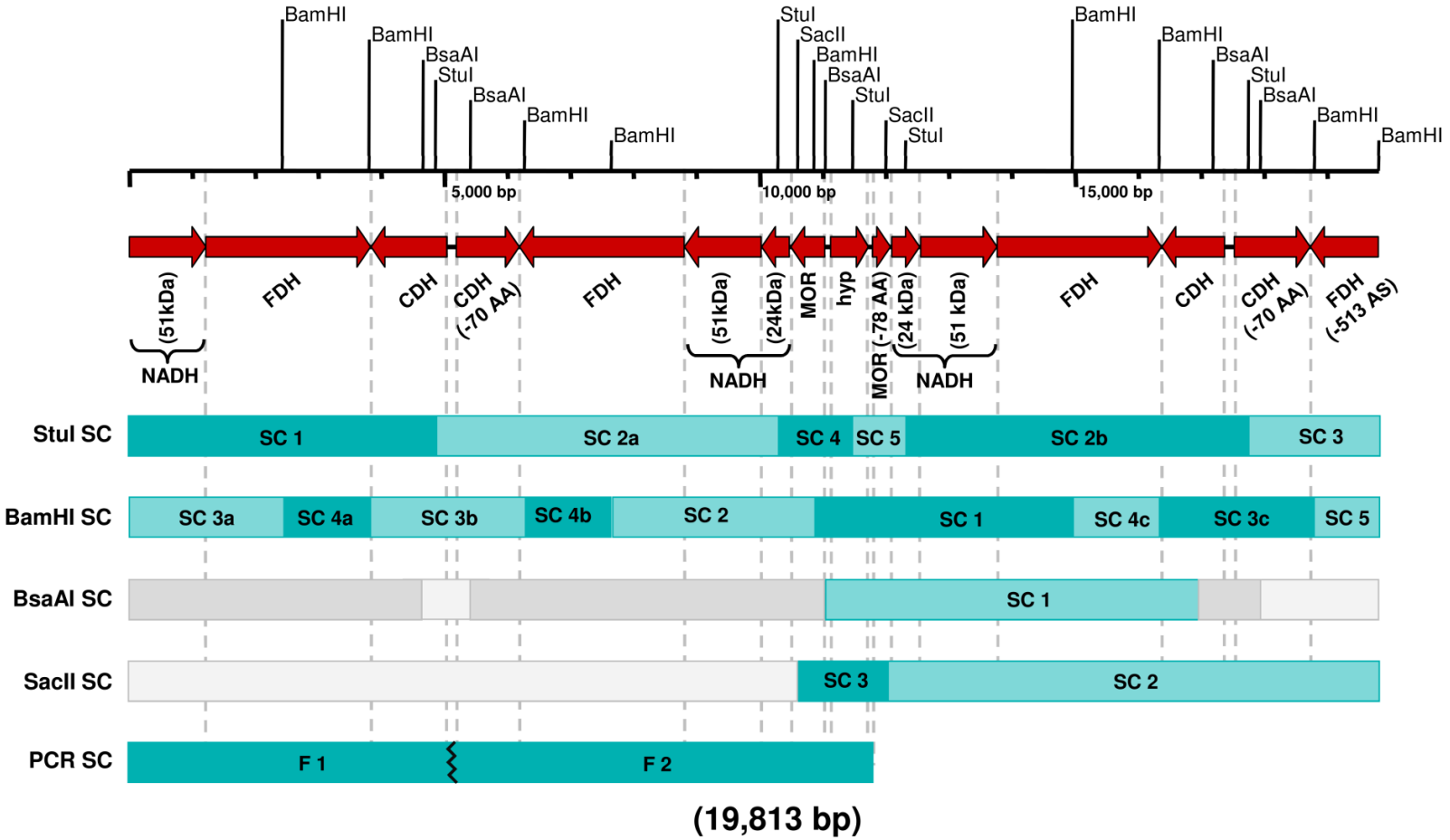


Figure 19: Subcloning of pRS44::SP-P18F11 (P18F11) to identify the ORFs, which mediated the H_2 uptake activity of this fosmid. The arrows represent the predicted ORFs and show the direction of the supposed gene transcription. P18F11 fragments were generated by digestion with several restriction enzymes and the StuI-, BamHI, BsaAI and SacII restriction sites are marked. Greenish bars illustrate the resulting fragments and the ORFs which are located on the corresponding fragment. Additionally, the amplified segments via PCR were illustrated equally (PCR SC). All created P18F11 segments were cloned into pRS44 and transferred to *S. oneidensis* Δ*hyaB* to test for restored ability to reduce Fe(III) and thereby for existing H_2 uptake activity. CDH = cytosine deaminase-like metal dependent hydrolase; FDH = formate dehydrogenase, alpha subunit archaeal-type; NADH (51kDa) = NADH:ubiquinone oxidoreductase, NADH-binding 51 kDa subunit; NADH (24kDa) = NADH:ubiquinone oxidoreductase, 24 kDa subunit; MOR = molybdopterin-guanine dinucleotide biosynthesis protein A; hyp = hypothetical protein AciM339_0272

The subclones pRS44::P18-SacII-SC2, pRS44::P18-BamHI-SC2, pRS44::P18-StuI-SC1, pRS44::P18-BsaI-SC1, pRS44::P18-Ppu21-SC1 and pRS44::P18-PCR-F1+2 were cultivated (2.8.2), protein crude extracts were prepared (2.8.3.1) and turnover rates for H_2 uptake activity were investigated as described (2.8.3.4). For no subclone H_2 uptake activity was detectable (data not shown).

3.5 Analysis of twelve further single fosmid clones with putative H_2 uptake activity found by the developed function-based screening

For identifying further fosmid clones with H_2 uptake activity of the constructed metagenomic hydrothermal deep-sea vent fosmid library from the massive sulfide chimney of the black smoker Sisters Peak, a second round of the developed screening procedure (3.3) with the remaining 5,184 clones were conducted by technical assistants. Thereby, 12 further single fosmid clones with putative positive tested H_2 uptake activity were identified (Table 6). These 12 fosmid clones which originated from the same Sisters Peak metagenomic library like the positive clones pRS44::SP-P05F01 and pRS44::SP-P18F11, were further examined as followed:

(1) Repeating the developed functional screen for seeking H_2 uptake active enzymes.
(2) Measuring H_2 uptake turnover rates of *S. oneidensis* Δ*hyaB* clones carrying the 12 identified fosmids.
(3) And sequencing the 5' / 3' insert ends of the 12 identified fosmids with putative positive tested H_2 uptake activity.

(1) The developed functional screen for seeking H_2 uptake active enzymes was repeated at least 3 fold, whereby each repeat contained six or more replica cultures for each examined *S. oneidensis* Δ*hyaB* fosmid clone with putative H_2 uptake activity (Table 6). A color change of the medium indicating an existing H_2 uptake activity couldn't be reliable confirmed for any of the 12 *S. oneidensis* Δ*hyaB* fosmid clones, because of the following observations: Onetime a color change was observable for nearly all parallel replica cultures of 3 examined *S. oneidensis* Δ*hyaB* fosmid clones with putative H_2 uptake activity (pRS44::SP-P11C02, pRS44::SP-P11H04 and pRS44::SP-P31A05), but this observation of a color change of nearly all replica cultures was unrepeatable for these fosmid clones. Instead several times for these three examined fosmid clones and the other investigated 9 fosmid clones a color

change could only be observed for single cultures of the replicas. Accordingly, a positive finding for H_2 uptake activity was unrepeatable by the conducted functional screen. These results indicated that the investigated fosmid clones were perhaps false positive results of the functional screen.

(2) As previously mentioned, a repeatability of positive results was not given for the 12 identified fosmid clones by the developed functional screen. Nevertheless, to validate the results of the function-based screen, the 12 putative positive *S. oneidensis* Δ*hyaB* fosmid clones were additionally analyzed by measuring turnover rates of putative existing H_2 uptake active enzymes encoded on the fosmid inserts. The H_2 uptake activities were determined for the corresponding membrane fractions of *S. oneidensis* Δ*hyaB* fosmid clone cultures. Therefore, corresponding *S. oneidensis* Δ*hyaB* fosmid clones were cultivated under lithotrophic conditions with H_2 as sole energy source and their membrane fractions were prepared as described (2.8.3). In general for all H_2 uptake activity assay measurements only minor yields of cells were available, with the exception of one cell preparation and H2 uptake measurement for the fosmids pRS44::SP-P11C02, pRS44::SP-P11H04 and pRS44::SP-P31A05. The reason for minor cell yields was due to the fact that corresponding *S. oneidensis* Δ*hyaB* fosmid clone cultures were only harvested and used for H_2 uptake activity assay, if a color change and thereby H_2 uptake activity was observable. But as previously mentioned a color change could only be observed for just a few *S. oneidensis* Δ*hyaB* fosmid clone cultures of the parallel cultivated replicas cultures resulting in the described minor yields of cells. All results of the H_2 uptake activity assay measurements for all 12 putative positive *S. oneidensis* Δ*hyaB* fosmid clones are listed in Table 11 and graphically shown in Figure 20.

For some putative positive *S. oneidensis* Δ*hyaB* fosmid clones minor activities on a level of the negative control (H_2 uptake activity of the *S. oneidensis* Δ*hyaB* deletion mutant: 25°C: 2.7 ± 2.2 nmol * min^{-1} * mg^{-1} total protein; 55°C: 9.6 ± 5.6 nmol * min^{-1} * mg^{-1} total protein; 70°C: 16.0 ± 2.3 nmol * min^{-1} * mg^{-1} total protein; Figure 17) were measured at certain assay temperatures (Table 11, Figure 20). The *S. oneidensis* Δ*hyaB* fosmid clones with H_2 uptake activities on the negative control level were pRS44::SP-P30F07 (55°C: 2.4 ± 3.3 nmol * min^{-1} * mg^{-1} total protein), pRS44::SP-P31A05 (55°C: 9.4 ± 7.7 nmol * min^{-1} * mg^{-1} total protein; 70°C: 5.5 ± 4.4 nmol * min^{-1} * mg^{-1} total protein), pRS44::SP-P44E07 (55°C: 8.3 ± 4.0 nmol * min^{-1} * mg^{-1} total protein) and pRS44::SP-P44H11 (25°C: 1.4 ± 0.1 nmol * min^{-1} * mg^{-1} total protein).

Table 11: Summery of H_2 uptake hydrogenase activity measurement results of all 12 identified *S. oneidensis* Δ*hyaB* fosmid clones with putative H_2 uptake activity.

	[nmol × min^{-1} × mg^{-1}]		
	25 °C [a]	**55 °C [a]**	**70 °C [a]**
pRS44::SP-P11C02	14.4 ± 0	157.4 ± 145.7	690.5 ± 56.6
pRS44::SP-P11H04	nd	137.5 ± 149.4	279.5 ± 267.1
pRS44::SP-P30B05	--	49.0 ± 20.6	nd
pRS44::SP-P30F07	nd	2.4 ± 3.3	--
pRS44::SP-P31A02	nd	170.3 ± 170.3	285.1 ± 279.7
pRS44::SP-P31A05	nd	9.4 ± 7.7	5.3 ± 4.4
pRS44::SP-P31B08	nd	nd	--
pRS44::SP-P35F04	nd	13.8 ± 7.8	--
pRS44::SP-P37B01	--	nd	--
pRS44::SP-P37B03	--	nd	--
pRS44::SP-P44E07	--	8.3 ± 4.0	--
pRS44::SP-P44H11	1.4 ± 0.1	25.2 ± 28.3	39.0 ± 29.5

[a] nd = not detectable, -- = no measurement

For the following putative positive *S. oneidensis* Δ*hyaB* fosmid clones no H_2 uptake activity was detectable at certain assay temperatures (Table 11, Figure 20): pRS44::SP-P11H04 (25 °C) pRS44::SP-P30B05 (70 °C) pRS44::SP-P30F07 (25 °C), pRS44::SP-P31A02 (25 °C), pRS44::SP-P31A05 (25 °C), pRS44::SP-P31B08 (25 °C and 55 °C), pRS44::SP-P35F04 (25 °C), pRS44::SP-P37B01 (55 °C) and pRS44::SP-P37B03 (55 °C).

For the following putative positive *S. oneidensis* Δ*hyaB* fosmid clones H_2 uptake activities were detectable at certain assay temperatures but having very high standard deviations indicating a possible missing H_2 uptake activity (Table 11, Figure 20): pRS44::SP-P11H04 (55°C: 137.5 ± 149.4 nmol * min^{-1} * mg^{-1} total protein; 70°C: 279.5 ± 267.1 nmol * min^{-1} * mg^{-1} total protein), pRS44::SP-P30B05 (55°C: 49.0 ± 20.6 nmol * min^{-1} * mg^{-1} total protein) and pRS44::SP-P31A02 (55°C: 170.3 ± 170.3 nmol * min^{-1} * mg^{-1} total protein; 70°C: 285.1 ± 279.7 nmol * min^{-1} * mg^{-1} total protein).

The conducted H_2 uptake activity assays (Table 11, Figure 20) indicated that a H_2 uptake activity was detectable for the 70 °C measurement with an activity of 690.5 ± 56.6 nmol * min^{-1} * mg^{-1} total protein for one of the 12 putative positive *S. oneidensis* Δ*hyaB* fosmid clones (pRS44::SP-P11C02). An important indicator for the assessment of this result was the fact that the measured H_2 uptake activity at 70 °C based only on two data points (n = 2, Figure 20). Additionally for the *S. oneidensis* Δ*hyaB* fosmid clone pRS44::SP-P11C02, a H_2 uptake activity with very high standard deviation was measured at 55 °C (157.4 ± 145.7 nmol * min^{-1} * mg^{-1} total protein) indicating a possible missing H_2 uptake activity. The putative positive *S. oneidensis* Δ*hyaB* fosmid clone pRS44::SP-P11C02 showed minor activities for the 25 °C measurements (25 °C: 14.4 ± 0 nmol * min^{-1} * mg^{-1} total protein), but for the assessment of this result it was important to keep in mind that the measured H_2 uptake activity at 25 °C based only on one data point.

(3) Nevertheless to validate the previous results of (1) repeating the developed functional screen and (2) measuring H_2 uptake turnover rates, the 12 identified *S. oneidensis* Δ*hyaB* fosmid clones with putative H_2 uptake activity were additionally analyzed by sequencing the fosmid insert 5'/3' ends (sequence data .

The sequence analysis of pRS44::SP-P11C02 showed, that the fosmid contained only a 2,000 bp insert with similarities to the *Detaproteobacteria Candidatus* Magnetoglobus multicellularis (pRS44::SP-P11C02 in Figure 20), but no ORF associated with hydrogenase genes could be found (pRS44::SP-P11C02 in Figure 20). Instead an ORF with similarities to an initiator RepB protein was found on the fosmid by a blastx analysis.

For the fosmid pRS44::SP-P11H04 sequencing was performed successfully with the primer for the 3' fosmid insert end (modT7_prom). The sequencing reaction for the 5' fosmid insert end (pCC1FOS_rev) was repeated several times, but each time sequencing reaction was interrupted after approximately 250 bp without any apparent reasons. The fosmid sequence was analyzed by blastx and a detected ORF showed similarities to the *Epsilonproteobacteria Sulfurovum* sp. NBC37-1 (pRS44::SP-P11H04 in Figure 20). The ORF on the fosmid insert showed similarities to an aminotransferase. No ORFs with similarities to hydrogenase genes or genes associated with hydrogenases were found (pRS44::SP-P11H04 in Figure 20).

For the fosmid pRS44::SP-P30B05 no ORF with similarities to hydrogenase genes could be found by sequencing the 5'/3' fosmid insert ends (pRS44::SP-P30B05 in Figure 20). The sequence analysis by blastx showed ORFs with similarities to a hypothetical protein of *Candidatus* Caldatribacterium califoniense, a 2-

methylthioadenine synthetase of the *Firmicutes bacterium* CAG-114 and an epsilon subunit of a DNA polymerase III originated from an uncultured bacterium (pRS44::SP-P30B05 in Figure 20).

Also for the fosmid pRS44::SP-P30F07 no ORF with similarities to hydrogenase genes could be found by sequencing the 5'/3' fosmid insert ends (pRS44::SP-P30F07 in Figure 20). Sequencing was performed with the primer for the 5' fosmid insert end (pCC1FOS_rev) and an ORF with similarities to a hypothetical protein of *Bacillus atrophaeus* was found by blastx analysis (pRS44::SP-P30F07 in Figure 20). For the sequence result with the primer for the 3' fosmid insert end (modT7_prom), no sequence similarity was found by blastx analysis to other database entries (pRS44::SP-P30F07 in Figure 20).

The sequence analysis of pRS44::SP-P31A02 showed, that the fosmid contained a 1,500 bp insert and on this insert an ORF with similarities to an hypothetical protein of the Euryarchaeota *Thermococcus kodakarensis* KOD1 was found by blastx analysis (pRS44::SP-P31A02 in Figure 20), but no ORF associated with hydrogenase genes (pRS44::SP-P31A02 in Figure 20).

For the fosmid pRS44::SP-P31A05 no ORF with similarities to hydrogenase genes could be found by sequencing the 5'/3' fosmid insert ends (pRS44::SP-P31A05 in Figure 20). The sequence analysis by blastx showed ORFs with similarities a trka-c domain-containing protein of *Sulfuricurvum kujiense* and an outer membrane efflux protein of *Sulfurovum* sp. NBC37-1 (pRS44::SP-P31A05 in Figure 20). Both organisms contain to the same class of *Proteobacteria*: the *Epsilonproteobacteria*.

No ORF with similarities to hydrogenase genes could be found by sequencing the 5'/3' fosmid insert ends of pRS44::SP-P31B08 (pRS44::SP-P31B08 in Figure 20). The sequence analysis by blastx showed ORFs with similarities a regulatory protein LacI of *Ammonifex degensii* and a putative replication protein of *Clostridium tetani* (pRS44::SP-P31B08 in Figure 20). Both organisms contain to the Firmicutes.

Also for the fosmid pRS44::SP-P35F04 no ORF with similarities to hydrogenase genes could be found by sequencing the 5' fosmid insert ends (pRS44::SP-P35F04 in Figure 20). For the sequence result with the primer for the 5' fosmid insert end (pCC1FOS_rev), no sequence similarity was found by blastx analysis to other database entries (pRS44::SP-P35F04 in Figure 20). The sequencing reaction for the 3' fosmid insert end (modT7_prom) was repeated several times, but each time sequencing reaction was interrupted after several bp without any apparent reasons.

For the fosmid pRS44::SP-P37B01 no ORF with similarities to hydrogenase genes could be found by sequencing the 5' fosmid insert ends (pRS44::SP-P37B01 in Figure 20). For the sequence result with the primer for the 5' fosmid insert end (pCC1FOS_rev), the blastx analysis showed an ORF with similarities to a hypothetical protein of the Firmicutes *Tepidanaerobacter acetatoxydans* (pRS44::SP-P37B01 in Figure 20). The sequencing reaction for the 3' fosmid insert end (modT7_prom) was repeated several times, but each time sequencing reaction was interrupted after several bp without any apparent reasons.

Also for the fosmid pRS44::SP-P37B03 no ORF with similarities to hydrogenase genes could be found by sequencing the 5' fosmid insert ends (pRS44::SP-P37B03 in Figure 20). For the sequence result with the primer for the 5' fosmid insert end (pCC1FOS_rev), the blastx analysis showed an ORF with similarities to an ABC transporter of the *Betaproteobacteria Cupriavidus basilensis* (pRS44::SP-P37B03 in Figure 20). The sequencing reaction for the 3' fosmid insert end (modT7_prom) was repeated several times, but each time sequencing reaction was interrupted after several bp without any apparent reasons.

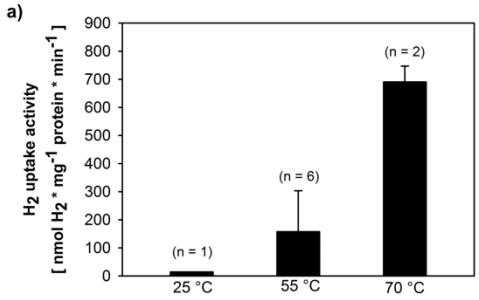
For the fosmid pRS44::SP-P44E07 no ORF with similarities to hydrogenase genes could be found by sequencing the 5'/3' fosmid insert ends (pRS44::SP-P44E07 in Figure 20). For the sequence result with the primer for the 5' fosmid insert end (pCC1FOS_rev), the blastx analysis showed an ORF with similarities to a hypothetical protein of the *Betaproteobacteria Burkholderia* sp. JPY347 (pRS44::SP-P44E07 in Figure 20). For the sequence result with the primer for the 3' fosmid insert end (motT7_prom), no sequence similarity was found by blastx analysis to other database entries (pRS44::SP-P44E07 in Figure 20).

The sequence analysis of pRS44::SP-P44H11 showed, that the fosmid contained a 1,272 bp insert and on this insert an ORF with similarities to an hypothetical protein of the *Epsilonproteobacteria Sulfurimonas* sp. AST-10 was found by blastx analysis (pRS44::SP-P44H11 in Figure 20), but the ORF is not associated with hydrogenase genes.

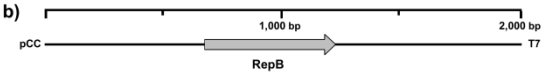
The sequencing results confirmed that no ORF with similarities to hydrogenase associated genes could be found on the examined fosmids.

pRS44::SP-P11C02

a)



b)

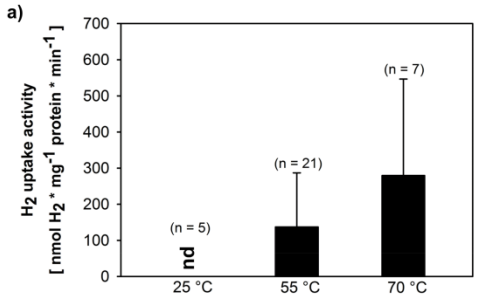


c)

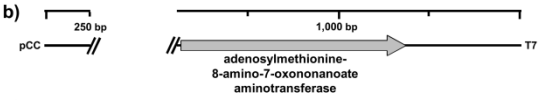
Description	Max score	Total score	Query cover	E-value	Identity
initiator RepB protein [*Candidatus* Magnetoglobus multicellularis str. Araruama]	154	154	29 %	2E-37	40 %

pRS44::SP-P11H04

a)



b)

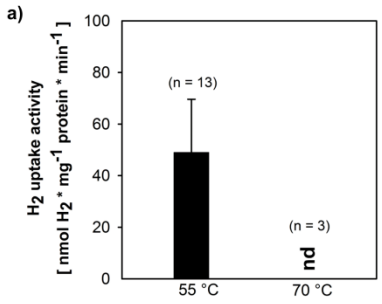


c)

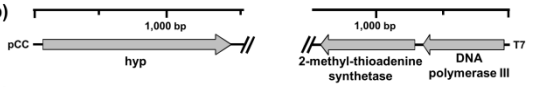
Description	Max score	Total score	Query cover	E-value	Identity
5' insert end pCC:					
No significant similarity found					
3' insert end T7:					
adenosylmethionine-8-amino-7-oxononanoate aminotransferase [*Sulfurovum* sp. NBC37-1]	160	241	47 %	2E-54	61 %

pRS44::SP-P30B05

a)



b)

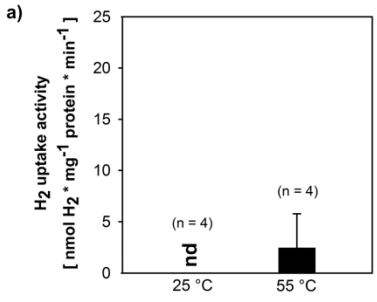


c)

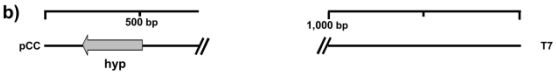
Description	Max score	Total score	Query cover	E-value	Identity
5' insert end pCC:					
hypothetical protein [*Candidatus* Caldatribacteirum californiense]	284	499	85 %	9E-132	60 %
3' insert end T7:					
DNA polymerase III, epsilon subunit [uncultured bacterium]	241	241	41 %	1E-72	56 %
2-methylthioadenine synthetase [*Firmicutes bacterium* CAG:114]	135	135	45 %	3E-32	37 %

pRS44::SP-P30F07

a)



b)

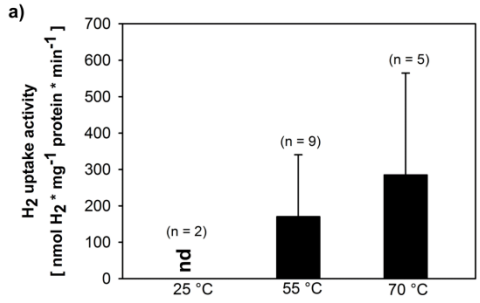


c)

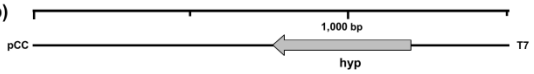
Description	Max score	Total score	Query cover	E-value	Identity
5' insert end pCC:					
hypothetical protein [*Bacillus atrophaeus*]	38	38	38 %	2.9	24 %
3' insert end T7:					
No significant similarity found					

pRS44::SP-P31A02

a)



b)

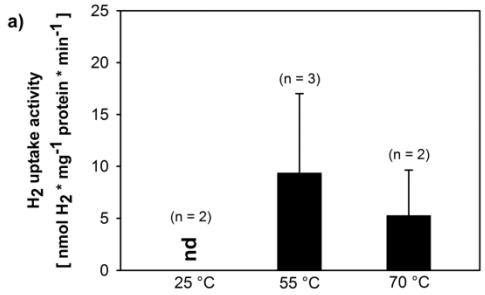


c)

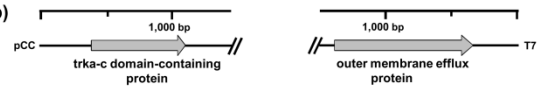
Description	Max score	Total score	Query cover	E-value	Identity
hypothetical protein TK0859 [*Thermococcus kodakarensis* KOD1]	110	110	31 %	4E-25	36 %

pRS44::SP-P31A05

a)



b)

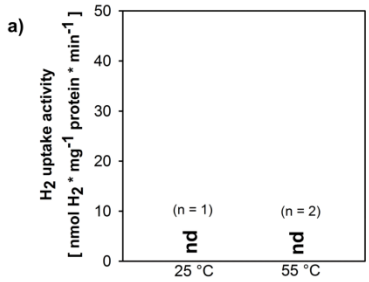


c)

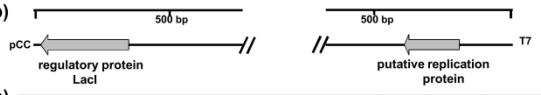
Description	Max score	Total score	Query cover	E-value	Identity
5‘ insert end pCC:					
trka-c domain-containing protein [*Sulfuricurvum kujiense* DSM 16994]	178	178	46 %	6E-49	44 %
3‘ insert end T7:					
outer membrane efflux protein [*Sulfurovum* sp. NBC37-1]	338	338	69 %	1E-107	53 %

pRS44::SP-P31B08

a)



b)

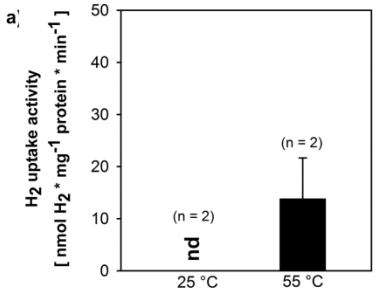


c)

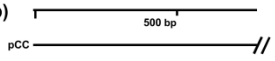
Description	Max score	Total score	Query cover	E-value	Identity
5‘ insert end pCC:					
regulatory protein LacI [*Ammonifex degensii*]	114	114	42 %	5E-26	50 %
3‘ insert end T7:					
putative replication protein [*Clostridium tetani*]	38.9	38.9	13 %	3.1	57 %

pRS44::SP-P35F04

a)



b)

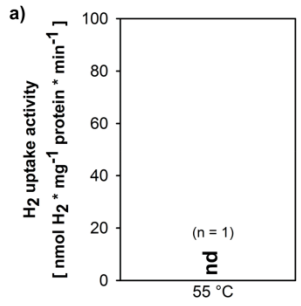


c)

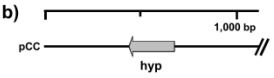
Description	Max score	Total score	Query cover	E-value	Identity
5‘ insert end pCC:					
No significant similarity found					

pRS44::SP-P37B01

a)



b)

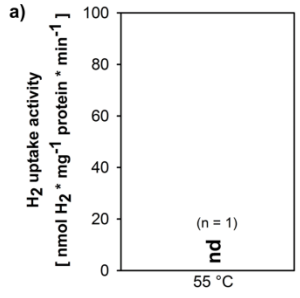


c)

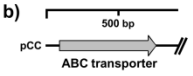
Description	Max score	Total score	Query cover	E-value	Identity
5‘ insert end pCC:					
hypothetical protein [*Tepidanaerobacter acetatoxydans*]	40.4	40.4	21 %	1.8	32 %

pRS44::SP-P37B03

a)



b)



c)

Description	Max score	Total score	Query cover	E-value	Identity
5‘ insert end pCC:					
ABC transporter [*Cupriavidus basilensis*]	273	347	73 %	3E-87	73 %

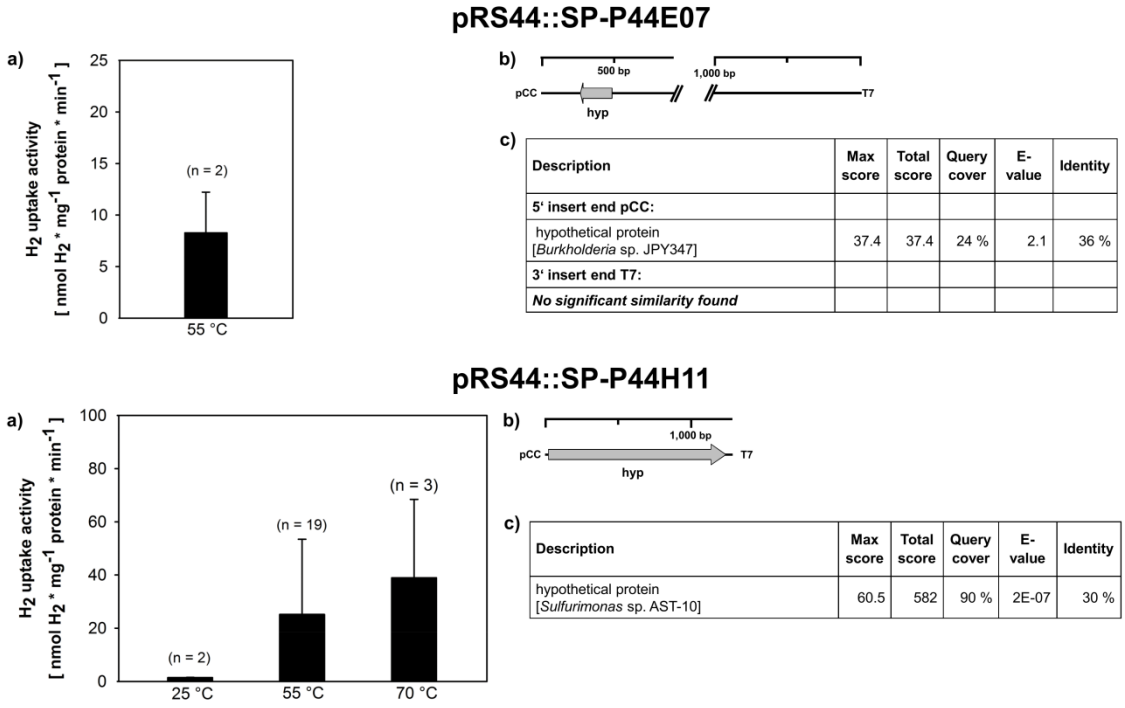


Figure 20: Graphic representation of H_2 uptake activity measurements and sequence analyses of all 12 identified *S. oneidensis* Δ*hyaB* fosmid clones with putative H_2 uptake activity. In this figure for each examined fosmid clone with putative H_2 uptake activity, the sequencing- and H_2 uptake activity assay results are summarized and listed in three separated parts: (a) H_2 uptake activity measurement results for prepared membrane fractions of *S. oneidensis* Δ*hyaB* cultures containing a corresponding fosmid, which were cultivated under lithotrophic conditions with H_2 as sole energy source. The H_2 uptake activities were measured spectrophotometrically by reduction of the artificial electron donor methyl viologen (from colorless to blue) at different temperatures. The numbers of performed measurements (n) are denoted above the bars. The exact values of all measurement results with standard deviations are listed in Table 11. (b) Schematic representation of the sequenced fosmid ends. The arrows represent a putative ORF and show the direction of the supposed gene transcription. (c) Blastx result of the detected ORFs. Listed are the best hits of a blastx search against the NCBI nr database. The identity based on amino acid similarities.

4 Discussion

4.1 Development of a function-based screening for seeking hydrogenase enzymes in metagenomic libraries

The primary aim of this study was to identify and study novel hydrogenase enzymes. Due to the large potential of hydrothermal deep-sea vent habitats for recovering novel hydrogenases and the powerful tools of metagenomics to gain access to the genetics of uncultured organisms to identify new enzymes, a function-based screening approach for metagenomic libraries to seek hydrogenases was developed. This new function-based screen was applied to investigate metagenomic broad-host range fosmid libraries from hydrothermal deep-sea vent habitats to identify corresponding hydrogenases encoded on the examined metagenomic material.

The screening was realized by using a generated *S. oneidensis* Δ*hyaB* deletion mutant as new heterologous host for metagenomic libraries. *S. oneidensis* MR-1 was chosen for several reasons:

(A) In the genome of *S. oneidensis* MR-1, only a sole [NiFe]-hydrogenase can be found, which is responsible for the capability to oxidize H_2. Only one directed deletion in an essential gen region of the [NiFe]-hydrogenase large structural subunit (e.g. the region where the [NiFe] metal center will be incorporated) is needed to obtain a mutant without H_2 uptake capabilities.

(B) A second putative hydrogenase, a [FeFe]-hydrogenase, was identified in the genome of *S. oneidensis* MR-1, but this hydrogenase did not seem to have any negative influences on the developed function based screening method (3.1.3, 3.1.5). An involvement on the metal reducing capability of *S. oneidensis* MR-1 is not described for the putative [FeFe]-hydrogenase in literature (Lovley *et al.*, 1989; Petrovskis *et al.*, 1994; Dawood and Brozel, 1998; Heidelberg *et al.*, 2002; Meshulam-Simon *et al.*, 2007) and could also not observed in this PhD study (3.1.3, 3.1.5).

(C) The functional screening should especially allow the detection of metagenomic hydrogenases originated from *Epsilonproteobacteria,* because they seem to be predominant hydrogen consumers in hydrothermal deep-sea vent habitats. Heterologous expression of metagenomic hydrogenases in a foreign host can be

difficult because several proteins are involved in assembly, maturation and activation of [NiFe]-hydrogenases. *S. oneidensis* MR-1 was chosen as new heterologous host for the functional screening due to their similarities of the [NiFe]-hydrogenase to those from *Epsilonproteobacteria*. Therefore, it could be assumed that the essential assembly and maturation apparatus of *S. oneidensis* MR-1 may be compatible for maturing hydrogenases from certain *Epsilonproteobacteria*.

(D) The *Gammaproteobacteria S. oneidensis* MR-1 is a metabolically versatile bacterium which can be cultivated easily under aerobic or anaerobic conditions. The organism can use diverse organic components or H_2 as an electron donor to gain energy required for its growth.

(E) Furthermore, this organism can be easily genetically manipulated by the tools developed for *E. coli*, for example *S. oneidensis* MR-1 can be transformed by pUC-type universal vectors (Sybirna *et al.*, 2008).

(F) Finally, *S. oneidensis* MR-1 is a “Dissimilatory Metal Reducing Bacteria” (DMRB) and can reduce Fe(III)citrate to Fe(II)citrate by utilization of H_2 as an electron donor under anaerobic conditions.

The described capability of *S. oneidensis* MR-1 to reduce metals by using H_2 as electron source was exploited to develop a function-based screening to seek hydrogenases as follows. The *S. oneidensis* MR-1 wt could reduce the yellow Fe(III)citrate to the colorless Fe(II)citrate by utilization of H_2 as an electron donor (3.1.3). This correlates with the observation of several previous studies, which describe that *S. oneidensis* MR-1 can utilize H_2 as an electron donor for metal reduction under anaerobic conditions (Dawood and Brozel, 1998; Myers and Myers, 2001; Liu *et al.*, 2002; De Windt *et al.*, 2005; Marshall *et al.*, 2008). In contrast for the generated *S. oneidensis* Δ*hyaB* deletion mutant no color change could be noticed under the same culture conditions (3.1.3). *S. oneidensis* Δ*hyaB* was complemented with a fosmid containing the [NiFe]-hydrogenase operon of the *S. oneidensis* MR-1 wt (pRS44::hyaB). This complemented mutant showed a restored ability for reducing the yellow Fe(III)citrate to the colorless Fe(II)citrate under anaerobic culture conditions with H_2 as sole energy source due to the reconstitution of an active [NiFe]-hydrogenase (3.1.3). This observation indicated that the utilization of H_2 catalyzed by the [NiFe]-hydrogenase is directly linked to the metal reducing capabilities of *S. oneidensis.* Additionally, the results in this study illustrated that an active [NiFe]-hydrogenase is directly detectable by the metal reducing capabilities of *S. oneidensis* due to an observed color change of the medium from the yellow Fe(III)citrate to the colorless Fe(II)citrate. Accordingly, a function-based screen for identifying hydrogenase enzymes encoded on metagenomic libraries could be realized

by a simple and directly optical detection of an existing or missing color change from yellow to colorless in serum bottle cultures without needing any additional tools.

The capability to detect hydrogenase enzymes by the developed function-based screen due to a color change of the medium was verified by analyzing [NiFe]-hydrogenase activity for *S. oneidensis* MR-1 wt, *S. oneidensis* Δ*hyaB* and *S. oneidensis* Δ*hyaB* complemented with pRS44::hyaB. Therefore, H_2 consumption rates, H_2 uptake activities and H_2 evolution rates were measured (3.1.5).

The H_2 consumption rate measurements confirmed the missing ability of *S. oneidensis* Δ*hyaB* to utilize H_2 and thereby the ability to reduce Fe(III)citrate in the medium (3.1.5). The measurements of *S. oneidensis* Δ*hyaB* complemented with the fosmid pRS44::hyaB showed a reestablished ability to consume H_2 (3.1.5). This corresponds to results in previous studies (Meshulam-Simon *et al.*, 2007; Marshall *et al.*, 2008; Shi *et al.*, 2011) and demonstrates that a functional [NiFe]-hydrogenase enables the consumption of H_2 and thereby the reduction of Fe(III)citrate.

The H_2 uptake activities and H_2 evolution rates of *S. oneidensis* MR-1 wt, *S. oneidensis* Δ*hyaB* and *S. oneidensis* Δ*hyaB* complemented with pRS44::hyaB (pRS44::hyaB = a *S. oneidensis* MR-1 genome fragment of ~35,000 bp including the whole [NiFe]-hydrogenase operon in pRS44) were quantified by spectrophotometric measurements or by a gas chromatograph using a prepared protein crude extracts of the *S. oneidensis* strains (3.1.5). In a previous study, the [NiFe]-hydrogenase of *S. oneidensis* MR-1 wt was purified and described as non-membrane bound periplasmatic enzyme (Shi *et al.*, 2011). Despite comparable investigations and methods, the assumption of Shi and colleagues (2011) about the [NiFe]-hydrogenase location is fundamentally opposed to the observed results of this PhD study. In contrast to the study of Shi *et al.* (2011), the results of this PhD study indicates that the [NiFe]-hydrogenase and therefore the H_2 uptake activity was always located in the *S. onsidensis* membrane fraction of prepared protein crude extracts (3.1.5 and 3.4.3) and no hydrogenase activity could be observed in the soluble fraction. The observation of this PhD study correlated with results of the program PSORTb v.3.0 (Gardy *et al.*, 2005). PSORTb is a program for computational prediction of subcellular location of proteins. Based on the amino acid sequences of the *S. oneidensis* MR-1 [NiFe]-hydrogenase, the PSORTb program predicted the location of the large-, small- and cytochrome b subunits in the cytoplasmatic membrane of *S. oneidensis* (Appendix 9.6). However, the question still remains: Why Shi and colleagues (2011) described the purified *S. oneidensis* [NiFe]-hydrogenase as non-membrane bound periplasmatic enzyme? It's possible that the location at Shi *et al.* (2011) was influenced by the added His-Tag to the N terminus of the large structural

subunit (*hyaB*) of the [NiFe]-hydrogenase? This perhaps resulted in the missed ability to anchor the large structural subunit together with the cytochrome B subunit in the cytoplasmatic membrane. Nevertheless, without further experiments or data from Shi and colleagues (2011), the fundamental different results in the [NiFe]-hydrogenase location of Shi *et al.* (2011) and the contrast to the results of this PhD study cannot be answered unequivocally.

The H_2 evolution rates were quantified by a gas chromatograph (2.8.5) for *S. oneidensis* MR-1 wt, *S. oneidensis* Δ*hyaB* and *S. oneidensis* Δ*hyaB* complemented with pRS44::hyaB (pRS44::hyaB = a *S. oneidensis* MR-1 genome fragment of ~35,000 bp including the whole [NiFe]-hydrogenase operon in pRS44). Therefore, the produced H_2 was measured using completely reduced methyl viologen as electron donor, whereby the reaction was catalyzed by added prepared protein crude extract fractions of the *S. oneidensis* strains (2.8.5). The detected H_2 evolution rates for the membrane protein crude extract fractions of *S. oneidensis* MR-1 wt and *S. oneidensis* Δ*hyaB* complemented with pRS44::hyaB illustrated that the [NiFe]-hydrogenase functions bidirectional (3.1.5). However, the measurements of membrane- and soluble protein crude extract fractions originated from *S. oneidensis* Δ*hyaB* showed no detectable H_2 production (3.1.5).

The result of no detectable H_2 production of *S. oneidensis* Δ*hyaB* disagree with the result of Meshulam-Simon *et al.* (2007), where a reduced but an existing H_2 formation rate could be measured for a *S. oneidensis* [NiFe]-hydrogenase deletion mutant. There are various possible reasons for this discrepancy. Indeed, in the study of Meshulam-Simon *et al.* (2007) the H_2 evolution rates were determined by a completely different method. The H_2 formations were calculated in vivo by measuring H_2 concentration shifts in the headspace of corresponding *S. oneidensis* cultures with a gas chromatograph. The cultures were grown anaerobically in minimal medium supplemented with the electron donor pyruvate or formate and the electron acceptor fumarate. Meshulam-Simon *et al.* (2007) postulated that the [NiFe]-hydrogenase provide the dominant hydrogenase activity for H_2 evolving in *S. oneidensis* MR-1 and that the residual H_2 formation of their *S. oneidensis* [NiFe]-hydrogenase deletion mutant originated from the putative [FeFe]-hydrogenase, identified on the genome of *S. oneidensis* MR-1 (Heidelberg *et al.*, 2002; Meshulam-Simon *et al.*, 2007; Marshall *et al.*, 2008; Shi *et al.*, 2011).

Other possible reasons for the missing H_2 formation for *S. oneidensis* Δ*hyaB* in this PhD study in contrast to the result of Meshulam-Simon *et al.* (2007) may originate from the incompatibility of the putative [FeFe]-hydrogenase to use reduced methyl viologen as an electron donor for H_2 formation. On the other hand, it might derive from a non-expressed

functional [FeFe]-hydrogenases under the selected culture conditions of *S. oneidensis* Δ*hyaB.* For the [FeFe]-hydrogenase it is postulated that required maturation enzymes are expressed and active under culture conditions with TMAO, DMSO or thiosulfate as sole electron donor (Sybirna *et al.*, 2008). In this study, such culture conditions were not chosen for generating *S. oneidensis* Δ*hyaB* cell material to conduct the H_2 evolution measurements. Perhaps therefore, no functional [FeFe]-hydrogenase could be measured with material of the *S. oneidensis* Δ*hyaB*.

The H_2 uptake activity rates for *S. oneidensis* MR-1 wt, *S. oneidensis* Δ*hyaB* and *S. oneidensis* Δ*hyaB* complemented with pRS44::hyaB (pRS44::hyaB = a *S. oneidensis* MR-1 genome fragment of ~35,000 bp including the whole [NiFe]-hydrogenase operon in pRS44) were quantified spectrophotometrically by a H_2 consumption assay (3.1.5). Results of previous studies showed that the [NiFe]-hydrogenase of *S. oneidensis* MR-1 is responsible for the H_2 uptake activity of the organism (Meshulam-Simon *et al.,* 2007; Marshall *et al.,* 2008). Corresponding H_2 uptake activity measurements in this PhD study confirmed the assumption of Meshulam-Simon et al. (2007) and Marshall et al. (2008). As expected, for the *S. oneidensis* Δ*hyaB* deletion mutant no H_2 uptake activity could be measured (3.1.5). Furthermore, a complementation of *S. oneidensis* Δ*hyaB* with a functional *S. oneidensis* MR-1 wild type [NiFe]-hydrogenase encoded on the fosmid pRS44::hyaB resulted in a restored H_2 uptake activity (3.1.5). These results of the PhD study confirmed the essential role of the [NiFe]-hydrogenase for the H_2 uptake activity in *S. oneidensis* MR-1. Additionally, the successful complementation of *S. oneidensis* Δ*hyaB* indicated that a complementation with a foreign H_2 uptake hydrogenase might be possible, which was an important requirement for the development of a function-based screening approach to identify and study novel hydrogenase enzymes encoded on metagenomic libraries.

The measurements for analyzing H_2 consumption rates, H_2 uptake activity and H_2 formation of *S. oneidensis* MR-1 wt, *S. oneidensis* Δ*hyaB* and *S. oneidensis* Δ*hyaB* complemented with pRS44::hyaB demonstrated, that the new developed activity-based screen worked and has a high potential to identify enzymes with H_2 uptake activity. In summary, the developed function-based screen was realized by the observation that the metal reducing capabilities of the constructed *S. oneidensis* Δ*hyaB* deletion mutant are restored, if functional hydrogenases are encoded on (metagenomic-) fosmids and are heterologously expressed in the surrogate host. The screening is easy to implement by anaerobic cultivation of *S. oneidensis* Δ*hyaB* clones in minimal medium supplemented with Fe(III)citrate and H_2 as sole energy source. Existing H_2 uptake activities are simply observable by a color change of the medium from yellow to colorless, due to the

reduction of Fe(III)citrate to Fe(II)citrate in corresponding *S. oneidensis* Δ*hyaB* cultures with functional H_2 uptake hydrogenases encoded on the (metagenomic-) fosmids.

4.2 Detection of H_2 uptake activities by the developed functional screening method

By using the developed function-based screening method, the constructed metagenomic fosmid library from the hydrothermal vent massive sulfide chimney of the black smoker Sisters Peak was examined for H_2 uptake hydrogenase activity. Initially, 2,592 clones were screened for H_2 uptake hydrogenase activity in the first round. With ca. one and a half year time distance, the remaining 5,184 clones were analyzed in a second round for H_2 uptake hydrogenase activity. The identified clones with putative H_2 uptake activity from the first- and the second round of the screening process were analyzed independently, due to the large time period between their identification.

During the first screening round with 2,592 clones, six clone pools containing 96 single clones (accordingly all clones of a positive tested 96 well microtiter plate) were identified by exhibiting the expected color change in the medium. Each clone pool might include one or more putative H_2 uptake active clones (clone pool 05, 11, 18 20, 23 and 24) (3.3). However, only two of these 96 clone pools (clone pool 05 and 18), active fosmids with supposed recombinant hydrogenase activity in the *S. oneidensis* Δ*hyaB* deletion mutant (pRS44::SP-P05F01 (P05F01) and pRS44::SP-P18F11 (P18F11)) could be identified (3.3).

In the second round of the functional screen with 5,184 clones, five further 96 clone pools were identified by observing the expected color and each clone pool might include one or more putative H_2 uptake active clones with H_2 uptake activity (clone pool 30, 31, 35, 37 and 44) (3.3). From these five clone pools, 10 putative active single fosmid clones were identified (P30B05, P30F07, P31A02, P31A05, P31B08, P35F04, P37B01, P37B03, P44E07 and P44H11) (3.3). Additionally, for one clone pool (clone pool 11) identified in the first round of the screening, two further putative active single fosmid clones were identified (P11C02 and P11H04) (3.3).

With the exception of the clone pools 05 and 18 and the corresponding fosmids P05F01 and P18F11, the results of the functional screen were ambiguous for the other clone pools and the identified putative active single fosmid clones during the second round of the functional screen. The possibility to detect false positive fosmids without

hydrogenase activity was assumed during performing a functional screen, but not in such large numbers as here observed.

For the clone pools 20, 23 and 24, no clear results were observed during the function-based screen for seeking putative single *S. oneidensis* Δ*hyaB* fosmid clones with a restored hydrogenase activity. Instead no color change for any clone of the pools or the exact opposite, a color change for nearly each determined clone of the 96 clone pools, was noticed.

The 12 identified *S. oneidensis* Δ*hyaB* fosmid clones with putative H_2 uptake activity, identified in the second round of the function-based screen were analyzed by repeating the screening process several times. Nevertheless, reliable and repeatable results have not been observed for any of these fosmids. In some cases it was observed that the half of all biological replicates change the color during the function-based screen, but the residual replicates stayed yellow. In other cases it could be observed that all parallel cultures stayed yellow. Therefore an existing H_2 uptake activity mediated by the encoded genes on the fosmids could not be predicted reliably for any of the 12 clones by performing the functional screen.

These results have illustrated the challenges of developing a functional screen. On one hand, a screen has to be easily performable in addition to exposing clear results. In principal both were given for the developed function-based screening method for seeking H_2 uptake activity and were proven by the successful identification of H_2 uptake activity for the fosmids P05F01 and P18F11. On the other hand, for a useable functional screen, the repeatability of their results must be given to verify positive findings. Therefore, it has to be important to minimize disturbing influences on the screening process. Consequently, the screen has to be implemented every time with identical parameters to ensure it is always comparable and has to be accomplished with great care. The described problem of missing repeatability was observed for the putative H_2 uptake activity of the 12 fosmids identified in the second round of the functional screen. To solve this described problem, the developed functional screen has to be modified by identifying possible reasons for the false positive hits to improve the repeatability of the screen in the future. The following points were recognized to possibly improve the functional screen:

Firstly, to exclude unwanted organic matter, a high accuracyin preparing the used minimal medium has to be ensured. It is shown that *S. oneidensis* can utilize many fermentative end products (e.g. lactate, acetate, pyruvate, formate, succinate) to gain energy (Petrovskis *et al.*, 1994; Dawood and Brozel, 1998; Heidelberg *et al.*, 2002; Serres and Riley, 2006). For the same reason the inoculated *S. oneidensis* cells seem

to be another origin for these fermentative end products. Pre-cultures of the investigated *S. oneidensis* cells are cultivated in LB medium before they are inoculated into the minimal medium of the functional screen. Therefore, in the *S. oneidensis* cell pellets have to be washed thoroughly to remove residual alternative energy sources originated from LB medium. Due to the possibility to reduce Fe(III)citrate also by utilization of many fermentative end products (e.g. lactate, acetate, pyruvate, formate, succinate) as alternative electron donors (Petrovskis *et al.*, 1994; Dawood and Brozel, 1998), unwanted organic matter in the prepared minimal medium and in the washed *S. oneidensis* cells for inoculation can efficiently inhibit the functional screen by capturing false positive results.

In addition, comparable low cell densities in the replicated *S. oneidensis* cultures can improve their repeatability during the implementation of the functional screening method. Consequently, washed *S. oneidensis* cells which are examined for H_2 uptake activity have to be adjusted to comparable cell numbers by measuring their optical density. During the function-based screening, the same volume of these prepared *S. oneidensis* cell solutions always has to be inoculated to the identical volume minimal medium resulting in replica cultures with comparable cell densities, controlled by measuring optical density.

Furthermore, it has to be ensured that the *S. oneidensis* cells in the functional screen are cultivated under O_2 deficient conditions. The capabilities of hydrogenases to utilize H_2 and the capabilities for dissimilatory metal reductions are inhibited in *S. oneidensis* cultures with too high O_2 concentrations in the minimal medium and / or headspace of the serum bottles (Dawood and Brozel, 1998; Myers and Myers, 2001; Liu *et al.*, 2002; De Windt *et al.*, 2005; Marshall *et al.*, 2008). An excessive O_2 concentration will result in the missing ability to reduce Fe(III)citrate and consequently results consequently in an unfeasible functional screen. Thus, the O_2 concentration in the medium and the serum bottle headspace has to be as low as possible. Unfortunately, the usage of reduction agents to reduce residual O_2 is excluded by the negative influences on the iron, which disturbs the functional screen (e.g. Fe(III)citrate is reduced by Na_2S to Fe(II)sulfide visible as a black precipitate). Hence, to obtain a working functional screen, prepared hot medium has to cool down to room temperature under purging with N_2 gas for at least a one hour to minimize the contained O_2. The usage of gas tight serum bottles with an intact septum and their filling with prepared minimal medium under a N_2 atmosphere ensures low O_2 concentrations in the minimal medium and serum bottle headspace.

The repeatability of the developed functional screen for H_2 uptake activity can be improved, if all these described improvements relating to the implementation of the screen are adhered to.

4.2.1 Analysis of the putative H_2 uptake active fosmid clones pRS44::SP-P05F01 and pRS44::SP-P18F11

The two identified fosmids pRS44::SP-P05F01 (P05F01) and pRS44::SP-P18F11 (P18F11) were investigated on the hydrogenase activity level in detail and were completely sequenced (3.4). Overall, the successful identification of the two fosmids P05F01 and P18F11, their exposed sequence and their hydrogenase activity analysis suggested, that the developed functional screen for detecting recombinant H_2 uptake active enzymes works. This is shown by the following observed restored abilities of complemented *S. oneidensis* Δ*hyaB* deletion mutant: (1) The ability to reduce the yellow Fe(III)citrate to the colorless Fe(II)citrate by utilization of H_2. (2) The capacity to consume H_2 under defined culture conditions. (3) The capability to oxidize H_2 during an H_2 uptake activity assay. (4) Finally, the ability to produce H_2 during an H_2 evolution assay.

4.2.1.1 Fosmid pRS44::SP-P05F01

The ORFs encoded on the insert DNA of the fosmid P05F01 resemble structural genes and genes encoding associated maturation proteins of a H_2 uptake [NiFe]-hydrogenase with similarities (82 % – 88 % identity, Figure 14, Table 7) to corresponding genes of *N. salsuginis* DSM 16511. *N. salsuginis* is an *Epsilonproteobacteria* of the Group F (classification according to Corre *et al.*, 2001). Only three cultured representatives of this group are described, which have contributed to our understanding of the physiological potential of this lineage. Besides *N. salsuginis*, the other two cultivated Group F affiliates are *Sulfurovum* sp. NBC37-1 and *Sulfurovum lithotrophicum* (Inagaki *et al.*, 2004; Nakagawa *et al.*, 2005; Nakagawa *et al.*, 2007). While *N. salsuginis* exclusively uses H_2 as an electron donor for its energy metabolism, *Sulfurovum* sp. NBC37-1 can additionally utilize oxidized sulfur compounds (Nakagawa *et al.*, 2005; Nakagawa *et al.*, 2007). In contrast, *S. lithotrophicum* cannot utilize H_2, but can utilize oxidized sulfur compounds (Inagaki *et al.*, 2004). Although the Group F lineage has only three cultivated representatives to date, a tremendous number of environmental 16S rRNA gene sequences are affiliated with this group, which appear widespread in hydrothermal deep-sea vent habitats (e.g., Corre *et al.*, 2001; Takai *et al.*, 2003; Perner *et al.*, 2011a) and for which the metabolic properties remain unknown. The identified environmental H_2

uptake hydrogenase, encoded on P05F01 and putatively from an *Epsilonproteobacteria* of the Group F, can consolidate the thesis of the importance that H_2 metabolism has for this group.

The DNA insert of the fosmid P05F01 has ORFs resembling for two [NiFe]-hydrogenases structural gene clusters with similarities to corresponding genes of *N. salsuginis* DSM 16511 (3.4.1, Table 7, Figure 14) (Anderson *et al.*, 2011). Each of the two gene clusters includes small and large subunits of the H_2 uptake [NiFe]-hydrogenase and a [NiFe]-hydrogenase cytochrome B subunit. Additionally, a further ORF is similar to the gene encoding for the small subunit of *N. salsuginis* H_2 uptake [NiFe]-hydrogenase. However, the genome of *N. salsuginis* contains only one [NiFe]-hydrogenase operon which includes one set of structural genes with a small subunit, a large subunit and a cytochrome B subunit of a H_2 uptake [NiFe]-hydrogenase (Anderson *et al.*, 2011). The discrepancy between the number of gene clusters found on the fosmid P05F01 and the genome of *N. salsuginis* DSM 16511 may be explained by one of the three following hypothesis: (1) The fosmid DNA insert may originate from an organism that has acquired the gene clusters from *N. salsuginis* by horizontal gene transfer twice: Multiple occurrence of [NiFe]-hydrogenases gene clusters is not unusual (found e.g. in the *Epsilonproteobacteria Sulfurovum* sp. NBC37-1 or *Nautilia profundicola*) and it has been postulated that multiple hydrogenases may reflect adaption's to variations in the concentration of H_2 and electron acceptors in the constantly changing deep-sea hydrothermal vents habitats (Nakagawa *et al.*, 2007; Campbell *et al.*, 2009). (2) Alternatively, *N. salsuginis* may have lost one [NiFe]-hydrogenase structural gene cluster over time. (3) Otherwise, the multiple gene clusters on the fosmid P05F01 originate from a bias due to the performed whole genome amplification (WGA) for generating high yield DNA in order to construct the metagenomic fosmid library. The WGA based on multiple displacement amplification (MDA) and is characterized by a high processivity and high strand displacement activity of the phi29 polymerase. For instance, this method still induces phi29 copy number biases (Cardoso *et al.*, 2004; Arriola *et al.*, 2007; Pugh *et al.*, 2008) and this bias may be a reason for the discrepancy.

The measured recombinant H_2 uptake activity of *S. oneidensis* Δ*hyaB* complemented with the fosmid P05F01 (219.5 ± 54.4 nmol * min^{-1} * mg^{-1} total protein at 28 °C, 2,677.0 ± 941.4 nmol * min^{-1} * mg^{-1} total protein at 55°C and 12,001.5 ± 1035.1 nmol * min^{-1} * mg^{-1} total protein at 70°C) corresponds to the H_2 uptake activity measured for *N. salsuginis* (850 ± 125 nmol * min^{-1} * mg^{-1} total protein at 37 °C) (Takai *et al.* 2005) when considering the different temperatures of performed enzyme assays. However, the comparable H_2 uptake activities of the hydrogenase on the fosmid P05F01 to the

hydrogenase described for *N. salsuginis* showed the feasibility to express successfully active epsilonproteobacterial H_2 uptake hydrogenase in the *S. oneidensis* Δ*hyaB* host. With respect to hydrogenase phylogeny, the hydrogenases of *N. salsuginis* and *S. oneidensis* show high similarities and therefore it can be assumed that the essential assembly and maturation apparatus may be compatible, resulting in a heterologous expressed active hydrogenase as observed.

4.2.1.2 Fosmid pRS44::SP-P18F11

The second fosmid which exhibited H2 uptake activity that was investigated in detail was the fosmid P18F11. The ORFs on this fosmid showed similarities to genes of the *Euryarchaeota Aciduliprofundum* sp. MAR08-339 (82 % – 96 % identity, Figure 15, Table 8). To date, the only cultured representatives of the "Deep-Sea Hydrothermal Vent Euryarchaeotic 2 lineage" (DHVE 2) are *Aciduliprofundum boonei* T469 and *Aciduliprofundum* sp. MAR08-339. Although DHVE 2 has been reported from actively venting sulfide deposits by 16S rRNA analyses several times, their metabolic properties and their role in the environment remain speculative (Takai and Horikoshi, 1999; Flores *et al.*, 2011). Both *Aciduliprofundum* strains have been described as a thermoacidophilic heterotroph (T_{opt} 70°C, pH 4.2-4.8), which can use ferric iron and elemental sulfur as electron acceptors (Reysenbach *et al.*, 2006; Reysenbach and Flores, 2008; Schouten *et al.*, 2008). Interestingly, neither *A. boonei* nor *Aciduliprofundum* sp. MAR08-339 is known to use H_2 as an electron donor (Reysenbach *et al.*, 2006; Reysenbach and Flores, 2008). But the membrane fraction of S*. oneidensis* Δ*hyaB* complemented with P18F11 exhibited detectable H_2 uptake activity (45.8 ± 9.1 nmol * min ^{1} * mg^{-1} total protein at 25°C, 292.6 ± 96.7 nmol * min^{-1} * mg^{-1} total protein at 55°C and 1331.5 ± 223.5 nmol * min^{-1} * mg^{-1} total protein at 70°C) (Figure 17). Since there are sequence discrepancies between the P18F11 fosmid ORFs and the genes of the genomes of *A. boonei* and *A.* sp. MAR08-339, both *Aciduliprofundum* strains may have lost its H_2 uptake ability. Alternatively, *A. boonei* and *A.* sp. MAR08-339 may simply not consume H_2 under the previously provided cultivation conditions (Reysenbach *et al.*, 2006) as shown by H_2 uptake activity measurements for both strains during this PhD study (Figure 17). Based on the findings of this PhD study, it is possible to say that some affiliates of the DHVE 2 linage appear to be able to metabolize H_2.

Despite H_2 uptake activity for the fosmid P18F11, none of the ORFs on the fosmid or on the respective genome sequence section of *A.* sp. MAR08-339 could be identified as hydrogenase genes so far. Nevertheless, two ORFs showed sequence identity of 74 % and 82 % to a 24 kDa and 51 kDa subunit of a respiratory-chain NADH:ubiquinone

oxidoreductase of *A.* sp. MAR08-339 (Figure 15; Table 8). Interestingly, it has been demonstrated that hydrogenases ([NiFe]- and [FeFe]-hydrogenases) are evolutionarily linked to NADH:ubiquinone oxidoreductase and share several homologous regions among electron-transferring modules and proton-pumping modules with NADH:ubiquinone oxidoreductase, also known as respiratory complex I (Friedrich and Scheide, 2000; Yano and Ohnishi, 2001; Mathiesen and Hagerhall, 2003; Hedderich, 2004; Calteau *et al.*, 2005; Vignais and Billoud, 2007). These resemblances demonstrate the exchange of redox protein modules among energy-conserving systems that occur during evolution (Vignais and Billoud, 2007). Based on this evolutionary link and the similarities between several homologous regions of hydrogenases and NADH:ubiquinone oxidoreductase, it can be postulated that the corresponding annotated ORFs may encode an enzyme, which interact with the other ORFs and mediate collectively the observed H_2 uptake activity.

However, corresponding results of functional analyses which confirm this hypothesis are currently missing. No generated subclone of the fosmid P18F11 could restore the ability of *S. oneidensis* Δ*hyaB* to reduce Fe(III)citrate by utilization of H_2 (3.4.4). This has included subclones containing the two ORFs with similarities to a 24 kDa and 51 kDa subunit of a respiratory-chain NADH:ubiquinone oxidoreductase. Also these subclones didn't restore the H_2 uptake activity of *S. oneidensis* Δ*hyaB* (i.e. pRS44::P18-BamHI-SC1, pRS44::P18-BamHI-SC2, pRS44::P18-BsaAI-SC1, pRS44::P18-SacII-SC2; Figure 19, Table 9). Theoretically, all ORFs present on the fosmid P18F11 were tested and all possible combinations of ORFs on the several generated P18F11 subclones showed negative results during the analyses for H_2 uptake activity. Feasibly the copy number of particular ORFs or the successions of ORFs are essential for restoring the H_2 uptake activity of *S. oneidensis* Δ*hyaB*, this theory has to be verified by further tests.

4.2.2 Analysis of 12 further fosmids with putative H_2 uptake activity found by the developed functional screen

In the second round of the functional screen, 12 further fosmids with putative H_2 uptake activity were detected.

As described before (4.2), the 12 fosmids indentified in the second round of the functional screen were analyzed by repeating the screen several times. Howewer, a reliable color change of the media indicating putative H_2 uptake activity by the reduction of Fe(III)citrate was not observed for any of these fosmids. Therefore, in contrast to the fosmid P05F01 and P18F11, an existing H_2 uptake hydrogenase activity mediated by the encoded genes on the fosmids has not been confirmed by the functional screen for any of these

investigated 12 fosmids. Possible reasons for the detected putative false positive 12 fosmids by the functional screen are discussed in 4.2. To seek metagenomic hydrogenases by the developed functional screen, the repeatability of positive screening results must be given in the future. Therefore, the functional screen has to be improved as discussed in 4.2.

Nonetheless, the 12 fosmids were investigated by H_2 uptake activity assays (3.5). Additionally, the 12 fosmids were analyzed by sequencing the fosmid 5'/3' insert ends. No ORFs with similarities to hydrogenases or related genes could be found by sequencing (3.5). The conducted H_2 uptake activity assays revealed no hints for restored H_2 uptake activities. This corresponds to the missing repeatability of the functional screen for the 12 investigated fosmids.

For *S. oneidensis* Δ*hyaB* cultures containing the fosmid pRS44::SP-P11C02, minor H_2 uptake activities with high standard deviations were measured with one exception: For the 70 °C measurements of lithotrophic grown cultures a H_2 uptake activity was detectable (690.5 ± 56.6 nmol * min^{-1} * mg^{-1} total protein). The total fosmid insert was sequenced and the sequence analysis verified that no ORF with a similarity to hydrogenase genes is encoded on the fosmid. The fosmid contained a 2,000 bp insert with an ORF showing similarities to an initiator RepB protein of *Candidatus* Magnetoglobus multicellularis. The sequencing results clarified the missing H_2 uptake activity of this fosmid due to the absent of ORFs with similarities to hydrogenases. Ignoring the described exception of the 70 °C assay measurements, the high standard deviations of the measured H_2 uptake activities are hints for missing H_2 uptake activities for the investigated fosmid. The portrayed 70 °C H_2 uptake activity measurements were measured only two-fold and the results compared with performed assays at 25 °C and 55 °C indicate that there were errors in the implementation of these both 70 °C measurements resulting in the detected H_2 uptake activity.

As for pRS44::SP-P11C02, the fosmid inserts of pRS44::SP-P44H11 and pRS44::SP-P31A02 were total sequenced and the sequence analysis verified that no ORF with similarities to hydrogenase genes is encoded on the fosmids. The 1,500 bp insert of pRS44::SP-P31A02 contained an ORF with similarities to a hypothetical protein of the Euryarchaeota *Thermococcus kodakarensis* KOD1. The fosmid pRS44::SP-P44H11 contained a 1,272 bp insert with an ORF showing similarities to a hypothetical protein of the *Epsilonproteobacteria Sulfurimonas* sp. AST-10. Turnover rates of the potential H_2 uptake activities were measured and the results showed activities with high standard deviations for pRS44::SP-P44H11 (e.g. 55 °C: 170.3 ± 170.3 nmol H_2 × min^{-1} × mg^{-1} total protein; 70 °C: 285.1 ± 279.7 nmol H_2 × min^{-1} × mg^{-1} total protein with membrane

fractions of lithotrophic grown cultures) and pRS44::SP-P44H11 (e.g. 55 °C: 25.2 ± 28.3 nmol H_2 × min^{-1} × mg^{-1} total protein; 70 °C: 39.0 ± 29.5 nmol H_2 × min^{-1} × mg^{-1} total protein with membrane fractions of lithotrophic grown cultures). The high standard deviations of these measurements confirmed missing H_2 uptake activities and supported the sequencing results which confirmed the missing of hydrogenase genes encoded on the fosmids.

Minor H_2 uptake activities were measured for membrane fractions of lithotrophic cultivated *S. oneidensis* Δ*hyaB* strains containing the fosmids pRS44::SP-P30F07 (55 °C: 2.4 ± 3.3 nmol H_2 × min^{-1} × mg^{-1} total protein), pRS44::SP-P31A05 (55 °C: 9.4 ± 7.7 nmol H_2 × min^{-1} × mg^{-1} total protein; 70 °C: 5.3 ± 4.4 nmol H_2 × min^{-1} × mg^{-1} total protein), pRS44::SP-P35F04 (55 °C: 13.8 ± 7.8 nmol H_2 × min^{-1} × mg^{-1} total protein) and pRS44::SP-P44E06 (55 °C: 8.3 ± 4.0 nmol * min^{-1} * mg^{-1} total protein). These measurements were in the same range as the turnover rates of the *S. oneidensis* Δ*hyaB* deletion mutant negative control (55 °C: 9.6 ± 5.6 nmol H_2 × min^{-1} × mg^{-1} total protein; 70 °C: 16.0 ± 2.3 nmol H_2 × min^{-1} × mg^{-1} total protein). The H_2 uptake activity assay results of the fosmids pRS44::SP-P30F07, pRS44::SP-P31A05, pRS44::SP-P35F04 and pRS44::SP-P44E06 confirmed their missing H_2 uptake hydrogenase activity.

For the investigated fosmids pRS44::SP-P31B08, pRS44::SP-P37B01 and pRS44::SP-P37B03 no H_2 uptake activities were detectable by the performed assays. Thereby, the H_2 uptake activity assay results confirmed missing H_2 uptake hydrogenase activity for these investigated fosmids.

H_2 uptake activities with high standard deviations were measured for membrane fractions of lithotrophic cultivated *S. oneidensis* Δ*hyaB* strains containing the fosmids pRS44::SP-P11H04 (25 °C: 0.3 ± 0.4 nmol * min^{-1} * mg^{-1} total protein; 55 °C: 137.5 ± 149.4 nmol * min^{-1} * mg^{-1} total protein; 70 °C: 279.5 ± 267.1 nmol * min^{-1} * mg^{-1} total protein) and pRS44::SP-P30B05 (25 °C: 49.0 ± 20.6 nmol H_2 × min^{-1} × mg^{-1} total protein). Additionally, for the fosmid pRS44::SP-P30B05 no H_2 uptake activity was detectable by the performed assay at 70 °C. As a result of this, the H_2 uptake activity assay results of the fosmids pRS44::SP-P11H04 and pRS44::SP-P30B05 confirmed their missing H_2 uptake hydrogenase activity.

4.3 Conclusions and outlook

The developed function-based screening is the first for seeking H_2 uptake active hydrogenases from metagenomic fosmid libraries based on restored H_2 uptake abilities

of the constructed *S. oneidensis* Δ*hyaB* mutant if hydrogenases are heterologous expressed in the surrogate host.

For seeking metagenomic hydrogenases by the developed functional screen in the future, the repeatability of positive screening results must be improved to gain reliable results therewith false positive hits are excluded and don't occur as observed for the second round of the screen with the metagenomic library from the Sisters Peak chimney massive sulfide sample.

Nonetheless, it could be demonstrated that the screen is useful for seeking H_2 uptake enzymes from metagenomic libraries making the large potential of hydrogenases from uncultured microbes accessible. It could be illustrated that a heterologous expression of active H_2 uptake hydrogenases is possible in *S. oneidensis* Δ*hyaB* beyond the class level (demonstrated by the heterologous expression of a putative epsilonproteobacterial [NiFe]-hydrogenase encoded on the fosmid pRS44::SP-P05F01) and even beyond the domain level (indicated by the heterologous expression of a putative archaeal hydrogenase encoded on the fosmid pRS44::SP-P18F11). One of the recovered H_2 uptake active fosmids contained a gene cluster resembling that of an Archaea, which is not known to utilize H_2. Since this archaeal group is widespread in hydrothermal vent environments, this finding is of significance for understanding the distribution and the role this lineage plays in these habitats.

5 References

Aakvik, T., Degnes, K.F., Dahlsrud, R., Schmidt, F., Dam, R., Yu, L. *et al.* (2009) A plasmid RK2-based broad-host-range cloning vector useful for transfer of metagenomic libraries to a variety of bacterial species. *FEMS Microbiol Lett* **296**: 149-158.

Adams, M.W., and Mortenson, L.E. (1984) The physical and catalytic properties of hydrogenase II of Clostridium pasteurianum. A comparison with hydrogenase I. *J Biol Chem* **259**: 7045-7055.

Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* **59**: 143-169.

Amend, J.P., and Shock, E.L. (2001) Energetics of overall metabolic reactions of thermophilic and hyperthermophilic *Archaea* and *Bacteria*. *FEMS Microbiol Rev* **25**: 175-243.

Anderson, I., Sikorski, J., Zeytun, A., Nolan, M., Lapidus, A., Lucas, S. *et al.* (2011) Complete genome sequence of *Nitratifractor salsuginis* type strain (E9I37-1(T)). *Stand Genomic Sci* **4**: 322-330.

Armstrong, F.A., and Hirst, J. (2011) Reversibility and efficiency in electrocatalytic energy conversion and lessons from enzymes. *Proc Natl Acad Sci U S A* **108**: 14049-14054.

Armstrong, F.A., Belsey, N.A., Cracknell, J.A., Goldet, G., Parkin, A., Reisner, E. *et al.* (2009) Dynamic electrochemical investigations of hydrogen oxidation and production by enzymes and implications for future technology. *Chem Soc Rev* **38**: 36-51.

Arriola, E., Lambros, M.B., Jones, C., Dexter, T., Mackay, A., Tan, D.S. *et al.* (2007) Evaluation of Phi29-based whole-genome amplification for microarray-based comparative genomic hybridisation. *Lab Invest* **87**: 75-83.

Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R., and Wolfe, R.S. (1979) Methanogens: reevaluation of a unique biological group. *Microbiol Rev* **43**: 260-296.

Blokesch, M., Magalon, A., and Bock, A. (2001) Interplay between the specific chaperone-like proteins HybG and HypC in maturation of hydrogenases 1, 2, and 3 from *Escherichia coli*. *J Bacteriol* **183**: 2817-2822.

Bock, A., King, P.W., Blokesch, M., and Posewitz, M.C. (2006) Maturation of hydrogenases. *Adv Microb Physiol* **51**: 1-71.

Bothe, H., Schmitz, O., Yates, M.G., and Newton, W.E. (2010) Nitrogen fixation and hydrogen metabolism in cyanobacteria. *Microbiol Mol Biol Rev* **74**: 529-551.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Breslauer, K.J., Frank, R., Blocker, H., and Marky, L.A. (1986) Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci U S A* **83**: 3746-3750.

Brugna-Guiral, M., Tron, P., Nitschke, W., Stetter, K.O., Burlat, B., Guigliarelli, B. *et al.* (2003) [NiFe] hydrogenases from the hyperthermophilic bacterium *Aquifex aeolicus*: properties, function, and phylogenetics. *Extremophiles* **7**: 145-157.

Calteau, A., Gouy, M., and Perriere, G. (2005) Horizontal transfer of two operons coding for hydrogenases between bacteria and archaea. *J Mol Evol* **60**: 557-565.

Campbell, B.J., Engel, A.S., Porter, M.L., and Takai, K. (2006) The versatile *Epsilonproteobacteria*: key players in sulphidic habitats. *Nat Rev Microbiol* **4**: 458-468.

Campbell, B.J., Smith, J.L., Hanson, T.E., Klotz, M.G., Stein, L.Y., Lee, C.K. *et al.* (2009) Adaptations to submarine hydrothermal environments exemplified by the genome of *Nautilia profundicola*. *PLoS Genet* **5**: e1000362.

Cardoso, J., Molenaar, L., de Menezes, R.X., Rosenberg, C., Morreau, H., Moslein, G. *et al.* (2004) Genomic profiling by DNA amplification of laser capture microdissected tissues and array CGH. *Nucleic Acids Res* **32**: e146.

Casalot, L., and Rousset, M. (2001) Maturation of the [NiFe] hydrogenases. *Trends Microbiol* **9**: 228-237.

Cord-Ruwisch, R., Seitz, H.-J.r., and Conrad, R. (1988) The capacity of hydrogenotrophic anaerobic bacteria to compete for traces of hydrogen depends on the redox potential of the terminal electron acceptor. *Arch Microbiol* **149**: 350-357.

Corre, E., Reysenbach, A.L., and Prieur, D. (2001) Epsilon-proteobacterial diversity from a deep-sea hydrothermal vent on the Mid-Atlantic Ridge. *FEMS Microbiol Lett* **205**: 329-335.

Courtois, S., Cappellano, C.M., Ball, M., Francou, F.X., Normand, P., Helynck, G. *et al.* (2003) Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl Environ Microbiol* **69**: 49-55.

Daniel, R. (2004) The soil metagenome--a rich resource for the discovery of novel natural products. *Curr Opin Biotechnol* **15**: 199-204.

Dawood, Z., and Brozel, V.S. (1998) Corrosion-enhancing potential of Shewanella putrefaciens isolated from industrial cooling waters. *Journal of Applied Microbiology* **84**: 929-936.

De Windt, W., Aelterman, P., and Verstraete, W. (2005) Bioreductive deposition of palladium (0) nanoparticles on Shewanella oneidensis with catalytic activity towards reductive dechlorination of polychlorinated biphenyls. *Environ Microbiol* **7**: 314-325.

Deming, J.W. (1998) Deep ocean environmental biotechnology. *Curr Opin Biotechnol* **9**: 283-287.

Dilling, W., and Cypionka, H. (1990) Aerobic Respiration in Sulfate-Reducing Bacteria. *FEMS Microbiol Lett* **71**: 123-127.

Felbeck, H., and Somero, G.N. (1982) Primary Production in Deep-Sea Hydrothermal Vent Organisms - Roles of Sulfide-Oxidizing Bacteria. *Trends in Biochemical Sciences* **7**: 201-204.

Ferrer, M., Beloqui, A., Timmis, K.N., and Golyshin, P.N. (2009) Metagenomics for mining new genetic resources of microbial communities. *J Mol Microbiol Biotechnol* **16**: 109-123.

Ferrer, M., Golyshina, O.V., Chernikova, T.N., Khachane, A.N., Martins Dos Santos, V.A., Yakimov, M.M. *et al.* (2005) Microbial enzymes mined from the Urania deep-sea hypersaline anoxic basin. *Chem Biol* **12**: 895-904.

Figurski, D.H., and Helinski, D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci U S A* **76**: 1648-1652.

Flores, G.E., Campbell, J.H., Kirshtein, J.D., Meneghin, J., Podar, M., Steinberg, J.I. *et al.* (2011) Microbial community structure of hydrothermal deposits from geochemically different vent fields along the Mid-Atlantic Ridge. *Environ Microbiol* **13**: 2158-2171.

Fouquet, Y., Cambon, P., Etoubleau, J., Charlou, J.L., Ondreas, H., Barriga, F.J.A.S. *et al.* (2010) Geodiversity of Hydrothermal Processes Along the Mid-Atlantic Ridge and Ultramafic-Hosted Mineralization: A New Type of Oceanic Cu-Zn-Co-Au Volcanogenic Massive Sulfide Deposit. *Diversity of Hydrothermal Systems on Slow Spreading Ocean Ridges* **188**: 321-367.

Friedrich, B., Friedrich, C.G., Meyer, M., and Schlegel, H.G. (1984) Expression of hydrogenase in *Alcaligenes spp.* is altered by interspecific plasmid exchange. *J Bacteriol* **158**: 331-333.

Friedrich, T., and Scheide, D. (2000) The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Lett* **479**: 1-5.

Gabor, E.M., Alkema, W.B., and Janssen, D.B. (2004) Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environ Microbiol* **6**: 879-886.

Gardy, J.L., Laird, M.R., Chen, F., Rey, S., Walsh, C.J., Ester, M., and Brinkman, F.S.L. (2005) PSORTb v.2.0: Expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics* **21**: 617-623.

Guezennec, J. (2002) Deep-sea hydrothermal vents: a new source of innovative bacterial exopolysaccharides of biotechnological interest? *J Ind Microbiol Biotechnol* **29**: 204-208.

Guiral, M., Tron, P., Belle, V., Aubert, C., Leger, C., Guigliarelli, B., and Giudici-Orticoni, M.T. (2006) Hyperthermostable and oxygen resistant hydrogenases from a hyperthermophilic bacterium *Aquifex aeolicus*: Physicochemical properties. *Int J Hydrogen Energ* **31**: 1424-1431.

Hallenbeck, P.C. (2009) Fermentative hydrogen production: Principles, progress, and prognosis. *Int J Hydrogen Energ* **34**: 7379-7389.

Handelsman, J. (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* **68**: 669-685.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., and Goodman, R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* **5**: R245-249.

Heath, C., Hu, X.P., Cary, S.C., and Cowan, D. (2009) Identification of a novel alkaliphilic esterase active at low temperatures by screening a metagenomic library from antarctic desert soil. *Appl Environ Microbiol* **75**: 4657-4659.

Hedderich, R. (2004) Energy-converting [NiFe] hydrogenases from archaea and extremophiles: ancestors of complex I. *J Bioenerg Biomembr* **36**: 65-75.

Heidelberg, J.F., Paulsen, I.T., Nelson, K.E., Gaidos, E.J., Nelson, W.C., Read, T.D. *et al.* (2002) Genome sequence of the dissimilatory metal ion-reducing bacterium *Shewanella oneidensis*. *Nat Biotechnol* **20**: 1118-1123.

Inagaki, F., Takai, K., Nealson, K.H., and Horikoshi, K. (2004) *Sulfurovum lithotrophicum* gen. nov., sp. nov., a novel sulfur-oxidizing chemolithoautotroph within the epsilon-Proteobacteria isolated from Okinawa Trough hydrothermal sediments. *Int J Syst Evol Microbiol* **54**: 1477-1482.

Iqbal, H.A., Feng, Z., and Brady, S.F. (2012) Biocatalysts and small molecule products from metagenomic studies. *Curr Opin Chem Biol* **16**: 109-116.

Jannasch, H.W., and Taylor, C.D. (1984) Deep-Sea Microbiology. *Annu Rev Microbiol* **38**: 487-514.

Kanai, T., Ito, S., and Imanaka, T. (2003) Characterization of a cytosolic NiFe-hydrogenase from the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1. *J Bacteriol* **185**: 1705-1711.

Karl, D.M., Wirsen, C.O., and Jannasch, H.W. (1980) Deep-Sea Primary Production at the Galapagos Hydrothermal Vents. *Science* **207**: 1345-1347.

Karyakin, A.A., Morozov, S.V., Karyakina, E.E., Zorin, N.A., Perelygin, V.V., and Cosnier, S. (2005) Hydrogenase electrodes for fuel cells. *Biochem Soc Trans* **33**: 73-75.

Kelley, D.S., Baross, J.A., and Delaney, J.R. (2002) Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annu Rev Earth Pl Sc* **30**: 385-491.

Kibbe, W.A. (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res* **35**: W43-46.

Kim, Y.C., Cramer, J., Battaglia, T., Jordan, J.A., Banerji, S.N., Peng, W. *et al.* (2013) Investigation of in Situ Surface Plasmon Resonance Spectroscopy for Environmental Monitoring in and around Deep-Sea Hydrothermal Vents. *Analytical Letters* **46**: 1607-1617.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., 2nd, and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175-176.

Lamle, S.E., Vincent, K.A., Halliwell, L.M., Albracht, S.P.J., and Armstrong, F.A. (2003) Hydrogenase on an electrode: a remarkable heterogeneous catalyst. *Dalton Trans*: 4152-4157.

Lassak, J., Henche, A.L., Binnenkade, L., and Thormann, K.M. (2010) ArcS, the cognate sensor kinase in an atypical Arc system of *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* **76**: 3263-3274.

Le Maitre, R.W., ., eds., and *et al.* (2002) *Igneous Rocks: A Classification and Glossary of Terms*: Cambridge University Press.

Lee, J.P., LeGall, J., and Peck, H.D.J. (1973) Isolation of assimilatroy- and dissimilatory-type sulfite reductases from *Desulfovibrio vulgaris*. *J Bacteriol* **115**: 529-542.

Li, Y., Wexler, M., Richardson, D.J., Bond, P.L., and Johnston, A.W. (2005) Screening a wide host-range, waste-water metagenomic library in tryptophan auxotrophs of Rhizobium leguminosarum and of Escherichia coli reveals different classes of cloned trp genes. *Environ Microbiol* **7**: 1927-1936.

Liu, C., Gorby, Y.A., Zachara, J.M., Fredrickson, J.K., and Brown, C.F. (2002) Reduction kinetics of Fe(III), Co(III), U(VI), Cr(VI), and Tc(VII) in cultures of dissimilatory metal-reducing bacteria. *Biotechnol Bioeng* **80**: 637-649.

Lojou, E. (2011) Hydrogenases as catalysts for fuel cells: Strategies for efficient immobilization at electrode interfaces. *Electrochim Acta* **56**: 10385-10397.

Lojou, E., Luo, X., Brugna, M., Candoni, N., Dementin, S., and Giudici-Orticoni, M.T. (2008) Biocatalysts for fuel cells: efficient hydrogenase orientation for H_2 oxidation at electrodes modified with carbon nanotubes. *J Biol Inorg Chem* **13**: 1157-1167.

Lovley, D.R., Phillips, E.J., and Lonergan, D.J. (1989) Hydrogen and Formate Oxidation Coupled to Dissimilatory Reduction of Iron or Manganese by *Alteromonas putrefaciens*. *Appl Environ Microbiol* **55**: 700-706.

Ludwig, M., Cracknell, J.A., Vincent, K.A., Armstrong, F.A., and Lenz, O. (2009) Oxygen-tolerant H_2 oxidation by membrane-bound [NiFe] hydrogenases of *Ralstonia* species. Coping with low level H_2 in air. *J Biol Chem* **284**: 465-477.

Luther, G.W., 3rd, Rozan, T.F., Taillefert, M., Nuzzio, D.B., Di Meo, C., Shank, T.M. *et al.* (2001) Chemical speciation drives hydrothermal vent ecology. *Nature* **410**: 813-816.

Maroti, G., Tong, Y., Yooseph, S., Baden-Tillson, H., Smith, H.O., Kovacs, K.L. *et al.* (2009) Discovery of [NiFe] hydrogenase genes in metagenomic DNA: cloning and heterologous expression in *Thiocapsa roseopersicina*. *Appl Environ Microbiol* **75**: 5821-5830.

Marshall, M.J., Plymale, A.E., Kennedy, D.W., Shi, L., Wang, Z., Reed, S.B. *et al.* (2008) Hydrogenase- and outer membrane c-type cytochrome-facilitated reduction of technetium(VII) by Shewanella oneidensis MR-1. *Environ Microbiol* **10**: 125-136.

Martinez, A., Kolvek, S.J., Yip, C.L., Hopke, J., Brown, K.A., MacNeil, I.A., and Osburne, M.S. (2004) Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl Environ Microbiol* **70**: 2452-2463.

Mathiesen, C., and Hagerhall, C. (2003) The 'antiporter module' of respiratory chain complex I includes the MrpC/NuoK subunit -- a revision of the modular evolution scheme. *FEBS Lett* **549**: 7-13.

McCollom, T.M. (2007) Geochemical constraints on source of metabolic energy for chemolithoautotrophy in ultramafic-hosted deep-sea hydrothermal systems. *Astrobiology* **7**: 933-950.

McCollom, T.M., and Shock, E.L. (1997) Geochemical constraints on chemolithoautotrophic metabolism by microorganisms in seafloor hydrothermal systems. *Geochim Cosmochim Acta* **61**: 4375-4391.

McMahon, M.D., Guan, C., Handelsman, J., and Thomas, M.G. (2012) Metagenomic analysis of Streptomyces lividans reveals host-dependent functional expression. *Appl Environ Microbiol* **78**: 3622-3629.

Meshulam-Simon, G., Behrens, S., Choo, A.D., and Spormann, A.M. (2007) Hydrogen metabolism in *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* **73**: 1153-1165.

Mullineaux, L.S., Peterson, C.H., Micheli, F., and Mills, S.W. (2003) Successional mechanism varies along a gradient in hydrothermal fluid flux at deep-sea vents. *Ecological Monographs* **73**: 523-542.

Myers, J.M., and Myers, C.R. (2001) Role for outer membrane cytochromes OmcA and OmcB of Shewanella putrefaciens MR-1 in reduction of manganese dioxide. *Appl Environ Microbiol* **67**: 260-269.

Nakagawa, S., Takai, K., Inagaki, F., Horikoshi, K., and Sako, Y. (2005) *Nitratiruptor tergarcus* gen. nov., sp. nov. and *Nitratifractor salsuginis* gen. nov., sp. nov., nitrate-reducing chemolithoautotrophs of the ε-Proteobacteria isolated from a deep-sea hydrothermal system in the Mid-Okinawa Trough. *Int J Syst Evol Microbiol* **55**: 925-933.

Navas, J., González-Zorn, B., Ladrón, N., Garrido, P., and Vázquez-Boland, J.A. (2001) Identification and mutagenesis by allelic exchange of choE, encoding a cholesterol oxidase from the intracellular pathogen *Rhodococcus equi*. *J. Bacteriol* **16**: 4796-4805

Nakagawa, S., Takaki, Y., Shimamura, S., Reysenbach, A.L., Takai, K., and Horikoshi, K. (2007) Deep-sea vent epsilon-proteobacterial genomes provide insights into emergence of pathogens. *Proc Natl Acad Sci U S A* **104**: 12146-12150.

Ogawa, J., and Shimizu, S. (1999) Microbial enzymes: new industrial applications from traditional screening methods. *Trends Biotechnol* **17**: 13-21.

Perner, M., Petersen, J.M., Zielinski, F., Gennerich, H.-H., and Seifert, R. (2010) Geochemical constraints on the diversity and activity of H_2-oxidizing microorganisms in diffuse hydrothermal fluids from a basalt- and an ultramafic-hosted vent. *FEMS Microbiol Ecol* **74**: 55-71.

Perner, M., Ilmberger, N., Köhler, H.U., Chow, J., and Streit, W.R. (2011b) Emerging Fields in Functional Metagenomics and Its Industrial Relevance: Overcoming Limitations and Redirecting the Search for Novel Biocatalysts. In *Handbook of Molecular Microbial Ecology II*: John Wiley & Sons, Inc., pp. 481-498.

Perner, M., Hentscher, M., Rychlik, N., Seifert, R., Strauss, H., and Bach, W. (2011a) Driving forces behind the biotope structures in two low-temperature hydrothermal venting sites on the southern Mid-Atlantic Ridge. *Environ Microbiol Rep* **3**: 727-737.

Perner, M., Hansen, M., Seifert, R., Strauss, H., Koschinsky, A., and Petersen, S. (2013) Linking geology, fluid chemistry, and microbial activity of basalt- and ultramafic-hosted deep-sea hydrothermal vent environments. *Geobiology* **11**: 340-355.

Petrovskis, E.A., Vogel, T.M., and Adriaens, P. (1994) Effects of electron acceptors and donors on transformation of tetrachloromethane by Shewanella putrefaciens MR-1. *FEMS Microbiol Lett* **121**: 357-363.

Prieur, D. (1997) Microbiology of deep-sea hydrothermal vents. *Trends in Biotechnology* **15**: 242-244.

Pugh, T.J., Delaney, A.D., Farnoud, N., Flibotte, S., Griffith, M., Li, H.I. *et al.* (2008) Impact of whole genome amplification on analysis of copy number variants. *Nucleic Acids Res* **36**: e80.

Rabausch, U., Juergensen, J., Ilmberger, N., Bohnke, S., Fischer, S., Schubach, B. *et al.* (2013) Functional screening of metagenome and genome libraries for detection of novel flavonoid-modifying enzymes. *Appl Environ Microbiol* **79**: 4551-4563.

Rees, H.C., Grant, S., Jones, B., Grant, W.D., and Heaphy, S. (2003) Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries. *Extremophiles* **7**: 415-421.

Reysenbach, A.L., and Flores, G.E. (2008) Electron microscopy encounters with unusual thermophiles helps direct genomic analysis of *Aciduliprofundum boonei*. *Geobiology* **6**: 331-336.

Reysenbach, A.L., Longnecker, K., and Kirshtein, J. (2000) Novel bacterial and archaeal lineages from an in situ growth chamber deployed at a Mid-Atlantic Ridge hydrothermal vent. *Appl Environ Microbiol* **66**: 3798-3806.

Reysenbach, A.L., Liu, Y., Banta, A.B., Beveridge, T.J., Kirshtein, J.D., Schouten, S. *et al.* (2006) A ubiquitous thermoacidophilic archaeon from deep-sea hydrothermal vents. *Nature* **442**: 444-447.

Rondon, M.R., Raffel, S.J., Goodman, R.M., and Handelsman, J. (1999) Toward functional genomics in bacteria: analysis of gene expression in Escherichia coli from a bacterial artificial chromosome library of Bacillus cereus. *Proc Natl Acad Sci U S A* **96**: 6451-6455.

Rousset, M., Casalot, L., Rapp-Giles, B.J., Dermoun, Z., de Philip, P., Belaich, J.P., and Wall, J.D. (1998) New shuttle vectors for the introduction of cloned DNA in Desulfovibrio. *Plasmid* **39**: 114-122.

Sako, Y., Takai, K., Ishida, Y., Uchida, A., and Katayama, Y. (1996) *Rhodothermus obamensis* sp. nov., a modern lineage of extremely thermophilic marine bacteria. *Int J Syst Bacteriol* **46**: 1099-1104.

Sambrook, J., and Russell, D.W. (2001) *Molecular cloning : a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Sapra, R., Verhagen, M.F., and Adams, M.W. (2000) Purification and characterization of a membrane-bound hydrogenase from the hyperthermophilic archaeon Pyrococcus furiosus. *J Bacteriol* **182**: 3423-3428.

Schallmey, M., Ly, A., Wang, C., Meglei, G., Voget, S., Streit, W.R. *et al.* (2011) Harvesting of novel polyhydroxyalkanaote (PHA) synthase encoding genes from a soil metagenome library using phenotypic screening. *FEMS Microbiol Lett* **321**: 150-156.

Schlapbach, L., and Zuttel, A. (2001) Hydrogen-storage materials for mobile applications. *Nature* **414**: 353-358.

Schouten, S., Baas, M., Hopmans, E.C., Reysenbach, A.L., and Damste, J.S. (2008) Tetraether membrane lipids of Candidatus "Aciduliprofundum boonei", a cultivated obligate thermoacidophilic euryarchaeote from deep-sea hydrothermal vents. *Extremophiles* **12**: 119-124.

Serres, M. H., and M. Riley (2006) Genomic analysis of carbon source metabolism of *Shewanella oneidensis* MR-1: predictions versus experiments. *J Bacteriol* **188**: 4601–4609.

Shi, L., Belchik, S.M., Plymale, A.E., Heald, S., Dohnalkova, A.C., Sybirna, K. *et al.* (2011) Purification and characterization of the [NiFe]-hydrogenase of *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* **77**: 5584-5590.

Shock, E., and Canovas, P. (2010) The potential for abiotic organic synthesis and biosynthesis at seafloor hydrothermal systems. *Geofluids* **10**: 161-192.

Simon, C., and Daniel, R. (2011) Metagenomic analyses: past and future trends. *Appl Environ Microbiol* **77**: 1153-1161.

Simon, C., Daniel, R., and Streit, W.R. (2010) Construction of Small-Insert and Large-Insert Metagenomic Libraries. In *Metagenomics Methods and Protocols*: Humana Press, pp. 39-50.

Streit, W., Bjourson, A.J., Cooper, J.E., and Werner, D. (1993) Application of subtraction hybridization for the development of a *Rhizobium leguminosarum* biovar *phaseoli* and *Rhizobium tropici* Group-Specific DNA-Probe. *FEMS Microbiol Ecol* **13**: 59-67.

Streit, W.R., and Schmitz, R.A. (2004) Metagenomics--the key to the uncultured microbes. *Curr Opin Microbiol* **7**: 492-498.

Streit, W.R., Daniel, R., and Jaeger, K.E. (2004) Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. *Curr Opin Biotechnol* **15**: 285-290.

Sybirna, K., Antoine, T., Lindberg, P., Fourmond, V., Rousset, M., Mejean, V., and Bottin, H. (2008) *Shewanella oneidensis:* a new and efficient system for expression and maturation of heterologous [Fe-Fe] hydrogenase from *Chlamydomonas reinhardtii*. *BMC Biotechnol* **8**: 73.

Takai, K., and Horikoshi, K. (1999) Genetic diversity of archaea in deep-sea hydrothermal vent environments. *Genetics* **152**: 1285-1297.

Takai, K., Inagaki, F., Nakagawa, S., Hirayama, H., Nunoura, T., Sako, Y. *et al.* (2003) Isolation and phylogenetic diversity of members of previously uncultivated epsilon-Proteobacteria in deep-sea hydrothermal fields. *FEMS Microbiol Lett* **218**: 167-174.

Takai, K., Campbell, B.J., Cary, S.C., Suzuki, M., Oida, H., Nunoura, T. *et al.* (2005) Enzymatic and genetic characterization of carbon and energy metabolisms by deep-sea hydrothermal chemolithoautotrophic isolates of *Epsilonproteobacteria*. *Appl Environ Microbiol* **71**: 7310-7320.

Tang, K.H., Tang, Y.J., and Blankenship, R.E. (2011) Carbon metabolic pathways in phototrophic bacteria and their broader evolutionary implications. *Front Microbiol* **2**: 165.

Thauer, R.K. (1998) Biochemistry of methanogenesis: a tribute to Marjory Stephenson *Microbiology* **144 (Pt 9)**: 2377-2406.

Thauer, R.K., Kaster, A.K., Goenrich, M., Schick, M., Hiromoto, T., and Shima, S. (2010) Hydrogenases from methanogenic archaea, nickel, a novel cofactor, and H2 storage. *Annu Rev Biochem* **79**: 507-536.

Tuffin, M., Anderson, D., Heath, C., and Cowan, D.A. (2009) Metagenomic gene discovery: how far have we moved into novel sequence space? *Biotechnol J* **4**: 1671-1683.

Tye, J.W., Hall, M.B., and Darensbourg, M.Y. (2005) Better than platinum? Fuel cells energized by enzymes. *Proc Natl Acad Sci U S A* **102**: 16911-16912.

Uchiyama, T., and Miyazaki, K. (2009) Functional metagenomics for enzyme discovery: challenges to efficient screening. *Curr Opin Biotechnol* **20**: 616-622.

Vargas, W.A., Weyman, P.D., Tong, Y.K., Smith, H.O., and Xu, Q. (2011) [NiFe] Hydrogenase from *Alteromonas macleodii* with unusual stability in the presence of oxygen and high temperature. *Appl Environ Microbiol* **77**: 1990-1998.

Venkateswaran, K., Moser, D.P., Dollhopf, M.E., Lies, D.P., Saffarini, D.A., MacGregor, B.J. *et al.* (1999) Polyphasic taxonomy of the genus *Shewanella* and description of *Shewanella oneidensis sp. nov*. *Int J Syst Bacteriol* **49 Pt 2**: 705-724.

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

Wait, A.F., Parkin, A., Morley, G.M., dos Santos, L., and Armstrong, F.A. (2010) Characteristics of Enzyme-Based Hydrogen Fuel Cells Using an Oxygen-Tolerant Hydrogenase as the Anodic Catalyst. *J Phys Chem C* **114**: 12003-12009.

Wang, C., Meek, D.J., Panchal, P., Boruvka, N., Archibald, F.S., Driscoll, B.T., and Charles, T.C. (2006) Isolation of poly-3-hydroxybutyrate metabolism genes from complex microbial communities by phenotypic complementation of bacterial mutants. *Appl Environ Microbiol* **72**: 384-391.

Warren, R.L., Freeman, J.D., Levesque, R.C., Smailus, D.E., Flibotte, S., and Holt, R.A. (2008) Transcription of foreign DNA in Escherichia coli. *Genome Res* **18**: 1798-1805.

Weiland, N., Loscher, C., Metzger, R., and Schmitz, R. (2010) Construction and screening of marine metagenomic libraries. *Methods Mol Biol* **668**: 51-65.

Wexler, M., Bond, P.L., Richardson, D.J., and Johnston, A.W. (2005) A wide host-range metagenomic library from a waste water treatment plant yields a novel alcohol/aldehyde dehydrogenase. *Environ Microbiol* **7**: 1917-1926.

Yano, T., and Ohnishi, T. (2001) The origin of cluster N_2 of the energy-transducing NADH-quinone oxidoreductase: comparisons of phylogenetically related enzymes. *J Bioenerg Biomembr* **33**: 213-222.

Yun, J., and Ryu, S. (2005) Screening for novel enzymes from metagenome and SIGEX, as a way to improve it. *Microb Cell Fact* **4**: 8.

6 Abstract / Zusammenfassung

The ability to use H_2 as an energy source is widely distributed among chemolithotrophic organisms. The interconversion between molecular H_2 and protons and electrons is catalyzed by enzymes called hydrogenases. The aim of this study was the identification and investigation of novel, preferably oxygen tolerant hydrogenase enzymes. The sought hydrogenases were searched in metagenomic material originated from hydrothermal deep-sea vent habitats, because hydrothermal deep-sea vents are described as habitats with a high physiologic and taxonomic richness including a large number of diverse H_2-oxidizing microorganisms. Consequently, these habitats have a great potential for a successful searching for novel hydrogenases. Therefore, a novel function-based screening approach for environmental metagenomic libraries was developed to investigate metagenomic material conserved in broad host range fosmid libraries to identify the searched hydrogen uptake active enzymes. For the realization of the developed function-based screen, an alternative host was chosen to optimize the capability of maturation and assembly of heterologous hydrogenases and to circumvent the tap of the tremendous resources of uncultivated hydrogen oxidizing microbes in hydrothermal deep-sea vent habitats. Hereby a constructed Δ*hyaB* deletion mutant of *Shewanella oneidensis* MR-1 (H_2 uptake ability is lost) is used as alternative heterologous host and is complemented with fosmids from a hydrothermal deep-sea vent metagenomic library. Within die developed function-based screen the *S. oneidensis* wild type phenotype is restored and H_2 is consumed again if a fosmid exhibits H_2 uptake activity and can complement the *S. oneidensis* Δ*hyaB* deletion mutant. This is detectable in the function based screen by a color change in the medium. This is the first described function-based screen for seeking H_2 uptake active hydrogenases from metagenomic fosmid libraries.

In this study 7.776 fosmids of a metagenomic library constructed from a deep-sea hydrothermal venting chimney sample (Sisters Peak) were function-based screened for H_2 uptake activity and 14 fosmids with putative H_2 uptake activity were recovered. From these 14 fosmids two identified fosmids, namely pRS44::SP-P05F01 (P05F01) and pRS44::SP-P18F11 (P18F11), were investigated on the hydrogenase activity level in detail and were completely sequenced. The fosmid P05F01 exhibited high similarities to genes encoded on the hydrogenase operon of the *Epsilonproteobacteria Nitratifractor salsuginis*. In *S. oneidensis* Δ*hyaB* the recombinant hydrogenase from the P05F01 fosmid enabled the consumption of H_2, exhibited an *in vitro* H_2 uptake activity at 25 °C, 55 °C and 70 °C (25 °C: 219.5 ± 54.4 nmol * min^{-1} * mg^{-1} total protein, 55 °C: 2677.0 ± 941.4 nmol * min^{-1} * mg^{-1} total protein, 70 °C: 12001.5 ± 1035.1 nmol * min^{-}

^{1} * mg^{-1} total protein) and showed a H_2 evolution rate of 1.15 ± 0.20 µmol H_2 * min^{-1} * mg^{-1} total protein at 28 °C. The second fosmid with H_2 uptake activity was fosmid P18F11, which showed similarity to six genes of the *Euryarchaeota Aciduliprofundum* sp. MAR08-339, not known to consume H_2. Although the fosmid P18F11 enabled the consumption of H_2, an *in vitro* H_2 uptake activity could be measured at 25 °C, 55 °C and 70 °C (25 °C: 45.8 ± 9.1 nmol * min^{-1} * mg^{-1} total protein, 55 °C: 292.6 ± 96.7 nmol * min^{-1} * mg^{-1} total protein, 70 °C: 1331.5 ± 223.5 nmol * min^{-1} * mg^{-1} total protein) and this strain showed a H_2 evolution rate of 0.29 ± 0.02 µmol H_2 * min^{-1} * mg^{-1} total protein at 28 °C, no classical hydrogenase gene was identified on the fosmid or on the homologous region of *A.* sp. MAR08-339's genome (based on known sequences) that could be allocated to the H_2 uptake function. The 12 further *S. oneidensis* Δ*hyaB* fosmid clones with putative H_2 uptake activity were also investigated in detail, but an existing H_2 uptake hydrogenase activity mediated by encoded genes on the fosmids has not been confirmed by repeating the developed functional screen, by measuring H_2 uptake turnover rates and by sequencing the 5'/3' fosmid ends. For seeking metagenomic hydrogenases by the developed functional screen in the future, the repeatability of positive screening results must be improved to gain reliable results therewith false positive hits are excluded and don't occur as observed for the 12 *S. oneidensis* Δ*hyaB* fosmid clones with putative H_2 uptake activity.

Nevertheless, this study illustrates that H_2 uptake hydrogenases from an environmental sample can be identified using the developed novel culture-independent screening procedure. It further demonstrates the great potential of recovering genes encoding novel H_2 uptake active enzymes previously not associated with this type of activity from environmental samples.

Die Fähigkeit zur Nutzung von H_2 als Energiequelle ist weit verbreitet innerhalb chemolithotropher Organismen. Die Spaltung und Produktion von molekularem H_2 in / aus Protonen und Elektronen wird hierbei von Enzymen, den Hydrogenasen, katalysiert. Das Ziel dieser Studie war die Detektion und Analyse von neuen, möglichst Sauerstoff-toleranten Hydrogenasen. In dieser Studie wurde eine Metagenombank, generiert aus Material einer hydrothermalen Tiefseequelle, durchmustert, um entsprechende Hydrogenasen zu detektieren. Da hydrothermale Tiefseequellen als Habitate beschrieben werden, an denen eine hohe physiologische und taxonomische Diversität inklusive einer hohen Anzahl an H_2-oxidierenden Mikroorganismen vorherrscht, haben diese Habitate ein hohes Potential für eine erfolgreiche Suche nach neuen, nicht beschriebenen Hydrogenasen. Hierfür wurde eine neue funktions-basierte Screening-Methode zur Durchmusterung von Metagenombanken aus Umweltproben entwickelt und anschließend genutzt, um die gesuchten Enzyme mit einer H_2-uptake Aktivität innerhalb des metagenomischen Materials einer hydrothermalen Tiefseequelle, konserviert in einer Metagenombank mit einem breiten Wirtsspektrum, zu identifizieren. Zur Realisierung der neuen funktions-basierten Screening-Methode wurde gezielt ein alternativer Wirt gewählt, um eine Optimierung der Fähigkeiten zur Maturierung und Assemblierung von heterolog exprimierten Hydrogenasen innerhalb des Wirtes zu erreichen und um damit die enormen Ressourcen der bisher unkultivierten H_2-oxidierenden Mikroorganismen innerhalb einer Probe effizienter zu erfassen und zu untersuchen. Dies wurde erreicht indem eine Δ*hyaB* Deletionsmutante von *S. oneidensis* MR-1 erstellt und als heterologer Wirt für die Metagenombank, erstellt aus Material einer hydrothermalen Tiefseequelle, genutzt wurde. Die *S. oneidensis* Δ*hyaB* Deletionsmutante ist nicht mehr fähig H_2 als Energiequelle zu nutzen. Innerhalb der entwickelten funktions-basierten Screening-Methode wird die Fähigkeit der *S. oneidensis* Δ*hyaB* Deletionsmutante zur Konsumierung von H_2 wiederhergestellt, sobald auf einem untersuchten Fosmid Gene für Enzyme mit einer H_2-uptake Aktivität kodiert und in *S. oneidensis* Δ*hyaB* heterolog exprimiert werden. Diese Komplementierung ist mittels der entwickelten funktions-basierten Screening-Methode über ein Farbwechsel innerhalb des Kulturmediums detektierbar. Dies ist die erste beschriebene funktionsbasierte Screening-Methode zur Identifikation von Enzymen mit H_2-uptake Aktivität aus metagenomischen Fosmidbanken.

In dieser Studie wurde eine Metagenombank mir 7.776 Fosmiden, konstruiert aus Material eines Schornsteins einer hydrothermalen Tiefseequelle (Sisters Peak), nach Enzymen mit H_2-uptake Aktivitäten funktions-basiert durchmustert. Hierbei wurden 14 Fosmide mit mutmaßlicher H_2-uptake Aktivität detektiert. Von diesen 14 Fosmiden wurden zwei Fosmide, nämlich pRS44::SP-P05F01 (P05F01) und pRS44::SP-P18F11

(P18F11), genauer im Hinblick auf ihre Hydrogenase-Aktivitäten untersucht und zudem komplett sequenziert. Die ORFs auf Fosmid P05F01 zeigten hohe Ähnlichkeiten zu Genen des Hydrogenase Operon von dem Epsilonproteobakterium *Nitratifractor salsuginis*. In *S. oneidensis* Δ*hyaB* ermöglichte die rekombinante Hydrogenase auf P05F01 eine Wiederherstellung des H_2 Konsums, es konnten *in vitro* H_2-uptake Aktivitäten bei 25 °C, 55 °C und 70 °C gemessen werden (25 °C: 219,5 ± 54,4 nmol * min^{-1} * mg^{-1} Gesamt-Protein, 55 °C: 2.677,0 ± 941,4 nmol * min^{-1} * mg^{-1} Gesamt-Protein, 70 °C: 12.001,5 ± 1035,1 nmol * min^{-1} * mg^{-1} Gesamt-Protein) und eine H_2 Produktion von 1,15 ± 0,20 µmol H_2 * min^{-1} * mg^{-1} Gesamt-Protein bei 28°C wurde ermittelt. Das zweite Fosmid mit H_2-uptake Aktivität war das Fosmid P18F11, welches auf Sequenzebene Ähnlichkeiten zu sechs Genen des *Euryarchaeota Aciduliprofundum* sp. MAR08-339 zeigte, wobei für diese Spezies die Fähigkeit H_2 als Energiequelle zu nutzen nicht bekannt ist. Das Fosmid P18F11 ermöglichte dem heterologen Wirt *S. oneidensis* Δ*hyaB* die Konsumierung von H_2, zudem konnte eine *in vitro* H_2-uptake Aktivität bei 25 °C, 55 °C und 70 °C gemessen werden (25 °C: 45,8 ± 9,1 nmol * min^{-1} * mg^{-1} Gesamt-Protein, 55 °C: 292,6 ± 96,7 nmol * min^{-1} * mg^{-1} Gesamt-Protein, 70 °C: 1331,5 ± 223,5 nmol * min^{-1} * mg^{-1} Gesamt-Protein) und es konnte für die komplementierte Deletionsmutante eine H_2 Produktion in Höhe von 0,29 ± 0,02 µmol H_2 * min^{-1} * mg^{-1} Gesamt-Protein bei 28 °C ermittelt werden. Allerdings konnte weder auf dem sequenzierten Fosmid, noch auf dem Genoms von *A.* sp. MAR08-339 ein Gen für eine klassische Hydrogenase gefunden werden, dass für die beobachtete Hydrogenase Aktivität verantwortlich sein könnte. Die 12 weiteren *S. oneidensis* Δ*hyaB* Fosmid Klone mit putativen H_2-uptake Aktivitäten wurden auch genauer untersucht. Allerdings konnte eine existierende H_2-uptake Aktivität, vermittelt durch die Gene enkodiert auf die entsprechenden Fosmiden, weder durch das mehrfache Wiederholen der entwickelten Screening-Methode bestätigt werden, noch durch Messung von *in vitro* H_2-uptake Aktivität, und auch nicht durch Sequenzierung der Fosmid 5‘ / 3‘ Enden. Diese Ergebnisse zeigten, dass zur zukünftigen Detektion metagenomischer Hydrogenasen die entwickelte funktionelle Screening-Methode weiter verbessert werden sollte, um zuverlässigere und wiederholbare positive Ergebnisse zu erhalten und zukünftig falsch positive Ergebnisse, wie sie für die 12 putativ positiven Klone beobachtet wurden, möglichst komplett auszuschließen.

Nichtsdestotrotz zeigt diese Studie, dass H_2-uptake Hydrogenasen aus Umweltproben über die neu entwickelte Kultur-unabhängige Screening-Methode identifiziert werden können. Des Weiteren demonstriert die Studie das große Potential zur Entdeckung von Genen aus Umweltproben, die für neuartige Enzyme mit H_2-uptake Aktivitäten kodieren, welche vorher nicht mit diesem Typ von Aktivität assoziiert wurden.

7 Acknowledgements

Many thanks to Jun.-Prof. Mirjam Perner for providing me with these interesting topic, having confidence and giving me the opportunity to work self-dependently in her working group and of course all the support in the last years.

Thanks to Prof. Dr. Wolfgang Streit for being the second evaluator of this thesis and for all the support in the last years.

Christian Utpatel, Jessica Grote, Frederike Haack, Jennifer Chow and Daniel Maughan for proofreading my thesis.

Many thanks to all members of the Perner Lab: Thanks Steffi, Moritz, Nicole, Dagmar, Wenke and Yuchen.

And thanks to the whole group of the last years for the fun, help and for some good advices: Mariita, Christian, Julia, Moritz, Frederike, Jessica, Jenny C., Jenny H., Dagmar, Sabine K, Sabine Z., Steffi, Nine, Sascha, Uli, Gesche, Angela, Christel, Ines, Gao, Boris, Patrick, Claudia, Ebrahim, Katja, Arek, Birhanu, Gabim, Martina, Uschi and Regina.

And last but not least I want to thank my parents, Karen, family and friends for the support, always listening to my complaints and the motivation I just needed sometimes.

8 Declaration on oath / Eidelsstattliche Versicherung

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, 22.12.2015

(Nicolas Rychlik)

9 Appendix

9.1 Sequence analysis of the PCR products bracketing the *hyaB* gene of the *S. oneidensis* MR-1 wt and the putative *S. oneidensis* Δ*hyaB* mutant.

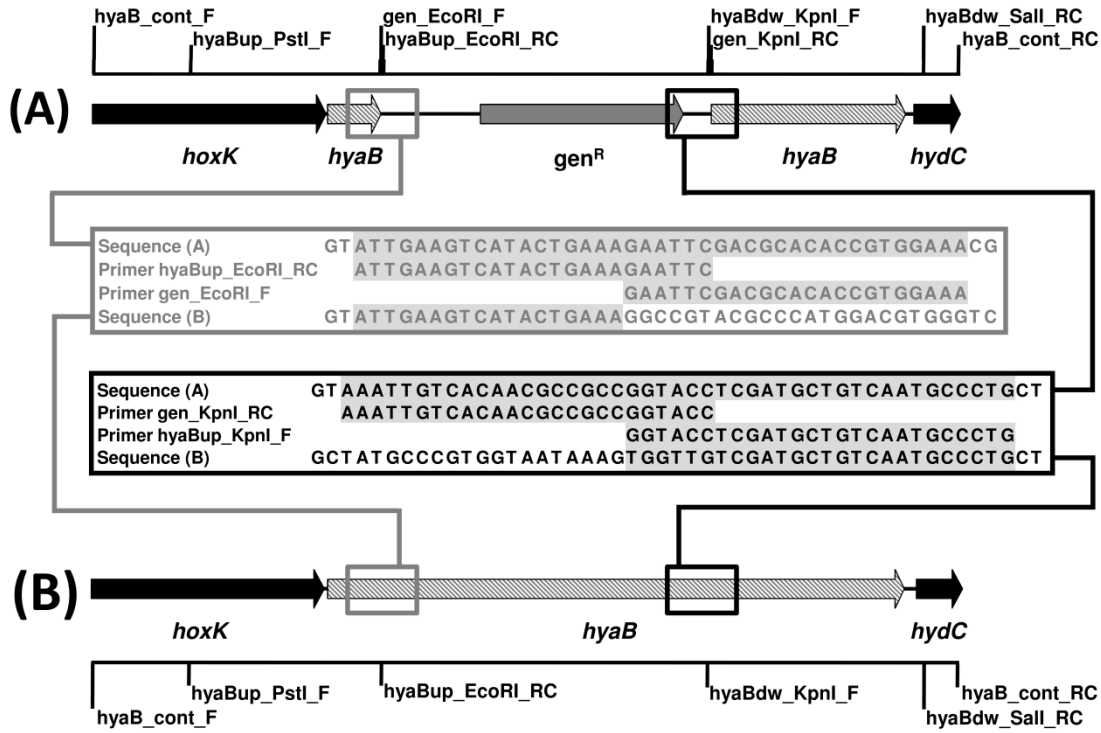


Figure 21: Sequence analysis of the PCR products bracketing the *hyaB* gene of the *S. oneidensis* MR-1 wt and the putative *S. oneidensis* Δ*hyaB* mutant. (A) Illustrate the gene arrangement expected for the *S. oneidensis* Δ*hyaB* mutant, whereas (B) shows the gene arrangement of the *S. oneidensis* MR-1 wt. The gray box exemplifies the sequenced N-terminus (5' region) of the *hyaB* gene originated from the *S. oneidensis* Δ*hyaB* mutant (sequence (A)) and *S. oneidensis* MR-1 wt (sequence (B)). The gray box shows for sequence (A) the successful replacement of an essential part for a functional HyaB subunit due to a gentamycin resistance cassette by identical sequences to the primer sequence of gen_EcoRI_F and to the inserted gentamycin resistance cassette sequence. The black box exemplifies the same as the gray box but for the C-terminus (3' region) of the *hyaB* gene originated from the *S. oneidensis* Δ*hyaB* mutant (sequence (A)) and *S. oneidensis* MR-1 wt (sequence (B)).

9.2 Determination of the extinction coefficient of methyl viologen at different temperatures

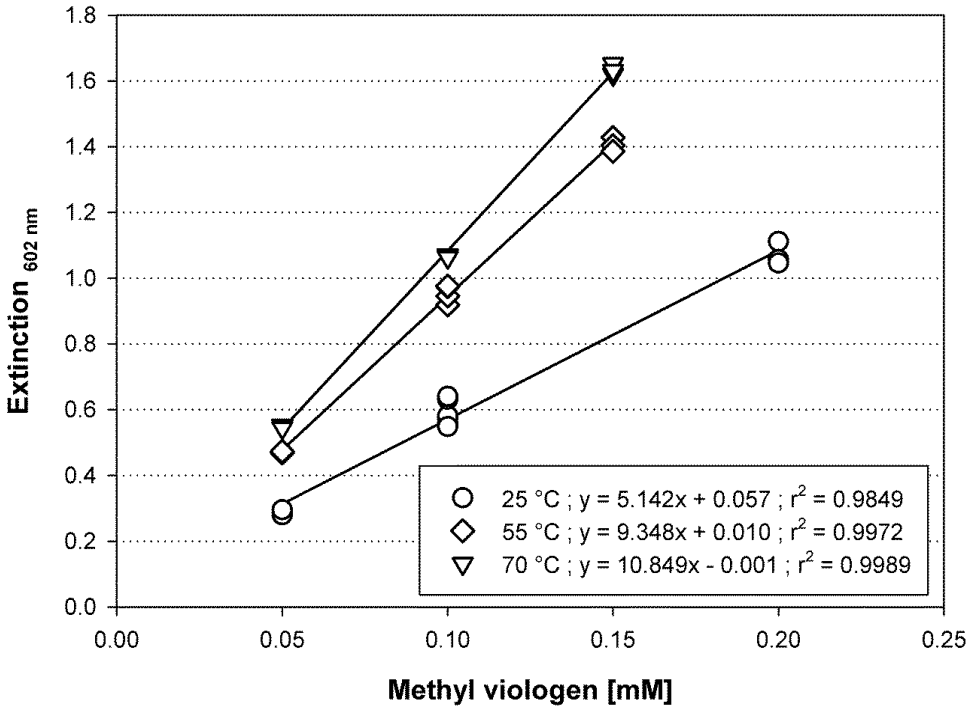


Figure 22: Determination of the extinction coefficient of methyl viologen at different temperatures. A linear correlation could be observed for the extinction at 602 nm and methyl viologen concentrations in a final reaction volume of 1 ml. The methyl viologen extinction coefficient was calculated via linear regression for 25 °C (circle), 55 °C (diamond) and 70 °C (triangle).

9.3 Sequence of P05F01 and P18F11

>P05F01, whole sequence

AATGGGAGGAAGCTCTAATACAGTGCTTCATATGCTTGCAATTGCAAATGAGGCTGAAGTTGATTTTAA
TATTGCAAATATTAATGAGATTAGCAAAAAAGTTAGCCATATTGCAAAGATATCTCCAAGTTTAAGCACA
GTTCATATGGAAGATATAAATAGAGCTGGTGGTGTTAGTGCTGTAATGAATGAGATGAATAAAAGAGG
AGAGGTTTTAATAAATAATCCTACAATTACTGGTGAGATGCTATTTGATAGAATAAAAGATGCAAAAATT
ATTGATACAAATATTATTCATACAATTGATAATCCATACTCTAAAGTTGGAGGACTTGCAATACTATTTG
GAAATTTAGCAGAACAAGGGTGCGTAATTAAAACAGCTGGAATTGTAGGAAGCAGACAATTTAAAGGA
AAAGCAGTTTGTTTTAACTCTCAAGATGAAGCAATAGCAGGAATTACAGGTGGTAAGGTAAAAGCTGG
AGATGTAGTAGTGCTAAGATATGAAGGACCAAAAGGTGGGCCTGGTATGCAAGAGATGCTATCTCCA
ACAAGCCTAATTATGGGAATGGGCTTAGGAGATAGTGTAGCATTAATTACAGATGGAAGATTTAGTGG
TGCTACAAGAGGTTTAAGTATTGGACATGTATCTCCTGAAGCTGCTGAAGGTGGATTAATTGCACTTG
TAGAAGATGGTGATGAGATATTTATTGATGTAGATGAGTATAAGATTGAACTTTTAGTAAGTGATGAAG
AGATTCAAAAAAGAAAAGAAAAATTCAAGCCAATACAAAAAGAGTTAAAAAGCAGATGGCTAAGACAAT
ATAGAGCCCTTGTTACAAATGCAAGTAATGGAGCTGTATTAGAGGCATAAAGTATATAACTTAAGGATA
ATTAACTATCCTTAATTTTTAGATAATTTTAACCTGTTACTCCTACTTCGAAAATTATATAAGAAAACATA
AGTTTAAACTCTCTTTTAAGTAAAGAAAAAATTTATATGTTATTATTCTGAATGACTATTCAATCGGAGG
TATTTTTATGGACAGACGCGATGCATTGATGAAGCTATTTAGTGCTAAGGGTGTTAGAGTTAACACCAA
TAGAGGTGAGGCTTATTATAAAAAACTTAAAGAGAGTATGCAAAAGAGACTAAATGAGTTAAGAGAGT
CTCCTGCTGCATCTAAAGTTGATATTCATAAAGTTTTAGAGTCAGAAGGCTTAACAAGAAGAGATTTTA

TGAAGTGGGCAAGTGCTGCTTGTGCTGCTTTGATGCTTCCTGCTTCGTTTACACCACTTGTTGCTGAA
GCAGCTGAGTTAATGAATAGAGTGCCTATTATTTGGATTGAGCTTCAAGATTGTGCAGGTAACTCTGA
GGCAATTTTAAGAAGTGATGCACCTACTATTGATGAGCTTATATTAGAGACAATTTCACTTGAATTTAAT
GAGACTCTAATGGCTGCAGCTGGACATCAAGCAGAAGAGCATCTTGATGAGGCTATGAAAACATTTGC
AGGTAAATATCTCTGTGTTGTTGAAGGTTCAATTCCTGTTGGAGCTGGTAAAGAGTGGTGCACAATTG
GTGCAAAGGGTGAGACTTTTGAAGAGCATTTAAAAAGAGTTGCAGGTAGTGCAGCAGCTATTGTAGCA
GTTGGAACATGTGCTACATTTGGTGGTGTTCCTGCAGCTGCTCCAAACCCTACTGGAGCTGTAGGAG
TGCAAGATATTATTAAGGGTAAGCCAATTATCAATATTCCGGCATGTCCAGCAAACCCAGCAAATATTA
CTGGGACAATTTTGCACTATGTATTAACTGGGCAAGTGCCAGAGCTTGACCACTTAAATAGACCAAAA
TTTGCATTTGGATATAGAATTCATGATAATTGCGAGAGAAGAGCACACTTTGATGCAGGTGAGTTTGTA
GAAGAGTGGGGTGATGAAGGTGCACAAAACAATTGGTGTTTATATAAAATGGGTTGTAAGGGACCAAT
GACATTTAACAACTGCTCAATTGTAAGATACAATAGTGGCACAAACTGGCCAATTGGTGCAGGACATG
GATGTATTGGATGTAGTGAACCACAATTTTGGGATAAGTATGCACAAGAGAGACCAATGGCAGATACA
CACTTTAAGGCACCAACTGGTGGAGTTGAAAAGAGTGTAGATGAGTTTGGTTTAGGTTTACTTACAGC
AGCAGGTATTGGTATTGGAATCCACGCAGCAATTAGTGCAGTTGCAGGCAAGAGAGATAAAGAGGAG
GCATAAAATATGGCAAAACATATAATAGTAGATCCAATTACAAGAATAGAAGGACACTTAAGAATAGAG
GCAGTTATTGATGATAATGGTGTAATTACAGATGCATATTCAAGCTCTACAATGTTTAGAGGAATAGAG
ACAATTCTAAAAGGAAGAGACCCAAGAGATGCTGGGCTTTTGGCTATGAGAATTTGTGGTGTTTGCAC
AGGAACACACTACCAAAGAAGTATTGAAGCAGTTGAGCATGCATTTAATGTAAAGATTCCAAAAAATGC
AAGACTTGTTAGAAATTTAATGCAAGGTGCATTGTATTTGCATGATCATGTTGTTCATTTCTATCACCTG
CATGCTCTTGATTGGGTTGATATTACAAGTGCTCTAAAAGCAGATGCTAAAGCAGCTGCTCAAGAGGC
ATTTAAATGGAGTGATAATCCAATAGGAGTTGGAGAAGGAGAGCTTAAAGCTATTCAAGAGAGATTAA
CAAAATTTGTAAAACAAGGAAGATTGGGTCTATTTGCAAATGCTTATTGGGGCAATAAGCACTACAAGC
TAACACCAGAGCAAAACCTAATTGCAGTAGCTCACTACTTGCAAGCTCTTGATATGCAAAGAGAGATG
TCTAAAATGATGGCAATTTATGGTGGTAAAATGCCTCATCCACAAAGTATTGTAGTTGGTGGTGTTACT
TGTGTTCAAGATATTCAAAATCCAGCAAGAAATGAGCAGTTTAAATCGTTACTGATAAAAGCAAGAGAT
TTTATAAAAAGAGCATATCTTGCAGATGTTTTAATGGCTGGAACAGTATATTCTGATGAAGCACTTGAT
GGAACAGGTGCTGGACTTAAAAATTATATGGCTTATGGTGGATTTAGACTTGATGATAATGATTTTTAT
AAAGCAAAAACACTATTTCCAAGTGGGATTGTATTAAATGGAGATATAAGTAAACTAATTCCACTAAAT
CAAGAAAAGATTGCAGAAGATGTAAGCCACGCTTGGTATAAAGGCGATAAACCACTTCATCCATATGA
GGGTAAAACAGAGCCAAATTATACTGGACTTGAAAAGAAAGAAGATGGATATGCATATTTAAAAACAAA
AGAGAAATACTCTTGGATAAAATCACCAATTTATGATGATAATAGAGTAGAGGTTGGTCCTTTAGCAAG
AGTTGTTGTAGGACTTGCTGCAAAAGATGAAAGAATTACAAAATATGCAACAAACTTTTTAGCAAAACT
TGGAGAGAAGTTAAATCTTGGCAAAGCAGCACCAATTGATGTGTTATTTTCAACAGTTGGAAGAACTG
CTGCAAGAGCTATTGAGACTGAAATGATGTCTGATGTGATGATGGAGTGGATTGATGAGTTAGCTAAA
AATGTTGCAAGTGGCGATCTTTCTACTTGGAGTGAGTTTGATTTTGATAAAGTTAGTAAAGATGCAAAA
GGTTATGGATTAGAAGAGGCACCAAGAGGTGCTTTAGGACACTGGGTAGTTATTAAAGATGGAAAAAT
TGAAAACTATCAAGCAGTTGTGCCATCTACTTGGAATGCAGCTCCAAGAGATTTCAAAAATAGAATGG
GAGCTTATGAAGCAGCATTAATTGGAACAAAAGTAGCAAACCCAGAAGAACCACTTGAAATTTTAAGA
ACAATTCATAGTTTTGATCCTTGTATTGCATGTGCTGTGCATATTGTTGATACAAAAGGTAAAGAGCTT
GGAGAGTTTAAAGTAAATACAAGCTGTTCGATTTAAGGAGGGTGCTATGAAACCTGGATATAAAAAGA
TAAAGAGAATGACAGCATTTATGCGAATTAACCACTGGGTGGTTGCTATCTCTATGGTGGCTGCTGTT
GTTACGGGGCTATATATTGGGCATCCATACTATCAGACTTTAATTAGTGAACCAGCAGTTGATAAGTAT
GTAATGGCTTGGAATAGATGGGTGCATTTAATGGTGGCAATTATATTTGATGTAAGCTCGATTGTAATT
GCATATCTTTACTTTTTTAGTAGATTTGAAAAGCCTTATAAAAAGATATTACCAACAGGTAAAAATATAA
AAGAGTTTATTGCAGTATTTATAAACTTAATTACATTTAATAGAAGTAAAAAGTTTGATAGCTCTCATGC
AGATAGCTTTAATACTGTTTACTTCTTAATATTTCACCTAATGTTAGCTTGGATGCTTTTAACTGGACTT
CAACTATATGTGCATGGCTTAGAGAGTGGGCATAGCTCTATTGGTGATTGGTGGCCAGCTATGCTTCA
TCTTGTTACAGATTGGACAATTGCATTTACAGGTGGCACAAATATGGATGTTAGAATATCACACCATAC
TACAATGTATTTTATAATTGTTTGGGTGATGTTTCATATATATTATCAAGTATGGAGAACAATTTTTTGG
CAAGAGGGTGATATTAATATCGTATTTGGCGGAAGTAAATATGTTAAAGAGAAGTAGTTAGTTTACTTT
TTTGATATTTATTTGGCTAAGATTATGGAAATTTTTTGGGGCTGAAGCCTCCGTTTCTCGAATGCGGTA
TGCCATCTACTTGGAATCGGTGGCTTTAGCCTCGCCATATAGATTAAGAAAAATAGTTTTTATTTCAAA
AAACAGATTTTTATAAGCCCTATTATCTCTTCGTAAGCTTGATAGCTATCTTATGTAGGGCTTTTAAAGA
TTTTAAGAAGGTAATAATATAAATGAGAATTGCTGTTGTTGGTGCAGGAACTATTATCTTTAGAGATGA
AGGTGTTGGTGTCTATGCACAAAGGTATTTAAGTGAAAATTTTGATGTAGAGGGTGATGTTACCTTTGT
TGATGGTGGTGTGCTTGGATTTAAGCTAATGACATACTATACAGATTATGACAAAGTTATTATCTTAGA
TACTATTACAATGCATAATGATAAAGCTGGAAGTATCTATAATATCCCTGGCGAAGAGCTTTTGGGTCT
TGGAAGCTACAAGCAAACAGCTCACGAGGTTGAAATTGTAGAGATGCTTGAGATTGCAGCACTAAATG
GAAATTTGAGTGATGTAAATATAATTGGCATTGTGCCTGAAGATATCTTAAGTGTACTTCTATTCTTGTA
ATTGGATCTACTATTATATGTTTTGCCATATTTTATGCCTCCTCTTTATCTCTCTTGCCTGCAACTGCAC
TAATTGCTGCGTGGATTCCAATACCAATACCTGCTGCTGTAAGTAAACCTAAACCAAACTCATCTACAC
TCTTTTCAACTCCACCAGTTGGTGCCTTAAAGTGTGTATCTGCCATTGGTCTCTCTTGTGCATACTTAT
CCCAAAATTGTGGTTCACTACATCCAATACATCCATGTCCTGCACCAATTGGCCAGTTTGTGCCACTAT
TGTATCTTACAATTGAGCAGTTGTTAAATGTCATTGGTCCCTTACAACCCATTTTATATAAACACCAATT
GTTTTGTGCACCTTCATCACCCCACTCTTCTACAAACTCACCTGCATCAAAGTGTGCTCTTCTCTCGCA
ATTATCATGAATTCTATATCCAAATGCAAATTTTGGTCTATTTAAGTGGTCAAGCTCTGGCACTTGCCC
AGTTAATACATAGTGCAAAATTGTCCCAGTAATATTTGCTGGGTTTGCTGGACATGCCGGAATATTGAT

AATTGGCTTACCCTTAATAATATCTTGCACTCCTACAGCTCCAGTAGGGTTTGGAGCAGCTGCAGGAA
CACCACCAAATGTAGCACATGTTCCAACTGCTACAATAGCTGCTGCACTACCTGCAACTCTTTTTAAAT
GCTCTTCAAAAGTCTCACCCTTTGCACCAATTGTGCACCACTCTTTACCAGCTCCAACAGGAATTGAA
CCTTCAACAACACAGAGATATTTACCTGCAAATGTTTTCATAGCCTCATCAAGATGCTCTTCTGCTTGA
TGTCCAGCTGCAGCCATTAGAGTCTCATTAAATTCAAGTGAAATTGTCTCTAATATAAGCTCATCAATA
GTAGGTGCATCACTTCTTAAAATTGCCTCAGAGTTACCTGCACAATCTTGAAGCTCAATCCAAATAATA
GGCACTCTATTCATTAACTCAGCTGCTTCAGCAACAAGTGGTGTAAACGAAGCAGGAAGCATCAAAGC
AGCACAAGCAGCACTTGCCCACTTCATAAAATCTCTTCTTGTTAAGCCTTCTGACTCTAAAACTTTATG
AATATCAACTTTAGATGCAGCAGGAGACTCTCTTAACTCATTTAGTCTCTTTTGCATACTCTCTTTAAGT
TTTTTATAATAAGCCTCACCTCTATTGGTGTTAACTCTAACACCCTTAGCACTAAATAGCTTCATCAATG
CATCGCGTCTGTCCATAAAAATACCTCCGATTGAATAGTCATTCAGAATAATAACATATAAATTTTTTCT
TTACTTAAAAGAGAGTTTAAACTTATGTTTTCTTATATAATTTTCGAAGTAGGAGTAACAGGTTAAAATT
ATCTAAAAATTAAGGATAGTTAATTATCCTTAAGTTATATACTTTATGCCTCTAATACAGCTCCATTACTT
GCATTTGTAACAAGGGCTCTATATTGTCTTAGCCATCTGCTTTTTAACTCTTTTTGTATTGGCTTGAATT
TTTCTTTTCTTTTTTGAATCTCTTCATCACTTACTAAAAGTTCAATCTTATACTCATCTACATCAATAAATA
TCTCATCACCATCTTCTACAAGTGCAATTAATCCACCTTCAGCAGCTTCAGGAGATACATGTCCAATAC
TTAAACCTCTTGTAGCACCACTAAATCTTCCATCTGTAATTAATGCTACACTATCTCCTAAGCCCATTCC
CATAATTAGGCTTGTTGGAGATAGCATCTCTTGCATACCAGGCCCACCTTTTGGTCCTTCATATCTTAG
CACTACTACATCTCCAGCTTTTACCTTACCACCTGTAATTCCTGCTTCGAAAATTATATAAGAAAACATA
AGTTTAAACTCTCTTTTAAGTAAAGAAAAAATTTATATGTTATTATTCTGAATGACTATTCAATCGGAGG
TATTTTTATGGACAGACGCGATGCATTGATGAAGCTATTTAGTGCTAAGGGTGTTAGAGTTAACACCAA
TAGAGGTGAGGCTTATTATAAAAAACTTAAAGAGAGTATGCAAAAGAGACTAAATGAGTTAAGAGAGT
CTCCTGCTGCATCTAAAGTTGATATTCATAAAGTTTTAGAGTCAGAAGGCTTAACAAGAAGAGATTTTA
TGAAGTGGGCAAGTGCTGCTTGTGCTGCTTTGATGCTTCCTGCTTCGTTTACACCACTTGTTGCTGAA
GCAGCTGAGTTAATGAATAGAGTGCCTATTATTTGGATTGAGCTTCAAGATTGTGCAGGTAACTCTGA
GGCAATTTTAAGAAGTGATGCACCTACTATTGATGAGCTTATATTAGAGACAATTTCACTTGAATTTAAT
GAGACTCTAATGGCTGCAGCTGGACATCAAGCAGAAGAGCATCTTGATGAGGCTATGAAAACATTTGC
AGGTAAATATCTCTGTGTTGTTGAAGGTTCAATTCCTGTTGGAGCTGGTAAAGAGTGGTGCACAATTG
GTGCAAAGGGTGAGACTTTTGAAGAGCATTTAAAAAGAGTTGCAGGTAGTGCAGCAGCTATTGTAGCA
GTTGGAACATGTGCTACATTTGGTGGTGTTCCTGCAGCTGCTCCAAACCCTACTGGAGCTGTAGGAG
TGCAAGATATTATTAAGGGTAAGCCAATTATCAATATTCCGGCATGTCCAGCAAACCCAGCAAATATTA
CTGGGACAATTTTGCACTATGTATTAACTGGGCAAGTGCCAGAGCTTGACCACTTAAATAGACCAAAA
TTTGCATTTGGATATAGAATTCATGATAATTGCGAGAGAAGAGCACACTTTGATGCAGGTGAGTTTGTA
GAAGAGTGGGGTGATGAAGGTGCACAAAACAATTGGTGTTTATATAAAATGGGTTGTAAGGGACCAAT
GACATTTAACAACTGCTCAATTGTAAGATACAATAGTGGCACAAACTGGCCAATTGGTGCAGGACATG
GATGTATTGGATGTAGTGAACCACAATTTTGGGATAAGTATGCACAAGAGAGACCAATGGCAGATACA
CACTTTAAGGCACCAACTGGTGGAGTTGAAAAGAGTGTAGATGAGTTTGGTTTAGGTTTACTTACAGC
AGCAGGTATTGGTATTGGAATCCACGCAGCAATTAGTGCAGTTGCAGGCAAGAGAGATAAAGAGGAG
GCATAAAATATGGCAAAACATATAATAGTAGATCCAATTACAAGAATAGAAGGACACTTAAGAATAGAG
GCAGTTATTGATGATAATGGTGTAATTACAGATGCATATTCAAGCTCTACAATGTTTAGAGGAATAGAG
ACAATTCTAAAAGGAAGAGACCCAAGAGATGCTGGGCTTTTGGCTATGAGAATTTGTGGTGTTTGCAC
AGGAACACACTACCAAAGAAGTATTGAAGCAGTTGAGCATGCATTTAATGTAAAGATTCCAAAAAATGC
AAGACTTGTTAGAAATTTAATGCAAGGTGCATTGTATTTGCATGATCATGTTGTTCATTTCTATCACCTG
CATGCTCTTGATTGGGTTGATATTACAAGTGCTCTAAAAGCAGATGCTAAAGCAGCTGCTCAAGAGGC
ATTTAAATGGAGTGATAATCCAATAGGAGTTGGAGAAGGAGAGCTTAAAGCTATTCAAGAGAGATTAA
CAAAATTTGTAAAACAAGGAAGATTGGGTCTATTTGCAAATGCTTATTGGGGCAATAAGCACTACAAGC
TAACACCAGAGCAAAACCTAATTGCAGTAGCTCACTACTTGCAAGCTCTTGATATGCAAAGAGAGATG
TCTAAAATGATGGCAATTTATGGTGGTAAAATGCCTCATCCACAAAGTATTGTAGTTGGTGGTGTTACT
TGTGTTCAAGATATTCAAAATCCAGCAAGAAATGAGCAGTTTAAATCGTTACTGATAAAAGCAAGAGAT
TTTATAAAAAGAGCATATCTTGCAGATGTTTTAATGGCTGGAACAGTATATTCTGATGAAGCACTTGAT
GGAACAGGTGCTGGACTTAAAAATTATATGGCTTATGGTGGATTTAGACTTGATGATAATGATTTTTAT
AAAGCAAAAACACTATTTCCAAGTGGGATTGTATTAAATGGAGATATAAGTAAACTAATTCCACTAAAT
CAAGAAAAGATTGCAGAAGATGTAAGCCACGCTTGGTATAAAGGCGATAAACCACTTCATCCATATGA
GGGTAAAACAGAGCCAAATTATACTGGACTTGAAAAGAAAGAAGATGGATATGCATATTTAAAAACAAA
AGAGAAATACTCTTGGATAAAATCACCAATTTATGATGATAATAGAGTAGAGGTTGGTCCTTTAGCAAG
AGTTGTTGTAGGACTTGCTGCAAAAGATGAAAGAATTACAAAATATGCAACAAACTTTTTAGCAAAACT
TGGAGAGAAGTTAAATCTTGGCAAAGCAGCACCAATTGATGTGTTATTTTCAACAGTTGGAAGAACTG
CTGCAAGAGCTATTGAGACTGAAATGATGTCTGATGTGATGATGGAGTGGATTGATGAGTTAGCTAAA
AATGTTGCAAGTGGCGATCTTTCTACTTGGAGTGAGTTTGATTTTGATAAAGTTAGTAAAGATGCAAAA
GGTTATGGATTAGAAGAGGCACCAAGAGGTGCTTTAGGACACTGGGTAGTTATTAAAGATGGAAAAAT
TGAAAACTATCAAGCAGTTGTGCCATCTACTTGGAATGCAGCTCCAAGAGATTTCAAAAATAGAATGG
GAGCTTATGAAGCAGCATTAATTGGAACAAAAGTAGCAAACCCAGAAGAACCACTTGAAATTTTAAGA
ACAATTCATAGTTTTGATCCTTGTATTGCATGTGCTGTGCATATTGTTGATACAAAAGGTAAAGAGCTT
GGAGAGTTTAAAGTAAATACAAGCTGTTCGATTTAAGGAGGGTGCTATGAAACCTGGATATAAAAAGA
TAAAGAGAATGACAGCATTTATGCGAATTAACCACTGGGTGGTTGCTATCTCTATGGTGGCTGCTGTT
GTTACGGGGCTATATATTGGGCATCCATACTATCAGACTTTAATTAGTGAACCAGCAGTTGATAAGTAT
GTAATGGCTTGGAATAGATGGGTGCATTTAATGGTGGCAATTATATTTGATGTAAGCTCGATTGTAATT
GCATATCTTTACTTTTTTAGTAGATTTGAAAAGCCTTATAAAAAGATATTACCAACAGGTAAAAATATAA

AAGAGTTTATTGCAGTATTTATAAACTTAATTACATTTAATAGAAGTAAAAAGTTTGATAGCTCTCATGC
AGATAGCTTTAATACTGTTTACTTCTTAATATTTCACCTAATGTTAGCTTGGATGCTTTTAACTGGACTT
CAACTGGGATC

>pRS44-SP-P18F11, whole sequence:

ATCCCACGGGAGAGGAAAAGGGCGCTTGAGGAATGATGCTCCTTCGCAATCTTCCAGAATTGCAGGA
TTATATGGATGCTGGTGGCTTTTCAGCCCTGCGAGATGCTCTCAGTATGAACCAGGGGGAAATTGTTG
ATATGATTGAAAGATCATCTCTTGTTGGAAGGGGCGGCGCGGCGTTTCCCACGGGGCTGAAGATGAG
ATTTGTTGCGGAGAGAAATGATACTCCAAAGTACATCGTTGCCAACGCTGACGAGGGTGAGCCGGGC
ACGTTCAAGGATAGGGTGCTAATGGAGAACAATCCGTACCAGATAATAGAGGGTATGATCATAGCCG
CCTATGCCATTGGCGCATCCAAGGGATTTTTTTACATCCGCTACGAATACCGCGAACTGGCAAGGAGA
ATGGAGAGGATAATTGAAGAGCTTCGGAGAAATGGAATTCTGGGCACAAACATTTACGGCTCGAATTT
CTCATTTGATGTTGAACTTGTGCTTGGCGCCGGCTCCTATGTGTGCGGCGAAGAGACCGCGCTTATG
GAGAGCATTGAGGGTAAACGCAGCTTTCCCAGAATCAAGCCACCATATCCTGCGCAGCGTGGGCTCT
GGGGAAAACCAACGCTTATCAACAATGTGGAGACCCTTGCCAATATTTCCGTGATATTGAAAATCGGC
TCGGAGGAATACGCGAAGATGGGCATGAAATTCTGCTCGGGCCCAAAGTTGTTCAGCGTGAGCGGCT
TCGTTCACAAGCCGGGAGTTTACGAGGAGATTTTAGGAGCGGTGACGATTCGGGACATGATTGAAAG
AGCTGGCGGCGTGGATGGAGAGCTCAAGGGCATGATGCTCGGCGGCGGAGCATCTGGGTACATTGT
TGATCCCTCATTTCTTGACATGAAACTCTGCTTCACAGATGCCAAGGAGCGCGGCATCTCTCTCGGAA
CTGGAGATATAATCGTTTTCAATCACACGGTTGATCTCTGGAAGGTTCTTCTCAACATAGCAGGATTCT
TTGAGCACGAGTCCTGCGGACAGTGCTTCCCTTGCAGATACGGCACAAGGAAGATGCGGGAGATAAT
TGAGAAAATCGTTAGGGGTGAGGGGACGAAGGACGACATTTCCCGTCTGGATAGAATTGCAAGGGCC
ATGAAAATTGCGTCCCTGTGCCCCCTGGGAACAAGCGCGATGCTGGCTTATGAATCGGCCATGCGCT
ATTTCCGCGATGAGCTCATGGGGGTGATAAGATGAAGGTGACGCTCAACGGCCGTGAGATTGATGTT
CAGGGTGAGAAAACCATACTTCAGGTTGCCAGGGAGAATGGAATTTCCATACCCGCGATATGCTACAT
GGAGAACCATAAACCCATCGGCTCCTGCGGGTTATGCGTCGTGAAGGTCACGCAGAATTCCCGTGTT
TCTTATCCACGAGCCTGTGCAACCCGGGTGCAGGAAGGAATGCTGATAGAGACCCACGCTCCAGAG
GTGGAGGAGGCTCGCAGGACCATACTTGAGATGCTGGCGGAGGATTATTCCGGTGATGGGGAATTC
AATGATATTCTGGAGCAGTATGGTATAAAACCTGCTGGTAGTGCAAAGATAAAGGACGATAATCCCTT
CTTCATCAGGGATTACGGTGCTTGCATAAACTGCTTCCGATGCGTTCTCGCCTGCGATGAGATAAAGG
GCAACAATGTCCTCGTGCGCTCCGGGAGAGGTTCCGAAATAAGGATATTCGCAGGAGTTAATAATTCA
TTGATTGAATCGGGCTGTGCCTTCTGTGGCGCGTGCATTGATGCATGTCCAACAGGTGCGATGAGGG
AGAAAACATATGAGAGAGAAGGGAAGAGGATAACAACCATATGCCCCTACTGCGGCGTGGGCTGCG
CGATGGATTACTATGTGAAGGATGGCCGTATCGTATACGCGAGGGGCAATCCAGACGGTGTTAACAG
AGGAGACGATTGCATCAAGGGCCGTTTTGGCTGGGAGTTCGTGCACAGTGATGAGAGGTTGAAAAAA
ACCGCTGATAAAGAGAAATGGTGAATTCGTGGAGGTATCATGGGAGGAGGCGCTGGAGTACATAGCG
CGAAGATTGAATGAGATAAAGGAGAAATACGGGCCCGATAGCATCGCCGGTTTATCATCTGCGAAAT
GCACCAACGAGGAGAACTACCTGTTCCAGAAGTTCATGCGCTCGGTTATCGGCACCAACAATGTGGA
TCACTGCGCCCGTCTATGCCATGCTTCCACCGTTACGGGCCTGATGAAGGTGTTCGGTGCGGGAGC
GATGACGAACAGCATTGAAGACCTTGCCAAGGCGGAGGTTTTCTTCGTTATCGGCTCAAACACGACA
CATGCGCATCCGGTAATTGGCACAATGATAAAGAGGGCGGTGAGAAACGGAGCTGGGTTGATAGTTG
CGGATCCTAGGGATATTGAACTATCAAAATATGCCACGGTGCATCTCCGTCAGAGACCTGGCACGGA
TGTCGCGCTAATCAACGCCATGATGCATGTCATAATCAGGTACGGGCTCGTGGATGAAGAATTTGTAA
AGAACCGCACCGAGGGATATGAGGAATTAAAAAGAATAGTGGAGAAATACCCGCCAGAATATGCCGA
GAAGATAACAGGAGTTCCCGCTGATGATATTATACGGGCTGCGATTCTCTATGCGACGCATCGCTCC
GCAATAATTTACGCCATGGGGATAACCCAGCACACAAGTGGCACGGCGAATGTTGCCTCTCTTGCCA
ATCTAGCATTGTTAACTGGCAACGTGGGTAAGGAGGGCACGGGTGTGAATCCGCTTCGCGGTCAGAA
CAATGTGCAGGGTGCAGGTGACATGGGTGCGTTGCCCAACCTCCTGCCAGGATATATTAGGCTAAAC
TCGCCAAAGGTTGAGGAATTTGAGAGAGCATGGCAGGCTAAGATACCCAGAGAACCCGGGCTTACTG
TGATTGAGATGATGCATGCCATAGGGAATGAAATAAAGGCAATGTACATAATGGGCGAAAATCCCGCT
GTTAGCGACCCGGACCAGAAACATGTTCTTTCCGCTCTCAAAAATCTTGAATTTCTGGTGGTTCAGGA
CCTGTTTATGAGCGAAACCGCGCACCTTGCGGATGTGGTTCTTCCAGCAGCGGCATCGCTTGAGAAG
GAGGGCACATTCACCAACACCGAGCGCAGGGTACAGTTGCTTCGCAAAGTAGTTGATCCTCCGGGAG
AAGCCAAGCCGGACTGGTGGATTCTCACAGAACTTGCGAGAAGAATGGGCGCTGAGTGGAATTATGG
CTCACCGAAGGATGTGTTTGAAGAGATTCGCCGCCTCGTGCCACAGTATCGTGGCATAACATACAGG
AGAATTGAGAAATTCGGATTGCAGTGGCCATGCCCGGACGAGGAGCATCCCGGTACGAAAATCCTGC
ATGTGGATAGATTTTCCAGACCCAATGGTCTGGCGAGATTCTACCCATATGAATATAAACCTCCTGCG
GAGGAGCCGGATGACGAATACCCATTCGTGCTCACCACGGGGCGCACATACGAGCATTTTCATACCG
GCACTTTAACGAGGAGAATTGAAGGTTTCAATCAGCTTGTGCCCGAAGCATTTGTGGAGATAAATAAG
GAGGATGCGAAATCTCTTGGGTTAGAGGAGGGAGATTGCGTCGTTGTGGAATCTCGCAGAGGAGCAA
TAAAGGTGAAGGTGAAAATTGCCGGGATAAAAAAGGGCACAGTGTTCATCCCATTCCACTTCGCAGA
GAGCGCCGCGAATGCCCTTACAATCGATGCGCTGGATCCTGAGGCAAAGATTCCCGAATTGAAGGTT
GCAGCGGTGCGCATTGAAAAATGCTAAATTTTCCCATATCTTATTTTTCCATTCACAATTGTATATTCTA
CATTGAGTCCGTTGAGAGCGTAGACTATATTGGATACAATGGTATCCTTCTTTAGAGGCAATCCGTTG
GGCTTGGGATTTAGCACAACAATATCTGCTAGTTTTCCTTCCTCTACGCTCCCTGCATTTAGGCCCAG
TGCCTTGGGAGCATTCACCGTAGCGAAATCGAGAATCATCTGAGCATTTGTCACGCTAGCATCCCATC
TCTCATTCTTGTGTAGTAGTGCGGCGAATTTCATAACCTCAAACATGTCTAGGTTGTTGTTGCTCACCG

CACTATCTGTCCCGAGGGTTATTGTTATTCCTTCATTGATCATCTCCACCAAGGGCATTGAGCCACCG
TTTCCAAGTTTCATGTTGCTCGTTGGATTGTGGGATGCCTTAACCCCATTTCTTGCAAGCATTCTTATT
TCGTTGAGGGTGAGCCACACAAGATGCACACCAATAAGCCTGGGAGAGAGAAATCCAATTTTATCCA
GATACTCCACCGGTCTTAAACCCGTTTTTTTGCGATGTTCATAAACCTCCTTTCTCGTTTCGGAAACAT
GCATAGTCATAAGCGTATTGTACCTATCGGCAATTTCCTTTGCTTTAAGGAGGGTCTCTTCAGAGCAA
GTGTACACACCATGGGGAGCAACCATGGGAGTTATCAAATCCTCCTTTAAATATTCCTGTACAAAATTT
TCAGCGTTTTTTAGGGGTGAACCCTTCTGGGTGGTGATATCCTCATCCACAACTGCCCAGCCCAGAAA
TGCACGTATTCCCTTCTCCCTCGCGATCTCAGCCACCACATCTTCATCTGAGTACATATCCACGAAGG
TGGTTGTTCCAGTGCGAAGCATCTCATCTATGCCAAGTGCAGCACCATGGCGGAGATCCTCTCTTGTT
CTCTTAGATTCTAGGGCCCACATTTTATTTAGGAATTCCTCCAGATCAACATCGTCAGCAAGGCCTCTG
AGATCGGTCATTGGTATGTGGGTGTGCGTGTTTATCAGCCCAGGCATAACTATCTTGTTTTTTCCATC
GATTATGATATCTGCCTCCACTTTGACCCTTCCAACATCGGAAATCCTGTTATCTTCAACATAAACATC
TCCCTTAATTATGTCCCTGCGTTCATTCTGGGTGATGATCCAAGCATCACGGATGAGAATGCTCATAC
TCCGTAATCATTCGCATCCTAAATATGTTTTTTGTCCACACCAAAAATTAAATCCTATCAGAGTATTAGC
CCTACAGTGACAAGGATAACTTTTTTAGGAACTGGCGGCGGGCGGTTCACCACCATATATCAGATTCG
AGCCACAGGAGGAGTTTACTTAGATGTTGATCTGGAGGAATTCCTAAATAAAATGTGGGCCCTAGAAT
CTAAGAGAACAAGAGAGGATCTCCGCCATGGTGCTGCACTTGGCATAGATGAGATGCTTCGCACTGG
AACAACCACCTTCGTGGATATGTACTCAGATGAAGATGTGGTGGCTGAGATCGCGAGGGAGAAGGGA
ATACGTGCATTTCTGGGCTGGGCAGTTGTGGATGAGGATATCACCACCCAGAAGGGTTCACCCCTAA
AAAACGCTGAAAATTTTGTACAGGAATATTTAAAGGAGGATTTGATAACTCCCATGGTTGCTCCCCATG
GTGTGTACACTTGCTCTGAAGAGACCCTCCTTAAAGCAAAGGAAATTGCCGATAGGTACAATACGCTT
ATGACTATGCATGTTTCCGAAACGAGAAAGGAGGTTTATGAACATCGCAAAAAAACGGGTTTAAGACC
GGTGGAGTATCTGGATAAAATTGGATTTCTCTCTCCCAGGCTTATTGGTGTGCATCTTGTGTGGCTCA
CCCTCAACGAAATAAGAATGCTTGCAAGAAATGGGGTTAAGGCATCCCACAATCCAACGAGCAACATG
AAACTTGGAAACGGTGGCTCAATGCCCTTGGTGGAGATGATCAATGAAGGAATAACAATAACCCTCG
GGACAGATAGTGCGGTGAGCAACAACAACCTAGACATGTTTGAGGTTATGAAATTCGCCGCACTACTA
CACAAGAATGAGAGATGGGATGCTAGCGTGACAAATGCTCAGATGATTCTCGATTTCGCTACGGTGAA
TGCTCCCAAGGCACTGGGCCTAAATGCAGGGAGCGTAGAGGAAGGAAAACTAGCAGATATTGTTGTG
CTAAATCCCAAGCCCAACGGATTGCCTCTAAAGAAGGATACCATTGTATCCAATATAGTCTACGCTCT
CAACGGACTCAATGTAGAATATACAATTGTGAATGGAAAAATAAGATATGGGAAAATTTAGCATTTTTC
AATGCGCACCGCTGCAACCTTCAATTCGGGAATCTTTGCCTCAGGATCCAGCGCATCGATTGTAAGG
GCATTCGCGGCGCTCTCTGCGAAGTGGAATGGGATGAACACTGTGCCCTTTTTTATCCCGGCAATTTT
CACCTTCACCTTTATTGCTCCTCTGCGAGATTCCACAACGACGCAATCTCCCTCCTCTAACCCAAGAG
ATTTCGCATCCTCCTTATTTATCTCCACAAATGCTTCGGGCACAAGCTGATTGAAACCTTCAATTCTCC
TCGTTAAAGTGCCGGTATGAAAATGCTCGTATGTGCGCCCCGTGGTGAGCACGAATGGGTATTCGTC
ATCCGGCTCCTCCGCAGGAGGTTTATATTCATATGGGTAGAATCTCGCCAGACCATTGGGTCTGGAAA
ATCTATCCACATGCAGGATTTTCGTACCGGGATGCTCCTCGTCCGGGCATGGCCACTGCAATCCGAA
TTTCTCAATTCTCCTGTATGTTATGCCACGATACTGTGGCACGAGGCGGCGAATCTCTTCAAACACAT
CCTTCGGTGAGCCATAATTCCACTCAGCGCCCATTCTTCTCGCAAGTTCTGTGAGAATCCACCAGTCC
GGCTTGGCTTCTCCCGGAGGATCAACTACTTTGCGAAGCAACTGTACCCTGCGCTCGGTGTTGGTGA
ATGTGCCCTCCTTCTCAAGCGATGCCGCTGCTGGAAGAACCACATCCGCAAGGTGCGCGGTTTCGCT
CATAAACAGGTCCTGAACCACCAGAAATTCAAGATTTTTGAGAGCGGAAAGAACATGTTTCTGGTCCG
GGTCGCTAACAGCGGGATTTTCGCCCATTATGTACATTGCCTTTATTTCATTCCCTATGGCATGCATCA
TCTCAATCACAGTAAGCCCGGGTTCTCTGGGTATCTTAGCCTGCCATGCTCTCTCAAATTCCTCAACC
TTTGGCGAGTTTAGCCTAATATATCCTGGCAGGAGGTTGGGCAACGCACCCATGTCACCTGCACCCT
GCACATTGTTCTGACCGCGAAGCGGATTCACACCCGTGCCCTCCTTACCCACGTTGCCAGTTAACAAT
GCTAGATTGGCAAGAGAGGCAACATTCGCCGTGCCACTTGTGTGCTGGGTTATCCCCATGGCGTAAA
TTATTGCGGAGCGATGCGTCGCATAGAGAATCGCAGCCCGTATAATATCATCAGCGGGAACTCCTGT
TATCTTCTCGGCATATTCTGGCGGGTATTTCTCCACTATTCTTTTTAATTCCTCATATCCCTCGGTGCG
GTTCTTTACAAATTCTTCATCCACGAGCCCGTACCTGATTATGACATGCATCATGGCGTTGATTAGCGC
GACATCCGTGCCAGGTCTCTGACGGAGATGCACCGTGGCATATTTTGATAGTTCAATATCCCTAGGAT
CCGCAACTATCAACCCAGCTCCGTTTCTCACCGCCCTCTTTATCATTGTGCCAATTACCGGATGCGCA
TGTGTCGTGTTTGAGCCGATAACGAAGAAAACCTCCGCCTTGGCAAGGTCTTCAATGCTGTTCGTCAT
CGCTCCCGCACCGAACACCTTCATCAGGCCCGTAACGGTGGAAGCATGGCATAGACGGGCGCAGTG
ATCCACATTGTTGGTGCCGATAACCGAGCGCATGAACTTCTGGAACAGGTAGTTCTCCTCGTTGGTGC
ATTTCGCAGATGATAAACCGGCGATGCTATCGGGCCCGTATTTCTCCTTTATCTCATTCAATCTTCGCG
CTATGTACTCCAGCGCCTCCTCCCATGATACCTCCACGAATTCACCATTTCTCTTTATCAGCGGTTTTT
TTCAACCTCTCATCACTGTGCACGAACTCCCAGCCAAAACGGCCCTTGATGCAATCGTCTCCTCTGTT
AACACCGTCTGGATTGCCCCTCGCGTATACGATACGGCCATCCTTCACATAGTAATCCATCGCGCAG
CCCACGCCGCAGTAGGGGCATATGGTTGTTATCCTCTTCCCTTCTCTCTCATATGTTTTCTCCCTCATC
GCACCTGTTGGACATGCATCAATGCACGCGCCACAGAAGGCACAGCCCGATTCAATCAATGAATTATT
AACTCCTGCGAATATCCTTATTTCGGAACCTCTCCCGGAGCGCACGAGGACATTGTTGCCCTTTATCT
CATCGCAGGCGAGAACGCATCGGAAGCAGTTTATGCAAGCACCGTAATCCCTGATGAAGAAGGGATT
ATCGTCCTTTATCTTTGCACTACCAGCAGGTTTTATACCATACTGCTCCAGAATATCATTGAATTCCCC
ATCACCGGAATAATCCTCCGCCAGCATCTCAAGTATGGTCCTGCGAGCCTCCTCCACCTCTGGAGCG
TGGGTCTCTATCAGCATTCCTTCCTGCACCCGGGTTGCACAGGCTCGTGGATAAGAAACACGGGAAT
TCTGCGTGACCTTCACGACGCATAACCCGCAGGAGCCGATGGGTTTATGGTTCTCCATGTAGCATAT
CGCGGGTATGGAAATTCCATTCTCCCTGGCAACCTGAAGTATGGTTTTCTCACCCTGAACATCAATCT

CACGGCCGTTGAGCGTCACCTTCATCTTATCACCCCCATGAGCTCATCGCGGAAATAGCGCATGGCC
GATTCATAAGCCAGCATCGCGCTTGTTCCCAGGGGGCACAGGGACGCAATTTTCATGGCCCTTGCAA
TTCTATCCAGACGGGAAATGTCGTCCTTCGTCCCCTCACCCCTAACGATTTTCTCAATTATCTCCCGCA
TCTTCCTTGTGCCGTATCTGCAAGGGAAGCACTGTCCGCAGGACTCGTGCTCAAAGAATCCTGCTAT
GTTGAGAAGAACCTTCCAGAGATCAACCGTGTGATTGAAAACGATTATATCTCCAGTTCCGAGAGAGA
TGCCGCGCTCCTTGGCATCTGTGAAGCAGAGTTTCATGTCAAGAAATGAGGGATCAACAATGTACCCA
GATGCTCCGCCGCCGAGCATCATGCCCTTGAGCTCTCCATCCACGCCGCCAGCTCTTTCAATCATGT
CCCGAATCGTCACCGCTCCTAAAATCTCCTCGTAAACTCCCGGCTTGTGAACGAAGCCGCTCACGCT
GAACAACTTTGGGCCCGAGCAGAATTTCATGCCCATCTTCGCGTATTCCTCCGAGCCGATTTTCAATA
TCACGGAAATATTGGCAAGGGTCTCCACATTGTTGATAAGCGTTGGTTTTCCCCAGAGCCCACGCTGC
GCAGGATATGGTGGCTTGATTCTGGGAAAGCTGCGTTTACCCTCAATGCTCTCCATAAGCGCGGTCT
CTTCGCCGCACACATAGGAGCCGGCGCCAAGCACAAGTTCAACATCAAATGAGAAATTCGAGCCGTA
AATGTTTGTGCCCAGAATTCCATTTCTCCGAAGCTCTTCAATTATCCTCTCCATTCTCCTTGCCAGTTC
GCGGTATTCGTAGCGGATGTAAAAAAATCCCTTGGATGCGCCAATGGCATAGGCGGCTATGATCATA
CCCTCTATTATCTGGTACGGATTGTTCTCCATTAGCACCCTATCCTTGAACGTGCCCGGCTCACCCTC
GTCAGCGTTGGCAACGATGTACTTTGGAGTATCATTTCTCTCCGCAACAAATCTCATCTTCAGCCCCG
TGGGAAACGCCGCGCCGCCCCTTCCAACAAGAGATGATCTTTCAATCATATCAACAATTTCCCCCTGG
TTCATACTGAGAGCATCTCGCAGGGCTGAAAAGCCACCAGCATCCATATAATCCTGCAATTCTGGAAG
ATTGCGAAGGAGCATCATTCCTCATACCCCCCTGAGAATCTCTATCGCACTTTTTGGAGTGAGATTTC
CGTATGCCCTATCATCCACCAGCATTGCAGGCGGGCCCTCGCAGCGCCCGATGCATTCCACCACTTC
AAGTCTGAACAGGCCATCCTCTGTGGTTTCCCCATCCTTTATGCCGTAGTGAGATTCTATCGCATGGA
TTATTTTCTCAGAACCTTTAAGGTTGCAAACTATGCCGTTGCATACCCGTATGACATGCTTTGCAGGCC
TGTCAAAATGAAAGAATGAGTAGAAAGAGGCGGTATCGTGTATCTCCCCCGGCTTTTTTTTCCAATTTC
CTCGGCGACAATCTTCATGGCCTTCTTCGGGATATATCCGAATTCGTCCTGGATTGCATGGAGAATTG
CAAGGAATTTTGAATCCTTTTTAGAATGATCCTGAGTGATTTCCCTGATAATCTGCTCCATGGGATGAT
ATTGGAATCTTTTTATTTATCTTTTGCCGCAAGAATTCTACTTAGCATTTCCATATCTTCCCTCGTATTCA
CGTTGAAAAACGTATCCTCGGGCATTTTTTCCGCGGGTACAAAATCCACATAGGGACAACTTCGCAGT
CCGAGAGTTAAACTCCTGGCACCTTCGAGGCATTTTAGTGCCGAACGAAAGTAGAGGGCATGAAGGG
GTTCAAGATACCTCGTTGAACTCCACCTTGGAACGATGCTCTTGTGCTTTGCATTTTTCAAGAGAATAT
CCACAGCCTCCTTCCTCACGAGGGGCATATCGCCTCCAAAGAGAAGAAAGGATTCATCCATCTTAAG
GAGTGAGACAACGGATTCCAGGATCCAGGAGGATTCGTCGTGAAGAACTGGAACATCCATATCCAGA
GGTATTTTTGAGTAAACAACGACATCCAATCCCAAGTTTTGAAGGTTTTCCACCACAAGGTCTATGAGC
CTCTTCCCAGAGATAACCTCCATCGTGTGCTTTCCGGGAAGTCGCCTTGAGAGTCCAACGAGCACGT
ATGCACGCACATGGTTGCAATGCTCAATCCCCTATATTATTTTCCTTGCGTGAATATGCAATGACAAAA
CCTATTTATAGGGTGTTGCAAATCACATTTCAGATTTTCATGGGTAAGGAGATTGTGGAGAAGCCTCTC
AAGGACCTAATAATCGAAATCTTGGGCAAGGATGGAAGGTCTATAAACGACCTTTCCAAGGAGTTGGA
AAAGAGGGGAATAAGAAAACACAGGTTGATTCTAACAGGTTATCTTCAGGCCATGGCCGATTTGGGAA
TTCTGAAGGAGAGGTATATCAAGCCGGCAAAGGTTTACAGCGTGCAGCAGAATAATAAAAGGAGCAT
CTATGAAATCGTTGGAAACAAGGTAAAGGAGATGAATGAGGAGGAAGCTCCAGATATTGCTCTTTACA
TCCTGTACAAACTTTTCAACAGGCCTATTTTTATGAGGGAGATTGAGAAATGCAATGTGGGCATCCCC
CAGTACAGGGAAAAAATTGTGGGTGCGGAGAGGAAAAGGGCGCTTGAAATCCTCTTATCCCAGGGGT
TTCATATCCCCAGAAACAATTCCGCATACCTGCCCACAAAGGATTTCAACAGCGAATTCATCCAAGTG
CTAAGCGAACTGGTTATTGATGGATACGACCTGAAACATCTTGTAAATAGAGAAAGGAAACAAACAAA
ATTGGATTACGATTGATCACTTTATTATTCCGTAGCCTATAAGACGCCATCTGTTCATTATTCTTCTTCC
TATAGCAATTCTCTGCCCCTCGTGAGGAAGGAGGCTGTGGATATTCTCTTGAAAAATGCAAAGCACAA
GAGCATCGTTCCAAGGTGGAGTTCAACGAGGTATCTTGAACCCCTTCATGCCCTCTACTTTCGTTCGG
CACTAAAATGCCTCGAAGGTGCCAGGAGTTTAACTCTCGGACTGCGAAGTTGTCCCTATGTGGATTTT
GTACCCGCGGAAAAAATGCCCGAGGATACGTTTTTCAACGTGAATACGAGGGAAGATATGGAAATGC
TAAGTAGAATTCTTGCGGCAAAAGATAAATAAAAAGATTCCAATATCATCCCATGGAGCAGATTATCAG
GGAAATCACTCAGGATCATTCTAAAAAGGATTCAAAATTCCTTGCAATTCTCCATGCAATCCAGGACGA
ATTCGGATATATCCCGAAGAAGGCCATGAAGATTGTCGCCGAGGAAATTGGAAAAAAAAGCCGGGGGG
AGATACACGATACCGCCTCTTTCTACTCATTCTTTCATTTTGACAGGCCTGCAAAGCATGTCATACGGG
TATGCAACGGCATAGTTTGCAACCTTAAAGGTTCTGAGAAAATAATCCATGCGATAGAATCTCACTACG
GCATAAAGGATGGGGAAACCACAGAGGATGGCCTGTTCAGACTTGAAGTGGTGGAATGCATCGGGC
GCTGCGAGGGCCCGCCTGCAATGCTGGTGGATGATAGGGCATACGGAAATCTCACTCCAAAAAGTG
CGATAGAGATTCTCAGGGGGGTATGAGGAATGATGCTCCTTCGCAATCTTCCAGAATTGCAGGATTAT
ATGGATGCTGGTGGCTTTTCAGCCCTGCGAGATGCTCTCAGTATGAACCAGGGGGAAATTGTTGATAT
GATTGAAAGATCATCTCTTGTTGGAAGGGGCGGCGCGGCGTTTCCCACGGGGCTGAAGATGAGATTT
GTTGCGGAGAGAAATGATACTCCAAAGTACATCGTTGCCAACGCTGACGAGGGTGAGCCGGGCACG
TTCAAGGATAGGGTGCTAATGGAGAACAATCCGTACCAGATAATAGAGGGTATGATCATAGCCGCCTA
TGCCATTGGCGCATCCAAGGGATTTTTTTACATCCGCTACGAATACCGCGAACTGGCAAGGAGAATG
GAGAGGATAATTGAAGAGCTTCGGAGAAATGGAATTCTGGGCACAAACATTTACGGCTCGAATTTCTC
ATTTGATGTTGAACTTGTGCTTGGCGCCGGCTCCTATGTGTGCGGCGAAGAGACCGCGCTTATGGAG
AGCATTGAGGGTAAACGCAGCTTTCCCAGAATCAAGCCACCATATCCTGCGCAGCGTGGGCTCTGGG
GAAAACCAACGCTTATCAACAATGTGGAGACCCTTGCCAATATTTCCGTGATATTGAAAATCGGCTCG
GAGGAATACGCGAAGATGGGCATGAAATTCTGCTCGGGCCCAAAGTTGTTCAGCGTGAGCGGCTTCG
TTCACAAGCCGGGAGTTTACGAGGAGATTTTAGGAGCGGTGACGATTCGGGACATGATTGAAAGAGC
TGGCGGCGTGGATGGAGAGCTCAAGGGCATGATGCTCGGCGGCGGAGCATCTGGGTACATTGTTGA

TCCCTCATTTCTTGACATGAAACTCTGCTTCACAGATGCCAAGGAGCGCGGCATCTCTCTCGGAACTG
GAGATATAATCGTTTTCAATCACACGGTTGATCTCTGGAAGGTTCTTCTCAACATAGCAGGATTCTTTG
AGCACGAGTCCTGCGGACAGTGCTTCCCTTGCAGATACGGCACAAGGAAGATGCGGGAGATAATTGA
GAAAATCGTTAGGGGTGAGGGGACGAAGGACGACATTTCCCGTCTGGATAGAATTGCAAGGGCCATG
AAAATTGCGTCCCTGTGCCCCCTGGGAACAAGCGCGATGCTGGCTTATGAATCGGCCATGCGCTATT
TCCGCGATGAGCTCATGGGGGTGATAAGATGAAGGTGACGCTCAACGGCCGTGAGATTGATGTTCAG
GGTGAGAAAACCATACTTCAGGTTGCCAGGGAGAATGGAATTTCCATACCCGCGATATGCTACATGG
AGAACCATAAACCCATCGGCTCCTGCGGGTTATGCGTCGTGAAGGTCACGCAGAATTCCCGTGTTTC
TTATCCACGAGCCTGTGCAACCCGGGTGCAGGAAGGAATGCTGATAGAGACCCACGCTCCAGAGGT
GGAGGAGGCTCGCAGGACCATACTTGAGATGCTGGCGGAGGATTATTCCGGTGATGGGGAATTCAAT
GATATTCTGGAGCAGTATGGTATAAAACCTGCTGGTAGTGCAAAGATAAAGGACGATAATCCCTTCTT
CATCAGGGATTACGGTGCTTGCATAAACTGCTTCCGATGCGTTCTCGCCTGCGATGAGATAAAGGGC
AACAATGTCCTCGTGCGCTCCGGGAGAGGTTCCGAAATAAGGATATTCGCAGGAGTTAATAATTCATT
GATTGAATCGGGCTGTGCCTTCTGTGGCGCGTGCATTGATGCATGTCCAACAGGTGCGATGAGGGAG
AAAACATATGAGAGAGAAGGGAAGAGGATAACAACCATATGCCCCTACTGCGGCGTGGGCTGCGCG
ATGGATTACTATGTGAAGGATGGCCGTATCGTATACGCGAGGGGCAATCCAGACGGTGTTAACAGAG
GAGACGATTGCATCAAGGGCCGTTTTGGCTGGGAGTTCGTGCACAGTGATGAGAGGTTGAAAAAAAC
CGCTGATAAAGAGAAATGGTGAATTCGTGGAGGTATCATGGGAGGAGGCGCTGGAGTACATAGCGC
GAAGATTGAATGAGATAAAGGAGAAATACGGGCCCGATAGCATCGCCGGTTTATCATCTGCGAAATG
CACCAACGAGGAGAACTACCTGTTCCAGAAGTTCATGCGCTCGGTTATCGGCACCAACAATGTGGAT
CACTGCGCCCGTCTATGCCATGCTTCCACCGTTACGGGCCTGATGAAGGTGTTCGGTGCGGGAGCG
ATGACGAACAGCATTGAAGACCTTGCCAAGGCGGAGGTTTTCTTCGTTATCGGCTCAAACACGACACA
TGCGCATCCGGTAATTGGCACAATGATAAAGAGGGCGGTGAGAAACGGAGCTGGGTTGATAGTTGCG
GATCCTAGGGATATTGAACTATCAAAATATGCCACGGTGCATCTCCGTCAGAGACCTGGCACGGATGT
CGCGCTAATCAACGCCATGATGCATGTCATAATCAGGTACGGGCTCGTGGATGAAGAATTTGTAAAGA
ACCGCACCGAGGGATATGAGGAATTAAAAAGAATAGTGGAGAAATACCCGCCAGAATATGCCGAGAA
GATAACAGGAGTTCCCGCTGATGATATTATACGGGCTGCGATTCTCTATGCGACGCATCGCTCCGCA
ATAATTTACGCCATGGGGATAACCCAGCACACAAGTGGCACGGCGAATGTTGCCTCTCTTGCCAATCT
AGCATTGTTAACTGGCAACGTGGGTAAGGAGGGCACGGGTGTGAATCCGCTTCGCGGTCAGAACAAT
GTGCAGGGTGCAGGTGACATGGGTGCGTTGCCCAACCTCCTGCCAGGATATATTAGGCTAAACTCGC
CAAAGGTTGAGGAATTTGAGAGAGCATGGCAGGCTAAGATACCCAGAGAACCCGGGCTTACTGTGAT
TGAGATGATGCATGCCATAGGGAATGAAATAAAGGCAATGTACATAATGGGCGAAAATCCCGCTGTTA
GCGACCCGGACCAGAAACATGTTCTTTCCGCTCTCAAAAATCTTGAATTTCTGGTGGTTCAGGACCTG
TTTATGAGCGAAACCGCGCACCTTGCGGATGTGGTTCTTCCAGCAGCGGCATCGCTTGAGAAGGAGG
GCACATTCACCAACACCGAGCGCAGGGTACAGTTGCTTCGCAAAGTAGTTGATCCTCCGGGAGAAGC
CAAGCCGGACTGGTGGATTCTCACAGAACTTGCGAGAAGAATGGGCGCTGAGTGGAATTATGGCTCA
CCGAAGGATGTGTTTGAAGAGATTCGCCGCCTCGTGCCACAGTATCGTGGCATAACATACAGGAGAA
TTGAGAAATTCGGATTGCAGTGGCCATGCCCGGACGAGGAGCATCCCGGTACGAAAATCCTGCATGT
GGATAGATTTTCCAGACCCAATGGTCTGGCGAGATTCTACCCATATGAATATAAACCTCCTGCGGAGG
AGCCGGATGACGAATACCCATTCGTGCTCACCACGGGGCGCACATACGAGCATTTTCATACCGGCAC
TTTAACGAGGAGAATTGAAGGTTTCAATCAGCTTGTGCCCGAAGCATTTGTGGAGATAAATAAGGAGG
ATGCGAAATCTCTTGGGTTAGAGGAGGGAGATTGCGTCGTTGTGGAATCTCGCAGAGGAGCAATAAA
GGTGAAGGTGAAAATTGCCGGGATAAAAAAGGGCACAGTGTTCATCCCATTCCACTTCGCAGAGAGC
GCCGCGAATGCCCTTACAATCGATGCGCTGGATCCTGAGGCAAAGATTCCCGAATTGAAGGTTGCAG
CGGTGCGCATTGAAAAATGCTAAATTTTCCCATATCTTATTTTTCCATTCACAATTGTATATTCTACATT
GAGTCCGTTGAGAGCGTAGACTATATTGGATACAATGGTATCCTTCTTTAGAGGCAATCCGTTGGGCT
TGGGATTTAGCACAACAATATCTGCTAGTTTTCCTTCCTCTACGCTCCCTGCATTTAGGCCCAGTGCCT
TGGGAGCATTCACCGTAGCGAAATCGAGAATCATCTGAGCATTTGTCACGCTAGCATCCCATCTCTCA
TTCTTGTGTAGTAGTGCGGCGAATTTCATAACCTCAAACATGTCTAGGTTGTTGTTGCTCACCGCACTA
TCTGTCCCGAGGGTTATTGTTATTCCTTCATTGATCATCTCCACCAAGGGCATTGAGCCACCGTTTCC
AAGTTTCATGTTGCTCGTTGGATTGTGGGATGCCTTAACCCCATTTCTTGCAAGCATTCTTATTTCGTT
GAGGGTGAGCCACACAAGATGCACACCAATAAGCCTGGGAGAGAGAAATCCAATTTTATCCAGATAC
TCCACCGGTCTTAAACCCGTTTTTTTGCGATGTTCATAAACCTCCTTTCTCGTTTCGGAAACATGCATA
GTCATAAGCGTATTGTACCTATCGGCAATTTCCTTTGCTTTAAGGAGGGTCTCTTCAGAGCAAGTGTA
CACACCATGGGGAGCAACCATGGGAGTTATCAAATCCTCCTTTAAATATTCCTGTACAAAATTTTCAGC
GTTTTTTAGGGGTGAACCCTTCTGGGTGGTGATATCCTCATCCACAACTGCCCAGCCCAGAAATGCAC
GTATTCCCTTCTCCCTCGCGATCTCAGCCACCACATCTTCATCTGAGTACATATCCACGAAGGTGGTT
GTTCCAGTGCGAAGCATCTCATCTATGCCAAGTGCAGCACCATGGCGGAGATCCTCTCTTGTTCTCTT
AGATTCTAGGGCCCACATTTTATTTAGGAATTCCTCCAGATCAACATCTAAGTAAACTCCTCCTGTGGC
TCGAATCTGATATATGGTGGTGAACCGCCCGCCGCCAGTTCCTAAAAAAGTTATCCTTGTCACTGTAG
GGCTAATACTCTGATAGGATTTAATTTTTGGTGTGGACAAAAAACATATTTAGGATGCGAATGATTACG
GAGTATGAGCATTCTCATCCGTGATGCTTGGATCATCACCCAGAATGAACGCAGGGACATAATTAAGG
GAGATGTTTATGTTGAAGATAACAGGATTTCCGATGTTGGAAGGGTCAAAGTGGAGGCAGATATCATA
ATCGATGGAAAAAACAAGATAGTTATGCCTGGGCTGATAAACACGCACACCCACATACCAATGACCGA
TCTCAGAGGCCTTGCTGACGATGTTGATCTGGAGGAATTCCTAAATAAAATGTGGGCCCTAGAATCTA
AGAGAACAAGAGAGGATCTCCGCCATGGTGCTGCACTTGGCATAGATGAGATGCTTCGCACTGGAAC
AACCACCTTCGTGGATATGTACTCAGATGAAGATGTGGTGGCTGAGATCGCGAGGGAGAAGGGAATA
CGTGCATTTCTGGGCTGGGCAGTTGTGGATGAGGATATCACCACCCAGAAGGGTTCACCCCTAAAAA

ACGCTGAAAATTTTGTACAGGAATATTTAAAGGAGGATTTGATAACTCCCATGGTTGCTCCCCATGGT
GTGTACACTTGCTCTGAAGAGACCCTCCTTAAAGCAAAGGAAATTGCCGATAGGTACAATACGCTTAT
GACTATGCATGTTTCCGAAACGAGAAAGGAGGTTTATGAACATCGCAAAAAAACGGGTTTAAGACCGG
TGGAGTATCTGGATAAAATTGGATTTCTCTCTCCCAGGCTTATTGGTGTGCATCTTGTGTGGCTCACC
CTCAACGAAATAAGAATGCTTGCAAGAAATGGGGTTAAGGCATCCCACAATCCAACGAGCAACATGAA
ACTTGGAAACGGTGGCTCAATGCCCTTGGTGGAGATGATCAATGAAGGAATAACAATAACCCTCGGG
ACAGATAGTGCGGTGAGCAACAACAACCTAGACATGTTTGAGGTTATGAAATTCGCCGCACTACTACA
CAAGAATGAGAGATGGGATGCTAGCGTGACAAATGCTCAGATGATTCTCGATTTCGCTACGGTGAATG
CTCCCAAGGCACTGGGCCTAAATGCAGGGAGCGTAGAGGAAGGAAAACTAGCAGATATTGTTGTGCT
AAATCCCAAGCCCAACGGATTGCCTCTAAAGAAGGATACCATTGTATCCAATATAGTCTACGCTCTCA
ACGGACTCAATGTAGAATATACAATTGTGAATGGAAAAATAAGATATGGGAAAATTTAGCATTTTTCAA
TGCGCACCGCTGCAACCTTCAATTCGGGAATCTTTGCCTCAGGATCCAGCGCATCGATTGTAAGGGC
ATTCGCGGCGCTCTCTGCGAAGTGGAATGGGATGAACACTGTGCCCTTTTTTATCCCGGCAATTTTCA
CCTTCACCTTTATTGCTCCTCTGCGAGATTCCACAACGACGCAATCTCCCTCCTCTAACCCAAGAGAT
TTCGCATCCTCCTTATTTATCTCCACAAATGCTTCGGGCACAAGCTGATTGAAACCTTCAATTCTCCTC
GTTAAAGTGCCGGTATGAAAATGCTCGTATGTGCGCCCCGTGGTGAGCACGAATGGGTATTCGTCAT
CCGGCTCCTCCGCAGGAGGTTTATATTCATATGGGTAGAATCTCGCCAGACCATTGGGTCTGGAAAAT
CTATCCACATGCAGGATTTTCGTACCGGGATGCTCCTCGTCCGGGCATGGCCACTGCAATCCGAATT
TCTCAATTCTCCTGTATGTTATGCCACGATACTGTGGCACGAGGCGGCGAATCTCTTCAAACACATCC
TTCGGTGAGCCATAATTCCACTCAGCGCCCATTCTTCTCGCAAGTTCTGTGAGAATCCACCAGTCCGG
CTTGGCTTCTCCCGGAGGATCAACTACTTTGCGAAGCAACTGTACCCTGCGCTCGGTGTTGGTGAAT
GTGCCCTCCTTCTCAAGCGATGCCGCTGCTGGAAGAACCACATCCGCAAGGTGCGCGGTTTCGCTCA
TAAACAGGTCCTGAACCACCAGAAATTCAAGATTTTTGAGAGCGGAAAGAACATGTTTCTGGTCCGGG
TCGCTAACAGCGGGATTTTCGCCCATTATGTACATTGCCTTTATTTCATTCCCTATGGCATGCATCATC
TCAATCACAGTAAGCCCGGGTTCTCTGGGTATCTTAGCCTGCCATGCTCTCTCAAATTCCTCAACCTTT
GGCGAGTTTAGCCTAATATATCCTGGCAGGAGGTTGGGCAACGCACCCATGTCACCTGCACCCTGCA
CATTGTTCTGACCGCGAAGCGGATTCACACCCGTGCCCTCCGGTGGGATCC

9.4 Sequencing data of random fosmids from the constructed Sisters Peak metagenomic braod host range and large insert library

>pRS44_SP_B1_T7
CGNNCGTACCGGGGATCCACATAGTTGTCTAAAATAGTCTATCCTGCCTTGCCTTCAGCCCAGAATCC
AGTCAGAGGACACTTTAAGGCTGTATTTGAGCGATTTACCGGCCCTTTTGAAGCTTATCCGGGTACTT
TTGACTGTTCTATCATAAATCTCCATGTTTGAAAGACTTAACTTACCATAGGTGAGTTGTGGAGAAATA
CTGAGCCTGAGGCTGTCGATCATAAGGTTTGTGTTTTTAAATGATATGCCGGATAGGGATGCTGTGAA
CATTGAATCATAGCCAGCCTTTATTCTCTCAATGCAGAGTATGCTGCTGCAGACCTCGCCATTCTCTAT
TTTAATACTGTTGCCGATTCTGTAAATATTCGCCTTGAGAAACTCAAGCCTTAAATCTTTTGCATGCGC
TTCTTCAAGAGAAAGGCGGGCGAAAAAATTTCGTGTTGTTCCTCTAAAGTCTGAGCGAAGGCTCTNTA
TATAAAAAGTGTCAGCTCCATGATAAATAATGCAGTTTTCAATTTGAAGAACAGCAAGGAAAAAGGGGA
GTTCTGGTAAAGTAAATTCCTTTATAGAGGACTGCATTTTTTGTTTTTGTGGTCTTAAAGGGAGCGAAT
NTGTATATATGCGTATGCCTTTTATTTTAAAGAAACTGGAGACGAAAGCGTATAAGTTCGAGTGGAGA
GGCTGATATACTCNAGGGTATCTATCTCAACAAATGGGGCATGNGCCCCNGTAATATTTNACTTAANC
GGCCCTGAGGATGTAACCTTGGTNATTTAAACCTCTGTCCCNGCAGTN
>pRS44_SP_C1_T7
NNCTGGTACCCGGGATCCACATGTTAGGTCTAAATGCGTCGGAAAAACGCTTTTACGCACGACTGATA
AACAAACGGGTAAAAAAATTGAACTTTCTGTCTCTAAAAAAGGGTATTTCTACAACAAACTTGATTTAGA
TTGGAAGATTACAAATAGTCGTGCTAATGATATTTGGGAGTTATTGCAGTCTAAAGGCATATGCCTTGA
AGCCAGGTAATGCTTTTGGACGCGGATTAAAAATGTTCTTATGGAGCTATAATATACTGACATTGCGTA
TTTGCTACCCTACTTGTTTAATCCGCCCCTGCTGTTGCTAGATTCCAGAGCGATAGCACATTAAAACG
CACAAATGATATTTTAAGCTATCCACAGTTTTTATTTATGAGGATGTTAACTTTTCCCCCTATATGCAAA
GAAAATGGCAGACACATAGCAACGCGTAGCGGTGCGGTGGCGTTTGAACCTGCAACAAGCTTTTGAG
ATTGTTGCAGTGGTTCAGGGCGACACGCGACAGCCCGCAGGGCTGACGGTGTGTCTGCCCCTTGAT
TATATATTATAAAACTTAGGCACTTTAAATTAAGAAGGGTTAAAGTTTATATTGAAAAATAGTGTTGAAA
ATTCTTTAAAAAAATTAGCTTCTCTCATGTGCGCGCACGCGCTCTAAGTTTTAATACTTTNAGTTTTTCT
CTGTTTTCGGGGTCATNNTGCCCCCCATAAAATCGGGCCTTTNGTGANGGGTTAACTGTGTACCAAAA
TCCNCCGTTNAACTGTGTACAAAATCCTCGGTTAATTGTGTACAAAATCCCCGTGAAATG
>pRS44_SP_D1_T7

NTCGCNCGTACCNGGGATCCACTAATTTAAATAATAACTAAGGCTGAATAAATAGAATAAATATAAAAT
ATAATATTAGAATTCGCAAGCAGTCTTATTAACAGTAACTCAAAAACTTTGGCGGTGATAGATGTATGC
TTGAACTGCTTGACGTTTTGTTTAATCCAGACGGGTTCTTTAAAAATGAAGTATCATTAGAAAAACCAAT
ATGGAGGGGAATAATAGTCGTTACAATTGCAGCCCTCCTTTCTTTTGTAAGTGCGTACATTAGTGCAC
CAATTGTAGCTAAAGCTACCTACAATATGTTAATCAGTAAAGGTTTAAACGAGCAAGTTGCTGAAGTCA
TTGCTACAAACATAAAAACGACTGGTGTTGCTCTCTCCGCTGTAATTGTTTTTGTAGCGTGGATTTTGA
TCGCACTAATACTTTACGGGCTATCTGCTCTATTCAACGGAAGCGGGTGTTTCCTTGACACCCTAAAA
GTTACCGCGTACAGTTTCATCCCGTCTATAGTAGTCTATCCATTTAGATATTACATAACAATTTCAAAAA
TAAGCATTATAAAGGCTGAAGGGCTAATGGCATTAGAACACTCCGGCATTGCTATTGCCGAATTGATT
ATAAATCTGGCAGTTCTCGTATGGCAGTTTCTACTCTGGAGGTATGGNATAAAGTATGCAAGAAGGCT
AAACGACAGGGATTCGATAATAGTAGCAGCTATACCAGCTGTTGCTTTCGCTCTACTTATGATTGCCT
CAACTATAAAAAGAATTTAAGAATTTTCATATTCNAAAATTTTTAATTTTTAANTCAT

>pRS44_SP_F1_T7

AATCGACTCGTACCCGGGATCCACAGGNCAGGTGGATAGCTCACAAAATGGCTTCTGAGCTTGAGAG
TTCCTCAAGTAGAGGGGCATACGAAAATCCGAGAGAGTTAATAGAAAACCTCTCTTCTCCAGTCCATC
AACTTGTAGCGGAGTTCCAATGGCAGGGAGGGTTGAGGATTCACGAAGTCCAGCAGATAAAGGAGAG
CTCCCTAAAGAGTTTCACAAATTCGTGCATGCAGGAATTTAAAGGGATTCTGGTTAGAGGAAAAGGAG
GCTATGAAAGGCTTATAGAGCTCCCTAATGGATTGTGGGAAGAGTTGAAAGACCACTTANAGCAACAC
GGTAGGCTTGAGTTCAACTACAAGGAATACCTTAACGATTTAAGACGTTCTGCTTGGGAGACTGGACA
GCCTTATCAAGGCTCTCACGGTTTACGCTACNACTTTGCGCAGGAAAAGTATGANNAGCTTATAAGAG
AGGGCTTANGACACANCGAGGCCTTAAAAGAGGTGAGNGAAATGCTGGGACATCACAGACCCGAAAT
AACTTTGCACTATTTGGGGAGGTAAACTTGAGAGAGCTTTTANAGAAAGAAAGAGAGCTCCTATCGGT
ATTCAGAAAATACCTTCNCTATGAAAAAGGCTTNGTCCTTTTCTGTAATGGAAAACTCGCCTTAAAGGA
GGATTTCCCNCCTCTCCATATCCTTCCGTAAAGGAATTTATTTCCTGANATANNATAGCTNTACTAACT
TTCATAAANCTNNCTTTTTGANATCTTTGTCCTTTNNNAAGNGATTTCCNAATACNAGNGANAGCTANT
NGGTT

>pRS44_SP_G1_T7

ANTCACTCGTACCCGGGTTCCACGGGCGATGTCGGTGCCGGGGGGCTTGCTAGCAAGCATAGCGGA
CCGCCGGCAGCAAATCGCGGATAATAGGAGCTTTCGTGGTGGACAAGCCCCTGGAAGGAAGCGCGA
TGGAATTTAGCACAAAAGCCTTGGATTGGAGCAAAGCGGGCCAAAATGGTTTCCTGGCAACCAAGAC
CGATTGTCTGGTCGTGGGCCTGTTTGAAGGACAGAGCCTGGCCGGTGTGGCCAAGGCGCTCGACGT
GGCAACCAAGGGTTTGGTCTCCCGGCTCGTCAAGCTGGGCGATTTCGAGGGCAAGCGCGGTACCCA
GCTGATGCTGCACGAAGTGGCCGGCGTGGGCGCTGCCCGCGTGCTGCTGGTGGGCTTCGGCAAGG
AAGCCGAGTTCGGCGACAAGGCGTTCGCCGAGGCCGTGCGCACCGCGCTGCGTGCCTTGTCCGGC
ACGCGCGCGGCCAATGTGCTGTGGTGCCTGGCCCAGCAGCAGGCCCCGCAGCAAGCCGGCAACGC
CCGCGACCTGGCCTGGCGCCTGATCACCACCGTCACCCTGATCCGCGAGGCGGGCTACCGCCTGCT
GGAGCGCCATCCGGGCCTCAAGCGCGCCAACGCCAATGCCAAGGGCAACGGCAACGGTGCCGACA
AGCCCACCGGCCTGCGCAAGGTGGTGCTTACCGTCGACGCGGCTGACGCCAAGGCGGCCGCCCAG
NCCGTGGTCCGCGGTGCGGCCATCGCCAACGGCATGGANCCTGACCAAAGAACCTGGGCAACCTTG
CCGTCNAANATCTGCACGCCNAACTTACNTGGC

>pRS44_SP_G1_T7

ANTCACTCGTACCCGGGTTCCACGGGCGATGTCGGTGCCGGGGGGCTTGCTAGCAAGCATAGCGGA
CCGCCGGCAGCAAATCGCGGATAATAGGAGCTTTCGTGGTGGACAAGCCCCTGGAAGGAAGCGCGA
TGGAATTTAGCACAAAAGCCTTGGATTGGAGCAAAGCGGGCCAAAATGGTTTCCTGGCAACCAAGAC
CGATTGTCTGGTCGTGGGCCTGTTTGAAGGACAGAGCCTGGCCGGTGTGGCCAAGGCGCTCGACGT
GGCAACCAAGGGTTTGGTCTCCCGGCTCGTCAAGCTGGGCGATTTCGAGGGCAAGCGCGGTACCCA
GCTGATGCTGCACGAAGTGGCCGGCGTGGGCGCTGCCCGCGTGCTGCTGGTGGGCTTCGGCAAGG
AAGCCGAGTTCGGCGACAAGGCGTTCGCCGAGGCCGTGCGCACCGCGCTGCGTGCCTTGTCCGGC
ACGCGCGCGGCCAATGTGCTGTGGTGCCTGGCCCAGCAGCAGGCCCCGCAGCAAGCCGGCAACGC
CCGCGACCTGGCCTGGCGCCTGATCACCACCGTCACCCTGATCCGCGAGGCGGGCTACCGCCTGCT
GGAGCGCCATCCGGGCCTCAAGCGCGCCAACGCCAATGCCAAGGGCAACGGCAACGGTGCCGACA
AGCCCACCGGCCTGCGCAAGGTGGTGCTTACCGTCGACGCGGCTGACGCCAAGGCGGCCGCCCAG
NCCGTGGTCCGCGGTGCGGCCATCGCCAACGGCATGGANCCTGACCAAAGAACCTGGGCAACCTTG
CCGTCNAANATCTGCACGCCNAACTTACNTGGC

>pRS44_SP_A2_T7

CTTCGCNCGTCCCGGGGNTTCCACAGGNATAACTGTTATTAGGAGTATTACTCCAACTATTAGGAAAA
CAATCCACAGTAAACCGCTAACCGTTATAAAGAGGGCATAGTTAAGGGAGAGAGCTCCCTTCCAAAG
GGAAGGTGAAAAGAGCTTAAAAAAGCCTAAAAACGCCTCTCCAAAGCTCTCCTTCTCCATTGCGTAGC
CAAAGGCTAATGGGTAAACGTAAAAGAACCAGGAGAGTACTATAAGGGTTAAGGCAAATACAAAAAAC
AGCTTTGTAAGTTTGCTTGTTAGCTCACTCTCGTTTAAAGAAAGGCCTGCAAGCAGTGCAAGCCCTAT
TAAAGATAAAACGAAAACGGTTCCAAGTATAAACTGGGCAAGGAAAATCCCTGCTGCAGTTTCAAACC
TCTCTGTAAAAACCTTCTTTACGGTTGTTTCATTTAAAAATTTTAAGAGCTCCTCCCTCTCACTCTTTCC
TCTGTAAAATGGGGTTGAGTAGTAAACCATTATCTGGGTTGAAATTAGCTGAATTAGGATATTTGAGAT
TATAGAGGTTATGTGGAAGAAGGAGAGAACTGTTAAAACTATGAAAATAACTGTATATAACCCTACAGC
CTTCCAGTTCTTTACTACGCTGTTNAAGGAGTAATCTAAAACAAGTGATAGCCTACTCATTTCTTCTCC
NAAAGTGGTGCTGGCGGGGGGACTTGAACCCCCGACCTGGGCATTACGAGTGCCCCGCTCTGCCCA

GCTGGNCCTACGCCAGCTTAANTCTTTTCCNTTCTNACCNNCCTTTTCNAAGTAAAGAAGGTNTTCCT
ACTT
>pRS44_SP_B2_T7
CNNCTCGTACCCGGGATCCACGTGCGCCCCATATTCCNTGAGAGGCCAAATACAAGCAGTAGGATTA
TACCGGCAACGGTGTGGGGAATCACAATTGGAAGGTCTATTATGCTTTCAACAACACTCTTTCCAGGG
AATTTTTTTCTTGCAAGTACATAGCCCAGGGGAAGGGCAAAGACGAGCGCTATGAGGGTTGAAGCGG
TGGATGCAACAAAACTCACCATTATGGCGTTTATGAGTGAGGAATCGTATATGGAAGAGGGAGAAAAA
TAAAGGAGCAAATTGAGGATTGGAAACAAAATAAATACAATGAAAAATGCGCCGGCAAGGTAGGCAG
CACGGTAGAAGTTATCCTCTTCCAAAATAAGGGTGATTGCCAGTATTGAGACCAGCCCAAGTATGGGG
TGAGTTGCAGCGGGCGGCAATTGCCAGGGCAGTCATTTTAAAGCCCAAACTTCTGCTTCTTGACGAG
CCCTTTAGTGCCCTTGACCCGAAGATGCGTGATTCTGCCCGAGCACTTGTAAGGGACTTCATAGCAG
AACTGGGAATTGATACCATTCATGTGACCCATGATTTCTCGGATGCCTGGGTCATGGGAAATAAACTG
GCAGTTATGAAAGATGGACAGATAGTAAGATCCGGTCCTGTGGATGAAGTATTTGCAGATCCGCAGAA
CGATTTTGTTGCATCTTTTCTTGGCTCTACAAACATTCTCAATGGAAAGGTTGAGGAGATTGATGGCCT
GACTAAAATAAAAATAANCGGANCATCNGTTTTCACGGTGGATTCTGCGCA
>pRS44_SP_D2_T7
GGCTNGTACCCGGGATCCACCGGGNGGCAGGATAGTAAACATCAAANAAATAAAAAATATTATTAAAG
AGAAAAAAAGTAAAAAAATTGGAGTATTTGGAGTAGTAGCTTCAGTAGTGATTAAAAAGATAGTAAATC
CAGATTTTGAATTTGGCTCAGTACATTCTGCCATAGACGCATCATTTAGAGGGATAGATACTATTCTCT
TTGTTTCTCATGAACTTTTGCCCTATACTCTTCAAGTAATTTCAGAAAGTAATGTGAGTTATGAAGTGAA
AAATATAAAGGATATTTAACTTTATCGGCGGGAAGTCCCCTTCCGAAAGGGAGGGGATGAGAGCCTAT
ACATTTATATAGTTGAGAGTAATTATATATAACTATGAAATATGTGTTAGATAAAGTTTCACATTCTGTTT
ATTCACTATATTACCATCTCATTTTTGTAGTGAAATATAGGAGAAAGGTNATAACTCCAGAAATTGATGA
TTTTCTAAAACAGAAAATAAGAGAAATTTCAGATNCATTTGAGGTTGAAATTTTAGCGATNGAAACAGA
TAAAGACCATGTCCATATTTTATTCAAAGCAACACCAAAACTTGATATTCCNAGATATATAAANGCACTC
AAAACGATAACTNCTCGTGNAATACAGAGAANGTTTCCNNAANTTAAAANGAAGTTNTGGAAAAATGTT
TTTTGGTCACCGTCCTTATTTCATTGCTACNTCGGGGTCAAGTAACANTAGATGTGCNGNAAAAGANA
TTGNTGAANGTCANGGTGAAAATAATGNNTGCGAAGTTTNTNCTT

9.5 Sequencing data of the 5’ / 3’ insert ends of the 12 identified fosmids with putative positive tested H_2 uptake activity

>pRS44_SP_P44H11_T7
ACCTATGGGGGCGAATTCCAGCTCGGTACCCGGGGATCCCACAAGAGCGCAAAGCGCGAGGGCATC
CCGCCGAGGGAAATACAGCGTTAGCGCAGTAAAACGGACTTTGTTCGTTTTGCGTCCTATTAAATAAA
AAGCGCAAAGCGCGAGAGCCCTCTGCTGAAGAGAACACAGCGTTAGCGTAGGCAAAGAGGAACTTG
TTCCTCTGCCGTCCTATTCAAGGAAAGAGCGCAAAGCGCGAGGGCATCCCGCCGAGGGAAATACAG
CGTTAGCGCAGTAAAACGGACTTTGTTCGTTTTGCGTCCTATTAAATAAAAAGCGCAAAGCGCGAGAG
CCCTCTGCTGAAGAGAACACAGCGTTAGCGTAGGCAAAGAGGAACTTGTTCCTCTGCCGTCCTATTC
AAGGAAAGAGCGCAAAGCGCGAGGGCATCCCGCCGAGGGAAATACAGCGTTAGCGCAGTAAAACGG
ACTTTGTTCGTTTTGCGTCCTATTAAATAAAAAGCGCAAAGCGCGAGAGCCCTCTGCTGAAGAGAACA
CAGCGTTAGCGTAGGCAAAGAGGAACTTGTTCCTCTGCCGTCCTATTCAAGGAAAGAGCGCAAAGCG
CGAGGGCATCCCGCCGAGGGAAATACAGCGTTAGCGCAGTAAAACGGACTTTGTTCGTTTTGCGTCC
TATTAAATAAAAAGCGCACAGCGCGAGAGCCCTCTGCTGAAGAGAACACAGCGTTAGCGTAGCAAAG
AGGAACTTGTTCCTCTGCCGTCCTATTCAAGGAAAGAGCGCAAGCGCGAGGGCATCCCGCCGAGGG
AAATACAGCGTTAGCGCAGTAAAACGGACTTTGTTCGTTTGCGTCCTATAAATAAAAGCGCAAAGCGC
GAGAGCCCTCTGCTGAGAGACACGTCGTTAGCGTAGGCCAAGAGGACTTGTTCCTCTGGCGTCCTAA
TTCAAGCATGAGCGCCAAAGCGGCGAGGGTCATCCGCGAGGTAAATCAGCGTAACTCATGA
>pRS44_SP_P30B05_pCC_PW2
AGGGGGTGAATTATTTATGAAGCGCAACAAGCAGAAGAGGTTGGTTATCGGGAAATTGTTCTATGTGG
AATTCACCTCGGCTTGTATAATCAAGAGAAAAATAACCTTCAATCTGATAAAAATTTAGCTTGGTTAATT
GCAAAAATCATCCAACGAACAGAAAAATTAAGAGTGAGAATTAGCTCTATTGAAGTTAATGACATTGGT
AAGGATTTAATTGATTTATTGCAAAAAGAAGAAAGATTTTGCAAGCATTTGCATATTCCATTACAAGGA
GGTGAAGATAAGTTGTTAAAAAAAATGAATCGTCCTTATAATATAAAATATTTTCAAGAAAAAATTTACA
AAATTAGAGAGGTGGCACCGGAAGTTGCCATAACCACTGATGTGATTGTTGGATTTCCGGGAGAAACT
GAAAAGCAATTTTAATCAACAATTGCTTTTATCTAGAAATTGTATTTTAGTTAATTTCATGTTTTT
>pRS44_SP_P11H04_T7_#3

CGCAAAAAGGGGGCCGATTCGAGCTCGGTACCGGGGATCCCACTCAAGAGAGATTGAGGCGGTATA
TAAAAAAAATGATGCGATTACACACAAGATTGACATATTTTTTATTTGGTATTTTTGTTGCATCTATAATT
TTACCATTTTTGGATGGCATCGGATCATTATCCCTTTCTATTGGATTGTTTATTTACCTTACCGTAGTCG
TTTCACTCTATTTTAAAAGAAAAAAAGATAGAAAAAAAATGAAACTTACTATCAAAAAAACAAAAGAGCA
TAAATCATTAGAAAAAATTTTCGACCAAAAAGATAAAGAAAAATCCCGTAAGCATGATGTGATCATCAA
TCAAATTGCCTATATCCAAGAGTTGTGGGATTTGAATAAAAATCAACAAAAAACTTTTAAAAATTTTGAA
GATAAACGTGCCTATAGTGTTCTTTATACTAAAATGACAAACTCCCTACTTCCCCAACTTATCAAATTGA
TTGAAATTTGTATCGATCAAAACAAAAGTGGATGTAAACGAAAAGTTACAAAAAGTATCAGTGATCTTG
TCATAATCATGAAATCTGACATTGCACACGCAAGGGACAAAACCTAAGATCATTTTGAAACACAAGCG
AAGTTTACGATAAACTTATCCCTGAAATAAAATACCTTAAAATATTACTTAGGTCTTTAAATTTTTGATGG
CATCTTATGCTACACTCTGTCATCTTTTGCATCTGCTCAGAAGAAACCATCATACGGTGGAATATAAAT
AGAACAACCATTTCCCTGGAGGTCATAACAATACGTCATG

>pRS44_SP_P31A02_pCC

TAAGGCAAGTTTTCACGGGAAAAGCTATGACCATGATTACGCCAAGCTATTTAGGTGAGACTATAGAA
TACTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACAACAACTCGGGCTATTTTTATATC
ATCCATATTTATTATTTTGTCGAACGTTGCGTTTGGATCACCCTCCATGAGCGACTCTACTGCCTTCAC
AACCCTCTTGCCCACGATTGGGTGATATTTTTTTGATAGTTCCGCTGCAACGGGCATAAGCATCTCGG
TAAAGAATATGAAAAAAGCGTTGTGCGGGTTTTTGGCATCCACAGCCATCCTGTAGAATATTCCATCTT
GCGTGGGTACATATGACCCGTCTCTTCGGATTTTATCAATTTTTTCAGCATCAAGACCTGTGAGGGCT
GCGATCTTCTTGTGTGTCTCGTATCCTTTGAGTATAGCTTCTGTTGCGGATAGCTTGCGGAAATCTTTG
TCGCTAATTCCTAATTTTTTCAGCCATTCTTCTTTTTTGTCCTTCTTTTTCCCAAAGCCGAATGGCACAT
CGCTCCCTCCTGCTTATGATTTTGTTCCCCAATACTACTTTTGGAGCTTTTTCCTTAAAAGTTGTTATCT
CTTGGCTTTATCCTAATCTGAAAACGTCCTCACAAAGACTCATCTTTCCTCCATAATGAGGAGTGAAAG
AACGTAACCACTCAAGCCCTGCTATTGAACGAGATAAGCCTTTCGAAGTCCTCCAAGTCCCAAACAGC
CCCTTCCCCGGGTTTTTCCTTACCTTCAACACATTTTGCAGAGAGGCCGTAACACCTCTCCCGGCCGC
AAGGCAACCTTTCAGCTCTGCTCAGCTTATACCGGGGAGTTTCTTTCCCGTTCTCTAAATATTTCCAGC
AGTTCTCTCGGAGAGTAAACTCTATGATCCGCTTGCATCTTTCTCCAAAGCCTGTCCCCGTGACCAGG
GACCGTGCCTGGTGGAAACACAAGGCAACGATAGGAGCTCTTTAACGTCCGCGGAAACTACATAGCA
GTCTACGCTCTGCAGTACTTTCTTGCACTGAAATATCATCTGAATCACATTCCATCGTGCAGTCGTCAC
TTCAGAGTTCTTGCTTCATCC

>pRS44_SP_P31A05_T7

TGCACTTTTGGGCGATTCGAGCTCGGTACCCGGGGATCCCACTTTGTGCTTGAAGACTTTGTAACTCA
TTTTGTGCTGCATTGGCTTCTGATGTTGCTTTTTGTATGCCTGTTTGTGCTTGTTCTATAAGATTTGGAA
TATCAACAGGATCAATTTCAACTAAAAGTTCTCCTTTTTTTACCCATTGACCTTCATCGGTGTGAATTTG
TAAAATACGACCACCTGTTTGGGCTGTGATATTATAAATGTCTTTTGCACCAAGATTTCCAATTCCAAAT
ACACGTACTTGCAAATCACCTCTTGTTGGTGTAAATTTTTCATAAGTTGTTTTGGTAATATAGATTTTTTT
ATAAAATATAAATCCTATAAGTATCAAAATAAATAGTGTAATGAGATATTTCATATATTTTTTCATTTAATT
TTTCCTTCGAGATAGAGGATACGATCGTAAGCAAGGCTTCTGGTATAATATGCTTGTAAGATAGCAAG
ACGGCTACTGAGTTGTTGCATAGATGCATCAAGTAGTTGAATATACGTACTCATACCTTCTTTGTATCG
TCCTTCAAGCACTTTTGCACTCTCTTTTGCTGCTGCAAGTTGTGCTTTTTTTGCTGCAATACTTTGTGTA
TATTGGTTGATATCTACTAGTAAAGATTTGAGTTCTTTTTTTAGTGAAATTTGTTTTGATAATTTACGTTC
TTTTGCAATTTGTGAACCAATTTTTGCTTTTTGTGTTTGGGCTGATAGTTTGCCTCCTGTATAAAGAGGA
AGGTTTAGAGAAATACCTATTACTTTACTATCATAGCTGTTAAGTGTATTTAAATGTGTAAAAGATCCTA
CAGCATCAATTGAACCATAGTGTGAAGCATGTGCTGCTTTATGTAAAGTTTTGTTCTTTTTCAATAGTAA
GATGTATCTATTTGAAGGTGTTGGGTTTCTTTTTAAAATAGTTTGGGGATAGTTGTTTTAATTCGTTTAG
GTGGCATTGT

>pRS44_SP_P37B01_pCC

CAATTCACGACGGCCCACTAACTCTTCTCCTCTATTTAATCATGATTACTCCAAGTTATTTATGTGAGA
CTATAGAATACTCAAGCTTCCCTGTAAAGCAGGTCGACTCTAGAGGATCCCACGAGCAGACGCGCAG
CGGATGATCTTCTACAGAGAGAAAAGTCAAGCACAACCGTGTTGTGCGCGGAGGTGCGTGCTCGCAC
CGGAGTTAATCTTAATCTTAATCCAAGAGGAACTTTGGGGTTTGGTCACAGGTCACCACTTCAGCCCA
GGTGTGGCGGTAGGGGAAAAGAAAAAGCTGGAGCTGAGATAGAAAGTGCCACCTGCCCCCAGAGCG
ATGCGCGAAAGCTCACGAAGTGAGTCGTGCAATCGCAAGACAAGCGAAGCGCGTCAGTGGTTAGCA
AGAGAGTTTACCACCCTTGATGAAAACCAAACTCGTGATCGTAGTCAAAGTCGTTGTCCTGGCTTTGG
CTTTGCTCTTTGATGCGTGCCTGCTCTTCTTCGTGAACGGATTGCCACTCTTCTAAAAATTCGGGTATG
AGTTGACGCATTTTAGATTTGAGTTTGTCTATGATCCTGTCAAGAGATCGTGAGACCGCTTCGATCGC
GGTACTCAATTCATACCCGACACCGCTTTCAAAATTTTCGTTATAGGCGCGCTCGGCAATTTCGATTTT
GAGGGGATCGCGTGATTTTGGCTTTTCATCGAACATCTCCATAAAAGACGCATTTATCTCCGCGTGAA
CGGGGTTTGGAAAAATAGTCTTGACAGCGGTCGCGATCGCCTCTTTTCGCTCGATCTCTTCAATTTCG
GCTTTGCCTCGATCCGAAATCAATTCAGAATACGGCGATTTAAGCGGTTTTTTTTCGATTTCAAGGGTT
AGGGTGCGGTTTTCGTCTTTTAAACGCTTATTTTCGCGTTTTAGGCTCTCTATTTCGCTTTTCAAACGT
GTTTCGATGTCGTCAATGGAGCGTTGTAGCTCCCCTATCGTCAAATCTTTTGATTTATGCGCTGCTTCA
ATCTCTGTGAACTGTTAAGCTGTGTGACTGGTCTCTACCGCACCCAGTTCCTTTAGT

>pRS44_SP_P44H11_pCC

TGTGGGAACTTTTCCCGGGAAAAGCTATGACCATGATTACGCCCAGCTATTTAGGTGAGACTATAGAA
TACTCAGGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACCCTTGAATAGGACGGCAGAGGAA
CAAGTTCCTCTTTGCCTACGCTAACGCTGTGTTCTCTTCAGCAGAGGGCTCTCGCGCTTTGCGCTTTT

TATTTAATAGGACGCAAAACGAACAAAGTCCGTTTTACTGCGCTAACGCTGTATTTCCCTCGGCGGGA
TGCCCTCGCGCTTTGCGCTCTTTCCTTGAATAGGACGGCAGAGGAACAAGTTCCTCTTTGCCTACGCT
AACGCTGTGTTCTCTTCAGGAGAGGGCTCTCGCGCTTTGCGCTTTTTATTTAATAGGACGCAAAACGA
ACAAAGTCCGTTTTACTGCGCTAACGCTGTATTTCCCTCGGCGGGATGCCCTCGCGCTTTGCGCTCTT
TCCTTGAATAGGACGGCAGAGGAACAAGTTCCTCTTTGCCTACGCTAACGCTGTGTTCTCTTCAGCAG
AGGGCTCTCGCGCTTTGCGCTTTTTATTTAATAGGACGCAAAACGAACAAAGTCCGTTTTACTGCGCT
AACGCTGTATTTCCCTCGGCGGGATGCCCTCGCGCTTTGCGCTCTTTCCTTGAATAGGACGGCAGAG
GAACAAGTTCCTCTTTGCCTACGCTAACGCTGTGTTCTCTTCAGCAGAGGGCTCTCGCGCTTTGCGCT
TTTTATTTAATAGGACGCAAACGAACAAAGTCCGTTTTACTGCGCTAACGCTGTATTTCCCTTCGGCGG
ATGGCCCTTGCGGCTTTGCCGCTTCTTTCTGAAATAGGACGGGCAGAAGAAACCAGTTCCTCCTTTGC
CTACGCTAACGCTGTGTTCTCTTCAGCAGAGGCTCCTCGCGCTTGGCGCTTTTATTTATAAGACGCAA
ACCGACCAGGTCCGGTTTACTGGCCCTACCTGTATCCCTCGCGGAAGTCCCCTCCGCTTGGACT

>pRS44_SP_P44E07_T7

ACAATAGGGGCGCGAATTTCAAGCTCGGTACCCGGGGATCCCACGTGTGAGTCTCTCATTCAGGGAT
TGAGGTGCTTAATGTCAGGGTAGTTTCTTTTCTGTTTTTTTCTCCTTGCCAAGTTACTGCTCCTATAGAA
CGAGGGCTACTCCCTTTGTCAGAGTTTAGTGCTTTTCGGTAGATTTTTCAAGTTTTTTGTTTTTTAAAAA
GGGTGTCCCCTAAAATTTGCGTTATATTCTGTCTTCATTACTAAGGTGTCGTCGGTGCCAGTCTTGAGT
CCCTGCAAAGTTACGCCCTATCGGGCGTTGCCCTGTCGGGCAAGGGTGAGAAATTGCTTCGAAGTTC
TATCCGCCTCCGGCGGATTGCGATTGCAAGCAATCGCAGTTGCCTGCTTGAGTTTTTTGCTGTTTCTT
TAAGTTTTTGCTAAAATAGAAGTGATTGGGGGTTAATTTGAGTTTTCTTGAGATTGACTTTGAGAGATA
CAGAGAATTTAAGGAGAGAATTTTCAAGGTTGCTAAACAGAAAAACAGTCATCGCTTTTTTTTGCTGAT
TGACATCTCCGAGCAGTTGGCAGAGAAGTATTTGCCAAAAAGGAAGTGGAACAACAGAAGAAAAAATC
AGAGACCTTATCTTGTGTATATTCCTTTGCCTTATGAGAGATTGTTAGTTTTAGGCAATCATAAGAGCA
AGATAGATTTTCTTAAGGATATTTCCCTTTTGCTTGAGCTGAGTGAGGATTATGTTGTTGGCAATTTATC
GAGAGGTGCGTCTATTGACAATGTTTTTTTGCTGAAAAGAGGAAGCACGGAGATTACACATTCAACCA
CACCCAAAGGCTATCAGGAGTATGTTTGGAGAGATAAGAAGGATTTTAGATGAGAGTAATTTAAGGGA
TGTGTATAAGACAATTTACAGGGGTTAGTGAAGTTTAAGTATCACTTTTATGAAGGGATACAGAGACCA
TATATCGGGTTGTCTCAGGTTGACTGCTCGTGGACATGAGAATCAGATGAAAATTAAGCCGA

>pRS44_SP_P44E07_pCC

TTGGGGCAAATTTTTCCCTTGGAAAATCTATGACCATGATTACTCTACGCTTTTTCGGTGAGACTATCC
AATACTCAAGCTTGAATGCCTGCAGGTCGACTCTAGAGGATCCCACGTGCAAGGAATAAAAAACAAAA
AACTGACTCAATCCCTTTACGAATCAATTAAATCATCAATGGATAAACTTTAAATAAACCTTATGGGGG
AAGTGGGGGCCCCCGTCTATTGCTGTAGCTGTTTTCTGTTAGAAAAGCCCTGTGCTCTGAAAGGAGC
TTGCGGGGTTTGGGGGGGGTTTGGATAGTATATTCCTCCCAAAAGGAATTGGCTATCCTTCCTACCCA
ACAATCACTTAAGAAGACGGGCGAGCCGAAAGCGCGTATAAAAGCGTTGCCCACGGCAGTGAACCC
CAAAGGCACGGCTGAAGCTCCATCATTCGTCCCCCTGAGCCGTCCCGTAAGCCACTCATCCGCGAG
GATGGGGGTGAAGGGATAAGCCCTCTTTTCTGAGGGTGAATGTAGGCTGTGGAGGAAAGTCCACAG
CTGAAATAGTTCGGCGAACAAGCACCTGAAAACGGTGCGAGGGGAGTTTATAGCTCACCTTCAAAAC
GAATAAGGTTAAAATTGCAGGATATGGTGACGCACAGAAAGGCCGTGTGAAGCCCGCTGGTGGCAG
GCGGATGATCTTCTGTCTTCGTAGTCCGCTTATGCCTCACAGGAAAAGGAGCTGATACTGTGAGGAT
GTCACCTTACTCTCCCCACGAGGACACTGGTGGACGGCCCACCTAAGTCCTGCAAGAGCCTATAACG
AGCTGCCGTGATGGATAGGCTGAAGCCTCTTGAGCTCTCTTGTTCAAGACAGTTAGCCCTTGGGGAT
CAGCACAAAAGCTTTCGAAACTACGGCTACTCCAGGTGAGGAGAAGGGCATTAGGGCTAAACCGTGT
GAAAACACGGTAGGATGAAAATCCCGTCCGAGGCCTTTGACTCTAACGAAAAAACAGGCACAGTGAT
AG

>pRS44_SP_P37B03_pCC

ATTTGGGAAAAATTTTACACGGAACAGCTATGACCTGATTACGCCAGCTATTTAGGTGAGACTATAGAA
TACTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACATGGGGGGGGGGCGGCCTCCCA
CCAGGCGGCCCGGCATAGGGAGCAGGCAGCATGAGCGCGCCAAGCGAGATGATTGCGGCACGCGG
CGTGCATGCCTGGTATGGCAGCAGCCACGTCCTGCGCGGGATCGATTTCAGCGTCGGCACGGGCGA
GACCGTGGGCTTGCTGGGGCGCAACGGCATGGGCAAGAGCACGCTGCTGCGTACTCTCCTGGGCC
ATGTGCGCCAGCGTGAGGGTACCATCCTCGCCGGCGGGCGCGATGTCTCCCGTGCGCAACCGTATG
AAGTGGCACGCCTGGGCGTGGCCTATGTGCCCGAAGGGCGAGGCATCTTCCCCAATCTCTCCGTGC
GCGAGAACCTCTTGATGGCGGCGCGCAAAGGGCCGTCTGGCCGGATGGACTGGGACGAGGCAAGG
GTGCTGGACCTGTTCCCGCGCCTCAAGGAGCGCCTTGGCCCACGGCGGCCAGCAACTGTCCGGCG
GCGAGCAGCAGATGCTGTCGATCGGACGGGCCCCTGATGACCAAACCCGGACTGCCTGATCCTCGA
TGAGGCGACCGAAGGCCTGCGCCACGCATTGTCGCCGAGATCTGGCAGTCATTGCCACCCGTGCGC
GCCACCGGCATTGCATCCGTGGTGGTGGATCGCAACTACCGCACGGTGCCTGGCCAATGCCGATCG
CGTGTGATGCTGGAGAAGGGCATGGTCGTCATGACCGGGCCTGCCGCCCGAAACTGGCGGTGGAG
CCCGGGCCTGCCTGGATCGCTATCCTTAGGCGGTT

>pRS44_SP_P35F04_pCC_#2

CATAAGGTTCAAATTTACACGGGAAAAGCTTGACCTGATTACGCCAAGCTACCCCTAGGTGAGACTAT
AGAATACTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACACGATATGTATGCTTGATTT
TGCCATTGCGCACGCTCCTTTTTGTGGTAAATAATTGAGGGGCTTGTCCCCCAATTATACCACAAAAA
CCCACACACATCTCGGAGAGGTGTGTAGTGGGTACGCACTACACACCCTTCGGGAGTGTGTGCGAAA
TTCTGCACCGCGAGGGTGCGAGAATTTGCTTTGCCAATCATAGCGAAAAGTCAGGCGGAAAATTGAT
ACTAACAAATCCAAAAAAGGATTTTGTTTCTATCTCCCACTTTTTTGTTAATAGGAAAAAAGTGGGAGAT

ATTAGGAAGCCTTAGAAACCGGTTTCCTAATATTCCCCCTTTTTTTCCTATTTCCCATTCAATCGTAAAG
TGTGGGGTCAAAATTAAGCAATGTGGGATTAGTGAAAATCGGCAATTTGTCTCTCATTTTGAACAGCTT
TTTCCAGCAAGATTTCGCATATTCTTCTGGGTGATATTTTCGCTCGATAGTTTCTTTGAGCGGGTCAAC
TATCACATAGTAGAGCCCTACCCTCTTCATCGCAGTAGGTCATTTTTCCCATCGAACTCAAGAAGCTTG
AGATCGTGAGTAGTCCCTCCTTGATGAGTTCTTGGAACTCATTAAGCTCTTTGTCGTTTGTTTTGATGA
TAAATCTGATCGCATAGACAGCTTTTCCACGCTTTTAATTTCTTCAAACTCGAACGTAAATGAAGATTAT
CTTTGAATTTCTTTCAAGCCTTCGTTCTATGGCACGTTTTTTTAAAACATAGATAAGTTATA

>pRS44_SP_P35F04_pCC

AGAAAAAAATCTCTCCACCGGGAAACACGCTATGACAATGATTACGCCAAGCCCTATTTAGGTGAGAC
TATAGAATACTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACACGATATGTATGCTTGA
TTTTGCCATTGCGCACGCTCCTTTTTGTGGTAAATAATTGAGGGGCTTGTCCCCCAATTATACCACAAA
AACCCACACACATCTCGGAGAGGTGTGTAGTGGGTACGCACTACACACCCTTCGGGAGTGTGTGCGA
AATTCTGCACCGCGAGGGTGCGAGAATTTGCTTTGCCAATCATAGCGAAAAGTCAGGCGGAAAATTG
ATACTAACAAATCCAAAAAAGGATTTTGTTTCTATCTCCCACTTTTTTGTTAATAGGAAAAAAGTGGGAG
ATATTAGGAAGCCTTAGAAACCGGTTTCCTAATATTCCCCCTTTTTTTCCTATTTCCCATTCAATCGTAA
AGTGTGGGGTCAAAATTAAGCAATGTGGGATTAGTGAAAATCGGCAATTTGTCTCTCATTTTGAACAG
CTTTTTCCAGCAGATTTCGC

>pRS44_SP_P31B08_T7

TACGTATTAGGGCGAATTCGAGCTCGGTACCCGGGGATCCCACGATGCTACTACGGAAAGGCTGAAA
GCGTGAACCACTCCCATTTTTCCTTTTTCCCTGCCAGCCTGTCAACAACAAAACACAAGGCGAGAGCT
CCCCGCAGTCAAGGTGGAGCTCCCGCAGGCGAAGCCGAGGAAGAACGACCTTGACGAGGAGAGCT
CCCGCCTGAGATTGAAAATTCAAGGCTGGCAGGACAGGTCAAAAACATCAACATATTCGCCCTATATT
CAAAAGAACCCTCTTCCCGAGTACCCAAACCTTTTTTTAACGAAAAATCGCCTCAGAGAGGCTGAGGA
TAGCTCATAAAGGGTATTTTGAAAACCTATCGGAGGAGAAAAGGGACAAAAAAAATTTTTTGCGCGCG
CCTTTGAAATTGTGGATAACTATTGAATTTCAATTACTTGATAACTTTTCCAAAGCCTTCAGAGAAAGAG
AACTCTAACACTGATACAAAACTCCAAAAACTTTTGAAAATCAAAACTTTTCGGGCAAACGGTATAACC
TACTATCGTAGGGGAGGGTATATATCAATGAAAGTTGACACTTTTATGTCAACTTTCAAACTGCCCCAA
AGCCTTGCAAATCAAAGAATTACTTAAAGAGATGGCCGTAGGTTTTATATAAATTCTGACGAGTAATTC
CCAACATCTCAGCAATTTCTGTTTTTGTATAAC

>pRS44_SP_P31B08_pCC

ACAGACAGTTTACACCTGAGTCGCTTGCAGGCTTACTAGCCCGAGCTATTTAGGTGAGACTATAGAAT
ACTCGACCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACCTGGGGCTCACATTATCAAGGTCA
ACTACAAACGCTCTTACCGTTTCTAAAGTCCTTGAATGGTAGTAATCCTTTTTGAAAGAGCAGTAGGGT
ATATAAGCATTTGAACGCCAAGCTATATCCCAAACTTCACAGAACTTTATTCTTCCTACCGTCCCATTC
TCTATGATGTTTCCCTTTTTGTCCTTTATAGTCCCGAAAACAACGATTAAATTGTCCTCATCTCCTTTAT
CGGGGAAGATTGTTTCAAGGTATTCCATAGGGGAGAGCTCAGGAAACCTTGACCTGAGAAGCTCGTT
TTTGGCTTCAAATTCCTTTTGTAGTTCGTTTTGAGGTTCGTATTGTCTTGCAAGGTTTTTATCCATCTCG
CTCCTCCGCTTTCGGTTTTTGCGGAGGAGCTACCACTTCTTGAATTCCACCTAAAGAGATATAATTTTT
TTATAGCTTGTCTTTAGGTGGAGTGGTAGCTCCTCGTTCTTCGGGTGGAGGAGGTGGTGGCTCCTCC
ATCTCTACTTTCCCAAACTTTTAGGATTTTCCAAAGGTAAGGCTTGGATACTTTTAGCCTGTGAGTATG
CCAAAGGTATTTGG

>pRS44_SP_P31A05_T7_PW1

TTTTGCTGCATACTTTGTGTATATTGGTTGATATCTACTAGTAAAGATTTGAGTTCTTTTTTTAGTGAAAT
TTGTTTTGATAATTTACGTTCTTTTGCAATTTGTGTAGCAATTTTTGCTTTTTGTGTTTGGGCTGATAGTT
TGCCTCCTGTATAAAGAGGAAGGTTTAGAGAAATACCTATTACTTTACTATCATAACTGTTAAGTGTATT
TAAATGTGTAAAAGATCCTACAGCATCAATTGAACCATAGTGTGAAGCATGTGCTGCTTTATGTAAAAG
TTTGTTTTTTTCAATAGTAAGTGTATCTATTTGAAGTGTTGGGTTTTCTTTTAAAATAGTTTGGGATAAGT
GTTTTATTCGTTTAGTTGGAATTTGTTTTAATATATTTTTTTGTAAAAATAAATCATTGGGTAAATCTTCA
TTCATATATAAAGAGAGTGTCAATTTGGCTTTTTGATAAGTTGCTTTTGCTATGGATAAATTATTTTTAG
CAGTATAAACTTCTGCTAAAAAACGACTCTCATCAGCTTTTGTTTTAAGACCTTGTTTGTATAGTGCTTG
GGCTTGTTGATAGTAAGCTTCTTTACGTTTGAGATCTTCTTTTTGAACTTTAATCGCTTCTTTTTGTACG
AGCATTAATTTGTAAAGTGATTTAACTTGATAAGCAAGCAATGCTTTAGCATCTTTTAAGGAAAGCAAA
GAAATTTCTTGATCTAATTTTGCTGCTTTTATTTTTGAATCAGTGAGAGAAAAATCCCAAAGTTTTTCTTT
GATTCCAATTCCTATGTTCCAGCTATCTTCATCAACAGTATGAAAACGACCATTAATAGGAAGAGCAAA
AGTTTGGATCGGATTATTAATTTCGCTTGTAATGGA

>pRS44_SP_P31A05_pCC_PW1

GTAGGGTGATAAAATATATTGTTTATTGTGCTCTTTGTGATGAAGCAGAAAATATATTTTTAACAATTTC
TTTGCGTATGGCTTTTGCTGATATTTCGATTATTGCATTGGCTAGTGATAGTGAACACGCAACAAAATT
AAAGATGGCTGGAGCAGACAAAACGATAGTTACCGAACAAATCACAGCAAATTTACTTTATGAACGTA
TTAGTAATCCTACTGCATCTGAACTTTTGCAAAAAATTCTTTATGAAAACAGTGATATAAAAATTGCTCA
AATTACTGTTAAAAAGCGTAGTGAATTATTGGGGCAAAGTATTTCTGCTACACAGAATATTCAAGATCA
TTATGGTGTGATTATTTTAGCGATTATTGATAGCCAAATGCAAACACTTTTTAGTTTTACACACAAAGGT
TTAGCCCATACAATACAAGAGGGTGATATTCTTATTGTCGTTGCTTTGCACAAAAATATTGAAGCATTT
AAAGAAGCTACAGGTGAACATTACAGTGTTGATTGGCATCACTGGAAATGGTTTAACGATTAAATAAAA
TGTAGGAATAATTATGCAAATAACTTCATGTTATCCGCTTTGCGAAGGGGAAATCCTTAATGGCTTTAA
ATCAGAGTATCCTCTTTTTGATGTGAGTATCCATGGCAGTCAAAGAGAGAGAATTAAACTCTCCAAGC

GTTTGGTTTGTGCGGTGAAACACTTACCACTTCGTTTACGTATTTGTTATGAGAAGGATACTTTAAAAT
CTATCGAAAAAGGTGTACAAAAAGATCCAACGTTAATTTTGAAAGATAAAAATATTGGCAGAGGGCTTG
GTTAGTGCTGAATCTTTGACTGAAAATA

>pRS44_SP_P31A05_pCC

CATGTATTGTTCCACTGGCAACTGCTATGTCCATGATTACGCCAAGCTATTTAGGTGAGACTATAGGA
ATACTGCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACGTGGCTTTTGAGGTGTTACATT
TGTTAAGGAGTGAACACTTAGGGGCAAAAATAGAAGAATTTTATATCGATTCTTGGATTGCTAATCGTT
TTGATACACTTGTTTCACTTGAATATAGGCGATTTCATTTGATTTTAATGGGAATAGTTAGGAAAGATGT
ATTTCTTTTCAAACCAGATTTAAATACCAAAATTCAAACAGGAGATATTTTGATCATGATAGGGGAAAAG
TCTCTTTTATTGGAATTTGAGCATTTTGTGCATTCAACAAAGGTATAAAAGGTGAAAGATAGAAATTTTT
TAATTTTTGGATGTGGTGAATATGCTATGCAAATTGCTTACAATTTACACAAAAGTGGACAGCGTGTAC
GTTTTTTTAGTTTTGATTCTGAAGAACTTGCAATGGTCCGTGAAAATGGTATTGAAGCCCTAGATTTTG
ATCTTGGTGATGATTGGGAAGCGTTAGGTATAAAAGATGAAAAAAATTATATTGTTTATTGTGCTCTTT
GTGATGAAGCAGAAAATATATTTTTAACAATTTCTTTGCGTATGGCTTTTGCTGATATTTCGATTATTGC
ATTGGCTAGTGATAGTGAACACGCAACAAAATTAAAGATGGCTGGAGCAGACAAAACGATAGTTACCG
AACAAATCACAGCAAATTTACTTTATGAACGTATTAGTAATCCTACTGCATCTGAACTTTTGCAAAAATT
CTTTATGAAAACAGTGATATAAAATTGCTCAATTACTGTTAAAAGCGTAGTGAATTATTGGGGCAAGTA
TTTCTGCTACACAGAATATTCAGATCATTATGGTGTTGATGATGATAGCGTTAACTGATATCGCATAGG
CAAATGCAAACTAC

>pRS44_SP_P31A02_T7_PW1

GAACTCTCCGGAGAACTGCTGGAATATTGAGGAACGGAAAGAAACTCCCGGTATAAGCTGAGCAGAG
CCTGAAAGGTTGGCCTTGCGGCCGGGAGAGGTGTTACGGCCTCTCTGCAAAATGTGTTGAAGGTAAG
GAAAAACCCGGGGAAGGGGCTGTTTGGGACTTGGAGGACTTCGAAAGGCTTATCTCGTTCAATAGCA
GGGCTTGAGTGGTTACGTTCTTTCACTCCTCATTATGGAGGAAAGATGAGTCTTTGTGAGGACGTTTT
CAGATTAGGATAAAGCCAAGAGATAACAACTTTTAAGGAAAAAGCTCCAAAAGTAGTATTGGGGAACA
AAATCATAAGCAGGAGGGAGCGATGTGCCATTCGGCTTTGGGAAAAAGAAGGACAAAAAAGAAGAAT
GGCTGAAAAAATTAGGAATTAGCGACAAAGATTTCCGCAAGCTATCCGCAACAGAAGCTATACTCAAA
GGATACGAGACACACAAGAAGATCGCAGCCCTCACAGGTCTTGATGCTGAAAAAATTGATAAAATCCG
AAGAGACGGGTCATATGTACCCACGCAAGATGGAATATTCTACAGGATGGCTGTGGATGCCAAAACC
CGCACAACGCTTTTTTCATATTCTTTACCGAGATGCTTATGCCCGTTGCAGCGGAACTATCAAAAAAAT
ATCACCCAATCGTGGGCAAGAGGGTTGTGAAGGCAGTAGAGTCGCTCATGGAGGGTGATCCAAACG
CAACGTTCGACAAAATAATAAATATGGATGATATAAAAATAGCCCGAGTTGTTGGCAAGCATATAGAAG
ACGTCAGAGGTAAGCCTCTATCGGAATGAAGAATGGCCACCA

>pRS44_SP_P31A02_T7

GCAACTTTTGGGCGAATCGAGCTCGGTACCCGGGGATCCCACGGGTGAAGATTCCGGAGGAAGTTG
GGAAGTACTTGAACGGAAGGCTGATGGAGGATATCAAGCTTGGGATACTCAGGGAGTATCAGGAGAG
AGCGGAAAAAGCACGAAGACTGCTTGAAATCTTTAAAAAGAGTCAACTCAGTGAGGAAGACGCCAGG
AAGCTTGAGGAGGAGCTAAAGACGGCACTTGCAAAGCGCCATGGGGTGCTCTGATGGAAGTGGTATT
GGACTACAACGTTGTCTTTTCGGCACTGCAGAAGAAGGGACTCGCCCACGAGCTCTTGGCCCTTATC
GTCTTCAGCGACTCCATTAACGCCATTGTCCCAGAGCACTTCTGGGTGGAATTTGAAAAGCACAAAAG
GAAACTCGTGAAGTTGAGCAGACTCGACGAGCGGGATTTTGAGGATATGGTTGGAATTCTCAGGAGT
GAAATAATCTCAGTACCAAAGAAGATGTACTGGAAAGAGCTGAAGACTGCTATGAAAGTATGCCGCGA
CGTAAAGGACGCTCCATACGTTGCCCTTGCTCTCCACAGGCACGTTCCCCTCGTCACGGGGGACAAG
GCTTTGAGAAAAGATGCAAGCGGGATCATAGAGGTTTACTCTCCGAGAGAACTGGTGGGAAATATTG
AGGAACGGAAAGAAACTCCCGGTATAAGCTGAGCAGAGCCTGAAAGGTTGGCCTTGCGGCCGGGAG
AGGTGTTACGGCCTCTCTGCAAATGTGTTGAAGGTAAGGAAAAACCCGGGGAAGGGGCTGTTTGGGA
CTTGGAGGACTTCGAAAGGCTTATCTCGTTCAATAGCAGGCTTGAGTGTTACGTTCTTTCACTCCTCAT
TATGGAGGAAAGAATGAGTCTTTGTGAGGACGTTTTCAGATTAGGATAAAGGCAAGAGATAACCACTT
TTGAGGGAAAAAGCTCCAAAGTAGTATTGGGGAACAAAATCATAGCAGGAGGGAGCGATGCGTCAAT
CGGCTTTGGAAAAGGAGGACAAAAAGAAGATGCCTGAAAAATTAAGAATTAGCGACAAAGATTCGCAG
CT

>pRS44_SP_P31A02_pCC_PW1

TATAATGAGGTGTGAAAGACGTAACCCTCAAGCCCTGCTATTGAACGAGATAAGCCTTTCGAAGTCCT
CCAAGTCCCAAACAGCCCCTTCCCCGGGTTTTTCCTTACCTTCAACACATTTTGCAGAGAGGCCGTAA
CACCTCTCCCGGCCGCAAGGCCAACCTTTCAGGCTCTGCTCAGCTTATACCGGGAGTTTCTTTCCGTT
CCTCAATATTTCCAGCAGTTCTCTCGGAGAGTAAACCTCTATGATCCCGCTTGCATCTTTTCTCAAAGC
CTTGTCCCCCGTGACGAGGGGAACGTGCCTGTGGAGAGCAAGGGCAACGTATGGAGCGTCCTTTAC
GTCGCGGCATACTTTCATAGCAGTCTTCAGCTCTTTCCAGTACATCTTCTTTGGTACTGAGATTATTTC
ACTCCTGAGAATTCCAACCATATCCTCAAAATCCCGCTCGTCGAGTCTGCTCAACTTCACGAGTTTCC
TTTTGTGCTTTTCAAATTCCACCCAGAAGTGCTCTGGGACAATGGCGTTAATGGAGTCGCTGAAGACG
ATAAGGGCCAAGAGCTCGTGGGCGAGTCCCTTCTTCTGCAGTGCCGAAAAGACAACGTTGTAGTCCA
ATACCACTTCCATCAGAGCACCCCATGGCGCTTTGCAAGTGCCGTCTTTAGCTCCTCCTCAAGCTTCC
TGGCGTCTTCCTCACTGAGTTGACTCTTTTTAAAGATTTCAAGCAGTCTTCGTGCTTTTTCCGCTCTCT
CCTGATACTCCCTGAGTATCCCAAGCTTGATATCCTCCATCAGCCTTCCGTTCAAGTACTTCCCAACTT
CCTCCGGAATCTTCACCCCTGATCACAACTTCAACCCTCTCCCATGCC

>pRS44_SP_P30F07_T7
TCCTTAGGGGGGCGAATTCACCCTCTGTCCCGGGGATATTATAGGAGATAGTAATCTTAGTTTTTGGG
GGGGCGGGTTATAATAATACCCCCACTAACCATAGGGGGACTAAAAGAAAAAAACTTATTTTATTTTTT
TAACGAATACTTTAGGGATGGGGGGGCCTGTTTAGCTTTTATTCCCGGTACCCTAAGAAAGTTTCCCG
AGGGGAAGACCCCCGTTTTCCTTTTTGAATTGGCTGAAATAAGGGCCGGGTAGGTGTTAAAAGTTTTT
TACTCGTTTTCTTGGCCCTTTTTCTTTTTTTTTTTTTGTAGGGAGCTTTGGTCGGATTATCTGGCCCCTT
TGGCAGCGGGGTATTTTCATTTTTTCTTTAAAAAAAAATGTAATACTTTGCGGCACACCGCACGGATCG
GGCTTCGCTTTTTTTGCCTGATCGTACAAATCCGTTTTTTAAGTAACGGCGAGAGGTGATGGCTGGCC
CTTCACGGGTTTATTATTCGTCTTCTACCCCCCAATATTGTAACACAAAAAGCAAAAGGGACCTTGATC
CCCCACCAATTGGCCTATTTAATGCAGTATCGGGATTGAGGTCCAAGGCTAGAAACCTCCTAAATTAA
TGAGGAACCCCGTCCTAATACTATTTTCCGAGTTGGTTTCCAATGGAACGAGGGTCCTGATATTGATT
GATACCACTAATCCATGTGGGAACCGTCCTCTAGGACTGGCTTGCCTTTTTTGCGTGCCAAAGGACCC
CAGTTTCTGCAAGAAAAAACAAGCCCGGCGGCGGATAGTGACAATAGTTGCTTGTGGCTGGCGAACC
TCACATACAATCGTTCTGTTACTCCCATCCAACACATCGACTATTCGTGTGATTGTGACCGCGC
>pRS44_SP_P30F07_pCC
AAGTGTTAAAATTTACACGGAAAAGCTTGACCTGATTACGCCAAGCTATTTAGGTGAGACTATAGAATA
CTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACGGGATCTTATGGGATCTACCGCAGC
CAACAGCAATCTACTACGATAGACCGCCGGCAGGCGTGCTCTGCTGGCAAATATCGGGCTATTTCGG
CACGGAGAAAGGGCAAGCCTACCTAAGGGACGCTATACGACGATTAGAGCTAACAATCAACGGCAAC
GACGCTAAGGAGATTGGAAAACCAAGGGGAGAAGCGAGAATAGAGAGGGAACGGGTCATAAAGCTA
AGGGAGCTGGTCAGGCATTGCAAATCATTACCGAGACACGATAAACCAGCACAAAGGGCGGACAGCT
ACGCAGATTATGCCTTTTGGGCTATCAAGCTGTGGGTAGAAGACGAAATAAGACGACAAGGCGAAGG
GCTACCGATCAGCTACGACCGTTTAGAACGCTGGGCTCTGGAAGAGTTCAGGCAAAAAGAGCGAAGC
ACGATCAGGGCGAAGGGCCGAAGCGTATGGAACTGGTATGAGAAGCAAAAACTGGAAACTACCCACA
GGGAGGCGTTACGAGATGACCAGACAAGAGCACGCAAAAAAAATAGCAAAAAGAAAGAGCACAAAGA
AATCGGGAAAAAGTCATTAACACAGTAACCGGACTTTATGCAGACCCAATACAGGAAGAAAAGCGGG
GCTTGGCACTTCGGGAAAGATTTCTGAGGCTACTGGGATAAGCTCTAAACAGTAGCCCAAAATCCTTA
AAGACTGGGAAAAGCCATAAAATGGGTTTTTGTATGAACCTTACTATGGTTAAGTGCGGGTATTACTAA
TAACCCGGAACACAGAAAAACTTAAGGCTTCCTC
>pRS44_SP_P30B05_T7_PW2
CTGGCATATTATTTATGAAGCGCAACAAGCAGAAGAGGTTGGTTATCGGGAAATTGTTCTATGTGGAA
TTCACCTCGGCTTGTATAATCAAGAGAAAAATAACCTTCAATCTGATAAAAATTTAGCTTGGTTAATTGC
AAAAATCATCCAACGAACAGAAAAATTAAGAGTGAGAATTAGCTCTATTTGAAGGCTTCCCCTTTATTA
ATGATGAATTGATTTATCGCCCCCCTGTCGAGGTTCTTAGCGCTCCTCGTCCTCGCCTCCTCTCGTT
>pRS44_SP_P30B05_T7_PW1
TTGTTGTATGGCTGTTGCAAACAAAAAGAAAGAGGGGACGAAGATTGGATTTATACTTTGGAACATTTT
TTGAAAAAATAATTTTATTGAAGAGAGTTTTGTTTTAAAAAATATTTATTTGTTTTTTTGATTGATGTTAAC
TTTTAAGATATATACTTTGGGTTGTAAGGTTAATCAATACGATTCCCAGAGTTTAAAAAATAAGCTACTT
GCCACCGGTTTTAAGCTTGCCAAGAAAAAACAACAAGCTGATTTGGTTATTATAAATACTTGTTCAGTG
ACTAAAAATGCTCACCGTAAAGATAAGCTAGCAGTAAAGAAAGCTAAAAAGGAAAATCCTCTAGCTAAA
GTAGTCGTAATGGGCTGTTGGTATAAAGTTTATCCGGAAATTAAAAAAGAGTTAAAAGCAGATCTATTT
TGGTCAGTAGGAAAATTAGATCAATTGGTAAAGAAAATACAACAAATTCTATTATTACAAAAATATAAGC
AAAAATACCAAGGCAAGACCTCGGTATTTTTAATGAGGCCTCCGGGACAGGGAGAGCGTTCGCGGTA
TTTCATAAAAGTGCAAGATGGTTGTCAACAGTTTTGTTCTTATTGTGTTATTCCTTATTCTCGAGGAAAG
TTGCAGAGTAGAGGTTGGGAAGAAATTATTTATGAAGCGCAACAAGCAGAAGAGGTTGGTTATCGGG
AAATTGTTCTATGTGGAATTCACCTCGGCTTGTATAATCAAGAGAAAAATAACCTTCAATCTGATAAAAA
TTTAGCTTGGTTAATTGCAAAAATCATCCAACGAACAGAAAATTAAGAGTGAGAATTAGCTCTATTGAA
GTTAATGACATTGGTAGGATTTAATTGATTTTATTGCA
>pRS44_SP_P30B05_T7
TCCCCCTTACGGCGATTCGAGCTCGGTACCCGGGGATCCCACGATGAAAGGGATAGGAATGGAAAAA
TAATTCGGAAAGGACAAATTGTTACCAAAAACTTTAAACCACCAGTGCCTATTTCGGTAAAAGCTATGA
GTGTAACTCATATTACTAATAAAATGGTGGAAAATGAAGAAAATTTTAAAAATAGTTCAACCGCCAAAG
AATTACAAGAAATTTTGAACAATAATATTTTGGTCGCTCATAATGCTAAATTTGATATAGCTATGTTAAG
GGCGGAAGGAGTTTCAGTCCCCAGAAGTATTTGTACATTGCGTTTAGCGCGTTACCTTGATAAGGAG
GGGGAAATTCCTGAATATTCTTTACAGTTTTTGCGTTATTATTATGATATTGAGGTTGAAGCGCAACCC
CATAGCGCTGAAGGAGATGTGGTTGTTTTGGAGGCAATTTTTAATCATCTTTATAAACAAGCTGAAAAA
ATATCTAAGTCACCTAACAGGAAGACAATATTACAGAGAATGATTGAGATTTCCTCTAAACCTTCATTA
ATCAGAAGGTTTGTGTTTGGAAAATATCAGGGTCGTCAGATAGCGGAAATTGCCATGGAAGATAGAAA
TTATTTGGAATGGCTGTTGCAAACAAAAAGAAAAGAAGGGGACGAAGATTGGATTTATACTTTGGAAC
ATTTTTTGAAAAAATAATTTTATTGAAGAGAGTTTTGTTTTAAAAAATATTTATTTGTTTTTTTGATTGATG
TTAACTTTTAAGATATATACTTTGGGTTGTAAGGTTAATCAATACGATTCCCAGAGTTTAAAAATAAGCT
ACTTGCCACCGGTTTTAAGCTTGCAAGAAAAACAACAAGCTGATTTGGTATTAAAAATACTTGTTCGTG
ACAAAAATGCTCCCCGTAAAGATAAGCTGCAGTAAGAAAAGCTAAAAGGAAATCCCTCA
>pRS44_SP_P11H04_T7_#2
TAAATTGGGGCGATTCGAGCTCGGTACCCGGGGATCCACTCAAGAGAGATAGAGGCGGTATATAAAA
AAAATGATGCGATTAGACACAAGATTGACATATTTTTTATTTGGTATTTTTGTTGCATCTATAATTTTACC

ATTTTTGGATGGCATCGTATCATTATCCCTTTCTATTGGATTGTTTATTTACCTTACCGTAGTCGTTTCA
CTCTATTTTAAAAGAAAAAAAGATAGAAAAAAAATGAAACTTACTATCAAAAAAACAAAAGAGCATAAAT
CATTAGAAAAAATTTTCAAACAAAAAGAGATAGAAAAATCCAATAAACATGATGCAATCAACAATCAAAT
TGCCTATATCCAAGAGTTGTGGGATTTGAGTAAAAATCAACAAAAAACTTTTAAAAATTTTGTAGATAAA
CGTGCCTATAGTGATCTTTATACTAAAATGACAAACTCCCTACTTCCCCAACTTATCAAAATGATTGAAA
TTTGTATCGATCAAAACAAAAGTGGATGTAAACGAGAAGTTACAAAACGTATCAATGAGCTTGTCAAAA
TCATGAAATCTGAAATTGCAAAAACAAAAAACAAAACCAAAGATCATTTTGAAACAACTAGTGAAGTTTA
CGATAAACTTATCAATGAAATAAAATAACTTAAAATTTTACTTATATCTTTAAATTTTTTGATGGCATCAT
ATGCTACATCTGTCATCTTTTTGTATCTGCTCCAGATGAAACAATATACGGTGGCATACAATAAACACC
ATTTCCTAGAGGTCTTAATAATACTTCATGAGAGAGTGCATATTTGGAAATTTCCAAACCAATACGTTCT
TTTTGGATGATATCCTTTTAATTCAACTGCTCCAAATCATCCCCTCCTGTG

>pRS44_SP_P30B05_pCC_PW1

ATAGTTGAAAGAGGACAAATTATTTCCGGAATCGGGGAAACTCATCTTTCCGTTATCATGGAAAAACTT
AAAAAACGTTATAACGTTAATATAGAACTTACAACTCCGAAGGTTCCTTATAGGGAAACAATCACAAAG
AAAACAAAAGTCGAAGGTAAATACAAAAAACAGACAGGAGGACACGGGCAGTACGGACATTGCTGGC
TGGAAGTTGAACCTCTTGAACGCGGAAAGGAATTTGAATTTGAAAGCAAAATCTTCGGTGGTGCAATT
CCAAAGCAGTATATACCGGCGGTGGAAAAAGGCGTGCGTGAGGCCATGGAGCAAGGAATACTTGCA
GGATACCCGGTAGTTGACATAAAAGTAACTGTTGTGGACGGTTCTTACCACCCGGTTGACTCTTCGGA
GCTTGCATTTAAGGTGGCAGGCTCGATGGCTCTTAAAAAAGCTCTTGAACAGGGTAATGCAATAATTC
TTGAGCCTATTGAAAAACTTGAAATTATTGTTGACGAGCAGTATATGGGAGATGTAATCGGTGATTTAA
ACAGTAAAAGAGGTAAAGTTCTCGGAATGAAGGCGGAAAACGGAAAACAAATTATAGAAGCACTTGTT
CCGCTGTCAGAAGTTCTTACTTATGCAAATGACCTGAATTCCATTACGCAGGGAACTGGAGAATTCAA
AATGTCCCATTCACATTATGAACCTGCTCCTCCAAAAGTTGCCGAGAAGGTAATTGCGGAATCGAAGA
GAGAGGAGAACGAATAGTCGGAGGTGAAAACTTTTGAAGAAAATATATTGAAGCGTTAAGATATTGCT
CGCGTCAAACGAGGACATATTAAAATGGTCCCACGGGGAAGTAAAAAAGGCTGAACGTATAGTTACA

>pRS44_SP_P30B05_pCC

TCGGTCCCTTTCCACGGAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGAGACTATAGAATA
CTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACTCCGATTCTCGCAGGTTCTGCAACC
GCTCTTATAGGAATCGACGAACTCCTCGATTTTATTGTTGACTATATGGCGTCCCCTGATGTTTTCAGT
CCTGTAAAAGGGAAAAATCCTAAGACAGGGGAAGAGGAAGAGAGAAAGGTTTCTCCGGATGAAAACT
TCTCTGCATTTGTATTTAAAACTTTAAGTGATCCGTATGTAGGTAAACTCTCTTTTGTTAAGGTTTATTC
GGGCAAGCTTTCTTCGAACACGCAGGTTTACAACAGCTCGCGAGATACGATGGAAAAAGTCTCTCAG
CTTCTATTTTTGAGAGGCTCAAATCAGGAGCCTACGGACAGCATAGTTGCCGGAGATATAGGTGTCGT
TACAAAGCTTGTTTCCACTACTACGAACGATACTCTTTCTGATAAGAGTGCGCCCATTGTTTATCCTCC
GATTGAGTTCCCCGAACCCATGCTTTCAAAGGCAATGAGGCCGAAATCCAAAGCGGATGAAGACAGG
ATGAGTATAGGCTTTGCCAAAATAATGGACGAAGATCCGACGGTGAAAGTTTCAAAGAACGTTGAAAC
AAAGGAACAAATTATTTCCGGAATCGTGGGAAACTCATCTTTCCGTTATCATGGAAAAAACTTAAAAAA
CGTTATAACGTTAATATAGAACTTACAACTCCGAAGGTTCCTTATAGGGAAACATCACAAAGAAACAAA
AGTCGAAGGTAATACAAAAACCGACCAGAAGACACCGGCAGTATGGACTTTCGCTGTGGGAAGTCAA
CCCTCTGAACGCGGAAAGGAATTTGATTGAAAGC

>pRS44_SP_P11H04_T7_PW2

GCTTACCACGTTATAACCTTGTGAATCCAAAAAAGATTTTCCTTCACCATAATCGCAATAAAAAGCACT
ATAGATATCATCAGTCATCATAACCACAGAAAGTGGCAAATAACCGCCACTAATACCTTTTGATAAACA
CAAAAAATCAGGTGTTACTTCTGCTTGTTCACAAGCGAACATAGTTCCCGTTCTACCAAATCCCACCG
CAATCTCATCTGCAATCAAGAAAACATCATACACATCACAAAGTATACGCAATTGTTGAATATAGCTCG
CATCATACATCGCCATAGAACCAGCACACTGAACCAGTGGCTCTATAACCACCGCAGAAATTTTTCCT
TTATGCTTTTTAAAGAGTTCTTCCATAGACGCTAATGCTTCATACTCTTCTACCAAAGCCGGTGATTTTG
CCTGTATCGTTTTGATCAAAATCTCATCATAAATCTCTTTATATAGTGACACATCACCTATAGCCAAAGC
ACCCATTGTTTCGCCATGATATGAATTTTCCAATGAAATAAAGATATCTCTTACTTCACCTTGATTTTTA
TAATATTGAAAACTCATTTTCAATGCAATTTCAATCGCACTAGAGCCATTATCAGCAAAAAAAAACTTTT
TGAAGTTTTGAGGGTGTTAGGTCAACCAATCTTTGTGCTAAAGTGATTTGCAGGTTTCGTGGCTAAAC
CCTGCAAATATTAACATGTTCTAGTGTATCTAATTGCTCTTTTATTTTTTCATTTATATATGGATTGGAAT
GACCCGAAAATATTGACCCCACCCATGGAACTGATAGCCATCAAAGATATTCATTTCCATCAAAGTCAT
ATAAATAAAGCTCCTTTTTTGCGGATTTTATTAGGGCATAAAAAGGAAGGACTTCTTCGGTATCTTCTC
AT

>pRS44_SP_P11H04_T7_PW1

TACCCAAAAAACTTTTAAAAATTTTTGTAGATAAACGTGCCTATAGTGATCTTTATACTAAAATGACAAA
CTCCCTACTTCCCCAACTTATCAAAATGATTGAAATTTGTATCGATCAAAACAAAAGTGGATGTAAACG
AGAAGTTACAAAACGTATCAATGAACTTGTCAAAATCATGAAATCTGAAATTGCAAAAACAAAAAACAA
AACCAAAGATCATTTTGAAACAACTAGTGAAGTTTACGATAAACTTATCAATGAAATAAAATAACTTAAA
ATTTTACTTATATCTTTAAATTTTTGATGGCATCATATGCTACATCTGTCATCTTTTGTATCTGCTCAGAT
GAAACAATATACGGTGGCATAAAATAAACAACATTTCCTAAAGGTCTTAACAATACTTCATGAGAGAGT
GCATATTTGTAAATTTCTAAACCAATACGTTCTTTTGGATGATATCCTTTTAATTCAACTGCTCCAATCA
TCCCTCTTTGCCTCACATCAGCTACATTTTGCAAATCACGAAACCTTTCAAGAAAACCATCCAAAAGAT
CGATTTTTACTTTATTGGTTTCAATAGTATGTTCTTTTTCAAAAAGATCCAATGTTGCATTTGCCGCACT
ACATGCCAATGCATTACCAGTATAACTATGTGAATCCAAAAAAAGATTTTTCCTTCACCATAATCGCAAT
AAAAAGCACTATAGATATCATCAGTCATCATACCACAGAAAAGTGGCCAATAACCGCCACTAATACCTT

TTGATAAACACAAAAAATCAGGTGTTACTTCTGCTTGTTCAACAAGCGACATAGTTTCCGGTCTACCAA
ATCCCACCGCAATCTCATCTTGCATCCAGAAAACCTTCTACACCTCACAAAT

>pRS44_SP_P11H04_T7_#4

GCCCGTGGTAGGCAACAGTCAGATATCTATTTAAGGCGGGATCCCCTCAGAAGAAGAGGCGGTATAT
AAAAAAAATGATGCGATTAAACACAAGATTGACATATTTTTTATTTGGTATTTTTGTTGCATCTATAATTT
TACCATTTTTGGATGGCATCGTATCATTATCCCTTTCTATTGGATTGTTTATTTACCTTACCGTAGTCGT
TTCACTCTATTTTAAAAGAAAAAAAGATAGAAAAAAAATGAAACTTACTATCAAAAAAAGAAAAAATCAT
AGATCATTAGAAAAAATTTTCAAACAAGAAGAGAAAAAAAAATCTAAAAAACATGATGCCCTCAACAAT
CAAATTGCCTATATCCAAGAGTTGTGGGATTTGAATAAAAATCAACAAAAAACTTTTAAAAATTTTGTAG
ATAATCGTGCCTATAGTGATCTTTATACTAAAATGACAAACTCCCTACTTCCCCAACTTATCCAAATGAT
TGAATTTTGAATCGATCAAAACCAAATTGGATGTAAACGAGAAATTACAAAACGTATCTTTGACCTTGT
CAAATTCATGAATCTGATATTGCCAAAACAAAAAACATACCTAAGATCTTTTTGAAACAACTAGTGAACT
TTACGATACAATTTATCAATGAAATAAAATAACTTAAAATTTAACTTATATCTTTTGTTTCTTGATGGCAT
CATATGCTACATCTGTCATCTTTTGCATCTGCTCAAAGGAAACAATATAC

>pRS44_SP_P11H04_T7

TTCCCCTTATTTGGGCGATTCGAGCTCGGCACCGGGGATCCCCCTCAAGGAGAGATAGAGGCGGTAT
ATAAAAAAAATGATGCGATTAGACACAAGATTGACATATTTTTTATTTGGTATTTTTGTTGCATCTATAAT
TTTACCATTTTTGGATGGCATCGTATCATTATCCCTTTCTATTGGATTGTTTATTTACCTTACCGTAGTC
GTTTCACTCTATTTTAAAAGAAAAAAAGATAGAAAAAAAATGAAACTTACTATCAAAAAAACAAAAGAGC
ATAAATCATTAGAAAAAATTTTCAAACAAAAAGAGATAAAAAAATCCAAAAAACATGATGCCCCAACAAT
CAAATTGCCTATATCCAAGAGTTGGGGGATTTAAGTAAAAATCAACAAAAAACTTTTAAAAATTTTGTAG
ATAAACGTGCCTATAGTGATCTTTATACTAAAATGACAAACTCCCTACTTCCCCAACTTATCAAAATGAT
TGAAATTTGTATCTATCAAAACAAAACTGGATGTAAACGAGAAGTTACTAAACGTATCGATGAACTTGT
CAAAATCATGAAATCTGAAATTGCAAAAACAAAAAACAAAACCAAAGATCATTTTGAAACAACTAGTGA
AGTTCACCGATAAACTTATCAATGAATTAAAATAACTTAAAATTTTACTTTTATCTTTAAATTTTTGATGG
AATCTTATGCTACATCTGTCATCTTTTGTATCTGCTCAGATGAAACAATATACCGGTGGAATAAAATAAA
CAACCTTTTCCTAAAGGTCTTACCAATACTTCATGACAGAGTGCATTTTGAAAATTGCTAAACCAATAC
GTTCTTTTGGATGGATATCCTCTTAATTCAACTGCTCCAATCACCCTACTGTGACCTCTAATCATCCAC
GATTTTGGCAATCACGT

>pRS44_SP_P11H04_pCC_#3

CCGATCTAATTGTTTTCCCACGGGAAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGAGACT
ATAGAATACTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACAACGGCTTGGGTCGTGT
ATTGCCAACGATAGTTGCTCCCCTGTGAGATAGTGAGCTGAAAGCTCAATAGGTCACGGGGGTGCAA
CGATAGACTAAACGGCATAACAACCAACTTGTTGGTTGTATTATTTTGGTGCGTAAGCACCTTATAGAG
GGTGTTTCTGAAAGAAACACCAACCTATTGAATGGA

>pRS44_SP_P11H04_pCC_#2

CCAGGTTTTTTTTCCCTCGGAAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGAGACTATAGA
ATACTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACAACGGCTTGGGTCGTGTATTGC
CAACGATAGTTGCTCCCCTGTGAGATAGTGAGCTGAAAGCTCAATAGGTCACGGGGGTGCAACGATA
GACTAAACGGCATAACAACCAACTTGTTGGTTGTATTATTTTGGTGCGTAAGCACCTTATAGAGGGTG
TTTCTGAAAGAAACACCAACCTATTGAATGGAATAT

>pRS44_SP_P11H04_pCC

GTTTTTAATTTTTTTTTCCCTGGAAACTCCTATGACCATGATTACGCCAAGCTATTTAGGTGAGACTATA
GAATACTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACAACGGCTTGGGTCGTGTATT
GCCAACGATAGTTGCTCCCCTGTGAGATAGTGAGCTGAAAGCTCAATAGGTCACGGGGGTGCAACGA
TAGACTAAACGGCATAACAACCAACTTGTTGGTTGTATTATTTTGGTGCGTAAGCACCTTATAGAGGGT
GTTTCTGAAAGAAACACCAACCTATTGAATGGAA

>pRS44_SP_P11C02_T7_#2

AAACCTTAAGGGCGAATTCAAGCTCGGTACCCGGGGATCCCACTGTTAGGTCTAAATGCGTCGGAAA
AACGCTTTTACCCCCGACTGATAAACAAACGGGTAAAAAAATTGAACTTTCTGTCTCTAAAAAAGGGTA
TTTCTACAACAAACTTGATTTAGATTGGAAGATTACAAATAGTCGTGCTAATGATATTTGGGAGTTATTG
CAGTCTAAAGGCATATGCCTTGAAGCCAGGTAATGCTTTTGGACGCGGATTAAAAATGTTCTTATGGA
GCTATAATATACTGACATTGCGTATTTGCTACCCTACTTGTTTAATCCGCCCCTGCTGTTGCTAGATTC
CAGAGCGATAGCACATTAAAACGCACAAATGATATTTTAAGCTATCCACAGTTTTTATTTATGAGGATG
TTAACTTTTCCCCCTATATGCAAAGAAAATGGCAGACACATAGCAACGCGTAGCGGTGCGGTGGCGTT
TGAACCTGCAACAAGCTTTTGAGATTGTTGCAGTGGTTCAGGGCGACACGCGACAGCCCGCAGGGCT
GACGGTGTGTCTGCCCCTTGATTATATATTATAAAACTTAGGCACTTTAAATTAAGAAGGGTTAAAGTT
TATATTGAAAAATAGTGTTGAAAATTCTTTAAAAAATTAGCTTCTCTCATGTGCGCGCACGCGCTCTAA
GTTTTAATACTTTAGTTTTCTCTGTTTTCGGGGTCATTTGCCCCCATAAAATCGGGCTTTCGTGAGGGT
TAACTGTGTACAAAATCCCCGTTTAACTGTGTACAAAATCCCCGATAACTGTGTACAAAATCCCCGTGA
AATGAGACATTGACAAATAGTGTACACTTATGCTATGTTGTACGAAACTAATGAGTGTCACAATGGAAA
ATCTTGTCGTAAAAAAAATCCTTTGATTGAGTCTCGTACAACTTTCCCCGCTTGAGAAAGATTTGTTAT
CTGTGTGGGCATGGTTGATAAGATGATCTGATTTTTGACTTTATCATGTAGTGTTAGGATTGGCGACTA
CAAAGAAATTAGAGGTTCT

>pRS44_SP_P11C02_T7
CACAAAATGGGAAGAAGTGGAGATCGGCTCCCTGGTTTTTCTGCAGTATGTGTCCAAGGACTTTAAGC
TATGAATATGTCACAACTGACACACAATTGGGGTATGAATCGAAGGTTCTGTTGGTTTCAAAGGTTTCA
ATGCGTCTTCTTACGTGACTAGTGATGAAACCGGCCCTGGAGTGCTGTTATTGAATTCGGGCTGTTGT
GAATGACATTTTGAAAGACACGGGATCACAAATCTTGTTTCTTTTTTTTTTTTTTTAAGTTTGTTTTGGGT
TCTTTAATTCATTGACATTTATTGAATGTTACAATTGAGGCATAATATACTCAGGGTGCTGCAAGCCCTT
TAGAAATCCGGCCGCCCCCCCTTAATTATGGCTAATTATTATGTATTATTAGATGAGGACGCCACCCC
ACCCGCCTTATGG
>pRS44_SP_P11C02_T7_PW1
CCGCGCAGCCGCAGGCTGACGGTGTGTCTGCCCCTTGATTATATATTATAAAACTTAGGCACTTTAAA
TTAAGAAGGGTTAAAGTTTATATTGAAAAATAGTGTTGAAAATTCTTTAAAAAATTAGCTTCTCTCATGT
GCGCGCACGCGCTCTAAGTTTTAATACTTTAGTTTTCTCTGTTTTCGGGGTCATTTTGCCCCCATAAAA
TCGGGCTTCGTGAGGGTTAACTGTGTACAAAATCCCCGTTAACTGTGTACAAAATCCCCGTTAACTGT
GTACAAAATCCCCGTGAAATGAGACATTGACAAATAGTGTACACTTATGCTATTGTTGTACGAAACTAA
ATGAGTGTACACAAATGGAAAATCTTGTCGTTAAAAAAAATCCTTTGATTGAGTCTCGTTACAAACTTTC
CCCGCTTGAGATAAAGATTTTGTTATCTGTTGTGGCAATGGTTGATAAAGATGACTCTGATTTTTGGAC
TTATCATGTACGTGTTAAGGATTTGGGCGACTACAAAGAGATTAAGAGGTCTTGTAAAAGCTTGATAGA
TCGTTCTTTTGAAATTAGAGATGGCACTAACTGGCTGATGTTTAGTTGGTTTAGTTATATCGGCTATAA
GCACAAAGAGGGGATTATAGAGTTTTCTTTTGATCCCCGCTTGAAGCCTTATTTGTTGCAGATTAAGGA
CAATTTTACTTCTTATACGCTTGAGTCTGTTTTACCGATGCGGTCGGTTTATAGTATTCGCATTTATGAG
CTATTGAAGCAATATGAGCCGATAGGTACGCGTGTTTTTAATCTCAACGATCTTCAGGATTTATTGCAA
GTCCCAAAGAGTTATAAAGTCTGGCGTGCTTTTAGTCGTCGTGTACTTGAACCTG
>pRS44_SP_P11C02_pCC_PW1
CTATAAAATGCGAATACTATAAACCGACCGCATCGGTAAAACAGACTCAAGCGTATAAGAAGTAAAATT
GTCCTTAATCTGCAACAAATAAGGCTTCAAGCGGGGATCAAAAGAAAACTCTATAATCCCCTCTTTGTG
CTTATAGCCGATATAACTAAACCAACTAAACATCAGCCAGTTAGTGCCATCTCTAATTTCAAAAGAACG
ATCTATCAAGCTTTTACAAGACCTCTTAATCTCTTTGTAGTCGCCCAAATCCTTAACACGTACATGATAA
GTCCAAAAATCAGAGTCATCTTTATCAACCATTGCCACAACAGATAACAAAATCTTTATCTCAAGCGGG
GAAAGTTTGTAACGAGACTCAATCAAAGGATTTTTTTTAACGACAAGATTTTCCATTTGTGTACACTCAT
TTAGTTTCGTACAACAATAGCATAAGTGTACACTATTTGTCAATGTCTCATTTCACGGGGATTTTGTACA
CAGTTAACGGGGATTTTGTACACAGTTAACGGGGATTTTGTACACAGTTAACCCTCACGAAGCCCGAT
TTTATGGGGGCAAAATGACCCCGAAAACAGAGAAAACTAAAGTATTAAAACTTAGAGCGCGTGCGCG
CACATGAGAGAAGCTAATTTTTTAAAGAATTTTCAACACTATTTTTCAATATAAACTTTAACCCTTCTTAA
TTTAAAGTGCCTAAGTTTTATAATATATAATCAAGGGGCAGACACACCGTCAGCCCTGCGGGCTGTCG
CGTGTCGCCCTGAACCACTGCAACAATCTCAAAAGCTTGTTGCAGGTTCAAACGCCACCGCACCGCT
ACGCGTTGCTATGTGTCTGCCATTTTTCTTTGCATAATAGGGGGAAAAGTTACC
>pRS44_SP_P11C02_pCC_#2
CCCCTCCTTCGTTTACCGGAAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGAGACTATAGA
ATACTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACAGGATAGCTTAAAATATCATTTG
TGCGTTTTAATGTGCTATCGCTCTGGAATCTAGCAACAGCAGGGGCGGATTAAACAAGTAGGGTAGC
AAATACGCAATGTCAGTATATTATAGCTCCATAAGAACATTTTTAATCCGCGTCCAAAAGCATTACCTG
GCTTCAAGGCATATGCCTTTAGACTGCAATAACTCCCAAATATCATTAGCACGACTATTTGTAATCTTC
CAATCTAAATCAAGTTTGTTGTAGAAATACCCTTTTTTAGAGACAGAAAGTTCAATTTTTTTACCCGTTT
GTTTATCAGTCGTGCGTAAAAGCGTTTTTCCGACGCATTTAGACCTAACTAATACTTTGTATCGGGTAA
GAGGTAGTTTTTGCAAATTAGGGGCAATTCTGAAACTTACAGCAGTATACTTACGAACAGTCTTAATCG
GTTCATAGTCAATTTTTATATCAGAAAAGGTGTTAATCTCTTTTTTTGCAGGTTCAAGTACACGACGACT
AAAAGCACGCCAGACTTTATAACTCTTTGGGACTTGCAATAAATCCTGAAGATCGTTGAGATTAAAAAC
ACGCGTACCTATCGGCTCATATTGCTTCAATAGCTCATAAATGCGAATACTATAAACCGACCGCATCG
GTAAAACAGACTCAAGCGTATAAGAAGTAAAATTGTCCTTAATCTGCAACAAATAAGGCTTCAAGCGG
GGATCAAAAGAAAACTCTATAACCCCCTCTTTGTGCTTATAGCCGATATAACTAAACCAACTAAACATC
AGCCAGTTAGTGCCATCTCTAATTTCAAAAGACGATCTATCAAGCTTTTACAGACCTCTTAATCTTCTTT
GTAGTCGCCAAATCCTTAACACGTACATGATAGTCCAAAAATCAGAGTCATCTTTATCACCATTGCCAC
AACAGGATAACAAAATCCTGATCT
>pRS44_SP_P11C02_pCC
AACACCCCTAACATTCATTTTCACGGTAACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGAGAC
TATAGAATACTCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCACAGGATAGCTTAAAATAT
CATTTGTGCGTTTTAATGTGCTATCGCTCTGGAATCTAGCAACAGCAGGGGCGGATTAAACAAGTAGG
GTAGCAAATACGCAATGTCAGTATATTATAGCTCCATAAGAACATTTTTAATCCGCGTCCAAAAGCATT
ACCTGGCTTCAAGGCATATGCCTTTAGACTGCAATAACTCCCAAATATCATTAGCACGACTATTTGTAA
TCTTCCAATCTAAATCAAGTTTGTTGTAGAAATACCCTTTTTTAGAGACAGAAAGTTCAATTTTTTTTAC
CCGTTTGTTTTATCAGTCGTGCGTAAAGCGTTTTTCCGACGCATTTAGACCTAACTAATACTTTGTATC
GGGTAAGAGGTAGTTTTTTGCAAATTAGGGGGCAATTCTGAAACTTACAGCAGTATACTTACGAACAG
TCTTAATCGGGTTCATAGTCAATTTTTTATATCAGGAAAAGGTGTTAATCTTCTTTTTTTTGCAGGTTCA
GGTACACGACGACTAACAGCACGCCCAGGACTTATATACTCTTTGGGACCTTGCAATCAAATCCTGGA
GAATCGGTGGAGATTAAAAACACGCCGTACCTTATCGGCTCATATTTGCTTCAATAGCCTCAATAATG
CATATACTATAGAGCGGACCCGCCATCCGGTTAATCCGGACTCAGGCGTATAAGAAGGTAAACTTGTC

CTTAATTCCGGCCACAAAGTAGGCCTTCAGCGGGGGAATTCAAAAGAAAGCTCTAATCCCTTCTTTGG
GCTATAGCGGGATTACTAACCCATCTAACACTCCGCAGGTAGGTCCACGTCTAGTTCAGAGCGCGAC
TTCATTACG

9.6 Location of [NiFe]-hydrogenase of *S. oneidensis* MR-1 analyzed by PSORTp

```
SeqID: gi|24373659|ref|NP_717702.1| periplasmic [Ni-Fe] hydrogenase
small subunit HyaA [Shewanella oneidensis MR-1]
  Analysis Report:
    CMSVM-       CytoplasmicMembrane    [No details]
    CytoSVM-     Unknown                [No details]
    ECSVM-       Unknown                [No details]
    ModHMM-      Unknown                [1 internal helix found]
    Motif-       Unknown                [No motifs found]
    OMPMotif-    Unknown                [No motifs found]
    OMSVM-       Unknown                [No details]
    PPSVM-       Unknown                [No details]
    Profile-     Unknown                [No matches to profiles found]
    SCL-BLAST-   CytoplasmicMembrane    [matched 1346498]
    SCL-BLASTe-  Unknown                [No matches against database]
    Signal-      Unknown                [No signal peptide detected]
  Localization Scores:
    Cytoplasmic            0.00
    CytoplasmicMembrane    10.00
    Periplasmic            0.00
    OuterMembrane          0.00
    Extracellular          0.00
  Final Prediction:
    CytoplasmicMembrane    10.00
-------------------------------------------------------------------------

SeqID: gi|24373658|ref|NP_717701.1| periplasmic [Ni-Fe] hydrogenase
large subunit HyaB [Shewanella oneidensis MR-1]
  Analysis Report:
    CMSVM-       Unknown                [No details]
    CytoSVM-     Unknown                [No details]
    ECSVM-       Unknown                [No details]
    ModHMM-      Unknown                [No internal helices found]
    Motif-       Unknown                [No motifs found]
    OMPMotif-    Unknown                [No motifs found]
    OMSVM-       Unknown                [No details]
    PPSVM-       Unknown                [No details]
    Profile-     Unknown                [No matches to profiles found]
    SCL-BLAST-   CytoplasmicMembrane    [matched 1346496]
    SCL-BLASTe-  Unknown                [No matches against database]
    Signal-      Unknown                [No signal peptide detected]
  Localization Scores:
    Cytoplasmic            0.15
    CytoplasmicMembrane    9.82
    Periplasmic            0.01
    OuterMembrane          0.01
    Extracellular          0.01
  Final Prediction:
    CytoplasmicMembrane    9.82
-------------------------------------------------------------------------
```

```
SeqID: gi|24373657|ref|NP_717700.1| periplasmic [Ni-Fe] hydrogenase
cytochrome b subunit HyaC [Shewanella oneidensis MR-1]
  Analysis Report:
    CMSVM-       CytoplasmicMembrane    [No details]
    CytoSVM-     Unknown                [No details]
    ECSVM-       Unknown                [No details]
    ModHMM-      CytoplasmicMembrane    [5 internal helices found]
    Motif-       Unknown                [No motifs found]
    OMPMotif-    Unknown                [No motifs found]
    OMSVM-       Unknown                [No details]
    PPSVM-       Unknown                [No details]
    Profile-     Unknown                [No matches to profiles found]
    SCL-BLAST-   CytoplasmicMembrane    [matched 399323]
    SCL-BLASTe-  Unknown                [No matches against database]
    Signal-           Unknown           [No signal peptide detected]
  Localization Scores:
    Cytoplasmic            0.00
    CytoplasmicMembrane    10.00
    Periplasmic            0.00
    OuterMembrane          0.00
    Extracellular          0.00
  Final Prediction:
    CytoplasmicMembrane    10.00
--------------------------------------------------------------------------
```