Genomic and Functional Analysis of Organohalide-Respiring *Firmicutes*

Thomas Kruse

Thesis committee

Promotors:
Prof. Dr Hauke Smidt
Personal chair at the Laboratory of Microbiology Wageningen University
Prof. Dr Willem M. de Vos
Professor of Microbiology Wageningen University

Other members:
Prof. Dr T. Abee - Wageningen University
Prof. Dr D. Springael - KU Leuven, Belgium
Prof. Dr D. Leys - University of Manchester, United Kingdom
Prof. Dr J. R. van der Meer - University of Lausanne, Switzerland

This research was conducted under the auspices of the Graduate School for Socio-Economic and Natural Sciences of the Environment (SENSE)

Genomic and Functional Analysis of Organohalide-Respiring *Firmicutes*

Thomas Kruse

Thesis

Submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis committee appointed by the the Academic Board
to be defended in public
on Monday 30th of November 2015
at 13:30 p.m. in the Aula

Thomas Kruse

Genomic and Functional Analysis of Organohalide-Respiring *Firmicutes*
261 pages.

Ph.D. thesis, Wageningen University, Wageningen, NL (2015)
With references, with summary in English

ISBN 978-94-6257-578-3

Table of Contents

Chapter 1	General introduction and thesis outline	1-11
Chapter 2	Complete genome sequence of *Dehalobacter restrictus* PER-K23	13-28
Chapter 3	The restricted metabolism of the obligate organohalide respiring bacterium *Dehalobacter restrictus*: lessons from tiered functional genomics	29-56
Chapter 4	Functional genomics of corrinoid starvation in the organohalide-respiring bacterium *Dehalobacter restrictus* strain PER-K23	57-76
Chapter 5	*vanI*: a Novel D-Ala-D-Lac Vancomycin Resistance Gene Cluster found in *Desulfitobacterium hafniense*	77-94
Chapter 6	Genomic, proteomic and biochemical analysis of the organohalide respiratory pathway in *Desulfitobacterium dehalogenans*	95-118
Chapter 7	The *Desulfitobacterium* genome	119-174
Chapter 8	Comparative genomics and transcriptomics of organohalide respiring bacteria and regulation of *rdh* gene transcription	175-199
Chapter 9	General discussion	201-210
	References	211-240
Appendices		
I	Summary	241-242
II	Acknowledgements	243-248
III	About the author	249
IV	List of publications	251-252
V	Overview of completed training activities	253-254

Chapter 1

General introduction and thesis outline

Organohalides are produced and released into the environment by both biotic and abiotic processes. Previously, the production of organohalides by biological processes was thought to be negligible, however, in recent years numerous organisms from soil, sediments and aquatic environments have been shown to synthesize a large range of organohalides [1]. A diverse range of abiotic sources of organohalides has been identified, such as volcanoes, spontaneous geochemical processes, and meteorites [1, 2]. Organohalides have been and are still used for a wide range of applications such as wood preservation, degreasing of metallic surfaces, as pesticides or lubricants, and many of these compounds have been proven as, or are suspected to be, toxic for humans and other animals [3, 4]. Many highly chlorinated compounds, such as tetrachloroethene (PCE), have proven to be recalcitrant to aerobic degradation, but are often degraded under anaerobic conditions [3, 5]. This degradation can be either co-metabolical, or via metabolic processes by bacteria that are able to couple reductive dechlorination to energy conservation in a respiratory manner, previously referred to as (de)halorespiration and more recently termed organohalide respiration (OHR) [3, 5, 6]. The use of organohalide respiring-bacteria (OHRB) as a potential cost efficient non-disruptive tool for cleaning up aquifers and other sites contaminated with organohalides, has been intensively studied. For recent reviews see e.g. [7, 8].

The first reported OHRB was isolated in 1984 and named *Desulfomonile tiedjei* DCB1^{T} in 1990. This anaerobic bacterium was shown to utilize 3-chlorobenzoate as terminal electron acceptor with pyruvate as electron donor [9, 10]. Since then a large number of OHRB have been described (Figure 1) [5, 8]. Members of the genera *Dehalococcoides*¸and *Dehalobacter* and *Dehalogenimonas lykanthroporepellens* BL-DC-9^{T} are obligate OHRB using hydrogen as sole electron donor [11-13]. *Dehalococcoides*, *Desulfitobacterium* and to a lesser extent *Dehalobacter* are the three most thoroughly studied genera of OHRB. Especially members of the genus *Dehalococcoides* have been intensively studied due to their ability to dehalogenate a very large range of organohalides and are currently the only known organisms capable of complete dehalogenation of PCE to ethene [14, 15]. In addition, *Desulfitobacterium* spp., which mostly include facultative OHRB, have received considerable attention, as these are fast growing and easy to maintain in pure culture in comparison to *Dehalococcoides mccartyi* and *Dehalobacter* spp. [11, 12, 16]

Due to their ability to degrade a very large range of organohalides [14], it is not surprising, that the first two complete genomes of OHRB to be sequenced in 2005 were those of *Dehalococcoides mccartyi* strains CBDB1 and 195^{T} (previously *D. ethenogenes* 195^{T}) [17, 18]. These were followed in 2006 by the complete genome of the facultative OHRB *Desulfitobacterium hafniense* Y51 [19]. Currently (as of March 2015) 23 (partial and complete) genomes of OHRB have been formally published, and many more are listed in the Genomes Online database (GOLD; https://gold.jgi-psf.org/) as planned or on-going projects (Figure 1) [20]. The availability of complete genome sequences has boosted our understanding of the physiology and evolution of OHRB. The genome size gives a crude indication of an organism's metabolic capacity. Organisms with a diverse metabolism therefore generally have larger genomes than organisms inhabiting more restricted ecological niches. In line with this assumption, the obligate OHRB *Dehalococcoides mccartyi* and *Dehalogenimonas lykanthroporepellens* have very small genomes with an average size of 1.4 Mbp, which is among the smallest for free living bacteria [11, 17, 21], whereas *Dehalobacter* spp. have an intermediate genome size of 3 Mbp, which is rather large for an organism with such a restricted metabolism [22-24]. In contrast, facultative OHRB in general have larger genomes ranging from 3.2 to 6.5 Mbp, reflecting their more versatile metabolism [19, 25-30]. The availability of an increasing number of genomes from *Dehalococcoides mccartyi* has allowed comparisons between strains and thereby gaining an improved understanding of their evolution and ecology. Currently the genomes of twelve *Dehalococoides mccartyi* strains have been published, and their comparison revealed highly conserved genomes. The inter-strain variability is mainly found in two genomic regions located at either side of the origin of replication, often referred to as 'high plasticity regions' [17, 31]. Most of the *rdh* genes that code for the key enzymes of OHR, the reductive dehalogenases, as well as accessory functions, are located in these regions, leading to the speculation that the 'high plasticity regions' serve as evolutionary hotspots with frequent horizontal gene transfers, allowing rapid evolution of *Dehalococcoides mccartyi* strains while at the same time preserving core metabolic functions [31]. One noteworthy finding is that the genomes of the obligate OHRB *Dehalococcoides mccartyi* and *Dehalogenimonas lykanthroporepellens* encode incomplete *de novo* corrinoid synthesis pathways, and that the organisms thus are dependent on salvaging these cofactors from the environment [21, 32, 33].

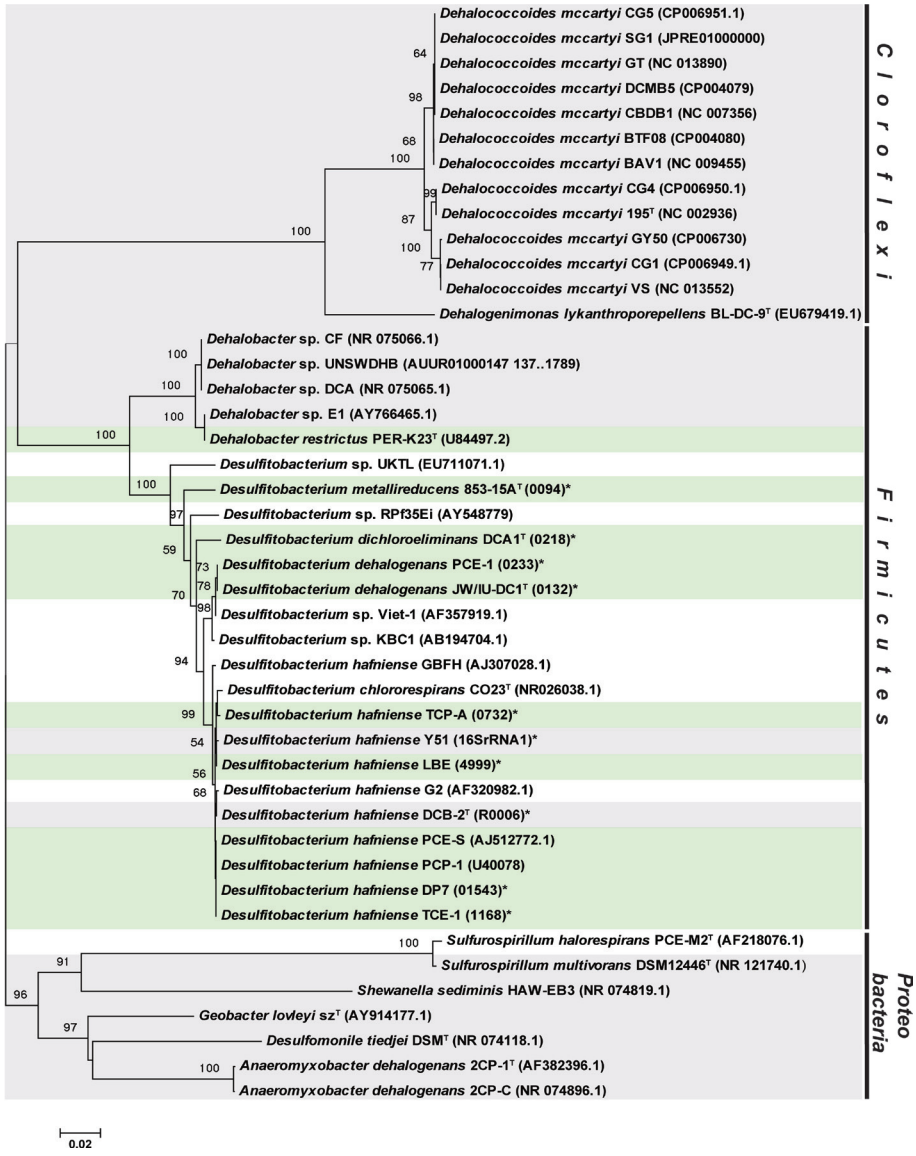


Figure 1. Phylogenetic tree of OHRB. The tree is based on 16S rRNA sequences that were either retrieved from Genbank (NCBI, accession numbers are given in parentheses), or if marked with an asterisk from the JGI–IMG server [34] with locus tags without the prefix given in parentheses. Background colour indicates strains from which a full genome sequence is available. Grey background indicates sequencing by other groups, whereas light green indicates strains sequenced as part of this thesis. Phylogenetic analysis was done using the MEGA5 software package [35]. Sequences were aligned using the MUSCLE algorithm before a neighbour joining tree was constructed and validated with 1,000 bootstraps [36, 37]. The reference bar indicates 2% sequence divergence. Phyla are shown to the right in the figure.

The genomes of several OHRB from other genera have also been sequenced (Figure 1). However, only one or a few genomes have been elucidated for each genus and thus extensive comparative genomics has not yet been possible. These include especially the genus *Desulfitobacterium* and to a lesser extent *Dehalobacter*. Both genera belong to the same phylum, the *Firmicutes*, and are Gram-positive, rod shaped, motile bacteria. These are especially interesting as they represent contrasting lifestyles in spite of their close phylogenetic relatedness.

Desulfitobacterium spp. are facultative OHRB, capable of coupling the oxidation of a range of organic and inorganic electron donors to the reductive dehalogenation of one or more organohalides [16]. In contrast, most isolated *Dehalobacter* spp. are obligate OHRB, which only grow by coupling the oxidation of H_2 to the reduction of a narrow range of organohalides [38-40]. However, it has been shown in enrichment cultures that at least some *Dehalobacter* strains are able to ferment dichloromethane, suggesting that this genus has a more extended metabolic repertoire than previously assumed [41, 42].

The first *Desulfitobacterium* strain was isolated in 1992 and named strain DCB-2 due to its similarity to *Desulfomonile tiedjei* DCB-1^T [43]. The publication of *Desulfitobacterium dehalogenans* JW/IU-DC1^T introduced the *Desulfitobacterium* genus [44]. In 1996 strain DCB-2 was classified as the type strain of *Desulfitobacterium hafniense* [45]. In the following years many more strains representing a number of different *Desulfitobacterium* spp. have been isolated. *Desulfitobacterium hafniense* strains PCP-1, TCE-1, DP7, TCP-A and G2 were originally published as belonging to a distinct species, *D. frappieri*, however, it was later shown that both *D. hafniense* and *D. frappieri* belong to the same species, and they have thus been merged to a single species, *D. hafniense* [16, 46]. The genus *Desulfitobacterium* contains at least six different species, five of which are currently listed in the "List of Prokaryotic names with Standing in Nomenclature" (LPSN), namely *D. hafniense*, *D. dehalogenans*, *D. metallireducens*, *D. aromaticivorans* and *D. chlororespirans* [47]. *D. dichloroeliminans* has not yet been included in LPSN although the species name is generally recognised [16, 48].

All but two of the characterized *Desulfitobacterium* spp. isolates are described as facultative OHRB [16]. The two exceptions are *Desulfitobacterium hafniense* DP7 and *Desulfitobacterium aromaticivorans* UKTL^T. The former was isolated from a human fecal sample, and has been shown to be incapable of degrading any chlorinated substrates [49]. The latter was isolated from a former coal gasification site, using toluene and ferric citrate as electron donor and acceptor [50].

The ability for OHR was not tested, however, it should be noted that 16S rRNA gene sequences with high similarity to that of *Desulfitobacterium aromaticivorans* UKTL^T were reported to be dominating in sediments where OHR of PCE did take place [51].

The first representative of the genus *Dehalobacter* was published in 1998 after its isolation from Rhine river sediments near Wageningen, and named *Dehalobacter restrictus* PER-K23^T. This name was chosen due to its extremely restricted metabolism, as it only has been reported to grow by coupling the oxidation of H_2 to the reduction of PCE or trichloroethene (TCE) [12]. Currently *Dehalobacter* sp. TEA, TCA1 and TCP1 are the only other *Dehalobacter* strains that were isolated and characterised [38-40]. Like *Dehalobacter restrictus* PER-K23^T, they are all obligate OHRB that only grow by coupling the oxidation of H_2 to the reduction of a narrow range of organohalides [38-40].

No genomes from a pure culture of *Dehalobacter* sp. had been published prior to the start of the work presented in this thesis. In 2013 Deshpande and co-workers published the draft genome of *Dehalobacter* sp. UNSWDHB, reported to grow by respiration using chloroform as terminal electron acceptor [52]. This strain has, however, not been further characterised. Draft genomes of three strains, *Dehalobacter* sp. E1, CF and 11DCA, have been obtained from metagenomes of mixed cultures [22, 23]. The genomes were reported to consist of 2.6-3.2 Mbp, encoding 10-17 predicted reductive dehalogenases. *Dehalobacter* sp. E1 was originally reported as using β-HCH as electron acceptor [53]. Analysis of the draft genome of *Dehalobacter* sp. E1 revealed an incomplete *de novo* corrinoid synthesis pathway and a gene with high amino acid similarity to the PCE/TCE reductive dehalogenase PceA. Preliminary data from cultures exposed to TCE showed release of chloride [54].

The genomes of *Desulfitobacterium hafniense* strains Y51 and DCB-2^T have previously been published [19, 26], consisting of 5.7 and 5.3 Mbp, and encoding one and seven reductive dehalogenases, respectively. Their relative large genome size is well in line with their versatile metabolism. Remarkably, it was reported that *D. hafniense* Y51 and DCB-2^T encode the largest numbers of molybdopterin oxidoreductase encoding genes, 53 and 57, respectively, of any sequenced bacterial genome. The presence of a complete *de novo* corrinoid synthesis pathway was also reported [19, 26]. Laboratory experiments have confirmed that *Desulfitobacterium hafniense* TCE1 and Y51 both are capable of *de novo* corrinoid synthesis [55, 56].

The reductive removal of the halide from organohalides is catalysed by reductive dehalogenases, termed RdhA, or RDase when characterised [57]. Reductive dehalogenases contain a twin arginine translocation (TAT) signal sequence, two Fe/S clusters and require the incorporation of a corrinoid co-factor for activity, although the type of the incorporated corrinoid may vary [58-63].

The *rdhA* genes are with very few exceptions found next to a small gene encoding a hydrophobic protein believed to act as a membrane anchor for RdhA (Figure 2) [21, 58, 64, 65]. This minimal *rdhAB* gene cluster is frequently accompanied by a variable set of accessory genes, for most of which the exact function is still not known. Some of these have been shown to encode proteins regulating expression of *rdhAB* [28, 66, 67] or chaperones probably assisting the folding of RdhA [68-70]. Three different types of transcriptional regulators have been proposed to regulate the expression of reductive dehalogenases (Figure 2). The type of regulators employed seems to follow the phylogeny of the OHRB. In *Firmicutes* genes encoding cAMP receptor protein/fumarate and nitrate reduction (CRP/FNR) regulators, *rdhK*, are often found in close proximity to reductive dehalogenase encoding genes. One of these regulatory proteins, CprK, has been characterised in detail and shown to induce expression of the chlorophenol reductive dehalogenase encoding *cprA* in the presence of its substrate Cl-OHPA in *D. hafniense* DCB-2 [71]. Genes encoding multiple antibiotic resistance regulator (MarR)-type regulators often are found in close proximity of reductive dehalogenase encoding genes in *Dehalococcoides mccartyi*. One of these regulators was recently shown to regulate negatively the transcription of reductive dehalogenase genes in *Dehalococcoides mccartyi* CBDB1 [72]. Two component transcriptional regulators have also been observed in both *Sulfurospirillum multivorans* and *Dehalococcoides mccartyi*, although not functionally characterised [29, 72]. Another potential regulatory mechanism has been described for the PCE reductase encoding gene *pceA*, which is constitutively expressed in *Desulfitobacterium hafniense* Y51 and TCE-1. The gene product PceA was found to be only translocated across the cell membrane when its substrate PCE was present in the medium [56]. This finding could suggest the presence of a post-translational regulation mechanism. It has also earlier been speculated that the NosrR/NirI like protein RdhC may play a role as a transcriptional regulator [73].

The localization of the reductive dehalogenase and the topology of the electron transport pathway from electron donor to electron acceptor have been a matter of debate. In early models, it was suggested that the electron-donating reaction takes place outside the cell, after which electrons would be transported across the cell membrane to an intracellular reductive dehalogenase, lead-

ing to formation of a proton gradient without vectorial proton translocation [74-76]. However, genetic data refuted this [3], and two studies have convincingly demonstrated that the reductive dehalogenases are located at the outside face of the cell membrane [56, 77].

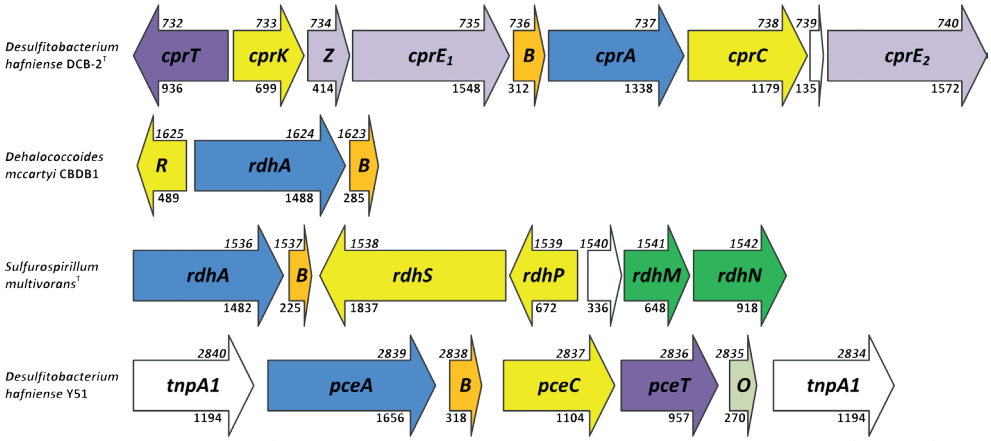


Figure 2. Organisation of reductive dehalogenase gene clusters respresenting the four different regulator systems. Top to bottom; Cl-OHPA reductase encoding *cprA*, regulated by the CRP/FNR type transcriptional activator CprK [71]; *rdhA*, negatively regulated by the MarR type regulator RdhR [72]; *rdhA* speculated to be regulated by the two component transcriptional regulator RdhSP [29]; the constitutively expressed PCE reductase gene *pceA* [19]. Colour coding of genes is as follows: ■: reductive dehalogenase, ■: membrane anchor, ■: Trigger factor like protein, ■: chaperone, ■: predicted transcriptional regulator, ■: membrane bound menaquinone dehydrogenase; ■: Chemotaxis protein. Empty arrows indicate hypothetical proteins or transposases. Nomenclature follows the guidelines given in **chapter 8** of this thesis. Numbers under arrows indicate size in bp, numbers above arrows show locus tags in the corresponding genomes without prefix. Missing prefixes are *Desulfitobacterium hafniense* DCB-2^{T}:Dhaf_; *Dehalococcoides mccartyi* CBDB1: cbdb_; *Sulfurospirillum multivorans*^{T}: SMUL_ *Desulfitobacterium hafniense* Y51: DSY.

Early attempts on heterologous expression of reductive dehalogenase encoding genes, and more specifically *pceA* from either *Desulfitobacterium hafniense* Y51 or *Sulfurospirillum multivorans* (formerly *Dehalospirillum multivorans*) using *E. coli* as expression host only led to the production of inactive enzyme [65, 78]. The lack of success was attributed to the fact that *E coli*, does not encode a cobalamin synthesis pathway, or did not salvage and incorporate exogenous supplied cobalamins into the active site of the PceA. Very recently the successful heterologous production of functional Rdases was reported in three studies, using two different approaches. Two studies used the cobalamide producing *Shimwellia blattae* or *Bacillus megaterium*, as expression hosts. This approach resulted in the production of the PceA and $RdhA_3$ reductive dehalogenases from *Desulfitobacterium hafniense* Y51 and *Desulfitobacterium hafniense* DCB-2^{T}, respectively, in an active form using *Shimwellia blattae* as expression host [69]. Employing a similar approach a

previously uncharacterised reductive dehalogenase, NprdhA, from *Nitratireductor pacificus* pht-3B, was produced in an active form using *Bacillus megaterium* as expression host [79]. The vinyl chloride reductase VcrA, from *Dehalococcoides mccartyi* VS, was produced in an active form using a fundamentally different approach. First VcrA was produced in an inactive form using *E. coli* as expression host, and subsequently iron, sulfur and cobalamin cofactors were incorporated chemically, leading to a reconstituted active form of VcrA [80]. These advances in heterologous production of RdhA has already led to significant progress in their analysis and resulted in solving the crystal structure of PceA from *Sulfurospirillum multivorans* and NpRdhA from *Nitratireductor pacificus* pht-3B [79, 81]. These findings represent a major breakthrough in the quest towards an understanding of the mechanism of reductive dehalogenation, including the role of the cobalamin co-factor.

The electron transport chain from electron donor to reductive dehalogenase has been proposed to involve both menaquinones and an intermediate electron carrier, as menaquinones are unable to directly reduce reductive dehalogenases [6, 82]. These electron-transferring elements have, however, not yet been identified. Recently two genes, *rdhMN*, encoding a membrane bound NapGH like menaquinone dehydrogenase, were identified in *Sulfurospirillum multivorans*^{T}. These are found in close proximity to two reductive dehalogenase encoding genes, and their expression was induced in the presence of PCE (Figure 2) [29]. However no other *rdhMN* like genes have been described to be associated with the genes for reductive dehalogenases, prior to the publication of this thesis.

Aim and outline

The genera *Dehalobacter* and *Desulfitobacterium* are both known to comprise OHRB. The obligate OHRB *Dehalobacter* spp. and the facultative OHRB *Desulfitobacterium* spp. both belong to the phylum *Firmicutes*. OHRB have attracted a great deal of interest due to their possible application for bioremediation. Their optimized application in environmental biotechnology, however, has been hampered by a lack of thorough understanding of their full metabolic capacity as well as their limitations with respect to e.g. co-factor synthesis and corresponding dependency on other microbial community members, and how this is linked to their genome organisation and evolution. In order to fill this knowledge gap, we have used a combination of complementary approaches, including genome sequencing, comparative and functional genomics, physiology and biochemistry.

In **chapter 2**, we describe the genome of *Dehalobacter restrictus* PER-K23^T, the first genome from a pure culture of a member of the *Dehalobacter* genus. The genome of *Dehalobacter restrictus* PER-K23^T consists of 2.9 Mbp, which is rather large for a bacterium with such a restricted metabolism. The most notable findings were the presence of the coding capacity for a *de novo* corrinoid synthesis pathway, which was rendered incomplete through the deletion of a small portion of the *cbiH* gene. The genome was predicted to contain 25 RdhA-encoding genes, the majority of which are full length and were found to be located in two genomic regions, resembling the high plasticity regions described for *Dehalococcoides mccartyi.* In **chapter 3**, we used a combination of proteomics and quantitative RT-PCR to follow the metabolic pathways operating in *Dehalobacter restrictus* PER-K23^T during exponential, late exponential and stationary growth phase. We found the PCE reductase-encoding gene to be highly expressed throughout all growth phases, as well as the expression of genes encoding several other reductive dehalogenases at a lower level. We identified one of the three putative uptake hydrogenases as the most likely candidate as the hydrogenase active under OHR with H_2 as electron donor and PCE as electron acceptor. Furthermore, we found indications that the *de novo* corrinoid synthesis pathway is not functional in *Dehalobacter restrictus* PER-K23^T. **Chapter 4,** describes the response of *Dehalobacter restrictus* PER-K23^T to different concentrations of corrinoids in the growth medium. We confirmed that the *de novo* corrinoid synthesis pathway is not functional, identified the genes encoding corrinoid transporters, and were able to show that their expression was regulated by the availability of corrinoids in the medium. In **Chapter 5,** we describe a novel vancomycin-resistance gene cluster, *vanI*, encoded on the chromosome of all *Desulfitobacterium hafniense* strains and *D. chlororespirans* DSM11544^T, and determined the minimum inhibitory concentration of both vancomycin and its analogue teicoplanin. Furthermore, we demonstrated the usefulness of *vanI* as a convenient selection marker by separation of *Desulfitobacterium hafniense* TCE-1 from a *Sedimentibacter* contaminant. In **chapter 6**, we report the genome of *Desulfitobacterium dehalogenans* JW/IU-DC1^T, the type strain of this genus, and important model organism for elucidation of the genetic, biochemical and physiological basis of OHR. We used a combination of genome analysis, enzyme activity assays and proteomics to identify elements of the electron transport chain from formate to Cl-OHPA. Unexpectedly, we identified a flavoprotein to be highly increased in abundance when Cl-OHPA was used as terminal electron acceptor, suggesting its involvement in OHR. **Chapter 7** describes the genomes of eight newly sequenced *Desulfitobacterium* spp., among which the first genome of *Desulfitobacterium dichloroeliminans* and *Desulfitobacterium metallireducens*. We used whole genome comparisons to assign strains to

species that had previously been described as *Desulfitobacterium* sp. Moreover, we identified genes associated with metabolic traits previously described in literature and compared the organisation of reductive dehalogenase encoding gene clusters across the genus. **Chapter 8** provides a review of the current state of the art regarding our understanding of the metabolism of OHRB obtained either directly or indirectly from genome sequencing. We discuss the (predicted) function of genes associated with reductive dehalogenases, propose a unifying nomenclature for these, and describe the mechanism of transcriptional regulation of OHR. Finally, **chapter 9** provides a brief summary and discussion of the results described in this thesis, also taking into account recent advances in the field, and indicates perspectives and directions for future research into the biology and biotechnological application of this exciting group of OHRB.

Chapter 2

Complete genome sequence of *Dehalobacter restrictus* PER-K23

Thomas Kruse, Julien Maillard, Lynne Goodwin, Tanja Woyke, Hazuki Teshima, David Bruce, Chris Detter, Roxanne Tapia, Cliff Han, Marcel Huntemann, Chia-Lin Wei, James Han, Amy Chen, Nikos Kyrpides, Ernest Szeto, Victor Markowitz, Natalia Ivanova, Ioanna Pagani, Amrita Pati, Sam Pitluck, Matt Nolan, Christof Holliger, and Hauke Smidt

Published as: Kruse T, Maillard J, Holliger C, Smidt H, et al. (2013) Complete genome sequence of *Dehalobacter restrictus* PER-K23. Standards in Genomic Sciences.

Abstract

Dehalobacter restrictus strain PER-K23 (DSM 9455) is the type strain of the species *Dehalobacter restrictus. D. restrictus strain* PER-K23 grows by organohalide respiration, coupling the oxidation of H_2 to the reductive dechlorination of tetra- or trichloroethene. Growth has not been observed with any other electron donor or acceptor, nor has fermentative growth been shown. Here we introduce the first full genome of a pure culture within the genus *Dehalobacter*. The 2,943,336 bp long genome contains 2,826 protein coding and 82 RNA genes, including 4 16S rRNA genes. Interestingly, the genome contains 25 predicted reductive dehalogenase genes, the majority of which appear to be full length. The reductive dehalogenase genes are mainly located in two clusters, suggesting a much larger potential for organohalide respiration than previously anticipated.

Abbreviations: OHR- organohalide respiration, OHRB- organohalide respiring bacteria, RDH- reductive dehalogenase homologue, PCE- tetrachloroethene

Introduction

Dehalobacter restrictus strain PER-K23 (DSM 9455), is the type strain of the species *Dehalobacter restrictus* [12]. Currently two pure cultures of *D. restrictus* have been described, namely *D. restrictus* strains PER-K23 and TEA [12, 40].

We publish here the first full genome of a pure culture within the genus *Dehalobacter* and a preliminary comparison with a previously obtained metagenome from a co-culture containing *Dehalobacter* sp. strain E1 and *Sedimentibacter* sp. [22].

Organohalide respiration (OHR) is considered as a key process in bioremediation of sites contaminated with organohalides such as tetrachloroethene (PCE) and trichloroethene (TCE), leading to a great interest in understanding the physiology and metabolism of organohalide respiring bacteria (OHRB). Most OHRBs are facultative organohalide respirers, capable of dehalogenating a limited number of halogenated compounds, as part of a versatile metabolism. This group consists of several genera, including *Desulfitobacterium*, *Geobacter* and *Sulfurospirillum*. Other isolates are obligate OHRB, among which isolates and enrichments of different *Dehalococcoides mccartyi* strains are the best studied. They have been shown to degrade a large variety of halogenated compounds solely using H_2 as the electron donor. Until recently, the genus *Dehalobacter* had been thought to encompass exclusively obligate OHRB, however, at least some members of this genus have been described as able to ferment dichloromethane [41, 42]. *D. restrictus* strain PER-K23 is an obligate OHRB, and like *Dehalococcoides mccartyi*, uses H_2 as a sole electron donor. These similarities in physiology and ecology are noteworthy since *Dehalobacter* spp. are phylogenetically closely related to the metabolically versatile *Desulfitobacterium* spp.

D. restrictus strain PER-K23 was isolated from a packed bed column containing sediment from the river Rhine collected near Wageningen, the Netherlands, and granular sludge from a sugar refinery. This column had been fed with PCE for a prolonged period, prior to isolation of *D. restrictus* strain PER-K23 [83].

D. restrictus strain PER-K23 was chosen for genome sequencing because it is the type strain of the *Dehalobacter restrictus* species. Studying the genome gives an improved insight into the physiology and evolution of the genus *Dehalobacter* and may ultimately lead to unlocking its full potential for bioremediation.

Classification and features

Dehalobacter restrictus is a member of the phylum *Firmicutes*, class *Clostridia*, order *Clostridiales*, and family *Peptococcaceae* [12], (Table 1). *D. restrictus* is closely related to the newly sequenced *Dehalobacter* sp. strain E1 [22], but grows in pure culture. Both *Dehalobacter* spp. and *Desulfitobacterium* spp. belong to the family *Peptococcaceae* (Figure 1). All members of this family are anaerobes, constituting a diverse group with respect to their metabolism and morphology [84]. *D. restrictus* strain PER-K23 is a rod-shaped bacterium with a single lateral flagellum and has not been reported to form spores. It stains Gram-negative, even though it phylogenetically belongs to the Gram-positive *Firmicutes*, and does not have an outer membrane, indicating that it should be considered a Gram-positive [12]. *D. restrictus* strain PER-K23 grows by coupling the oxidation of H_2 to the reduction of PCE or TCE, growth has not been observed with any other electron donor or acceptor, nor has fermentative growth been shown [12, 83]. *D. restrictus* strain PER-K23 requires iron as a trace element, the vitamins thiamine and cyanocobalamin, and the amino acids arginine, histidine and threonine for growth [12].

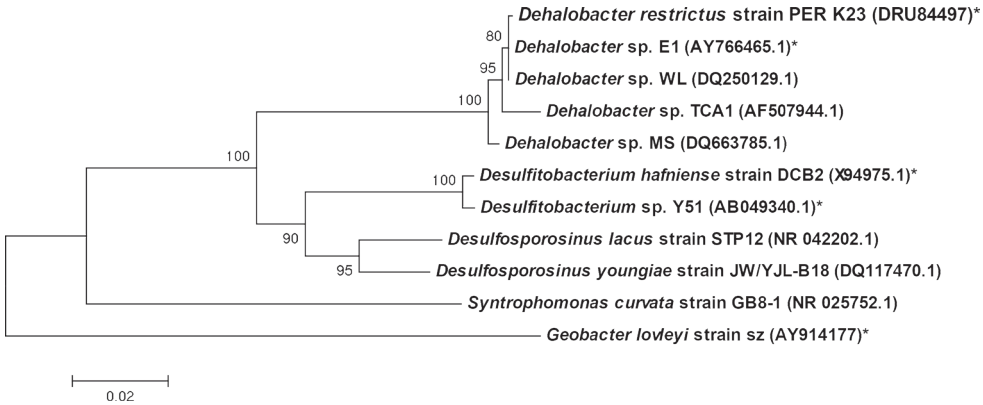


Figure 1. Phylogenetic tree highlighting the position of *Dehalobacter restrictus* relative to phylogenetically closely related organisms. 16S rRNA sequences were retrieved from Genbank (NCBI), and accession numbers are given in parentheses. Strains from which a full genome sequence are available are indicated with an asterisk. Phylogenetic analysis was done using the MEGA5 software package [35]. Sequences were aligned using the MUSCLE algorithm before a neighbor joining tree was constructed and validated with 1,000 bootstraps [36, 37]. The reference bar indicates 2% sequence divergence.

Table 1. Classification and general features of *D. restrictus* strain PER-K23 according to MIGS recommendations [85].

MIGS ID	Property	Term	Evidence code[a]
		Domain *Bacteria*	TAS [86]
		Phylum *Firmicutes*	TAS [87-89]
		Class *Clostridia*	TAS [90, 91]
	Current classification	Order *Clostridiales*	TAS [92, 93]
		Family *Peptococcaceae*	TAS[93, 94]
		Genus *Dehalobacter*	TAS [12, 95]
		Species *Dehalobacter restrictus*	TAS [12, 95]
		Type strain PER-K23	
	Gram stain	Negative	TAS [12]
	Cell shape	Straight rod	TAS [12]
	Motility	Motile	TAS [12]
	Sporulation	Not observed	TAS [12]
	Temperature range	10-37 °C	TAS [12]
	Optimum temperature	25-30 °C	TAS [12]
	Carbon source	Acetate, yeast extract	TAS [12]
	Energy source	H_2 as sole electron donor	TAS [12]
	Terminal electron receptor	PCE and TCE	TAS [12]
MIGS-6	Habitat	Anaerobic river sediment	TAS [12, 83]
MIGS-6.3	Salinity	Not tested	
MIGS-22	Oxygen	Strictly anaerobic	[12, 83]
MIGS-15	Biotic relationship	Free living	[12]
MIGS-14	Pathogenicity	None known	
MIGS-4	Geographic location	River Rhine, near Wageningen, The Netherlands	[12, 83]
MIGS-5	Sample collection time	1992	

[a] Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [96].

Genome project history

Table 2. presents the project information in compliance to MIGS version 2.0 [85].

MIGS ID	Property	Term
MIGS-31	Finishing quality	Closed genome
MIGS-28	Libraries used	Two genomic libraries, one paired-end 454 library and one Illumina library.
MIGS-29	Sequencing platforms	454 GS FLX Titanium and Illumina GAii
MIGS-31.2	Fold coverage	8.5 for 454 and 120 for Illumina
MIGS-30	Assemblers	Newbler version 2.3, VELVET, version 1.0.13 and phrap, version SPS – 4.24
MIGS-32	Gene calling method	Prodigal, GenePRIMP
	Genome Database release	December 28th, 2011
	Genbank ID	PRJNA66209
	Genbank Date of Release	
	GOLD ID	Gi05571
	Project relevance	Type strain, Bioremediation, Biotechnology

Growth conditions and DNA isolation

Dehalobacter restrictus strain PER-K23, DSM9455, was cultivated anaerobically as previously described [12]. DNA was extracted from bacterial pellets using the protocol recommended by the JGI. In brief, cell walls were digested with lysozyme before DNA was purified with hexadecyltrimethylammonium bromide, phenol and chloroform, and precipitated with isopropanol. Quality and quantity of the obtained DNA were checked by running aliquots on agarose gels using lambda phage DNA as mass standard and HindIII digested lambda phage DNA as a size marker.

Genome sequencing and assembly

The draft genome of *Dehalobacter restrictus* PER-K23 was generated at the DOE Joint genome Institute (JGI) using a combination of Illumina [97], and 454 technologies [98]. For this, genome we constructed and sequenced an Illumina GAii shotgun library which generated 77,929,756 reads totaling 5,922.7 Mb, and 1 paired end 454 library with an average insert size of 10 kb which generated 318,117 reads totaling 59.3 Mb of 454 data. All general aspects of library construction and sequencing performed at the JGI can be found at the JGI website [99]. The initial draft assembly contained 90 contigs in 1 scaffold. The 454 paired end data were assembled together with Newbler, version 2.3-PreRelease-6/30/2009. The Newbler consensus sequences were computationally shredded into 2 kb overlapping fake reads (shreds). Illumina sequencing data was assembled with VELVET, version 1.0.13 [100], and the consensus sequence were computationally shredded into 1.5 kb overlapping fake reads (shreds). We integrated the 454 Newbler consensus shreds, the Illumina VELVET consensus shreds and the read pairs in the 454 paired end library using parallel phrap, version SPS - 4.24 (High Performance Software, LLC). The software Consed [101-103] was used in the following finishing process. Illumina data was used to correct potential base errors and increase consensus quality using the software Polisher developed at JGI (Alla Lapidus, unpublished). Possible mis-assemblies were corrected using gapResolution (Cliff Han, unpublished), Dupfinisher [104], or sequencing cloned bridging PCR fragments with subcloning. Gaps between contigs were closed by editing in Consed, by PCR and by Bubble PCR (J-F Cheng, unpublished) primer walks. A total of 134 additional reactions were necessary to close gaps and to raise the quality of the finished sequence. The total size of the genome is 2,943,336 bp and the final assembly is based on 24.6 Mb of 454 draft data which provides an average 8.5× coverage of the genome and 348 Mb of Illumina draft data which provides an average 120× coverage of the genome.

Genome annotation

Genes of *D. restrictus* strain PER-K23 were identified using Prodigal [105] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [106]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) non-redundant database, UniProt, TIGRFam, Pfam, PRIAM, KEGG, COG, and InterPro databases. These data sources were combined to assert a product description for each predicted protein. Non-coding DNA and miscellaneous features were predicted using tRNAscan-SE [107], RNAMMer [108], Rfam [109], TMHMM [110], and signalP [111].

Genome properties

The genome consists of a single chromosome with a total size of 2,943,336 bp with 45% G+C content. A total of 2,908 genes were predicted, 2,826 of which are protein-coding genes. Genes with putative function corresponded to 76.7% (2,168), of all protein coding sequences with the remaining annotated as hypothetical proteins. In addition, 1,174 protein coding genes belong to 356 paralogous families in this genome. The properties and the statistics of the genome are summarized in Tables 3-5.

Table 3. Nucleotide content and gene count levels of the genome

Attribute	Value	% of total[a]
Size [112]	2,943,336	100.00
Coding region [112]	2,473,591	84.04
G+C content [112]	1,311,589	44.56
Total genes[b]	2,908	100.00
RNA genes	82	2.82
Protein-coding genes	2,826	97.18
Genes in paralog clusters	1,174	40.37
Genes assigned to COGs	2,127	73.14
Genes with signal peptides	756	26.00
Genes with transmembrane helices	755	25.96
Paralogous groups	356	40.37
Reductive dehalogenases[c]	25	0.86

[a] The total is based on either the size of the genome in base pairs or the total number of protein coding genes in the annotated genome.

[b] Also includes 143 pseudogenes.

[c] Including pseudogenes

Table 4. Number of genes associated with the general COG functional categories

Code	Value	%age	Description
J	149	6.4	Translation
A	--	--	RNA processing and modification
K	176	7.6	Transcription
L	180	7.8	Replication, recombination and repair
B	1	0.1	Chromatin structure and dynamics
D	44	1.9	Cell cycle control, cell division, chromosome partitioning
Y	--	--	Nuclear structure
V	39	1.7	Defense mechanisms
T	166	7.2	Signal transduction mechanisms
M	148	6.4	Cell wall/membrane biogenesis
N	71	3.1	Cell motility
Z	--	--	Cytoskeleton
W	--	--	Extracellular structures
U	73	3.1	Intracellular trafficking and secretion
O	84	3.6	Posttranslational modification, protein turnover, chaperones
C	169	7.3	Energy production and conversion
G	64	2.8	Carbohydrate transport and metabolism
E	166	7.2	Amino acid transport and metabolism
F	59	2.5	Nucleotide transport and metabolism
H	118	5.1	Coenzyme transport and metabolism
I	40	1.7	Lipid transport and metabolism
P	105	4.5	Inorganic ion transport and metabolism
Q	27	1.2	Secondary metabolites biosynthesis, transport and catabolism
R	241	10.4	General function prediction only
S	203	8.7	Function unknown
-	781	26.9[a]	Not in COGs

[a] Percentage of the total number of protein coding genes in the annotated genome.

Table 5. Reductive dehalogenase paralogs encoded in the genome of *D. restrictus* strain PER-K23

Locus tag[a]	Ortholog in *Dehalobacter* sp. strain E1[b]	Comment
Dehre_0785	No	
Dehre_0793	No	
Dehre_0806	No	
Dehre_0808	No	N-terminally truncated
Dehre_0815	No	N-terminally truncated
Dehre_0820	No	
Dehre_0826	No	
Dehre_0830	No	
Dehre_0832	(98.7; Dhb965)	
Dehre_0835	(98.7; Dhb968)	
Dehre_0990	(100; Dhb84)	
Dehre_1408	(99.6; Dhb1133)	
Dehre_2012	(96.6; Dhb1238)	C-terminally truncated[c]
Dehre_2022	No	
Dehre_2026	No	
Dehre_2031	No	
Dehre_2037	No	
Dehre_2039	No	
Dehre_2044	No	
Dehre_2052	No	
Dehre_2058	No	C-terminally truncated
Dehre_2064	No	
Dehre_2065	No	
Dehre_2398	(90.4; Dhb490)	PceA[d]
Dehre_2792	No[e]	Partial sequence

[a] RdhA paralogs are listed in order of their position in the genome. Light grey indicates RdhA paralogs belonging to rdh cluster A (Dehre_785-835) and dark grey rdh cluster B (Dehre_2012-2065).[b] Orthology defined as more than 90% pairwise identity at the amino acid level, as suggested in [57]. Identity percentage based on full length RDHs and locus tag of the corresponding genes in *Dehalobacter* sp. strain E1 are given in brackets [22]. Identity percentages were calculated using MatGat [113].[c] For the comparison, a manually curated version of Dehre_2012 was used, i.e. the entire gene without the annotated frame-shift mutation.[d] Dehre_2398 corresponds to the biochemically characterized PCE reductive dehalogenase (PceA) [114]. [e] The sequence is conserved between the two strains, but no gene is annotated at this position in *D.* sp. Strain E1.

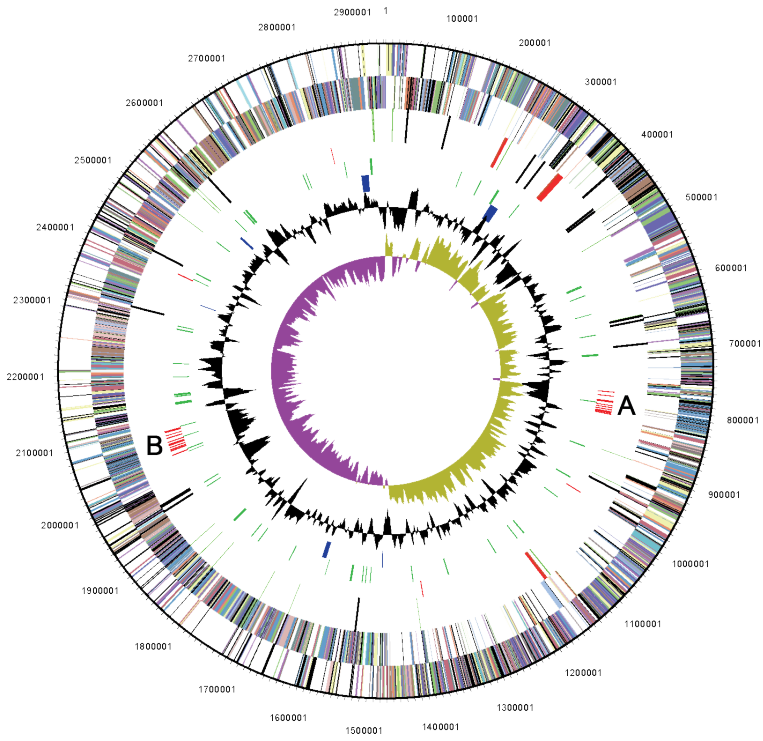


Figure 2. Circular map of the chromosome of *D. restrictus* strain PER-K23. Labeling from the outside circle towards the inside circle. Numbers outside the map indicate nucleotide positions; Circles 1 and 2: predicted coding sequences, including pseudogenes, on the forward and reverse strand, respectively (colored by COG categories); Circle 3: RNA genes (tRNAs green, rRNAs red, other RNAs black); Circle 4: Position of reductive dehalogenase genes, in red, both functional and truncated, A and B indicate two rdh clusters; Circle 5: Position of transposases including inactive derivatives, in green; Circle 6: Position of genes related to corrinoid synthesis and uptake, in blue; Circle 7: GC content (peaks out/inside the circle indicate above or below average GC content, respectively: Circle 8: GC skew, calculated as (G-C)/(G+C), purple or olive indicates values lower or higher than 1, respectively.

Insights from genome sequencing

Reductive dehalogenase paralogs

The genome of *D. restrictus* contains 25 loci predicted to code for proteins with sequence homology to reductive dehalogenases (RDHs). Among these 25 genes, one is a partial sequence and four are truncated due to possible frame-shift mutations (Table 5). This high number, is in contrast to those found to date for metabolically versatile organohalide respirers. These possess a limited number of RDHs typically in the range of 1 to7 [19, 26].

The number of RDHs in *D. restrictus* lies in the same range as seen in specialized organohalide respirers, such as *Dehalococcoides mccartyi* strains and *Dehalogenimonas lykanthroporepellens*, which have been predicted to possess between 10 and up to 36 RDHs [14, 21].

For *D. restrictus* however, this finding is intriguing since, PCE and TCE, currently, are the only electron acceptors known to be utilized by strain PER-K23 [12]. The identification of a total of 25 rdhA genes, suggest that *D. restrictus* possesses a much larger potential for OHR metabolism, than previously anticipated.

The majority of the *rdhA* genes are located in two clusters, one on each chromosome arm, with all but two RDHs being encoded on the leading strand. Cluster A is approximately 54 kb long, located on the right chromosome arm and contains 10 reductive dehalogenase genes including two truncated ones. Cluster B is approximately 61 kb long, located on the left chromosome arm and contains 11 reductive dehalogenase genes, of which two appear truncated (Table 5 & Figure 2).

The remaining three complete RDH genes and one partial RDH encoding gene are scattered throughout the genome (Table 5 & Figure 2). A similar pattern has previously been observed in the genomes of *Dehalococcoides mccartyi* strains, where the majority of the RDHs are located on each side of, and close to the origin of replication [31]. These regions were described as high plasticity regions, where frequent events of rearrangement and horizontal gene transfer are thought to occur. It was suggested that these regions enable fast adaptation to dehalogenation of new organohalides, while at the same time protecting key metabolic functions from being disrupted by horizontal gene transfer events [31].

We identified transcriptional regulators of the CRP/FNR type being encoded by genes in the vicinity of most of the RDH encoding genes, with PceA (encoded by Dehre_2398) as a notable exception [115]. A regulator of this type has been demonstrated to regulate the expression of the genes that code for chlorophenol reductive dehalogenase (cpr operon in *Desulfitobacterium dehalogenans* and *Desulfitobacterium hafniense* strain DCB-2^T [71]. The presence of transcriptional regulator genes close to almost all *rdhA* genes suggest that their transcription is regulated. This was confirmed by a recent study looking at transcription of *rdh* genes and the proteome of *Dehalobacter restrictus* strain PER-K23 growing in the presence of H_2 and PCE. In this study we found that PceA (encoded by Dehre_2398) was highly present at both RNA and proteomic level, whereas the remaining RDHs and the corresponding transcripts were either not detected at all or

at very low levels, suggesting that the RDH encoding genes are tightly regulated, and probably only expressed in the presence of their specific substrate [115].

Recently the draft genome of *Dehalobacter* sp. strain E1 was published [22]. This genome contains nine potentially functional *rdhA* genes, and one pseudogene. Six of these are conserved between *D. restrictus* strain PER-K23 and strain E1 (Table 5). Two of the conserved *rdhA* genes are located at the edge of cluster A and one at the edge of cluster B. Interestingly all four *rdhA* genes present outside cluster A or B are conserved between the two strains, which may indicate that both cluster A and B represent high plasticity regions unique to *D. restrictus* (Table 5). Currently, *pceA* (encoded by Dehre_2398) is the only RDH-encoding gene from *Dehalobacter restrictus* to be characterized in detail. The corresponding gene product PceA has been shown to catalyze the reduction of PCE to TCE and TCE to cis-DCE, the only two electron acceptors demonstrated to support growth of *D. restrictus* [12, 114]. The *pceA* gene belongs to a gene cluster, *pceABCT* (Dehre_2395-2398), which is highly similar to a gene cluster identified in a composite transposon structure identified in several *Desulfitobacterium* strains [116-118]. The transposon structure is not conserved in *D. restrictus* although the gene cluster is flanked by sequences resembling transposase genes in a late state of decay (Dehre_2394 and 2399). This combined with the fact that the *pceABCT* gene cluster including the cryptic transposases and the surrounding genomic context are conserved between *D. restrictus* and *D*. sp. strain E1 (data not shown) suggest that the presence of *pceABCT* is the result of an ancient horizontal gene transfer event.

Corrinoid synthesis and uptake

Corrinoid is the key cofactor in characterized RD catalytic subunits. *Dehalobacter restrictus* strain PER-K23 requires vitamin B_{12} in the medium for growth [12].

Therefore it is noteworthy to report the presence of a full set of corrinoid biosynthesis genes in the genome of *D. restrictus*, although *cbiH* (Dehre_2856) encoding precorrin-3B C17-methyltransferase displays a frame-shift mutation, and consequently is annotated as a pseudogene. The vitamin B_{12} synthesis pathway is encoded by two distinct gene clusters in *D. restrictus* strain PER-K23, where Dehre_2848-2865 encode enzymes of the upper pathway, and Dehre_1606-1615 the lower pathway. One additional gene (Dehre_1488) belonging to the lower pathway is located elsewhere in the genome (Figure 2) [115]. The genome encodes several gene clusters associated with corrinoid uptake and salvaging pathways. Preliminary studies of the proteome from cultures

grown at standard conditions or with partial vitamin B_{12} depletion showed that gene products encoded by one of the salvaging pathways (Dehre_0281-0291) were much more abundant in the vitamin B_{12} starved cells than in the cells grown under standard concentrations (J. Maillard and T. Kruse unpublished data). These findings suggest that the de novo corrinoid synthesis pathway is not functional and that *Dehalobacter restrictus* strain PER-K23 is dependent on salvaging corrinoids from the environment.

Hydrogenases

Another interesting feature is the presence of genes predicted to code for eight different hydrogenases. These include three periplasmic membrane-bound Ni/Fe uptake hydrogenases, consisting of three subunits: a catalytic unit, an Fe/S cluster protein and a membrane-bound b-type cytochrome (Dehre_551-553, 1061-1063 and 2405-2007), two six-subunits membrane-bound energy-conserving Ni/Fe hydrogenases (Dehre_1568-1573 and 1645-1650), and three Fe-only hydrogenases (Dehre_1739-1741, 2317-2320 and 2372-2374). The Fe-only hydrogenases consist of the catalytic subunit and two to three putative electron transferring subunits.

The presence of multiple uptake hydrogenases has also been observed in *Desulfitobacterium* spp., whereas *Dehalococcoides mccartyi* strains only have one uptake hydrogenase [8, 19, 26]. The two six-subunits Ni/Fe resemble the Hyc and Ech complexes found in *Dehalococcoides mccartyi* strain 195 [18], as well as the Hyc complex found in Desulfitobacterium spp [19, 26, 119].

Disrupting either one uptake hydrogenase or the six-subunits energy-conserving hydrogenase in *Desulfitobacterium dehalogenans* led to loss of the ability to grow using lactate or formate as electron donor and 3-chloro-4-hydroxyphenylacetate as electron acceptor, indicating that hydrogenases may play an important role in the electron transport chain to RD catalytic subunits, even when hydrogen is not used as the initial electron donor [119].

The role of the six-subunit hydrogenase complexes are still poorly understood. It has been speculated that they play a role in generating low potential electrons for OHR by reverse electron flow. However, this was considered as unlikely in one study where *Dehalococcoides mccartyi* strain 195 was cultivated in the presence of varying concentrations of hydrogen [120]. The exact role of the different hydrogenases in *Dehalobacter restrictus* strain PER-K23 still needs further studies.

The genome also encodes an intact Wood-Ljungdahl pathway (Dehre_0130-0155 and 2348-2351). The presence of a whole or partial Wood-Ljungdahl pathway has been observed in other OHRB. The closely related *Desulfitobacterium hafniense* strains Y51 and DCB-2^T both contain genes predicted to encode a full Wood-Ljungdahl pathway, and strain DCB-2^T has been shown to fix CO2 [42,43]. The more distantly related *Dehalococcoides mccartyi* strains have been shown to contain partial Wood-Ljungdahl pathways, but its exact role in the metabolism of these organisms remains unclear [121, 122].

The genome of *D. restrictus* contains 72 genes annotated as encoding transposases or inactive derivatives thereof, whereas it only contains few phage-associated genes despite the lack of a CRISPR phage immunity system.

Cells of *Dehalobacter restrictus* strain PER-K23 are motile [12]. The genome contains genes for synthesis of flagella and several genes predicted to be involved in chemotaxis. The role of chemotaxis in OHRB is currently understudied. Chemotactic behavior towards metals has been described for *Geobacter*, some members of this genus have been shown to be OHRB. Chemotactic behavior towards organohalides has, however, not been described for *Geobacter* spp [123-125].

Conclusion

The presence of an unexpectedly large number of putative RDH encoding genes suggests a far larger potential for use in bioremediation than previously anticipated, especially if *Dehalobacter restrictus* strain PER-K23 is attracted by organohalides in a chemotactic manner. The complete genome sequence of *Dehalobacter restrictus* strain PER-K23, the type strain of the genus *Dehalobacter*, represents a significant leap towards understanding the physiology, ecology and evolution of this specialized organohalide respiring group of bacteria. Current work focuses on obtaining a deeper understanding of the expression and regulation of the RDH genes, and thereby expanding the known organohalide substrate range of this organism. Shot-gun proteome analysis will aid in deciphering the metabolism of *D. restrictus* strain PER-K23 and allow generation of refined genome scale metabolic models of these dedicated degraders.

Acknowledgements

The work of Thomas Kruse and Hauke Smidt were supported by the Netherlands Genomics Initiative as well as the European Community's Seventh Framework Programme (FP7/ 2007-2013) through the Ecogenomics program and the BACSIN and METAEXPLORE projects (grant agreements No. 211684 and 222625), respectively. The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under contract No. DE-AC02-05CH11231.

Chapter 3

The restricted metabolism of the obligate organohalide respiring bacterium *Dehalobacter restrictus*: lessons from tiered functional genomics

Aamani Rupakula1[#], **Thomas Kruse**[#], Sjef Boeren, Christof Holliger, Hauke Smidt & Julien Maillard

# These authors contributed equally to the work.

Supplementary materials are available through the online version of this manuscript

Published as: Rupakula A, Kruse T, Boeren S, Holliger C, Smidt H, Maillard J (2013) The restricted metabolism of the obligate organohalide respiring bacterium *Dehalobacter restrictus*: lessons from tiered functional genomics. Philosophical Transactions of the Royal Society B: Biological Sciences 368 (1616). doi:10.1098/rstb.2012.0325

Abstract

Dehalobacter restrictus strain PER-K23 is an obligate organohalide respiring bacterium, which displays extremely narrow metabolic capabilities. It only grows via coupling energy conservation to anaerobic respiration of tetra- and trichloroethene with hydrogen as sole electron donor. *D. restrictus* represents the paradigmatic member of the genus *Dehalobacter*, which in recent years turned out to be a major player in the bioremediation of an increasing number of organohalides, both *in situ* and in laboratory studies. The recent elucidation of *D. restrictus* genome revealed a rather elaborate genome with predicted pathways that were not suspected from its restricted metabolism, such as a complete corrinoid biosynthetic pathway, the Wood-Ljungdahl pathway for CO_2 fixation, abundant transcriptional regulators and several types of hydrogenases. However, one important feature of the genome is the presence of 25 reductive dehalogenase genes, from which so far only one, *pceA*, was characterized on genetic and biochemical levels. This study describes a multi-level functional genomics approach on *D. restrictus* across three different growth phases. A global proteomic analysis allowed to consider general metabolic pathways relevant to organohalide respiration, while the dedicated genomic and transcriptomic analysis focused on the diversity, composition and expression of genes associated with reductive dehalogenases.

Introduction

Dehalobacter restrictus strain PER-K23 has been isolated from a tetrachloroethene (PCE) dechlorinating enrichment culture originally obtained from sediment of the Rhine River mixed with anaerobic granular sludge [12]. *D. restrictus* is a Gram-positive member of the *Firmicutes* growing exclusively via organohalide respiration (OHR) with H_2 as electron donor, PCE or trichloroethene (TCE) as sole electron acceptors, and acetate as carbon source. The key catalytic enzyme in OHR with PCE, the reductive dehalogenase PceA, has been purified and shown to harbour a corrinoid and two 4Fe/4S clusters [114]. In *D. restrictus*, the PceA enzyme is encoded by a gene that is part of the *pceABCT* gene cluster which has been shown to be highly conserved in several other OHR strains belonging to the genus *Desulfitobacterium* [116, 117]. The newly available genome sequence of *D. restrictus* was obtained and revealed a high number of 25 predicted reductive dehalogenase homologue (*rdhA*) encoding genes [24], though only PCE and TCE have been recognized as physiological substrates. This observation clearly raises the question of the true bioremediation potential of *D. restrictus*.

Two other *Dehalobacter* isolates have been reported: *Dehalobacter* sp. TEA able to dechlorinate PCE and TCE [40], and *Dehalobacter* sp. TCA1 dechlorinating *1,1,1*-trichloroethane to chloroethane [39], both strains being however not yet characterized in detail on biochemical and genetic level. Many studies have described cocultures or enrichment cultures where *Dehalobacter* spp. have been considered as the key player in the dechlorination of several other organohalides. A coculture containing *Dehalobacter* sp. E1 and *Sedimentibacter* sp. B4 has been obtained for the dechlorination of b-hexachlorocyclohexane (β-HCH) to benzene and chlorobenzene [53]. The draft genome of *Dehalobacter* sp. E1 has been recently reconstituted and was shown to harbour 10 putative *rdhA* genes, including a gene cluster with high similarity to *pceABCT* present in *D. restrictus*, although strain E1 has not been shown to grow on PCE [22, 54].

Further organohalides such as dichloroethane, chloroform, dichlorobenzenes or *4,5,6,7*-tetrachlorophthalide were shown to be dechlorinated by enrichment cultures dominated by *Dehalobacter* spp [126-129], suggesting that the degradation potential of the genus *Dehalobacter* is largely beyond PCE and TCE. Finally, fermentation of dichloromethane by members of *Dehalobacter* has been shown [41, 42], suggesting that not necessarily all members of this genus are obligate OHR bacteria (OHRB).

The apparent redundancy in *rdhA* genes can be rather considered as a genuine property of OHRB that are otherwise restricted in their metabolism. For example genomes of members of the OHR-obligate *Dehalococcoides* genus for which 5 different genomes are already available (and 3 more pending) display between 10 and 36 *rdhA* genes [31], most of which have unknown substrate range. In contrast, completed genomes of members of the metabolically versatile *Desulfitobacterium* genus revealed the presence of only a limited number of *rdhA* genes with *D. hafniense* DCB-2 harbouring a maximum of 7 copies [19, 26]. While the composition of the genes associated with *rdhA* genes is strongly varying in the genomes of OHRB, *rdhA* subunits are almost invariably accompanied by a short open reading frame, *rdhB*, with the exception of the recently sequenced genome of *Dehalogenimonas lykanthroporepellens* [21]. Despite a very low level of sequence similarity, RdhB proteins display consensually 2 or 3 transmembrane helices strongly indicating a role in anchoring the catalytic subunit in the membrane.

Recently proteomics and transcriptomics studies were used to study the metabolism of the two OHRB, *Desulfitobacterium hafniense* strain TCE1 [130] and Y51 [131], respectively, under different growth conditions, both confirming the apparent lack of regulation of the *pceA* gene that was postulated earlier [116]. Most omics studies involving OHRB have however focused on members of the *Dehalococcoides* genus. This genus, although phylogenetically distant to *Dehalobacter*, inhabits similar ecological niches and is exclusively dependent on OHR metabolism with H_2 as electron donor. These studies have employed both transcriptomics using full genome microarrays and proteomics to identify key components of the metabolism of OHRB under different growth conditions or growth phases [32, 120, 132-136].

In addition to genes directly linked to reductive dehalogenation, the genome of *D. restrictus* furthermore encodes one formate dehydrogenase, eight hydrogenase complexes, among which three uptake hydrogenases (Hup-type) and one energy-conservation hydrogenase (Ech-type) and one hydrogenase-3 (Hyc-type) [24], similar to what has been described for *Dehalococcoides*. No data is yet available, however, concerning the role of these enzymes in the metabolism of *D. restrictus*.

Detailed studies of the metabolism of members of the *Dehalobacter* genus have so far been hampered by the lack of full genome information. Hence, the recently elucidated genome sequence of *D. restrictus* now provides the necessary basis for detailed studies of the metabolism of this obligate OHR bacterium using a tiered functional genomics approach.

Material & Methods

Bacteria and growth conditions

Dehalobacter restrictus strain PER-K23 (DSM 9455) was cultivated as described earlier [12, 114]. Anaerobic serum flasks were supplemented with hydrogen as electron donor, inoculated with 2% (v/v) inoculum, and finally 1% (v/v) of 2 M PCE solution in hexadecane was added as electron acceptor. Nine batch cultures of *D. restrictus* were cultivated in 300 ml medium at 30°C under agitation (100 rpm) and their growth was monitored by chloride production and not optical density as it is biased by precipitation of medium component. The true nature of organohalide respiration (i.e. the link between dechlorination and growth) was already demonstrated for *D. restrictus* [12]. Triplicate cultures were each harvested at three different growth stages of chloride release (20, 30, and 40 mM) that we have defined as the exponential (E), late-exponential (LE) and stationary (S) phases (Fig. S1). Aliquots of 50 ml culture were collected for transcriptomic analysis, while the rest of each culture was harvested for proteomic analysis. For RNA extraction, 50 ml was collected by 2 min centrifugation at 4600 *g* at 15°C, the pellet was readily resuspended in 1 ml of LifeGuard™ (MoBio, Carlsbad, CA, USA), incubated for 1 min and flash-frozen in liquid nitrogen. The remaining 250 ml of culture were centrifuged for 10 min as above for proteomic analysis. The pellet was washed in 10 mM Tris-HCl (pH 7.5) containing 1 mM EDTA, and then flash-frozen in liquid nitrogen. All biomass samples were stored at -80°C until use. *E. coli* DH5α was cultivated on standard liquid or solid LB medium containing 100 $\mu g \cdot l^{-1}$ ampicillin when transformed with derivatives of the pGEM-T Easy vector (Promega, Duebendorf, Switzerland).

Sequence analysis

All sequences mentioned in this study are taken from the recently published genome of *D. restrictus* strain PER-K23 [24]. The annotation of specific genes was verified using manual search with BlastP [137]. Rho-independent transcription terminators were identified with TransTerm from the Nano+Bio-Center of Kaiserslautern Technical University (http://nbc11.biologie.uni-kl.de) using default parameters. Protein sequences were aligned using ClustalX 2.0 [138]. The RdhA tree was built with MEGA5 [35].

RNA Extraction

RNA was extracted using the TRIzol method according to [139] with the following modification. The DNaseI treatment was stopped by adding 1× DNase stop solution and incubating for 10 min at 65°C. RNA concentration was estimated using the Nanodrop ND-1000 spectrophotometer (Thermo Scientific, Ecublens, Switzerland).

Reverse transcription

Two µg of RNA were added to 4.5 µg of random hexamer (Microsynth GmbH, Balgach, Switzerland) in a volume of 85 µl. This mixture was incubated at 70°C for 5 min and then placed on ice. A 75 µl of reverse transcription (RT) mix contained, 32 µl of 5× buffer, 8 µl of 10 mM dNTPs, 19.2 µl of 25 mM MgCl2, 4 µl of RNasin (40 U·µl-1) and 8 µl of ImProm-II reverse transcriptase (Promega). The RT was performed as follows: 25°C for 5 min, 42°C for 60 min, and 70°C for 15 min in a T3 Thermocycler (Biometra, Goettingen, Germany).

Primer design

Specific primers were designed for each *rdhA* gene present in the *D. restrictus* genome by targeting unique regions. The primers were chosen such that the amplified products would fall in a size range suitable for quantitative PCR (qPCR, see below). Primer sequences and expected amplicon sizes are given in Table S1.

End-point PCR approaches

Different PCR strategies were applied in this study: standard endpoint PCR (sPCR), multiplex endpoint PCR (mPCR) and quantitative PCR (qPCR, see below). Standard PCR reactions were carried out in 10 µl containing 4.25 µl ddH_2O, 1 µl 10× buffer, 0.3 µl dNTPs at 10 mM each, 0.4 µl 25 mM $MgCl_2$, 1 µl each primer at 10 µM, and 0.05 µl Taq polymerase at 5 $U \cdot \mu l^{-1}$ (Peqlab, Erlangen, Germany). Two µl of genomic DNA or cDNA were added as template. For mPCR, a solution with 8 different primers (4 targets) was prepared containing 10 µM of each primer. Two µl of that solution was added in the standard reaction mix. Standard PCR and mPCR were performed in a Thermocycler (Biometra) using the following conditions: 5 min of initial denaturation at 95°C, followed by 30 cycles of 1 min denaturation at 95°C, one primer annealing at 52°C, and 1 min elongation at 72°C. A final extension step of 10 min at 72°C was added at the end. The PCR products were routinely analyzed in 1.5% (w/v) agarose gels stained with GelRed

(Biotium, Hayward, CA, USA). DNA was visualized using the Syngene gel imaging system (Syngene, Cambridge, UK).

Cloning and sequencing of PCR products

PCR products were purified with the QIAquick PCR purification kit (Qiagen, Hombrechtikon, Switzerland) according to manufacturer's instructions. The products were then A-tailed following instructions from the pGEM T-Easy vector manual (Promega), and finally ligated into pGEM T-Easy overnight at 16°C. The ligated products were cloned by heat shock transformation of $CaCl_2$-competent *E. coli* DH5α. Transformants were screened using colony PCR with primers T7 and SP6, and positive clones were cultivated overnight at 37°C followed by plasmid preparation with the QIAprep Spin Miniprep kit (Qiagen). Plasmid inserts were verified by sequencing using the BigDye Terminator 3.1 kit on the ABI Prism 3130 Genetic Analyzer according to manufacturer's instructions (Applied Biosystems).

Quantitative PCR

Standards for qPCR were prepared from plasmids containing the gene targets as follows. One µg of plasmid DNA was digested with 5 units of ScaI restriction enzyme (Promega) for 2 h at 37°C. The linearized plasmid was dephosphorylated during 1 h at 37°C by adding 1 µl shrimp alkaline phosphatase (Takara, Clontech Laboratories, Mountain View, CA, USA), followed by purification with the QIAquick PCR purification kit (Qiagen). The DNA concentration was measured with the Nanodrop ND-1000 spectrophotometer (Thermo Scientific). Serial dilutions (10-1 to 10-8 copies·µl-1) of the purified sample were finally prepared and used as standards. A typical 10 µl qPCR reaction contained 5 µl of KAPA SYBR FAST Universal 2× qPCR master mix (KAPA Biosystems, Woburn, MA, USA), 0.2 µl of each primer at 10 µM, 2.1 µl of ddH2O, and 2.5 µl of template DNA (standards or samples). The reactions were performed in the Rotor Gene qPCR machine (RG-3000, Corbett Research, Qiagen) using the following program: 2 min of initial denaturation at 95°C, then 40 cycles of 30 s denaturation at 95°C, 30 s primer annealing at 58°C, and 20 s elongation at 72°C. Fluorescence was measured at the end of each elongation step. Each run consisted of triplicate reactions for both the standards and the samples. Run performances are given in Table S2 for each considered gene target. The obtained data were expressed as transcript copy number per µl of initial cDNA samples.

Protein extraction and SDS-PAGE

Cell pellets were transferred to two ml low binding micro centrifuge tubes (Eppendorf, Nijmegen, The Netherlands) prior to protein extraction. Protein extraction was done in 500 µl SDT-lysis buffer (100 mM Tris-HCl pH 7.6, 4% SDS, 0.1 M dithiotreitol). Cells were lysed by sonication, using a Branson sonifier equipped with a 3 mm tip (six pulses of 30 s with 30 s rest on ice in-between each pulse, strength of the pulse was increased stepwise from setting 2 to 4). Proteins were denatured by boiling for 5 min, followed by 10 min centrifugation at 15700 *g*. Protein concentrations were determined using the Bradford method [140]. Finally SDS-PAGE was performed with gels containing 10% acrylamide using a MiniProtean III system (Bio-Rad, Veenendaal, The Netherlands). Samples containing 10 µg protein were mixed with 2× loading buffer (100 mM Tris-HCl pH 6.8, 200 mM dithiotreitol, 4% SDS, 0.2% bromophenol blue and 20% glycerol) and briefly heated to 95°C before loading on gels. Gels were stained with Coomassie Brilliant Blue.

In-gel trypsin digestion

For the growth phase experiment each lane was cut in five slices of approximately equal size. Each slice was cut into approximately 1 mm^3 pieces and transferred to independent 500 µl low binding micro centrifuge tubes (Eppendorf). All solutions were prepared using 50 mM NH_4HCO_3 unless otherwise stated. Tubes were briefly centrifuged and the liquid phase removed between each step. Proteins were reduced by incubating in 50 mM dithiotreitol for 1 h at 60°C while slowly shaking, and alkylated by incubation in 100 mM iodoacetamide for 1 h in the dark at room temperature, washed once and incubated with 20 ng trypsin (sequencing grade, Roche Diagnostics, Almere, The Netherlands) over night at room temperature. Samples were sonicated in a water bath for 30 min before the supernatant was transferred to fresh 500 µl low binding micro centrifuge tubes. To increase the yield the gel pieces were covered with 10% trifluoroacetic acid in H_2O and sonicated for another 30 min. Then an equal volume of a solution containing 15% acetonitrile and 1% trifluoroacetic acid in H_2O were added. The samples were sonicated for 1 min, before supernatants were combined in the low binding micro centrifuge tubes mentioned above. Peptides were concentrated using StageTip C18 columns essentially as described in [141]. Finally the volume was reduced to 10 µl using a SpeedVac vacuum centrifuge, and increased to 25 ul with 0.1% (v/v) formic acid. Samples were measured by nLC-MS/MS with a Proxeon nLC and a LTQ-Orbitrap mass spectrometer as described in [142].

LC-MS data analysis

LC-MS runs with all MS/MS spectra obtained were analyzed with MaxQuant 1.2.2.5 [143] using default settings for the Andromeda search engine [144], except that extra variable modifications were set for de-amidation of N and Q.

A protein database was generated based on the genomes of *D. restrictus* and *Dehalobacter* sp. E1 [22, 24], using the Artemis genome browser, and combined with a database that contains sequences of common contaminants as for instance BSA (P02769, bovine serum albumin precursor), trypsin (P00760, bovine), trypsin (P00761, porcine), keratins K22E (P35908, human), K1C9 (P35527, human), K2C1 (P04264, human) and K1CI (P35527, human) [145]. The label-free quantification (LFQ) as well as the match between runs options (with ± 2 min retention time deviation) were enabled. De-amidated peptides were allowed to be used for protein quantification and all other quantification settings were kept default. Filtering and further bioinformatic analysis of the MAXQUANT/Andromeda workflow output and the analysis of the abundances of the identified proteins were performed with the PERSEUS v. 1.2.0.16 module (available at the MAXQUANT suite). Accepted were peptides and proteins with a false discovery rate (FDR) of less than 1 per cent and proteins with at least two identified peptides of which one should be unique. Also, quantification was carried out by the MAXQUANT software for which MS data of at least three isotopes per peptide are used [143], and at least two quantified peptides per protein. This method makes label-free relative quantification reliable and therefore possible [146, 147]. Reversed hits were deleted from the MAXQUANT result table as well as all results showing a LFQ value of 0 for both sample and control. Zero values for one of the two LFQ columns were replaced by a value of 5 to make sensible ratio calculations possible. Relative protein quantification of sample to control was conducted with PERSEUS v. 1.2.0.16 by applying a two sample t-test using the 'LFQ intensity' columns obtained with threshold 0.10 and S0 = 1.

Results

Proteomic analysis of *D. restrictus* along growth phases

The genome of *D. restrictus* strain PER-K23 was predicted to encode 2826 proteins [24]. Using a combined protein database generated from the genomes of *D. restrictus* and *Dehalobacter* sp. E1, we identified 1055 proteins by proteome analysis (Table S3 and Fig. S2), of which 15 have been previously annotated as pseudogenes in *D. restrictus*, and one was newly discovered (Table S4). Data obtained from biological triplicates taken at the designated exponential (E), late-exponential (LE), and stationary (S) phases (see Material and Methods) were used to calculate the relative abundance ratio of proteins at stationary versus exponential phase (S/E); late-exponential versus exponential phase (LE/E), and stationary versus late-exponential phase (S/LE). The S/E protein abundance ratios of only 38 proteins were considered as statistically different (with False Discovery Rate < 0.1), and corresponded to ratios between 25- to 3000-fold (Table 1). However, in a mere qualitative approach, we considered a three-fold in-/decrease in relative protein abundance as cut-off to define the proteins which differed between growth phases. This selection allowed investigating general trends in protein changes across the different growth phases. The largest differences were seen between stationary and exponential phases, where the production of 29% of all identified proteins seemed to be regulated. Comparing late-exponential and exponential phase, or stationary and late-exponential phase, only 16 and 18% of all identified proteins were produced at different levels, respectively. In the following, we focused on selected proteins and metabolic pathways most directly linked to the organohalide respiratory lifestyle of *D. restrictus* (Table 2), but the complete data set is given as Supplementary Material (Table S5). The housekeeping enzyme RNA polymerase (RpoB, Dehre_0495) was detected at stable levels throughout all growth phases.

Table 1. Detected proteins showing a significant increase/decrease in abundance (expressed as S/E ratio) during the transition from exponential (E) to stationary (S) phases. The S/LE and LE/E ratios are also indicated. (a) Ratios above 1 mean increase in protein anundance, ratios below 1 mean decrease. (*) gene not annotated

Locus tag (Dehre_#)	Annotated function	Protein abundance ratio[a]		
		S/E	S/LE	LE/E
Proteins displaying significant increase in S/E ratio				
1215	Late competence development protein ($ComF_B$)	3.0×10^{3}	1.8×10	1.7×10^{2}
0983	Cupin-domain protein	2.0×10^{3}	1.5×10	1.3×10^{2}
0318	Uncharacterized protein conserved in bacteria	8.4×10^{2}	4.2	2.0×10^{2}
2151	Aspartyl/glutamyl-tRNA (Asn/Gln) amidotransferase	6.6×10^{2}	7.1	9.4×10
0568	Similar to acyl-coenzyme A synthetase/AMP-fatty acid ligase	4.4×10^{2}	8.7	5.1×10
0109	Predicted transcriptional regulator	3.9×10^{2}	1.4×10	2.8×10
1963	Uncharacterized protein conserved in bacteria	3.7×10^{2}	1.1×10	3.3×10
0668	RelE-type toxin (TA system)	2.6×10^{2}	3.6×10	7.3
0856	Response regulator with CheY-like and AraC-type domains	2.6×10^{2}	2.4×10	1.1×10
2645	Uncharacterized domain 1 protein	2.3×10^{2}	1.6×10	1.5×10
2325	Hypothetical protein	2.3×10^{2}	9.0×10	2.5
1400	Nitrogen regulatory protein PII	2.3×10^{2}	1.7×10	1.4×10
0147	CODH/ACS, maturation factor	1.8×10^{2}	3.1	5.7×10
1786	Hypothetical protein	1.2×10^{2}	1.5×10	8.0
0146	CODH/ACS, maturation factor	1.2×10^{2}	4.5	2.7×10
1237	Acetate-CoA ligase	7.3×10	2.6	2.8×10
0264	Hypothetical protein	7.1×10	5.3	1.3×10
2205	YGGT protein family	5.1×10	1.1×10	4.8
2544	Hypothetical protein	5.0×10	7.8	6.5
0651	Predicted transcriptional regulator	4.9×10	2.2	2.2×10
1310	Transcription antitermination factor NusB	3.7×10	1.0×10	3.5
0143	Pterin-binding enzyme	3.3×10	2.5	1.3×10
2265	Response regulator containing CheY-like receiver	3.0×10	6.5	4.6
2560	Transcriptional regulator	2.7×10	8.8	3.1
0144	CODH/ACS, γ-subunit (CFeSP)	2.5×10	1.6	1.5×10
0198	Bacterial nucleoid DNA-binding protein	2.5×10	7.3	3.4
Proteins displaying significant decrease in S/E ratio				
2864	Cobalt transport protein	3.4×10^{-2}	1.5×10^{-1}	2.3×10^{-1}
1795	Predicted transcriptional regulator containing CBS domains	1.5×10^{-2}	2.9×10^{-2}	5.0×10^{-1}
2895	Uncharacterized protein conserved in bacteria	2.0×10^{-3}	1.6×10^{-1}	1.3×10^{-2}
0194	Transcription-repair coupling factor	2.0×10^{-3}	5.0×10^{-3}	5.2×10^{-1}
*	LSU ribosomal protein L34p	3.0×10^{-3}	4.0×10^{-3}	6.6×10^{-1}
0258	Excisionase-like DNA-binding domain	3.0×10^{-3}	9.0×10^{-3}	3.6×10^{-1}
2307	Uncharacterized protein conserved in bacteria	5.0×10^{-3}	1.0	5.0×10^{-3}
2254	Chemotaxis protein stimulating methylation of MCP proteins	6.0×10^{-3}	8.0×10^{-3}	7.6×10^{-1}
2873	Trypsin-like serine protease	6.0×10^{-3}	9.0×10^{-3}	7.2×10^{-1}
1962	Predicted Fe-S oxidoreductase	9.0×10^{-3}	3.8×10^{-2}	2.3×10^{-1}
2146	rRNA (uracil-5-)-methyltransferase (RumA)	9.0×10^{-3}	1.0	9.0×10^{-3}
2280	Hypothetical protein	1.0×10^{-2}	6.7	1.0×10^{-3}

Table 2. Proteomic analysis of selected metabolic pathways of *D. restrictus*

Locus tag (Dehre_#)	Protein	Annotated function	Protein abundance ratio		
			S/E	S/LE	LE/E
Proteins associated with organohalide respiration					
2022	RdhA14	Reductive dehalogenase	3.4×10^{-1}	5.4×10^{-1}	6.3×10^{-1}
2025	RdhK15	CPR/Fnr-type regulator	1.2	3.1	3.8×10^{-1}
2048	RdhK20	CPR/Fnr-type regulator	4.7×10^{-1}	4.0×10^{-1}	1.2
2395	PceT	Chaperone (Trigger factor)	1.2×10^{-1}	2.8×10^{-1}	4.3×10^{-1}
2396	PceC	FMN-binding domain	6.9×10^{-1}	8.0×10^{-1}	8.6×10^{-1}
2397	PceB	Membrane anchor	5.3×10^{-1}	6.8×10^{-1}	7.8×10^{-1}
2398	PceA	PCE reductive dehalogenase	8.0×10^{-1}	1.2	6.9×10^{-1}
Proteins associated with corrinoid synthesis and uptake					
0286		ABC-type iron transporter, substrate-binding	6.6×10^{-1}	1.3	8.3×10^{-1}
0289		Mg/Co protoporphyrin IX chelatase	6.6×10^{-2}	1.0	6.6×10^{-2}
0291	NodI	ABC-type Nod export system, ATP-binding	5.2×10^{-1}	2.9	1.5
1488	CobT	Nicotinate-nt-DMB phosphoribosyltransferase	2.0×10^{-1}	3.2	6.4×10^{-1}
1606	CobA	Cob(I)yrinic acid a,c-diamide adenosyltransferase	1.2	1.0	1.2
1607	CbiP	Cobyric acid synthase	6.8×10^{-1}	1.1	7.5×10^{-1}
1608		Phosphoglycerate mutase	8.3×10^{-1}	1.1	9.3×10^{-1}
1609	CobD	L-Thr-O-3-phosphate decarboxylase	9.9×10^{-1}	8.4×10^{-1}	8.3×10^{-1}
1610	CbiB	Adenosylcobinamide-phosphate synthase	3.3	3.8×10^{-1}	1.2
1611	CobC	Alpha-ribazole-5'-phosphate phosphatase	1.3	8.1×10^{-1}	1.0
1612	CobU/CobP	Cobinamide kinase/phosphate guanylyltransferase	7.4×10^{-1}	1.7	1.3
1614	CobU/CobP	Cobinamide kinase/phosphate guanylyltransferase	1.1	1.1	1.3
1615	CbiA	Cobyrinic acid a,c-diamide synthase	6.9×10^{-1}	1.0	7.0×10^{-1}
2535	BtuF	ABC-type Cbl/$Fe3^{+}$ transporter, substrate-binding	1.5	8.5×10^{-1}	1.2
2537	BtuD	ABC-type Cbl/$Fe3^{+}$ transporter, ATPase	5.8×10^{-1}	1.2	7.2×10^{-1}
2538	CbiZ	Adenosylcobinamide amidohydrolase	2.5	2.3	5.6
2848	CbiC	Precorrin-8x methylmutase	8.3×10^{-1}	9.4×10^{-1}	8.9×10^{-1}
2850	CbiX	Sirohydrochlorin cobalt chelatase	1.7	2.1	8.1×10^{-1}
2851	HemL	Glutamate-1-semialdehyde 2,1-aminomutase	3.2×10^{-1}	5.0×10^{-1}	6.4×10^{-1}
2852	HemB	D-aminolevulinic acid dehydratase	5.9×10^{-1}	7.6×10^{-1}	7.8×10^{-1}
2853	CysG/HemD	Uroporphyrinogen-III synthase/C-methyltransferase	8.6×10^{-1}	8.7×10^{-1}	9.9×10^{-1}
2854	HemC	Porphobilinogen deaminase	6.5×10^{-1}	7.2×10^{-1}	9.0×10^{-1}
2857	HemA	Glutamyl-tRNA reductase	6.9×10^{-2}	8.2	5.7×10^{-1}
2859	CbiF	Precorrin-4 C11-methyltransferase	1.1	9.0×10^{-1}	9.6×10^{-1}
2860	CbiL	Precorrin-2 C20-methyltransferase	1.1	1.0	1.1
2862	CbiO	ECF-type cobalt transporter, ATPase	4.0×10^{-2}	9.6×10^{-1}	3.8×10^{-2}
2864	CbiN	ECF-type cobalt transporter, bipartite component	3.4×10^{-2}	6.7	2.3×10^{-1}
Proteins belonging to the Wood-Ljundahl pathway					
0140		Predicted RNA-binding protein	3.8 x 10	1.1	3.5 x 10
0142		CODH/ACS, α-subunit	1.0 x 10	2.0	5.2
0143		Pterin-binding enzyme	3.3 x 10	2.5	1.3 x 10
0144		CODH/ACS, γ-subunit	2.5 x 10	1.6	1.5 x 10
0145		CODH/ACS, δ-subunit	1.1 x 10	1.3	8.0
0146		CODH/ACS, maturation factor	$1.2E \times 10^{2}$	4.5	2.7 x 10
0147		CODH/ACS, maturation factor	1.8×10^{2}	3.1	5.7 x 10
0148		CODH/ACS, β-subunit	5.4	1.5	3.5
0150		Pterin-binding enzyme	1.3	1.0	1.2
0151		Methylene-H_4F-DH/methenyl-H_4F cyclohydrolase	2.3	1.1	2.1
0152		Methenyl-H_4F cyclohydrolase	6.3×10^{-1}	5.3×10^{-1}	1.2

Reductive dehalogenases

The genome of *D. restrictus* contains 25 genes predicted to encode reductive dehalogenase homologues (*rdhA*) (see below). Overall, a total of 86 genes are potentially associated with reductive dehalogenase expression and maturation, including genes that are predicted to encode putative membrane anchors, transcriptional regulators, chaperones, and other *rdh* associated genes. Two of the reductive dehalogenase catalytic subunits (RdhA) were detected in the proteome: RdhA14 (Dehre_2022) and RdhA24 (PceA, Dehre_2398). The former shows a very high amino acid sequence identity (89%) with RdhA2 from *Desulfitobacterium hafniense* DCB-2 (Dhaf_0693) [26], while the latter is the biochemically characterized PceA [114] (Table S6). All four proteins encoded by the *pceABCT* gene cluster (Dehre_2398 to Dehre_2395) were also identified in the proteome (Table 2). PceA was among the most abundant proteins at all growth stages (data not shown). The protein abundance ratio of PceA, PceB and PceC remained within the three fold cut-off value when comparing any of the three growth phases considered. The absence of transcription regulators in the direct vicinity of the *pce* gene cluster (see below) suggests that PceA is constitutively expressed, although it needs to be further investigated. PceT, however, was the only member of the gene cluster that seemed to be regulated as the relative protein abundance ratios were 0.12, 0.43 and 0.28 for S/E, LE/E and S/LE, respectively. Although the value for LE/E did not exceed the cut-off value, the data suggests that PceT was most abundant at exponential phase and then became slightly less abundant at later growth stages (Fig. 1 and Table 2).

Hydrogenases

Hydrogen is the only electron donor that *D. restrictus* has been shown to utilize. The key role of hydrogenases is underscored by the fact that the genome of *D. restrictus* is predicted to encode eight multi-subunit hydrogenase complexes. Three of these (Dehre_0551-0553; 1061-1063 and 2405-2407) belong to the group of periplasmic membrane-bound Ni/Fe uptake hydrogenases [148] consisting of three subunits, a membrane-bound *b*-type cytochrome, a Fe/S cluster protein and the catalytic subunit (Table S5, Fig. 1). Two membrane-bound energy-conserving Ni/Fe hydrogenases (Dehre_1568-1573 and 1645-1650) resemble the Hyc and Ech clusters found in *Dehalococcoides mccartyi* 195 [18]. These two hydrogenase complexes each consist of six subunits, a large and small subunit, and four subunits resembling elements of the proton-translocating respiration complex I (Table S5, Fig. 1). The three Fe-only hydrogenases [146] consist of the catalytic unit and two or three subunits predicted to be involved in electron transfer. Unlike what

was observed in *D. mccartyi* 195 [18], none of the Fe-only complexes contains any predicted transmembrane region, which suggests that they are either located in the cytoplasm or form a complex with other membrane-bound proteins (Table S5, Fig. 1).

A constant amount of the large and small subunits from one of the Hup-type hydrogenases (Dehre_0552-0553) was detected throughout all growth phases. We did not detect the *b*-type cytochrome subunit (Dehre_0551), possibly as a consequence of its strong association with the membrane. Both putative energy-conserving hydrogenase complexes were detected in the cells. We detected the large and small subunit of the Hyc-type hydrogenase (Dehre_1568-1569), but none of the four subunits (Dehre_1570-1573) predicted to be involved in electron transfer and proton transport across the cell membrane. Interestingly, the small subunit (Dehre_1568) was most abundant at late exponential phase and least abundant in stationary phase with S/E, S/LE and LE/E ratios of 0.24, 0.03, and 9.74, respectively, whereas the abundance of the large subunit did not differ between growth phases (Table S5, Fig. 1). All but one (Dehre_1649) component of the Ech complex (Dehre_1645-1650) were detected. The only protein that differed in abundance between growth phases was Dehre_1647, predicted to encode an NADH-ubiquinone oxidoreductase. This protein became gradually more abundant at later growth stages with S/E, S/LE and LE/E ratios of 8.4, 1.8 and 4.8, respectively. We detected both 3-subunit Fe-only hydrogenases (Dehre_1739-1741 and Dehre_2372-2374) in the proteome, but none of the components of the 4-subunit complex (Dehre_2317-2320). The abundance of Dehre_1739-1741 did not change with the growth phases, whereas Dehre_2372-2374 showed a weak trend of decreasing abundance at later growth phases, most pronounced for Dehre_2373 with S/E, S/LE, and LE/E ratios of 0.30, 0.32 and 0.96, respectively (Table S5).

Corrinoid synthesis and uptake

The genome of *D. restrictus* encodes a seemingly complete *de novo* corrinoid biosynthesis pathway starting from glutamyl-tRNA (Fig. 1). This pathway is encoded by two distinct gene clusters in *D. restrictus*: cluster I (Dehre_2848-2865), the upper pathway, and cluster II (Dehre_1606-1615), corresponding to the lower pathway. One additional gene (Dehre_1488) belonging to the lower pathway is located elsewhere in the genome (Table 2).

Cluster I contains all genes necessary for the synthesis of cobyrinic acid starting from glutamyl-tRNA. This pathway, however, appears to be incomplete since *cbiH* (Dehre_2856) encoding pre-

corrin-3B C17-methyltransferase displays a frame-shift mutation, and consequently is annotated as a pseudogene. We identified several proteins of the corrinoid synthesis pathway until cobyrinic acid, except CbiH and all enzymes responsible for the conversion of cobalt-precorrin-5A to cobalt-precorrin-8. From the upper pathway, only HemA (Dehre_2857) and HemL (Dehre_2851) showed a decreasing relative abundance from exponential to stationary phases with S/E ratios of 0.07 and 0.32, respectively. Most enzymes of the lower corrinoid synthesis pathway encoded by cluster II were found in stable amounts throughout the growth phases with exception of CbiB and CobS. CbiB (Dehre_1610) is responsible for the conversion of adenosylcobyric acid to adenosylcobinamide and was found in slightly increasing amounts at stationary phase (S/E: 3.26), while CobS (Dehre_1613) which is responsible for the conversion of adenosylcobinamide-GDP to adenosylcobalamin was not detected at all.

The genome of *D. restrictus* contains several gene clusters predicted to be involved in cobalt and corrinoid uptake. One predicted ABC-type cobalt transporter (Dehre_0850-0852) and two ECF-type cobalt transporters (Dehre_0278-0280 and Dehre_2862-2865) are present in *D. restrictus*. While none of Dehre_0850-0852 or of Dehre_0278-0280 were detected in the proteome, we identified both CbiQ (Dehre_2862) and CbiN (Dehre_2864) proteins from the transport system encoded in corrinoid synthesis gene cluster I. Both showed a decreasing trend when going from exponential to stationary phases with S/E ratio of 0.04 and 0.03, respectively (Table 2). Two gene clusters (Dehre_0281-0292 and Dehre_2535-2538) are predicted to encode proteins possibly involved in uptake of various corrinoid precursors as part of salvaging pathways. From the first cluster three proteins (Dehre_0286, 0289 and 0291) were detected. Their protein abundance ratio did not change over time, except for Dehre_0289 which was only detected during exponential phase (Table 2), while from the second cluster, all proteins except the membrane-associated Dehre_2536, were detected. Only Dehre_2538 showed an LE/E ratio exceeding the 3-fold cut-off (5.63). Interestingly this protein is predicted to encode a CbiZ homologue which salvages cobinamides and converts it back to cobyric acid [149].

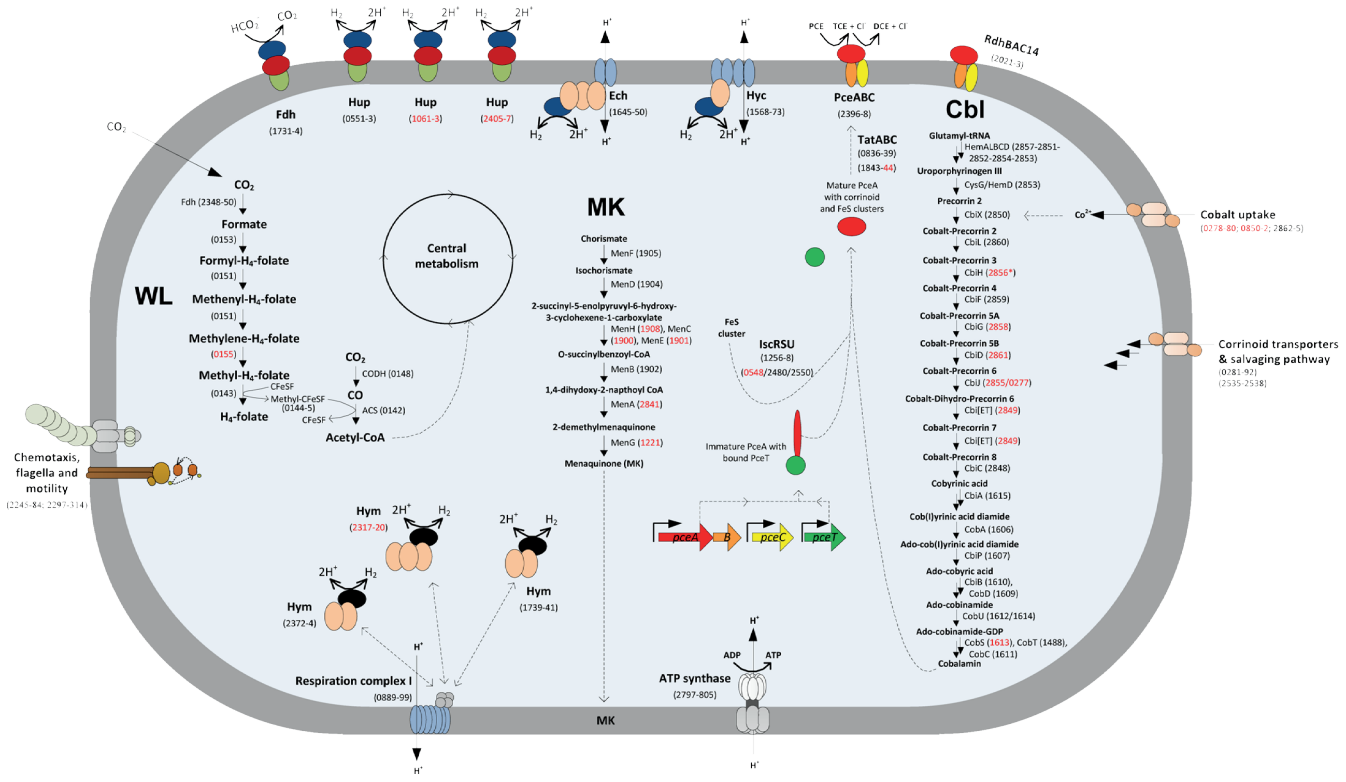


Figure 1. Metabolic map showing selected pathways of *D. restrictus*. Metabolic pathways, both predicted from *D. restrictus* genome and analysed by functional genomics, are presented in a simplified bacterial cell. The cytoplasmic membrane is depicted in grey, the cytoplasm in blue shading. *Dehalobacter restrictus* genomic loci (Dehre_#) are indicated in parentheses, with black label showing proteins detected in the proteomic analysis and those not detected indicated in red. Three important pathways are given in detail: the Wood–Ljungdahl pathway (WL), the menaquinone biosynthesis pathway (MK) and the corrinoid biosynthesis pathway (Cbl).

Additional elements of the general energy metabolism

Constant amounts of six proteins (Dehre_2797-2802) out of the ten subunits of the ATP synthase (Dehre_2797-2806) were detected in the proteome. The proton-translocating respiration complex I encoded in the genome of *D. restrictus* consists of 11 subunits (Dehre_0889-899) instead of the canonical 14 [24], lacking the components NuoEFG that usually receive electrons from NADH. Three subunits (NuoBCD, Dehre_890-892) were clearly detected in the proteome, whereas none of the membrane components could be seen.

The genome encodes enzymes of a putative Wood-Ljungdahl pathway for CO_2 fixation (Dehre_0130-0155 and Dehre_2348-2351). Most proteins belonging to this pathway were detected in the proteome (Table S5). They were observed at constant levels throughout the growth phases with the exception of proteins representing the carbonyl branch of the Wood-Ljungdahl pathway and the acetyl-CoA synthase/CO dehydrogenase (ACS/CODH) complex. Generally these proteins showed a gradual and significant increase in relative abundance towards later growth stages with S/E ratios between 25 and 175 (Table 1).

We also identified a putative three component formate dehydrogenase (Fdh), consisting of a membrane-bound *b*-type cytochrome, a Fe/S cluster protein and the catalytic subunit (Dehre_1730-1734), which were detected at all growth phases (Table S5, Fig. 1). The catalytic unit contains probably a selenocysteine as it is encoded by two in-frame genes (Dehre_1733-1734) separated by a UGA stop codon.

The genomic loci Dehre_2245-2284 and 2297-2314 contain large numbers of genes involved in the synthesis of flagella, motor proteins and chemotaxis (Fig. 1). In the proteome, we identified 32 out of 62 proteins encoded in these genomic regions. Sixteen of these were less abundant in stationary than in exponential phase, only three increased in abundance, and the remaining 13 were equally abundant during stationary and exponential phase (Table S5), indicating that the cells are reducing their motility when entering the stationary phase.

Proteins showing significant changes in abundance between stationary and exponential phase are displayed in Table 1. Generally, many proteins associated with regulation of transcription, chemotaxis, and sensing, were among the proteins displaying significant changes in their abundance. The protein showing the greatest change in abundance, with an S/E value of 2954, is Dehre_1215, annotated as $ComF_B$, an uncharacterized protein possibly involved in development

of late competence [150, 151]. The gene cluster containing the *comF*$_B$ gene (Dehre_1214-1220) in *D. restrictus* contains genes predicted to encode an RNA helicase, an ABC transporter, and two genes encoding proteins of unknown function. We detected the periplasmic component of the ABC transporter and one of the hypothetical proteins in the proteome, the latter increasing in abundance at later growth phases (Table S5). The genome of *D. restrictus* encodes other competence factors such as ComE$_A$ and ComE$_C$ (Dehre_0586-0587), and ComF$_A$ (Dehre_2784), suggesting that it is capable of natural competence. None of these additional proteins, however, were detected in the proteomic analysis. Two gene clusters encoding pili (Dehre_1166-1175 and Dehre_1272-1289) possibly involved in DNA uptake are also present.

Another protein (Dehre_0668) that was among those with the strongest increase in abundance in stationary phase (Table 1) has high similarity with RelE toxin and builds with Dehre_0667 a toxin/antitoxin addiction module system which could be involved in modulating the persistence of cell growth in unfavourable growth conditions [152]. The antitoxin component (Dehre_0667) was however never detected in the proteome. The direct vicinity of Dehre_0668 displays several phage- or plasmid-related genes, suggesting that Dehre_0667-0668 could have been acquired by horizontal gene transfer and represent a phage-like defence mechanism [153].

Diversity and composition of reductive dehalogenase gene clusters in *D. restrictus*

Multiple rdh gene clusters in *D. restrictus*. A thorough analysis of the *D. restrictus* genome [24] has revealed the presence of 25 reductive dehalogenase homologue encoding genes (*rdhA*), among which 20 are in full length, 4 harbour one or several frame-shifts (*rdhA04*, *05*, *13*, and *21*), and one is a partial gene (*rdhA25*) (see Table S6). The biochemically characterized reductive dehalogenase PceA [114] is encoded by *rdhA24*. While most *rdhA* genes are grouped in two genomic regions (*rdhA01-10* and *rdhA13-23*), a detailed analysis of the genetic structure around them allowed defining 13 clusters consisting of one to six *rdhA* surrounded by genes encoded on the same strand. It is however rather unlikely that these clusters represent actual operons as several rho-independent transcription terminators were predicted within the clusters (Fig. 2). Three general *rdh* genetic organizations can be considered here. Together with the well-characterized *pceABCT* cluster (*rdhA24*), two other *rdhA* are embedded in a similar configuration (*rdhA20* and *-22*), albeit harbouring an additional *rdhK* subunit at the 3'-end. Seven *rdhA* genes are accompanied by *rdhB* and *rdhC* subunits, five of them in the orientation *rdhABC* (*rdhA02*, *-05*, *-06*, *-13*, and *-17*) and two as *rdhBAC* (*rdhA14* and *-21*). Finally the remaining *rdhA* subunits

are only accompanied by their respective *B* subunit exclusively in the orientation *rdhBA*. Most of *rdh* gene clusters are also associated with one *rdhK* subunit in various orientations. The *rdhK* encoded proteins clearly belong to the large family CRP/Fnr regulatory proteins from which CprK members of *D. dehalogenans* and *D. hafniense* DCB-2 were extensively studied and represents the paradigmatic DNA-binding regulatory protein for the respective chlorophenol reductive dehalogenase (*cpr*) operons [66, 67, 71, 73, 154-158]. Screening of the genome of *D. restrictus* for RdhK proteins encoding genes revealed 25 paralogues from which 22 are located within the 13 *rdh* gene clusters, and the remaining 3 in their direct vicinity. This strongly suggests that RdhK are regulatory proteins dedicated to OHR metabolism.

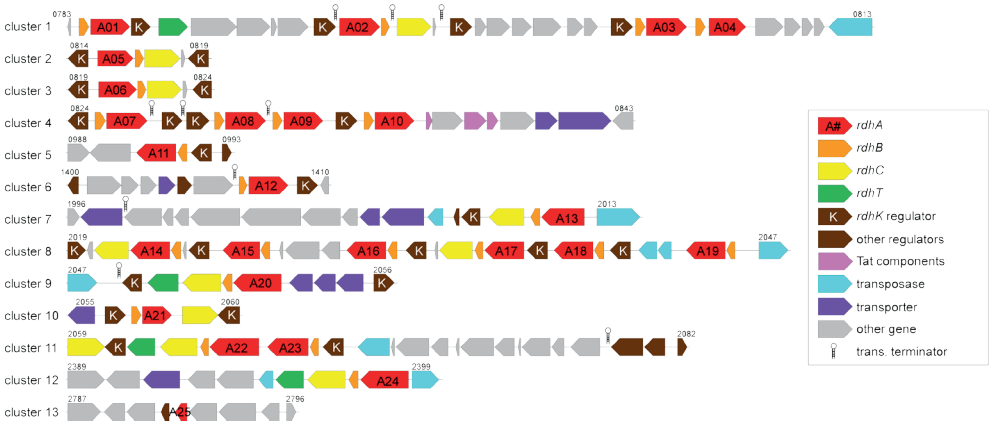


Figure 2. Genetic map of *D. restrictus* gene clusters containing reductive dehalogenase genes (*rdhA*, red numbered arrows). For each *rdh* cluster, all the genes present on the same DNA strand were considered together with the direct flanking genes in opposite orientation. The numbers indicated above each cluster are the corresponding loci in *D. restrictus* genome (Dehre_#).

Diversity of *D. restrictus* RdhA proteins

Protein sequence alignment of RdhA subunits of *D. restrictus* with selected sequences from other OHRB revealed several interesting features (Fig. 3). Firstly, a strong correlation could be established between the level of sequence identity (see also Table S7) and the genetic organization of the predicted *rdh* operons. Indeed the dominating group of 14 RdhA proteins encoded by minimal *rdhBA* operons forms a separate branch, which also contains the well-characterized chlorophenol reductive dehalogenase (CprA) of *D. dehalogenans*. All three *rdhABCT* predicted operons in *D. restrictus* also cluster together, however, with homology to enzymes with different substrate specificities. PceA (RdhA24) is highly similar to other PceA enzymes from members of the closely related genus *Desulfitobacterium*, but also highly similar (88% sequence identity) to DcaA of *D. dichloroeliminans,* as already reported [117, 159]. In contrast, both RdhA20 and -22 of *D. restrictus* have a rather strong sequence identity with CprA5 and RdhA3 of *D. hafniense* strain PCP-1 and strain DCB-2, respectively, which have been shown to use 3,5-dichlorophenol [26, 160], these two latter enzymes being encoded in a similar genetic structure (*rdhABCT*). Interestingly two pairs of RdhA proteins (RdhA03 with -04; RdhA16 with -19) show a very high level of sequence identity (Table S7). Another striking feature is the high conservation degree of RdhA proteins between *D. restrictus* and the newly available RdhA sequences identified in the metagenome of the β-HCH dechlorinating co-culture containing *Dehalobacter* sp. E1 (*Dhb*E1 in Fig. 3) [22]. Indeed 5 out of 9 *Dhb*E1 proteins have identical counterparts in *D. restrictus* (99-100% sequence identity), while 3 RdhA have highly similar homologues (70-92% identity) in *D. restrictus*. One last sequence (*Dhb*E1_1222) is partial.

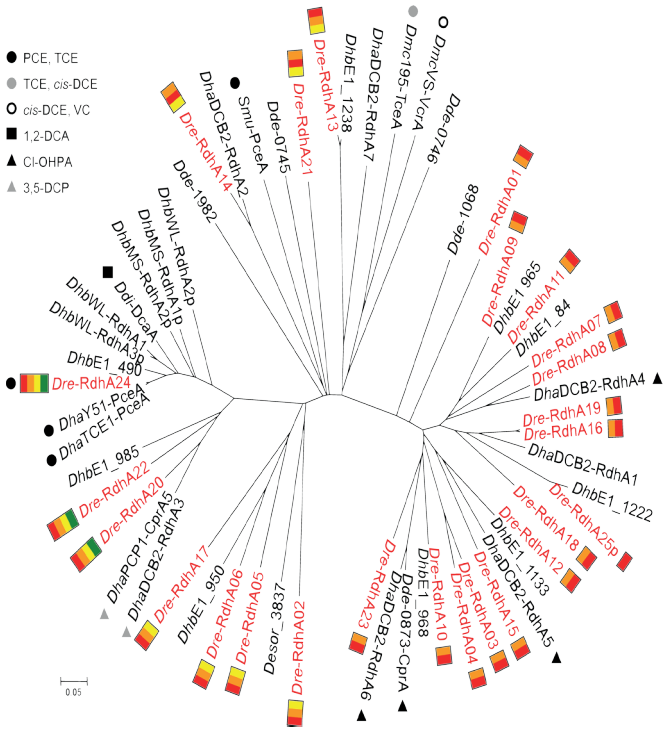


Figure 3. Diversity analysis of *D. restrictus* RdhA proteins. All *D. restrictus* RdhA proteins are indicated in red together with their genetic structure: *rdhA* (red box), *rdhB* (orange), *rdhC* (yellow), *rdhT* (green). Protein sequences were aligned with selected RdhA proteins from other OHRB. When known, the corresponding substrates are also indicated. Dre, *Dehalobacter restrictus*; Dhb, *Dehalobacter* spp. (strains E1, MS and WL); Dha, *Desulfitobacterium hafniense* (strains DCB-2, PCP-1, TCE1 and Y51); Dde, *Desulfitobacterium dehalogenans*; Ddi, *Desulfitobacterium dichloroeliminans*; Desor, *Desulfosporosinus orientis*; Dmc, *Dehalococcoides mccartyi* (strains195 and VS); Smu, *Sulfurospirillum multivorans*

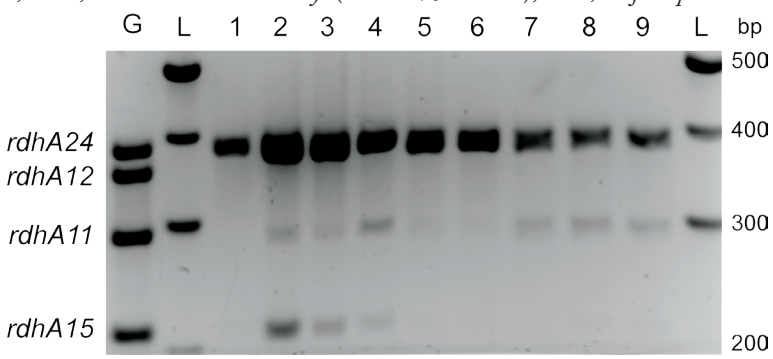


Figure 4. Growth phase dependent transcription level of *rdhA* genes in *D. restrictus* analyzed by RT-multiplex PCR. One multiplex PCR result is depicted as illustration. The targeted genes are indicated on the left of the gel. G: positive control on genomic DNA; L: 100 bp ladder from which the corresponding fragment sizes are indicated on the right. Samples 1-3, 4-6, and 7-9 were taken from triplicate cultures harvested in exponential, late exponential and stationary phases, respectively.

Transcriptomic analysis of *D. restrictus* reductive dehalogenase genes

Screening of rdhA gene transcription by RT-multiplex PCR

From the global proteomic analysis, only two RdhA proteins were clearly detected: the main PCE reductive dehalogenase (PceA) and RdhA14, albeit at a much lower abundance. A specific approach was then conducted in order to evaluate the transcriptional level of the 24 full-length *rdhA* genes in *D. restrictus* along the growth phases. First a RT-multiplex PCR method was developed allowing screening groups of *rdhA* genes at mRNA level in the triplicate cultures collected at the exponential (E), late-exponential (LE) and stationary (S) growth phases. Fig. 4 illustrates the qualitative data obtained for a combination of four *rdhA* genes using that method (the complete set of data is presented in Fig. S3). Five *rdhA* gene transcripts (*rdhA08*, *-14*, *-16*, *-19*, and *-24*) were strongly amplified, however, showing various transcription levels. The *pceA* gene (*rdhA24*) was clearly dominant and was still detected in the RNA samples collected in stationary phase (Fig. 4 and S2). Those five *rdhA* genes were further analyzed by RT quantitative PCR.

Quantitative assessment of selected *rdhA* gene transcription by RT-qPCR. Based on individual standards for each target gene, transcript copy numbers per µl of cDNA samples were measured for *rdhA08*, *-14*, *-16*, *-19* and *-24* (*pceA*) along with *rpoB* (Dehre_0495), which was chosen as a constitutively expressed housekeeping gene (Fig. 5, see Table S2 for qPCR parameters). A decrease in transcription level was generally observed for all genes along the growth phases, some of them dropping below the detection limit of the method applied. These data confirmed the trend observed by the qualitative multiplex PCR approach. In the exponential phase, the *pceA* gene (*rdhA24*) was highly transcribed in comparison with all other genes considered (between 51- and 3688-fold, depending on the gene, see Table S8 for details). Although decreasing, *pceA* remained strongly transcribed even at stationary phase. The level of transcription of the remaining *rdhA* genes decreased with the following order: *rdhA19* > *rdhA14* >> *rdhA16* > *rdhA08*.

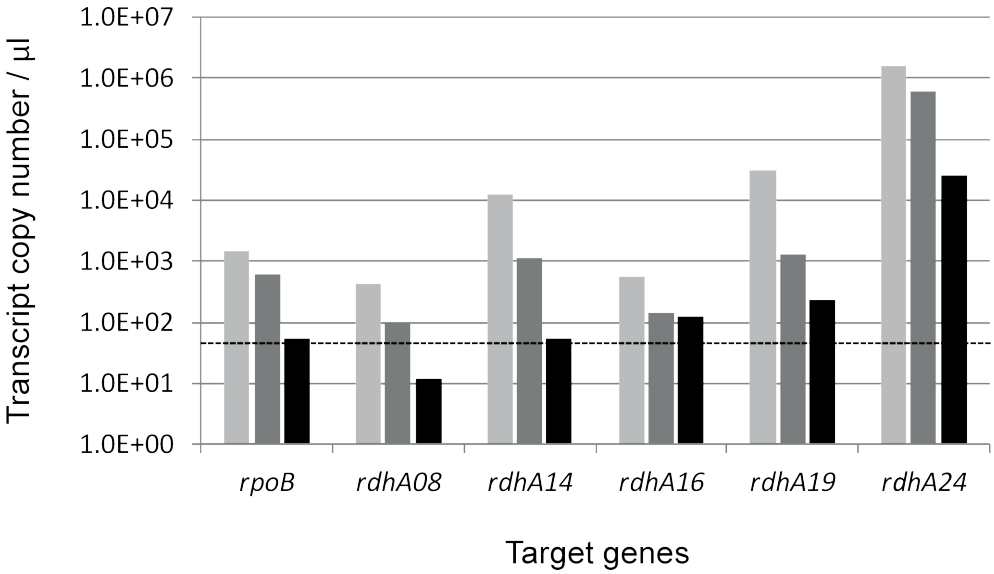


Figure 5. Growth phase dependent transcription level of selected *rdhA* genes in *D. restrictus* analyzed by RT-qPCR. The graph depicts the gene copy number of selected *rdhA* genes along with *rpoB* as control obtained from one culture replicate harvested in exponential (light grey), late exponential (dark grey), and stationary (black) phases, respectively. The same trend was observed for all replicates. Standard deviation of qPCR replicates was below 15% of the measured data. The dotted line (50 copies / ml) displays the lower detection limit that was generally considered for the data obtained.

Discussion

Although *D. restrictus* was among the first OHRB to be isolated, a significant part of its metabolism remained largely unresolved, mainly due to the lack of the genome sequence, but also due to the restricted conditions in which this bacterium has been found to grow, namely exclusively by anaerobic respiration with hydrogen as electron acceptor and PCE or TCE as unique terminal electron acceptors. We recently obtained the genome sequence of *D. restrictus* strain PER-K23 [24] which allowed us in the present study to consider general questions about its metabolism and a more specific investigation line focusing on the key players in OHR, the reductive dehalogenases.

The 2.9 Mb genome of *D. restrictus* can be considered to occupy an intermediate position among OHRB between the reduced genome size of the OHR obligate *Dehalococcoides* genus (~1.4 Mb) and the largely redundant genomes of the versatile *Desulfitobacterium* genus (> 5 Mb). Metabolically, however, *D. restrictus* is closer to *Dehalococcoides*, suggesting that, besides additional genetic information responsible for peptidoglycan synthesis and motility, some parts of *D. restrictus* genome may be not functional or encode for yet unsuspected metabolic pathways. A remarkable example is the presence of a complete cobalamin biosynthetic pathway, an essential cofactor for OHR metabolism. Indeed based on the anaerobic pathway described by Roessner [161] and the cobinamide salvaging pathway studied by Escalante-Semerena and co-workers [149, 162-164], all genes were clearly identified in *D. restrictus*, although it cannot grow without a supply of vitamin B_{12} in the medium ([12]; J. Maillard, unpublished data). The proteomic data obtained here showed that about half of the proteins of the biosynthetic pathway mostly from the upper pathway were not detected, indicating that under the growth conditions applied, *D. restrictus* used the corrinoid amended and possibly modified it according to its needs. On the genetic level, the frame-shift mutation observed in *cbiH* (Dehre_2856) needs to be confirmed, but could also be a reason why *D. restrictus* is not able to synthesize cobalamin *de novo*. Preliminary proteomic data obtained from cells which were partially depleted of vitamin B_{12} revealed that the production of corrinoid transporters and proteins of the salvaging pathway increased significantly rather than the biosynthetic proteins (J. Maillard & T. Kruse, unpublished data), suggesting that the biosynthetic pathway is not functional in *D. restrictus*.

Enzymes belonging to the Wood-Ljungdahl (WL) pathway for CO_2 fixation were clearly detected on the proteomic level in *D. restrictus*. A significant increase in the carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH/ACS) complex (Dehre_0143-8, corresponding roughly to the carbonyl branch of the pathway) was observed in the late-exponential and stationary phases when compared to the exponential phase. This suggested that acetate might be depleted in the medium already during the late-exponential phase, and that *D. restrictus* could partially assimilate CO_2 via the acetyl-CoA synthase. This is however contrasting with a previous data set where heterotrophic CO_2 assimilation (probably via the pyruvate:ferredoxin oxidoreductase, PFOR) has been postulated for an enrichment of *D. restrictus* [83]. Several homologues of the latter enzyme were also detected in the proteome (Table S5). of CO_2 have been questioned for other OHRB such as various isolates of *Desulfitobacterium* [19, 26, 130, 165] or of *Dehalococcoides* [121, 122, 166]. In *D. hafniense* strains, components of the WL pathway have been shown

to participate in the use of phenyl methyl ethers as electron donors [165]. There, however, the methyl branch was mainly used together with *O*-demethylases. For strain TCE1 it has been reported that components of this pathway increased in abundance when H_2 and PCE were used as a combination of electron donor and acceptor [130]. The work by Tang and co-workers suggested that the WL pathway was not involved in CO_2 fixation in *Dehalococcoides* [122]. As illustrated by these examples, the role of the WL pathway in OHRB might be diverse, and further dedicated experiments are required to fully understand why *D. restrictus* recruits it at late growth phases.

The presence of eight different hydrogenases underscores the central role of hydrogen in the metabolism of *D. restrictus*. The genomes of *Desulfitobacterium hafniense* DCB-2 [26] and Y51 [19] encode like *D. restrictus* three Hup-type hydrogenases. The fact that we only detected one of them (Dehre_0551-0553) at relatively constant abundance across growth phases indicates that (i) the three Hup complexes have different roles in the metabolism, and (ii) the detected Hup plays a role in the core metabolism. Concerning the three Fe-only hydrogenases [146], two complexes were present in stable abundance at all growth phases, whereas one was not observed at all. Unlike what is predicted for Hym in *Dehalococcoides mccartyi* 195 [18, 120], we did not find any membrane-associated components in the Hym-type hydrogenases in *D. restrictus*. We therefore suggest that these enzymes are located in the cytoplasm, where they might be involved in generating reducing equivalents (e.g. NADH, FADH) for biosynthetic reactions or maybe directly in generating a proton motive force with respiration complex I, as speculatively indicated in Figure 1. The respiration complex I in *D. restrictus* lacks the NuoEFG subunits that usually receive electrons from NADH. The electron donor for this type of respiration complex I is not yet known, but it has been speculated that they act as a docking station able to receive electrons from various electron donors [167]. It is interesting that we found two large membrane-bound putatively proton-translocating hydrogenase complexes, Hyc and Ech in *D. restrictus* like in *D. mccartyi* 195, whereas the more closely related *Desulfitobacterium hafniense* Y51 and DCB-2 only contain a Hyc homolog [18, 19, 26]. The role of these remains unclear, however, disrupting *hyc* in *Desulfitobacterium dehalogenans* resulted in loss of ability to utilize *3*-chloro-*4*-hydroxyphenylacetic acid and nitrate as electron acceptor when formate was used as electron donor, suggesting a role in the electron transport chain [119]. It has been suggested that Ech and Hyc may play a role in generating low potential electrons for OHR by reverse electron flow. It was however observed that the expression of both Hyc and Ech decreased when *D. mccartyi* 195 was cultivated under lower partial pressure of hydrogen. Since hydrogen is a stronger reductant at higher

partial pressure, the opposite would have been expected if they played a role in reverse electron flow [120]. Our findings suggest that different hydrogenases play specific and central roles in the metabolism of *D. restrictus*, but elucidating the exact role of the individual hydrogenases requires further studies.

Significant changes in the protein content between exponential and stationary phases were observed for various unrelated proteins for which the predicted function was often not clear. For example, $ComF_B$ (Dehre_1215) and a cupin-domain containing protein (Dehre_0983) were identified with more than 1000-fold increase in the stationary *vs.* exponential phase (Table 1). The former protein is predicted to play a role in the late development of competence, although no other competence protein was detected. Competence represents a general strategy for bacteria to survive in unfavourable conditions such as during stationary phase [168]. The latter protein has no clear predicted function, but might be part of an operon involved in the shikimate pathway responsible for the biosynthesis of aromatic amino acids. The list of proteins that in/decreased significantly between exponential and stationary phase clearly indicates that the cells are adjusting their metabolism when shifting from one growth phase to another.

The discovery of 25 reductive dehalogenase (*rdhA*) genes in the genome of *D. restrictus* was surprising given its currently known substrate range for reductive dehalogenation[24], but is in line with what has been observed in all available genomes of *Dehalococcoides mccartyi*. The detailed analysis of the *rdh* gene clusters we present here, together with the transcriptional and proteomic data on the components of these clusters, helped us to consider their diversity, evolution and function in *D. restrictus* during growth on PCE. Analysis of the sequence similarity of *D. restrictus* RdhA proteins along with the best characterized RdhA proteins revealed at least three groups of enzymes. The largest and relatively deep-branching first group contains 16 RdhA proteins which are affiliated to the characterized *ortho*-chlorophenol dechlorinating enzymes (CprA) of *Desulfitobacterium* isolates [64, 169]. Within this group, two gene duplication events must have occurred recently, as the couples RdhA16/19, and RdhA03/04 show 98 and 81% sequence identity, respectively (Table S7). Further synteny analysis revealed that the sequence conservation was extended to the corresponding *rdhB* genes (data not shown). A second deep-branching group of RdhA sequences contains *D. restrictus* PceA (RdhA24) and two slightly more distant members (RdhA20 and -22). These proteins form a closely related family together with some of the best characterized enzymes, namely PceA of several *D. hafniense* isolates [78, 116, 170], DcaA of *D. dichloroeliminans* [159] and CprA5 (dechlorinating 3,5-dichlorophenol) of *D. hafniense* PCP-1

[171]. The last 7 RdhA proteins build up a group of highly heterogeneous enzymes for which no characterized counterpart is yet available. Among them however, four *D. restrictus* RdhA proteins (RdhA02, -05, -06 and -17) show 45% sequence identity with a putative RdhA identified in the genome of *Desulfosporosinus orientis* (Desor_3837, [172]). The genetic organization around *rdhA* genes is tightly correlated with the sequence diversity of their encoded proteins. Indeed both deep-branching groups of *D. restrictus* RdhA show uniform genetic structures, *rdhBA* and *rdhABCT*, respectively. The rather heterogeneous third group is made of either *rdhABC* or *rdhBAC* operons. This strongly indicates a possible evolutionary line in which a few individual *rdh* operons might have been acquired by horizontal gene transfer, followed by several rounds of gene duplication.

Functional investigation of the *rdh* gene clusters along the growth curve of *D. restrictus* on PCE clearly revealed that the PCE reductive dehalogenase (PceA, RdhA24) was dominating both at transcriptional and proteomic levels, with only little changes along the growth phases. This can explain why only PceA could be purified from *D. restrictus* in earlier studies. On proteomic level RdhA14 was the only other reductive dehalogenase detected but at an estimated PceA/RdhA14 ratio of 212 during exponential phase. While all subunits encoded by the *pceABCT* operon were identified, neither RdhB nor RdhC belonging to the *rdhBAC14* operon were detected, possibly as a result of their lower expression and high hydrophobicity. On transcriptional level, the results are somehow contrasting. While *rdhA14* was also detected at a copy number ratio similar to the proteomic data (Table S8, B), other *rdhA* genes were also significantly transcribed, and among them *rdhA19* at a slightly higher level than *rdhA14*, although not detected in the proteome. Whether this is due to the sensitivity of the proteomic analysis or to a possible post-transcriptional regulation remains to be investigated. Similar to several omics studies on *Dehalococcoides* [120, 133, 135, 136, 173, 174], a relatively tight regulation seems to operate in *D. restrictus* for *rdhA* candidates, among which only a few of them are steadily expressed. In contrast to *Dehalococcoides* however, where mostly two-component systems and MarR-type regulators are likely to regulate the expression of *rdhA* genes [17], in *D. restrictus*, as well as in the closely related *Desulfitobacterium* isolates, numerous CprK activating regulators (so-called RdhK) are present in or in the direct vicinity of *rdh* gene clusters. Only two of them however were detected in the proteomic analysis (Dehre_2025 and 2048), suggesting that their expression level remains low in the cell or that they are themselves regulated.

Additional proteins encoded in *rdh* gene clusters were also detected in the proteome. The TatA and TatB components of two Tat systems (Dehre_0836, 0839 and 1843) were detected. Interestingly the former system is encoded directly downstream of *rdhA10* and surprisingly contains an ApbE homologue (Dehre_0837) involved in thiamine biosynthesis. Also possibly linked to the translocation of RdhA proteins across the cytoplasmic membrane, a SppA homologue (Dehre_0809) was detected. The corresponding gene is located directly downstream of *rdhA04* and its product is possibly involved in the degradation of signal peptides (such as the Tat signal peptides of RdhA proteins) after they have been cleaved from the mature proteins [175, 176].

Our multi-level study of *D. restrictus* metabolism revealed rather elaborate genomic and proteomic features despite its restricted physiology recognized so far, suggesting that there is much more to discover especially in the energy metabolism of this bacterium. In addition, the high number of reductive dehalogenase genes raises the question of a wider bioremediation spectrum via organohalide respiration for *D. restrictus*.

Acknowledgements

The Netherlands Genomics Initiative as well as the European Community's Seventh Framework Programme (FP7/ 2007-2013) are acknowledged for support to TK and HS through the Ecogenomics and BACSIN project (grant agreement No. 211684). Sequencing and annotation of the *D. restrictus* genome was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396. The Swiss National Science Foundation (SNSF) is acknowledged for support to AR, JM and CH in frame of the SNF project No. 31003A_138114.

Chapter 4

Functional genomics of corrinoid starvation in the organohalide-respiring bacterium *Dehalobacter restrictus* strain PER-K23

Aamani Rupakula[#], Yue Lu[#], **Thomas Kruse,** Sjef Boeren, Christof Holliger, Hauke Smidt, Julien Maillard

# These authors contributed equally to the work.

Supplementary materials are available through the online version of this manuscript

Published as: Rupakula A, Lu Y, Kruse T, Boeren S, Holliger C, Smidt H, Maillard J (2015) Functional genomics of corrinoid starvation in the organohalide-respiring bacterium Dehalobacter restrictus strain PER-K23. Front Microbiol 5. doi:10.3389/fmicb.2014.00761

Abstract

De novo corrinoid biosynthesis represents one of the most complicated metabolic pathways in nature. Organohalide-respiring bacteria (OHRB) have developed different strategies to deal with their need of corrinoid, as it is an essential cofactor of reductive dehalogenases, the key enzymes in OHR metabolism.

In contrast to *Dehalococcoides mccartyi*, the genome of *Dehalobacter restrictus* strain PER-K23 contains a complete set of corrinoid biosynthetic genes, of which *cbiH* appears to be truncated and therefore non-functional, possibly explaining the corrinoid auxotrophy of this obligate OHRB. Comparative genomics within *Dehalobacter* spp. revealed that one (operon-2) of the five distinct corrinoid biosynthesis associated operons present in the genome of *D. restrictus* appeared to be present only in that particular strain, which encodes multiple members of corrinoid transporters and salvaging enzymes. Operon-2 was highly up-regulated upon corrinoid starvation both at the transcriptional (346-fold) and proteomic level (46-fold on average), in line with the presence of an upstream cobalamin riboswitch. Together, these data highlight the importance of this operon in corrinoid homeostasis in *D. restrictus* and the augmented salvaging strategy this bacterium adopted to cope with the need for this essential cofactor.

Introduction

Corrinoids are essential cofactors for a wide variety of enzymes that facilitate reactions including rearrangements, methyl group transfers, and reductive dehalogenation [177]. A recent bioinformatic study has revealed that while 76% of 540 sequenced bacterial genomes contain corrinoid-dependent enzymes, only 39% of these genomes encode the complete corrinoid biosynthesis pathway, suggesting that the salvage of corrinoids from the environment is an important process for many bacteria [33]. Both aerobic and anaerobic corrinoid biosynthesis pathways have been described showing few but significant differences, notably in tetrapyrrole ring contraction and the step at which cobalt is inserted into the ring [178, 179]. This pathway is complex and consists of approximately 30 reactions (see [178] for a recent review).

Organohalide respiration (OHR) is an anaerobic bacterial respiration process of environmental interest, as many anthropogenic halogenated organic compounds can be used as terminal electron acceptors by organohalide-respiring bacteria (OHRB) [5]. OHRB are capable of removing halogens and therefore contribute to bioremediation of environments polluted with these compounds [3]. The key enzyme in OHR is the reductive dehalogenase (RDase) [57], which strictly depends on corrinoid cofactors for the dehalogenation reaction. Although the reaction mechanism has not yet been fully understood, RDases represent a particular family of corrinoid enzymes as they catalyze electron transfer rather than methyl transfer. Moreover the absence of a corrinoid binding motif in RDase sequences reflects the base-off/his-off conformation of the corrinoid in the enzyme [64, 180]. In recent years, corrinoid biosynthesis and salvaging in OHRB regained substantial interest in the scientific community as exemplified by the following studies: an unusual corrinoid cofactor (norpseudo-B_{12}) has been identified in the tetrachloroethene (PCE) RDase of *Sulfurospirillum multivorans* [60]; the lack of exogenous corrinoid had an effect on the RDase activity of *Desulfitobacterium hafniense* when cultivated with an alternative electron acceptor [56]; many essential corrinoid biosynthetic genes have been found on a plasmid in *Geobacter lovleyi* [28]; the involvement of the bacterial community accompanying members of *Dehalococcoides mccartyi* for corrinoid supply has been highlighted [63, 181-183]; modifying the lower ligand of the corrinoid had a severe effect on the activity of the PCE RDase of *S. multivorans* [62].

Contrasting situations have been observed regarding the ability of OHRB to produce corrinoid cofactors *de novo*. Both genome analysis and physiological studies have shown that the obligate OHR *D. mccartyi* is strictly dependent on exogenous corrinoid supply and that 5,6-dimethylbenzimidazole can serve as nucleotide loop in corrinoid cofactors [11, 61, 182-184]. On the contrary, the facultative OHRB *S. multivorans* strain K and *D. hafniense* strains encode the full corrinoid biosynthetic pathway in their genome and have been shown to grow without any supply of corrinoid in the medium[19, 26, 29, 55].

Dehalobacter restrictus strain PER-K23 is an obligate OHRB only able to grow by dechlorinating tetra- and trichloroethene (PCE and TCE, respectively). It was first isolated from Rhine river sediment and since then always cultivated in the presence of exogenous vitamin B_{12} (cyanocobalamin) [12]. The PCE RDase (PceA) of *D. restrictus* has been extensively studied and revealed a 60-kDa enzyme containing a corrinoid cofactor and two 4Fe-4S clusters with estimated redox potential of -350 mV ($Co^{1+/2+}$) and -480 mV ($4Fe\text{-}4S^{2+/1+}$), respectively, and a specific dechlorination activity of 250 nkat/mg [114, 180]. Spectrophotometric analysis of the corrinoids extracted from *D. restrictus* PceA with cyanide has shown a spectrum resembling the one of cyanocobalamin [180], although this method does not allow identifying corrinoid unambiguously. Analysis of the newly published genome of *D. restrictus* revealed the presence of a complete set of corrinoid biosynthetic genes where one gene, *cbiH*, is truncated due to a 101-bp deletion, likely responsible for the corrinoid auxotrophy of *D. restrictus* [24, 115] (Figure 1).

This present study aims to explore in detail the effect of corrinoid starvation on *D. restrictus* with a combination of comparative genomics, as well as transcription and proteome analysis.

Materials and methods

Bacteria, plasmids and growth conditions

Dehalobacter restrictus strain PER-K23 (DSM 9455) was cultivated as described earlier [12, 114, 115]. Anaerobic serum flasks of 500 mL were supplemented with hydrogen as electron donor, inoculated with 2 % (v/v) inoculum, and finally 1 % (v/v) of 2 M PCE dissolved in hexadecane was added as electron acceptor. Batch cultures of *D. restrictus* were cultivated in 300 mL medium at 30°C under gentle agitation (100 rpm), and chloride release was used as an indicator of growth. Chloride concentration was measured with a Chlor-o-counter (Flohr Instrument, Nieuwegein, The Netherlands) as described earlier [185]. Cultures for proteomic analysis were prepared in triplicate with high, mid and low concentration of cyanocobalamin corresponding to 250, 50 and 10 µg/L, respectively.

Escherichia coli DH5α was cultivated on liquid or solid LB medium containing 100 µg/L ampicillin after transformation with derivatives of the pGEM-T Easy vector (Promega, Duebendorf, Switzerland).

Sequence retrieval and genome analysis

All sequences mentioned in this study were taken from the recently published genome of *D. restrictus* strain PER-K23 [24] and from other *Dehalobacter* spp. genomes including *Dehalobacter* sp. E1 [22], *Dehalobacter* sp. DCA and sp. CF [23], *Dehalobacter* sp. FTH1 (RefSeq PRJNA199134, JGI genome project), *Dehalobacter* sp. UNSWDHB [52]. The original annotation of *D. restrictus* gene loci obtained in collaboration with the Joint Genome Institute (JGI project #402027) was used here as the present study is a follow-up study of two previous reports where the JGI annotation was used [24, 115]. Another version of *D. restrictus* genome was recently annotated by the automatic pipeline of the NCBI database and is available under accession number CP007033. Corresponding loci from both databases are given for the selected corrinoid proteome in Table S1. The annotation of selected genes was verified using a manual search with BLAST [137]. Protein sequences were aligned using ClustalX v.2.0 [138]. Sequence maximum likelihood tree analysis was done with MEGA5 [35]. Cobalamin riboswitches (Cbl-RS) were identified using Rfam [186] and initially aligned using ClustalX and then corrected manually as described earlier [55]. Comparative genome analysis was performed using the Artemis Comparison Tool [187].

Transcription analysis

RNA was extracted using the TRIzol method according to [139] with the following modification. The DNaseI treatment was stopped by adding the DNase stop solution and incubating for 10 min at 65°C. RNA concentration was estimated using the Nanodrop ND-1000 spectrophotometer (Thermo Scientific, Ecublens, Switzerland). Reverse transcription was performed as described in [115].

Primers targeting each gene present immediately downstream of the five Cbl-RS of *D. restrictus* were designed. PCRs and cloning using the pGEM-T Easy vector, cloning, clone selection, sequencing and quantitative PCR were performed as described earlier [115]. Primer sequences, amplicon sizes, plasmids and quantitative PCR specifications are given in Table S2.

Protein extraction and SDS-PAGE

Cells were harvested by 10 min centrifugation at 12000 × *g*, washed twice with 25 mL 20 mM Tris–HCl (pH 7.5), and then flash-frozen in liquid nitrogen. All biomass samples were stored at -80°C until use. Cell pellets were resuspended in 0.5 mL lysis buffer (100 mM Tris/HCl, pH 7.5, 4% sodium dodecyl sulfate, and 0.1 M dithiothreitol) and then transferred to 2-mL protein LoBind tubes (Eppendorf, Hamburg, Germany). Protein extraction was done as described earlier [115]. Protein concentration was determined with the Qubit® protein assay kit (Invitrogen, Eugene, OR, USA) following the manufacturer's instructions. Protein samples were stored at -20°C until use. SDS-PAGE was done following standard procedures [188]. In brief, 15 µg of proteins from each sample were loaded in separate lanes in gels containing 10% SDS. Gels were stained with Coomassie brilliant blue R250 (Merck, Darmstadt, Germany) and scanned using a GS-800 calibrated densitometer (Bio-Rad, Hercules, CA, USA). The Quantity One basic software package was used to quantify the intensity of lanes. Series of gels were prepared and analyzed until less than 5% differences in the intensity between any lanes were achieved.

Gel digestion and peptides purification

In-gel digestion of proteins and purification of peptides were done following a modified version of the protocol described earlier [115]. Disulphide bridges in proteins were reduced by covering whole gels with reducing solution (10 mM dithiothreitol, pH 7.6, in 50 mM NH_4HCO_3), and the gels were incubated at 60°C for 1 h. Alkylation was performed for 1 h by adding 25 mL of

iodoacetamide solution (10 mM iodoacetamide in 100 mM Tris-HCl, pH 8.0). Gels were thoroughly rinsed with dd H_2O water in between steps. Each lane of SDS-PAGE gels was cut into three equally sized slices, and each slice was cut into approximately 1 mm^3 cubes and transferred to separate 0.5 mL protein LoBind tubes (Eppendorf, Hamburg, Germany). Enzymatic digestion was done by adding 50 µL of trypsin solution (5 ng/µL trypsin in 50 mM NH_4HCO_3) to each tube, and incubating at room temperature overnight with gentle shaking. Extraction of peptides was performed with manual sonication in an ultrasonic water bath for 1 s before the supernatant was transferred to a clean protein LoBind tube. Additional peptides were recovered by adding 25 µL of 2.5% (v/v) trifluoroacetic acid to the gel pieces, which were sonicated for 2 s before the supernatant was combined with the first supernatant obtained. Peptides were purified with a C18 Empore disk as previously described [141]. Acetonitrile in the samples was removed by using a concentrator vacuum centrifuge. Finally, sample volume was adjusted to 50 µL with 0.1% (v/v) formic acid.

nLC-MS/MS and data analysis

Peptides derived from extracted and digested proteins were analyzed by nLC-MS/MS (Biqualys, Wageningen, The Netherlands) as described earlier [142]. MaxQuant v.1.3.0.5 [143] with default settings for the Andromeda search engine [143] in the label free quantitation mode was used to analyze MS and MS/MS spectra, except that extra variable modifications were set as described before [115]. A protein database of *D. restrictus* was generated from the whole genome sequence [24] using the Artemis genome browser (release 15.0.0). Also, a contaminant database including sequences of common contaminants like trypsin, BSA and human keratins [115, 145] was used. Further filtering and bioinformatics analysis was performed with Perseus software v. 1.3.0.4 as described before [189]. Also, protein groups with a logarithmic label-free quantitation (LFQ) intensity of zero for all treatments were deleted from the MaxQuant result table. Subsequently, remaining Log LFQ zero values were replaced by 5 (slightly below the lowest value measured) in order to make sensible ratio calculations possible. Students T-test was used to identify significant differences in the proteome when comparing logarithmic LFQ values obtained from two culture conditions.

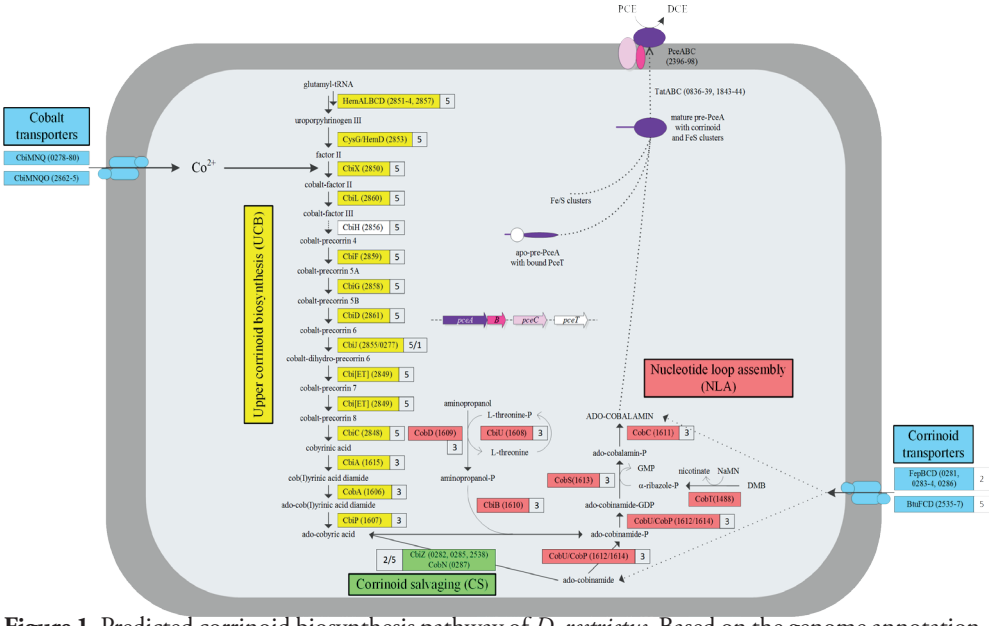


Figure 1. Predicted corrinoid biosynthesis pathway of *D. restrictus*. Based on the genome annotation, a map of the corrinoid biosynthesis was predicted including four pathways: the upper corrinoid biosynthesis (UCB, in yellow), the nucleotide loop assembly (NLA, in red), cobalt and corrinoid transporters (in blue) and corrinoid salvage (in green). The PCE reductive dehalogenase (PceA and associated proteins PceB and PceC) is also depicted. The enzymes catalyzing each reaction are given colour boxes indicating the protein name (when available) and corresponding gene loci (with the Dehre_# prefix removed). CbiH in *D. restrictus* is likely to be non-functional as the corresponding gene is truncated (white box). The operon number to which every locus belongs is indicated besides each protein. Note: the details of CbiH catalysed reaction can be found in [178].

Results

Corrinoids are essential as a growth factor for *D. restrictus* [12]. The corrinoid present in the PCE reductive dehalogenase (PceA) of *D. restrictus* is presumably similar to the type added to the medium, i.e. cobalamin [114]. Detailed analysis of the genome of *D. restrictus* revealed a seemingly complete corrinoid biosynthesis pathway. Compared to other *Dehalobacter* genomes, however, a 101-bp fragment was found to be missing in the *cbiH* gene of *D. restrictus* [24, 115] (Figure S1). The present study aimed specifically at obtaining a broader understanding of the corrinoid metabolism in *D. restrictus*.

Growth of *D. restrictus* under corrinoid-limiting conditions

The full corrinoid biosynthetic pathway was described earlier [115]. A modified and extended version of it is depicted in Figure 1. Briefly, the pathway can be divided in two branches, namely the upper corrinoid biosynthesis (UCB) and the nucleotide loop assembly (NLA), which are connected at the level of ado-cobyric acid.

In the present study, batch cultures were cultivated with addition of 250 µg/L cyanocobalamin to the growth media. An experiment was performed to assess the effects of lowering the initial corrinoid concentration in the medium (250, 50, 10, 1 µg/L and no corrinoid) on dechlorination, which for this obligate OHRB is also a good estimation for growth (Figure 2) [12]. The extent of PCE dechlorination was the same in cultures provided with 50 or 250 µg/L corrinoid demonstrating that the former was enough to reach the maximum dechlorination capacity. In contrast, the chloride release was only half of the maximum in cultures supplemented with 10 µg/L corrinoid, implying that availability of corrinoids was a limiting factor. Further lowering the corrinoid concentration to 1 or 0 µg/L resulted in negligible levels of dechlorination, and

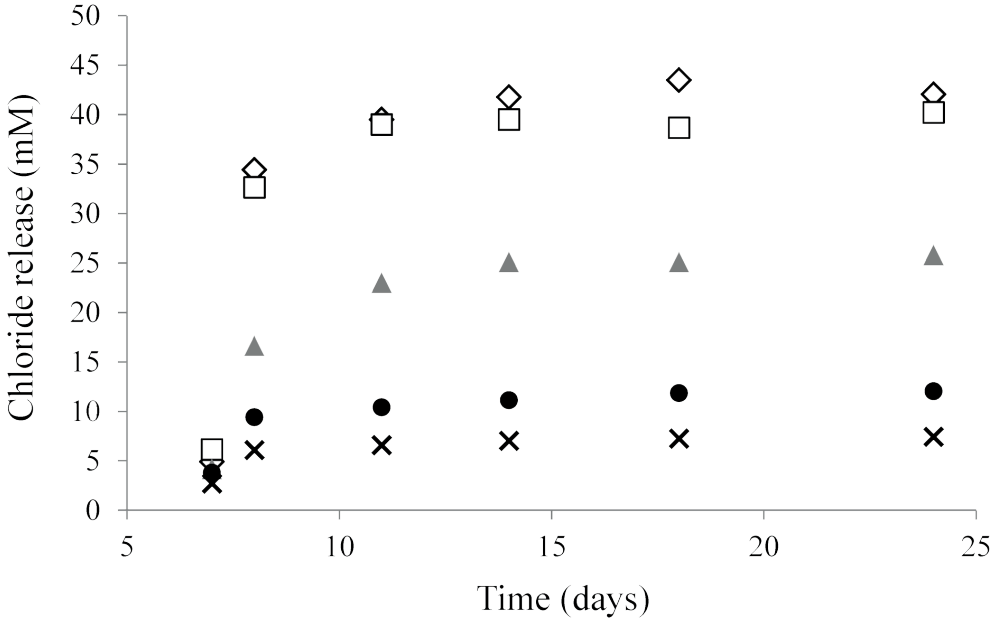


Figure 2. Corrinoid starvation effect on PCE dechlorination by *D. restrictus* by lowering the initial concentration of corrinoid supplemented into the medium. Corrinoid concentration: 250 µg/L (white diamonds), 50 µg/L (white squares), 10 µg/L (grey triangles), 5 µg/L (black circles) and 1 µg/L (black crosses).

therefore growth was assumed to be abolished in these cultures.

Corrinoid metabolic gene arrangement in *D. restrictus*

The complete corrinoid biosynthesis and uptake pathway is genetically encoded in D. restrictus. These genes can be divided into four functional groups depending on the part of the pathway they encode for (Figure 1). The first group denoted as the upper corrinoid biosynthesis (UCB) pathway genes contains genes required to synthesize ado-cobyric acid. The second group consists of genes required for synthesis and the nucleotide loop assembly (NLA) of corrinoids, and the third functional group comprises the corrinoid salvaging (CS) pathway, i.e. genes involved in remodeling corrinoid intermediates salvaged from the environment into ado-cobyric acid (*cbiZ*

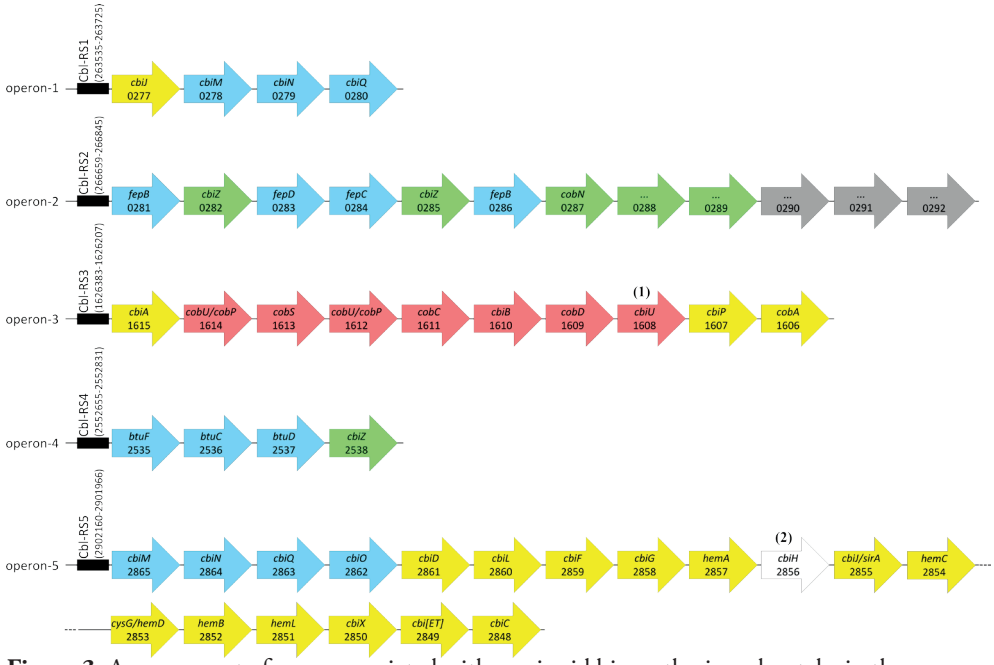


Figure 3. Arrangement of genes associated with corrinoid biosynthesis and uptake in the genome of *D. restrictus*. Genes involved in corrinoid metabolism are grouped in five operons (operon-1 to -5) located in different places of the genome. All five operons are predicted to be regulated by cobalamin riboswitches (Cbl-RS1 to 5, indicated by a black box) located directly upstream of the first gene of each operon. Cbl-RS coordinates are given based on their identification by Rfam. Genes are depicted as arrows in colours corresponding to four distinct pathways (see also Figure 1): upper corrinoid biosynthesis (UCB, in yellow), nucleotide loop assembly (NLA, in red), cobalt and corrinoid transporters (in blue) and corrinoid salvage (in green). Grey arrows depict genes that have no clear function in corrinoid biosynthesis. Gene names and corresponding gene loci (with the Dehre_# prefix removed) are given when available. Notes: [190] *cbiU* was newly annotated as a possible L-threonine kinase encoding gene. (2) a 101-bp deletion makes *cbiH* non-functional in *D. restrictus*.

gene family). The fourth group harbors both cobalt and corrinoid transporter encoding genes (CT).

Most genes associated with corrinoid metabolism are arranged in the genome of *D. restrictus* in five gene clusters (referred to as operon-1 to -5), which are roughly organized according to the function they play in corrinoid biosynthesis (Figure 3). Most proteins involved in the UCB pathway are encoded in operon-5 with the exception of the three last steps that are catalyzed by the product of genes present in operon-3. This latter operon also codes for the enzymes involved in the NLA pathway. Within operon-3, the locus Dehre_1608 was initially annotated as a phosphoglycerate mutase, but shows also sequence similarity with archaeal-type homoserine kinase (with conserved domain TIGR02535) involved in the synthesis of threonine. Here, we propose it could act as an L-threonine kinase (in analogy to PduX in *Salmonella* [191]), which might therefore be involved in the production of aminopropanol-phosphate. No *pduX* homolog could be identified in *D. restrictus*, suggesting that this function is fulfilled by the gene product of Dehre_1608. Hence, we propose to name it *cbiU*. Operon-1 contains a homolog of *cbiJ* (besides the *cbiJ*/*cysG* gene, Dehre_2855, present in the conserved biosynthesis operon-5), and a set of genes coding for the energy-coupling factor-type CbiMNQ cobalt transporter. An additional, albeit different *cbiMNQO* gene cluster is also present at the 5'-end of operon-5 together with the genes for the UCB pathway. Operon-2 harbors a combination of genes coding for transporters (with sequence similarity to FepBCD/BtuCDE ABC-type transporters) likely involved in corrinoid transport, the genes for two different salvaging enzyme (CbiZ) paralogues (Dehre_0282 and _0285), a gene cluster encoding the cobaltochelatase CobN (Dehre_0287), and several subunits of a magnesium chelatase complex. Finally, operon-4 contains a gene cluster coding for an ABC-type corrinoid transporter (BtuFCD) and another copy of *cbiZ* (Dehre_2538). Two additional genes potentially involved in corrinoid biosynthesis (*cobT*/Dehre_1488 and *cobB*/*cobQ*, Dehre_2360) are located elsewhere in the genome and not in one of the five operons.

Comparative genomics of corrinoid operons in *Dehalobacter* spp.

The genome of *D. restrictus* was compared with newly available genomes of *Dehalobacter* spp. strains DCA, CF, E1, FTH1, and UNSWDHB with regard to the organization of corrinoid operons (Table 1 and Table S1). Synteny maps for operon-1 and -2 (Figure S2), and for operon-3, -4 and -5 (Figure S3 to S5, respectively) are given as supplementary material. Operon-1 is con-

served in *D. restrictus* and *Dehalobacter* sp. E1 but absent in all other genomes. Operon-2, which is directly following operon-1 in *D. restrictus*, is lacking in all other *Dehalobacter* spp. for which genome sequences are available to date.

However, a detailed analysis of *Dehalobacter* sp. E1 suggests that operon-2 was lost in that strain as the sequence conservation with *D. restrictus* is extended slightly beyond operon-1 but is readily interrupted within the homolog of Dehre_0281 (the first gene of operon-2 in *D. restrictus*). This deletion in strain E1 includes all of the remaining genes of operon-2, as a 5'-truncated version of Dehre_0297 is again found in strain E1 (Figure S2, panel C). *D. restrictus* operon-3 to -5 are fully conserved in all *Dehalobacter* spp. with the exception of another deletion in the proximal region of operon-5 in strain E1 (Figure S5).

Table 1. Comparative genomics of corrinoid operons in *Dehalobacter* spp.

	Dehalobacter restrictus	**Dehalobacter sp. E1**	**Dehalobacter sp. DCA**	**Dehalobacter sp. CF**	**Dehalobacter sp. FTH1**	**Dehalobacter sp.UNSWDHB**
Operon-1	+	+	-	-	-	-
Operon-2	+	-	-	-	-	-
Operon-3	+	+	+	+	+	+
Operon-4	+	+	+	+	+	+
Operon-5	+	Partial	+	+	+	+
cbiH	Deletion	Intact	Intact	Intact	Intact	Intact

In *Desulfitobacterium hafniense* in contrast, the corrinoid biosynthesis genes are organized in two operons. The major operon (corresponding to DSY4057-4072 in *D. hafniense* strain Y51) encodes proteins of the UCB pathway and part of the NLA pathway, while a 3-gene operon (DSY2114-2116) encodes for the remaining NLA proteins [19, 55]. Some proteins encoded by *D. restrictus* operon-2 have their counterpart in other OHRB. For example, corrinoid transporters are present in most OHRB, however, corrinoid producers such as *Desulfitobacterium hafniense* and *Sulfurospirillum multivorans* do not harbor any *cbiZ* homologous gene. In *Dehalococcoides mccartyi* in contrast, multiple *cbiZ* genes are present in the genomes (Figure S6).

All proteins encoded in operon-2 of *D. restrictus* share between 50 and 77% sequence identity with homologous proteins present in the non-dechlorinating *Firmicute Acetobacterium woodii.* A high level of genetic synteny was further identified between operon-2 of *D. restrictus* and a part of the genome of *A. woodii* (GenBank NC_016894.1, [192]) (Figure S7). Significant sequence similarity of individual proteins of operon-2 was mostly found with homologs of some other members of *Clostridia* and a few *d-Proteobacteria* (data not shown).

Identification of cobalamin riboswitches in *D. restrictus*

Upstream of each of the five corrinoid biosynthesis-related operons in *D. restrictus* a distinct cobalamin riboswitch (Cbl-RS) was identified using Rfam. These five sequences were manually refined in a similar way as done previously for the Cbl-RS sequences of *Desulfitobacterium hafniense* [55]. The alignment of structurally conserved regions of *D. restrictus* riboswitches (Cbl-RS01 to -RS05) was compared to *E. coli btuB* Cbl-RS (Figure S8). In contrast to *E. coli* Cbl-RS, which is regulated at the level of translation [193], all five *D. restrictus* Cbl-RS sequences end with a predicted transcriptional terminator, suggesting that the regulation operates at the level of transcription.

Transcriptional analysis of corrinoid biosynthesis operons in *D. restrictus*

The transcription of genes located directly downstream of the Cbl-RS in D. restrictus was analyzed for cells cultivated in the presence of high (250 µg/L) and low (10 µg/L) corrinoid concentration, and after corrinoid replenishment from low to high concentrations (Figure 4). Quantitative PCR was applied on complementary DNA targeting the first gene located directly downstream of each cobalamin riboswitch. Analysis of corrinoid-starved D. restrictus RNA revealed a higher transcription level of these genes, confirming an active regulation of the respective riboswitches at transcriptional level. Two hours after corrinoid replenishment, transcription of all selected genes was again repressed to the same level as observed under high corrinoid concentration. However, individual responses were significantly different. Indeed, the most pronounced effect was observed for two genes, namely Dehre_0277 (73-fold repression) and _0281 (346-fold), corresponding to the first genes in operon-1 and -2 in D. restrictus, respectively.

Proteome analysis of corrinoid starvation in *D. restrictus*

Comparative whole-proteome analysis was done on *D. restrictus* PER-K23 cells cultivated in the presence of 250 (high), 50 (mid) or 10 (low) µg/L cyanocobalamin. A total of 1195 proteins were detected, corresponding to 42% of the predicted 2826 proteins encoded on the genome [24]. The majority of the detected proteins (1175) were identified in cells from all the tested cyanocobalamin concentrations (Table S4). Normalized LFQ protein intensities were used to compare relative abundances of proteins between different cyanocobalamin treatments. A minimal change of 3-fold in LFQ protein intensity was considered throughout the study. The abundance of 44 proteins showed significant difference ($P<0.01$) between high (250 µg/L) and low (10 µg/L)

corrinoid concentration, and the relative abundance of another 29 proteins showed more than 10-fold changes, albeit not significant due to high variation between triplicates (Figure S9). The results for protein abundance ratios between high and mid, and between mid and low are in the same range (see Table S3).

Proteins associated with cobalamin biosynthesis were further analyzed. A complete *de novo* corrinoid biosynthesis pathway was predicted in the genome of *D. restrictus* starting from glutamyl-tRNA to cobalamin (Figure 1) [24, 115]. All proteins required for biosynthesis of ado-cobyric acid from cobalt-precorrin 5B were identified in proteomic data including CbiD (cobalamin biosynthesis protein, Dehre_2861), an alternative CbiJ (precorrin-6x reductase, Dehre_0277) and Cbi[ET] (precorrin-6Y methyltransferase, Dehre_2849), which were not dectected in a previously analyzed proteome from *D. restrictus* [115]. However, CbiH (precorrin-3B C17-methyltransferase, Dehre_2856), CbiG (cobalamin biosynthesis protein, Dehre_2858) and CbiJ/CysG (precorrin-6x reductase, Dehre_2855) belonging to the UCB pathway and CobS (cobalamin 5'-phosphate synthase, Dehre_1613) of the NLA pathway were not found in the current proteome analysis. The lack of CbiH in the proteome is in line with the observation of a 101-bp deletion in *cbiH* likely leading to a non-functional gene (Figure S1), thus likely to explain why *D. restrictus* requires exogenous corrinoids supply to the growth medium.

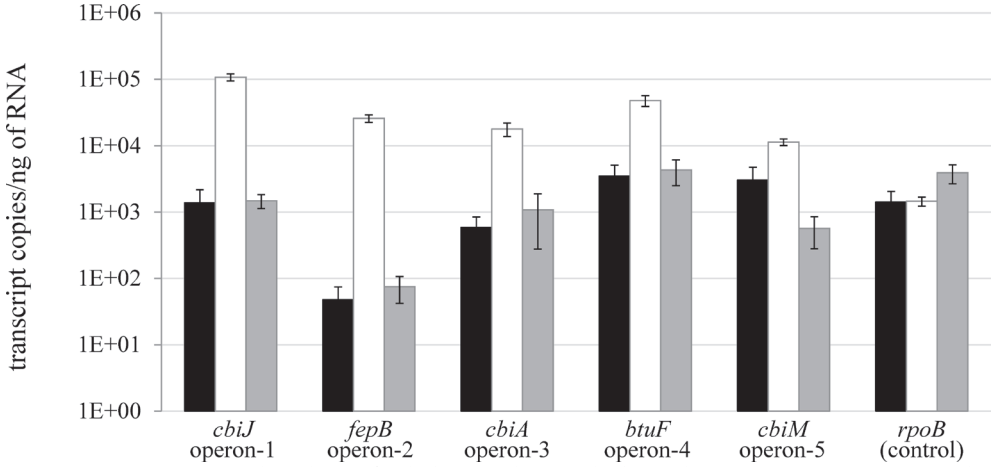


Figure 4. Transcriptional analysis of cobalamin riboswitch-dependent genes in *D. restrictus*. The transcription of the genes located directly downstream of the cobalamin riboswitches in operon-1 to -5 was analyzed. The housekeeping gene *rpoB* was used as unregulated control. The black bars indicate the transcriptional level under corrinoid standard conditions (250µg/L), the white bars under corrinoid starvation conditions (10 µg/L), and the grey bars show the transcriptional level 2 h after replenishment of the latter cultures with 250 µg/L cyanocobalamin. The data show the mean of triplicate cultures with standard deviation.

Previously, the presence of one ABC-type cobalt transporter (Dehre_0850-0852) and two energy-coupling factor-type cobalt transporters (Dehre_0278-0280 and Dehre_2862-2865) was predicted in the genomic study of *D. restrictus* [115]. Here, we identified CbiM (Dehre_0278), CbiQ (Dehre_0280), CbiO (Dehre_2862) and CbiN (Dehre_2864) in the proteome dataset obtained in this study (Table S4).

Relative abundance of proteins associated with corrinoid biosynthesis and salvaging pathways was further analyzed. Interestingly, nearly all proteins associated with corrinoid biosynthesis and salvaging pathways were up-regulated under corrinoid limiting growth conditions (Figure 5 and Table S5). As expected the overall largest change in the abundance of proteins related to corrinoid biosynthesis and salvaging pathways was observed when comparing the proteome of cells cultivated at high *vs*. low concentration of cyanocobalamin (Figure 5A). The corrinoid metabolism differed more strongly when comparing cells cultivated in the presence of high *vs*. mid than mid *vs*. low concentrations (Figure 5B and 5C).

Proteins encoded by operon-2 showed the largest change in protein abundance ratios with on average 46-fold up-regulation when comparing cells cultivated at low *vs*. high corrinoid concentrations (Figure 5A). Operon-2 encodes proteins predicted to be involved in corrinoid salvaging or corrinoid transport. Among these are two predicted CbiZ proteins, amidohydrolases required for salvaging the corrinoid precursor cobinamide, which were up-regulated 80-fold (Dehre_0285) and 58-fold (Dehre_0282) under corrinoid starvation, respectively (Table S4). Furthermore, proteins encoded in operon-1 including an energy-coupling factor-type cobalt transporter (Dehre_0278-0280) which is likely to be involved in the cobalt uptake process, and a precorrin-6x reductase (Dehre_0277), were on average 8-fold up-regulated when comparing cells cultivated in the presence of low *vs*. high corrinoid concentrations. Fewer changes were found for proteins encoded by the three remaining corrinoid-related operons under the different corrinoid conditions.

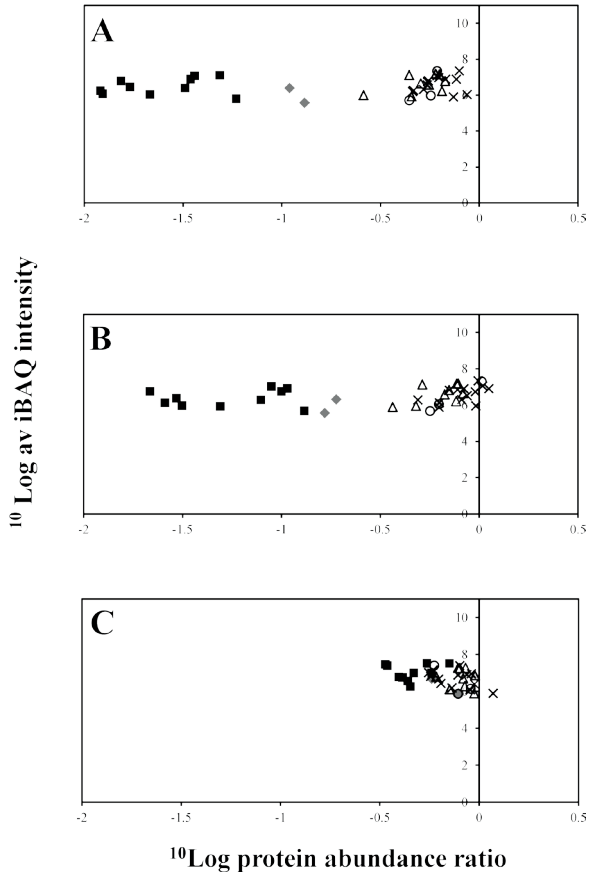


Figure 5. Proteins associated with corrinoid biosynthesis and uptake, extracted from *D. restrictus* cells cultivated in the presence of different cyanocobalamin concentrations. (A) Proteomics analysis of cells cultivated in the presence of high (250 µg/L) *vs*. low (10 µg/L) corrinoid concentrations; (B) high (250 µg/L) *vs*. mid (50 µg/L) corrinoid concentrations; (C) mid (50 µg/L) *vs*. low (10 µg/L) corrinoid concentrations. Proteins encoded by corrinoid operon-1 (grey diamonds), operon-2 (black squares), operon-3 (white triangles), operon-4 (white circles) and operon-5 (black crosses) are shown. Logarithmic average of iBAQ (Intensity based absolute quantitation) value is plotted against the log value of protein abundance ratio based on LFQ value. CbiM (Dehre_0278) was left out because it could not be identified in two replicates of cultures under low corrinoid concentration.

Discussion

Corrinoid biosynthesis of *D. restrictus* and other OHRB

In the present study, ≥ 50 µg/L of cyanocobalamin was required for *D. restrictus* to reach its maximum PCE dechlorination, in line with previous observation that this organism depends on externally supplemented corrinoids [12]. The genome of *D. restrictus* encodes a complete set of corrinoid biosynthesis genes, with the exception of a non-functional *cbiH* gene, suggesting that tetrapyrrole ring contraction does not occur here and represents a dead-end in the biosynthesis pathway. Comparative genomic analysis among other *Dehalobacter* spp. revealed that an intact *cbiH* gene is present in all other genomes. However, little is known about the capacity of other members of this genus to *de novo* synthesize corrinoids since they only have been studied under growth conditions with external addition of cyanocobalamin or in co-cultures [22, 52, 126, 129, 194]. Similarly, strains of *Dehalococcoides mccartyi*, which are also obligate OHRB, are corrinoid-auxotroph [11]. Unlike *D. restrictus*, the corrinoid auxotrophy in *D. mccartyi* strains is due to the lack of the complete biosynthetic pathway. Instead they rely on uptake of extracellular corrinoids via the salvaging pathway and on remodelling of incomplete or non-functional corrinoids in the presence of appropriate free lower ligands, among which 5,6-dimethlybenzimidazole plays a key role [61, 63, 182, 184, 195]. Interestingly, most facultative OHRB such as *Sulfurospirillum multivorans* [60], *Desulfitobacterium hafniense* [19, 55] or *Geobacter lovleyi* [28] are capable of *de novo* biosynthesis of corrinoids.

The genome of *D. restrictus* encodes five well-organized operons containing most of the corrinoid biosynthesis-associated genes. Comparing the genomes of the sequenced *Dehalobacter* spp. revealed that *D. restrictus* harbours an extra set of genes (operon-2) coding for putative corrinoid transporters and salvaging enzymes (CbiZ and cobaltochelatases), suggesting an augmented capacity for corrinoid uptake and remodelling compared to other *Dehalobacter* spp. The importance of *cbiZ* genes in remodelling corrinoids has been already demonstrated for *D. mccartyi* [17, 18, 184]. The role of operon-2 in *D. restrictus* was evidenced by the significant up-regulation of the corresponding enzymes when corrinoid concentration in the medium was lowered. This result clearly showed that *D. restrictus* has developed a particular strategy to cope, at least partially, with its lack of corrinoid biosynthesis under unfavourable corrinoid conditions.

The presence of additional *cbiZ* genes in operon-2 raises the questions of the functional redundancy *vs*. specificity of multiple CbiZ proteins within a single strain, and of the origin of the additional *cbiZ* genes present in *D. restrictus*. While *in vitro* biochemical investigations would be required to answer the first question, a detailed analysis of CbiZ sequence homology (Figure S6) revealed that the two additional CbiZ proteins in *D. restrictus* show a high level of sequence identity with CbiZ homologues present in *A. woodii*, a corrinoid-producing bacterium [196], which has been well-characterized for the Wood-Ljungdahl pathway that also requires corrinoids as an essential cofactor [197]. The high degree of genetic synteny identified between the operon-2 of *D. restrictus* and *A. woodii* suggests that *D. restrictus* - but not the other members of the *Dehalobacter* genus - most probably acquired operon-2 by horizontal gene transfer and successfully exploited this operon to partially alleviate the loss of a functional *cbiH* gene.

Effect of corrinoid starvation on *D. restrictus* metabolism

Reduction of cyanocobalamin amendment in the growth medium strongly inhibited PCE dechlorination by *D. restrictus*. It also had a profound effect on *D. restrictus* corrinoid metabolism both at the level of transcription and at the proteome level. While changing from high, (250 µg/L) corrinoid concentration to an intermediate concentration (50 µg/L), *D. restrictus* responded by up-regulating proteins associated with corrinoid transport and salvaging pathways encoded in operon-1 and -2, allowing the strain to reach the same PCE dechlorination level as observed during high corrinoid concentration. Decreasing the corrinoid concentration even further to 10 µg/L showed, however, that, while the extent of PCE dechlorination was strongly affected, the amount of corrinoid-associated proteins did not notably change when compared to cells cultivated in the presence of 50 µg/L. This indicates that at corrinoid concentrations as low as 10 µg/L, *D. restrictus* was not able to compensate the lack of externally provided corrinoids by increased corrinoid transport and salvaging.

The presence of cobalamin riboswitches directly upstream of the five corrinoid operons in *D. restrictus* already suggested an active repression at the level of transcription by cyanocobalamin. Similar to transcriptional studies on *D. mccartyi* [32] and *D. hafniense* [55], the cobalamin riboswitches of *D. restrictus* responded to addition of excess cyanocobalamin, and the level of repression of the gene located directly downstream of the riboswitches correlated well with the proteomic data, showing the strongest effect for *cbiJ* (Dehre_0277, operon-1) and for *fepB* (Dehre_0281, operon-2).

The sequence of individual cobalamin riboswitches is likely responsible for their differential responsiveness towards cobalamin concentration, as both their affinity to cobalamin and the strength which the expression platform exerts on transcriptional repression are sequence dependent. Such effects have already been shown for a few cobalamin riboswitches in *D. hafniense* [55],

D. mccartyi strain 195, another corrinoid-auxotroph, requires a concentration of 25 µg/L cyanocobalamin to support optimal TCE dechlorination rates and growth yield [198], a value that is similar to what was observed for *D. restrictus*. Therefore, and in addition to the ecogenomic biomarkers defined by Maphosa *et al.* [8], one could consider the physiological threshold of corrinoid concentration as a possible diagnostic tool to delineate the reductive dechlorination potential by corrinoid-auxotrophic OHRB in anaerobic environments. Meanwhile, the production of the PCE reductive dehalogenase (PceA, Dehre_2398) in *D. restrictus* showed no significant change under different corrinoid concentrations, which strongly suggests that the amount of available corrinoid and not of the apo-enzyme represents the main limiting factor for PCE dechlorination.

Taken altogether, our results support the hypothesis that, besides the partial deletion of *cbiH* in *D. restrictus* [24, 115], which already represents a crucial checkpoint in the corrinoid biosynthesis pathway, the energetic cost of *de novo* corrinoid biosynthesis might explain why *D. restrictus* has developed enhanced corrinoid transport and salvaging strategies. *D. restrictus* corrinoid metabolism represents an intermediate situation between the true corrinoid-auxotrophic and obligate organohalide-respiring *D. mccartyi*, which lacks the corrinoid biosynthesis pathway completely [63, 184, 198, 199], and the facultative OHRB able to produce corrinoids *de novo*.

List of abbreviations

Cbl-RS cobalamin riboswitch

CS corrinoid salvaging

CT corrinoid transporter

LFQ label free quantitation

NLA nucleotide loop assembly

OHR organohalide respiration

OHRB organohalide respiring bacteria

PCE tetrachloroethene (perchloroethylene)

RDase reductive dehalogenase

TCE trichloroethene

UCB upper corrinoid biosynthesis

Acknowledgements

The Swiss National Science Foundation (SNSF) is acknowledged for support to AR, JM and CH in frame of the SNF project No. 31003A_138114. We furthermore thank the European Community program FP7 (grants KBBE-211684; BACSIN, and KBBE-222625; METAEXPLORE) for financial support to HS and TK, and the China Scholarship Council for support to YL.

Author contribution

AR and YL performed the experiments, analysed the data and wrote the manuscript. TK analysed the data and revised the manuscript. SB performed the experiments and revised the manuscript. CH and HS revised the manuscript. JM designed the work, analysed the data and wrote the manuscript.

Chapter 5

vanI: a Novel D-Ala-D-Lac Vancomycin Resistance Gene Cluster found in *Desulfitobacterium hafniense*

Thomas Kruse, Mark Levisson, Willem M. de Vos and Hauke Smidt

Published as: Kruse T, Levisson M, de Vos WM, Smidt H (2014) *vanI*: a novel d-Ala-d-Lac vancomycin resistance gene cluster found in *Desulfitobacterium hafniense*. Microbiol Biotechnolology doi:10.1111/1751-7915.12139

Abstract

The glycopeptide vancomycin was until recently considered a drug of last resort against Gram-positive bacteria. Increasing numbers of bacteria, however, are found to carry genes that confer resistance to this antibiotic. So far, ten different vancomycin resistance clusters have been described. A chromosomal vancomycin resistance gene cluster was previously described for the anaerobic *Desulfitobacterium hafniense* Y51. We demonstrate that this gene cluster, characterized by its D-Ala-D-Lac ligase-encoding *vanI* gene, is present in all strains of *D. hafniense, D. chlororespirans* and some strains of *Desulfosporosinus* spp. This gene cluster was not found in vancomycin-sensitive *Desulfitobacterium* or *Desulfosporosinus* spp., and we show that this antibiotic resistance can be exploited as an intrinsic selection marker for *Desulfitobacterium hafniense* and *D. chlororespirans*. The gene cluster containing *vanI* is phylogenetically only distantly related with those described from soil and gut bacteria, but clusters instead with vancomycin resistance genes found within the phylum *Actinobacteria* that include several vancomycin-producing bacteria. It lacks a *vanH* homologue, encoding a D-lactate dehydrogenase, previously thought to always be present within vancomycin resistance gene clusters. The location of *vanH* outside the resistance gene cluster likely hinders horizontal gene transfer. Hence, the vancomycin resistance cluster in *D. hafniense* should be regarded a novel one that we here designated *vanI* after its unique D-Ala-D-Lac ligase.

Introduction

Antibiotic resistance is a two edged sword. On one hand it has proven extremely useful, as a convenient selection marker for biotechnological applications, and a fast way to eliminate the susceptible fraction from enrichment cultures [200, 201]. On the other hand the spread of antibiotic resistance among bacteria can lead to multi-resistant pathogens, making their inhibition one of the most serious challenges in modern healthcare [202].

The glycopeptides vancomycin and teicoplanin are widely used antibiotics for treatment of infections with Gram-positive bacteria. They act by binding to the D-alanyl-D-alanine (D-Ala-D-Ala) terminus of intermediates in peptidoglycan formation, thereby inhibiting cell wall cross-linking [203]. Since the late 1980s, vancomycin-resistant pathogens have emerged in hospitals worldwide. Vancomycin resistance results from the production of modified peptidoglycan precursors terminating in either D-alanyl-D-lactate (D-Ala-D-Lac) or D-alanyl-D-serine (D-Ala-D-Ser) to which vancomycin exhibits low binding affinities. Vancomycin resistance is classified into clusters based on the DNA sequence of the ligase gene *vanA* and its homologs that encode the key enzyme in the synthesis of D-Ala-D-Lac or D-Ala-D-Ser. Currently, ten different types of vancomycin resistance gene clusters have been identified. D-Ala-D-Lac type resistance is associated with clusters *vanA, vanB, vanD, vanF and vanM*, whereas D-Ala-D-Ser resistance is linked to clusters *vanC, vanE, vanG, vanL and vanN* [204-208]. The *vanA* cluster also confers a high level of resistance to teicoplanin whereas the other gene clusters of the D-Ala-D-Lac type give less or no resistance to teicoplanin [207, 209].

Three genes are considered essential for D-Ala-D-Lac mediated vancomycin resistance: *vanH*, *vanA* and homologs and *vanX*. These are found in three genes cluster forming the *vanHAX* core that is conserved across pathogenic bacteria, glycopeptide antibiotic producers, and other environmental bacteria (Fig. 1) [209, 210]and references cited therein). The *vanH* gene encodes a D-lactate dehydrogenase that converts pyruvate to D-lactate, *vanA* and homologs encode an ATP-dependent depsipeptide ligase involved in the synthesis of D-Ala-D-Lac, and *vanX* encodes a dipeptidase that cleaves D-Ala-D-Ala present in the cell, thereby enriching the cell wall in D-Ala-D-Lac. These key genes are found together with a flexible pool of accessory genes, such as those encoding D,D-carboxypeptidases and regulatory proteins (Fig. 1) [205, 211].

Desulfitobacterium spp. include strictly anaerobic bacteria that have mainly been isolated from environments contaminated with halogenated organic compounds. Most strains of *Desulfitobacterium* spp. isolated to date are able to degrade one or more halogenated compounds by organohalide respiration [16]. One of few exceptions is *D. hafniense* DP7, which was isolated from a human fecal sample [49]. Previously, a vancomycin-resistance gene cluster was identified in the tetrachloroethene (PCE) dehalogenating bacterium *Desulfitobacterium hafniense* Y51 [209]. The predicted *vanA* gene was not found in the usual context of three genes forming the *vanHAX* core, but rather was present in a cluster consisting of seven genes, referred to as *vanXmurFvanKWARS* (Fig. 1) [209]. This gene cluster does not contain a *vanH* gene, unlike clusters previously described for vancomycin-resistant bacteria. Functional expression of the genes in *D. hafniense* Y51 was confirmed by resistance to vancomycin, with a minimal inhibitory concentration (MIC) for growth of 64 μg/ml [209].

In the present study we investigated the distribution and phylogenetic affiliation of vancomycin resistance among *Desulfitobacterium* spp. of environmental and human origin. In addition, it was investigated whether vancomycin could be successfully used as a selective agent for isolation of *Desulfitobacterium* spp. from enrichment cultures.

Experimental procedures

Strains and culturing conditions

The bacterial strains used in this study are listed in Table 1. *Escherichia coli* XL1-blue was grown in Lysogeny Broth (LB) at 37°C. *Desulfosporosinus* spp. were grown in DSMZ medium DSM 641 supplemented with 50 mM pyruvate for growth of *D meridei*, or 12.5 mM pyruvate and 6.25 mM malate for growth of *D. youngia* or *D. orientis*.

Desulfitobacterium spp. were grown as 20 ml cultures in 120 ml serum bottles under an atmosphere of 100% N_2. *Desulfitobacterium dichloroeliminans* DCA1 was grown as described previously [48]. *Desulfitobacterium metallireducens* was grown in the medium recommended by the DSMZ (DSM 838). All other strains of *Desulfitobacterium* spp. were cultivated in basal medium as described previously [227], supplemented with 1 g/l yeast extract, and with additional modifications from the original medium as follows.

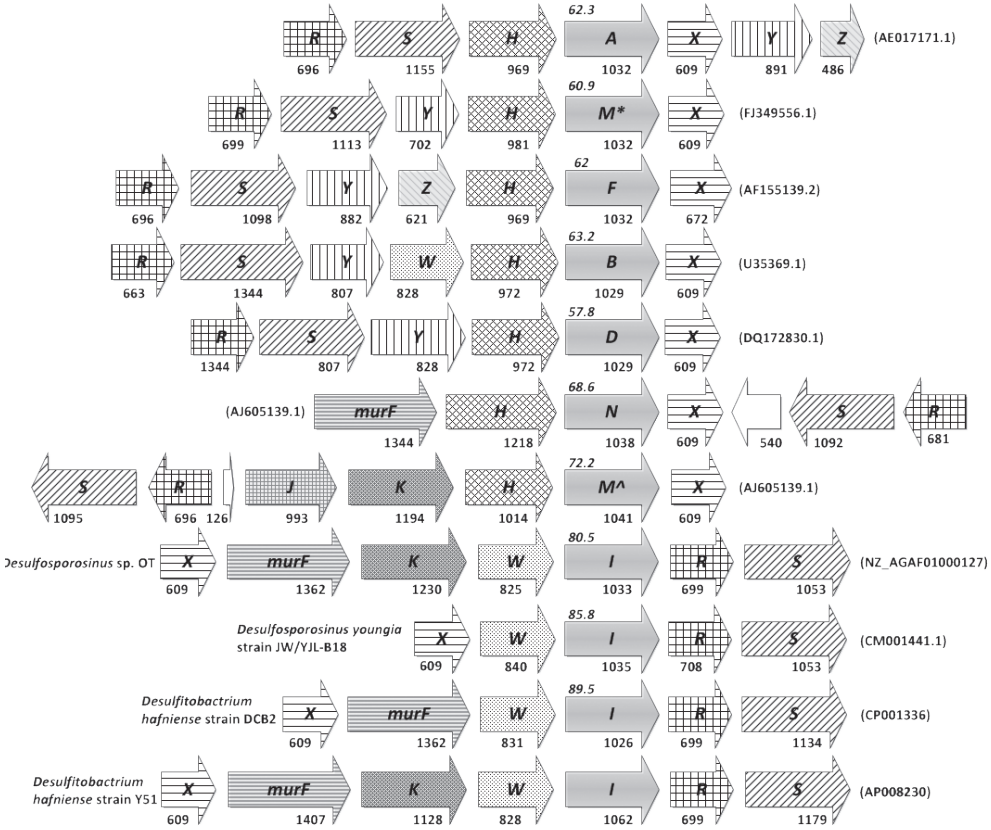


Figure 1. Alignment of vancomycin resistance gene clusters representing the currently known *van* gene clusters (see also Fig. 3). Arrows show direction of transcription, arrow size indicates gene size. Letters inside arrows denote gene names with the "*van*" prefix removed. Empty arrows are hypothetical genes. Light grey arrows represent the D-Ala-D-Lac ligases giving name to the gene clusters. Numbers under arrows indicate gene size in bp, and numbers in *italics* above D-Ala-D-Lac ligases show percentage amino acid identity to *VanI* from *Desulfitobacterium hafniense strain* Y51. Note that *vanI* previously was referred to as *vanA* by Kalan and co-authors [209]. Accession numbers are given in parentheses.

After autoclaving, the medium was supplemented with vitamins and trace elements from anaerobic filter sterilised (0.22 µm pore size) stocks giving final concentrations per litre medium of 1.34 mM EDTA; 10.06 mM $FeCl_2 \bullet 4H_2O$; 0.51 mM $MnCl_2 \bullet 4H_2O$; 0.8mM $CoCl_2 \bullet 6H_2O$; 0.51 mM $ZnCl_2$; 0.02mM $CuCl_2$; 0.04 mM $AlCl_3 \bullet 6H_2O$; H_3BO_3; 0.15 mM Na_2MoO_4; 0.1 mM $NiCl_2 \bullet 6H_2O$, 0.75 mM $CaCl_2 \bullet 2H_2O$; 0.5 mM $MgCl_2 \bullet 6H_2O$; 3 µg Na_2SeO_3; 9 µg $Na_2WO_4 \bullet 2H_2O$; 0.04 µl concentrated HCl; 50 µg biotin; 250 µg p-aminobenzoate; 50 µg panthothenate; 20 µg folic acid; 50 µg lipoic acid; 100 µg pyridoxine; 550 µg nicotinamide;

100 µg thiamine HCl; 50 µg riboflavine; 50 µg cyanocobalamine; and finally reducing and buffer solutions to a final concentration of 1mM $Na_2S•9H_2O$, 5.6 mM NH_4HCO_3 and 44.4 mM $NaHCO_3$.

Cultures were routinely grown fermentatively on 40 mM pyruvate at the temperatures recommended by the DSMZ, unless stated otherwise. Pyruvate, lactate and PCE were added from sterile anaerobic stocks to final concentrations of 40, 20 and 10 mM, respectively, when indicated.

Antibiotic susceptibility testing

All plating of *Desulfitobacterium* spp. and *Desulfosporosinus* spp. was performed in an anaerobic glove box containing an atmosphere of 96% N_2 and 4% H_2. Strains were streaked with cotton swaps, on plates prepared as described previously (Smidt, et al., 1999), with the modification that Difco™ agar noble (Becton Dickinson Le pont de Claix, France) was used for solidification. The medium contained 40 mM pyruvate, 6 mM malate and 10 mM SO_4 as carbon and energy source. Plates were prepared with various concentrations of vancomycin or teicoplanin (0, 1, 5, 10, 25, 50, 75, 150 and 250 µg/ml) and incubated anaerobically at 30 °C.

DNA isolation

Cells were harvested from 5 ml cultures by spinning at 6700 × g for 10 min. DNA was extracted from cell pellets using the Bio101 FASTDNA isolation kit according to manufacturers' protocols (MP Biomedicals, Solon, OH, USA).

PCR amplification, cloning and sequencing of vanI genes

All PCR reactions were performed using the GoTaq DNA Polymerase Kit (Promega, Leiden, the Netherlands) in a total reaction volume of 30 µl, with 0.5 µl DNA extract or cell lysate as template. PCR products were analysed on a 1% agarose gel using gene ruler 1 kb or 100 bp plus ladders (Fermentas, St. Leon-Rot, Germany).

A 987 bp fragment of the vanI gene was amplified using primers K46F/K47R (Supplemental Table 2) and the following temperature programme: 94 °C for 120 s, followed by 35 cycles of 94 °C for 30 s, 56 °C for 30 s and 72 °C for 50 s, and a final elongation step at 72 °C for 6 min. Resulting amplicons were purified prior to sequencing using the Zymo DNA Clean & Concentrator Kit (Zymo Research, San Francisco, CA, USA), sequenced in one direction using primer

K46 and initially analysed using nucleotide BLAST [137]. An overlapping fragment of 512 bp was obtained for all amplicons, and used for phylogenetic analysis.

Purity of Desulfitobacterium hafniense TCE1

Nearly full length 16S rRNA gene fragments were amplified from *D. hafniense* TCE1 using primers 27F/1492R (Supplemental Table 2) and the following temperature programme: 94 °C for 120 s, followed by 35 cycles of 94 °C for 30 s 52 °C for 40 s and 72 °C for 90 s, ending with a final elongation step of 72 °C for 5 min.

PCR products were purified as described above, cloning of 16S rRNA genes was done using the cloning vector pGEMT-Easy (Promega) and E. coli XL1-Blue as host following the manufacturer's instructions. Full length inserts were amplified from transformant cells, lysed by heating of 10 µl o/n culture in 90 µl TE (10 mM tris-HCl, 1 mM EDTA, pH 7.5) at 97 °C for 10 min. PCR amplification using primers T7 and SP6, targeting the flanking region of the pGEMT-Easy cloning site (Supplemental Table 2), was done as described above, using the following temperature programme: 94 °C for 120 s, followed by 35 cycles of 94 °C for 30 s 52 °C for 40 s and 72 °C for 90 s, ending with a final elongation step of 72 °C for 5 min.

Clones were analysed for the presence of *Desulfitobacterium* spp. derived 16S rRNA gene sequences using genus-specific primers DSB406F/619R (Supplemental Table 2), and the following temperature programme: 94 °C for 4 min, followed by 35 cycles of 94 °C for 30 s 58 °C for 20 s, 72 °C for 35 s, and a final elongation step of 72 °C for 10 min. Inserts from three randomly selected clones that did not give a product with *Desulfitobacterium* specific primers were partially sequenced and analysed using BLAST [137].

DNA sequence analysis

Alignments of DNA sequences were done using MUSCLE [36], and trees constructed using the neighbour joining method and tested by 1000 bootstraps using the MEGA 5 software package [35]. Sequences obtained in this study were first aligned with those used in the study of Aminov and Mackie (2007) and one sequence representing the recently described cluster *vanM* [207]. Finally sequences obtained from genome databases were added [216]. Identity matrixes were generated using MatGAT with default settings [113]. Full genome alignment and analysis was done using the Artemis comparison tool, version 10 [187].

Check for *vanI* homologues in (meta)genomes

All available (as of 4-2-2013; 1236 metagenomes covering a wide range of environments) metagenomes on the JGI IMG/M server [228] were screened for the presence of *vanI* homologues. We performed an initial blastp using the strictest possible cut off, E<e-50, against the entire img/m database using VanI (dsy_3690) from *Desulfitobacterium hafniense* Y51 as query. We obtained 1338 sequences with homology to VanI, created a local blast database containing the contigs harbouring the corresponding DNA sequences, and performed a tblastx analysis against all contigs, using the *D. hafniense* Y51 *vanI* gene sequence as query. This was done in order to correct for genes that potentially have not been incorporated correctly into the blastp database, due to for example frameshift mutations or assembly errors. The eight full length D-Ala-D-Lac ligase gene homologs with the best hits to *vanI* (E≤ 3.00E-135), were aligned and checked for phylogenetic affiliation with *vanI*.

vanI homologs in sequenced bacterial genomes outside the *Desulfitobacterium* genus

All available genomes (as of 19-5-2014) at the JGI/IMG server [216] were screened for the presence of *vanI* homologues. Initially, as for metagenome mining, BLASTp search was done using *D. hafniense* Y51 VanI (dsy3690) sequence as query against all available genomes. This was followed by manual check of sequences giving strong (E=0) hits to VanI. The *vanI* sequence identified in *Desulfosporosinus* sp. OT was found to be N-terminally truncated. A closer inspection of the genome sequence showed that this is due to either a frameshift mutation or sequencing error, by combining two reading frames we were able to assemble a 1032 bp fragment The corresponding predicted amino acid sequence showed 93.3% identity to the *vanI* homologue from *Desulfosporosinus youngiae* JW/Y JL-B18 (Supplemental Table 1).

Results

D-Ala-D-Lac ligases-in *Desulfitobacterium* spp.

Chromosomal DNA was isolated from 13 *Desulfitobacterium* spp. strains, covering the majority of the currently described species of this genus. The DNA sequences of the two D-Ala-D-Lac ligase-encoding genes from *D. hafniense* Y51 (DSY3690) and *D. hafniense* DCB-2 (Dhaf_1664) were used to design primers for the detection of homologous D-Ala-D-Lac ligase genes in other *Desulfitobacterium* spp. strains. Amplicons of the expected size of 987 bp were obtained for nine

of the thirteen tested strains (Table 1). All nine PCR-positive isolates were strains of *D. hafniense*, with the exception of *D. chlororespirans*, which is the closest relative of *D. hafniense* based on 16S rRNA gene similarity (Fig. 2). Sequence analysis of the obtained PCR products confirmed that they encode D-Ala-D-Lac ligases (data not shown).

Phylogenetic analysis

Previously, a phylogenetic analysis was performed using a data set of concatenated ORFs comprising *vanH*, *vanA* and its homologs, and *vanX* [212]. This analysis showed a very early branching event, leading to the formation of two large clusters, one of which consists of genes associated with vancomycin -resistant enterobacteria (VRE) and other vancomycin resistant gut and soil bacteria (cluster I). The second cluster consists of members of the *Actinobacteria,* some of which have been shown to produce vancomycin (cluster II). Furthermore, this phylogenetic analysis suggested that *vanHAX* and homologuous gene clusters behave as a single evolutionary unit [212].

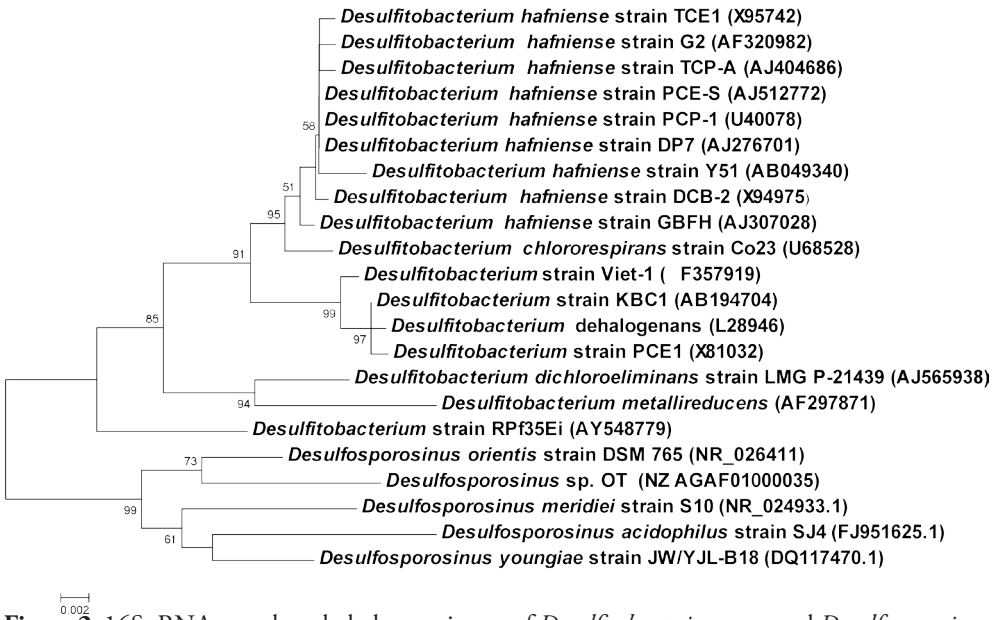


Figure 2. 16S rRNA gene-based phylogenetic tree of *Desulfitobacterium* spp. and *Desulfosporosinus* spp. Sequences were aligned with MUSCLE [36], and a neighbour joining tree was constructed and validated with 1000 bootstraps analysis using the MEGA5 software package [35]. Bootstrap values higher than 50% are given at corresponding nodes in the tree. The reference bar indicates 0.2% base substitutions per site. GenBank accession numbers of sequences used in this analysis are given in parentheses.

#	Strain	Presence of *vanI*	vancomycin MIC [µg/ml]	Teicoplanin MIC [µg/ml]
1	*D. hafniense* Y51[1]	Yes (AP008230.1)	> 250	>50<75
2	*D. hafniense* TCE-1[2]	Yes (HQ433581)	> 250	>50<75
3	*D. hafniense* DP7 (DSM[3] 13498)	Yes (HQ433580)	> 150 < 250	>25<50
4	*D. hafniense* DCB-2[T] (DSM[3] 10664)	Yes (CP001336.1)	> 150 < 250	>25<50
5	*D. hafniense* PCE-S (DSM[3] 14645)	Yes (HQ433585)	> 250	>25<50
6	*D. hafniense* PCP-1 (DSM[3] 12420)	Yes (HQ433583)	> 150 < 250	>50<75
7	*D. hafniense* TCP-A (DSM[3] 13557)	Yes (HQ433582)	> 250	>25<50
8	*D. hafniense* G2 (DSM[3] 16228)	Yes (HQ433584)	> 250	>50<75
9	*D. chlororespirans* CO23[T] (DSM[3] 11544)	Yes (HQ433586)	> 250	>50<75
10	*D. dehalogenans* (DSM[3] 9161)	No	< 1	<1
11	*D. sp.* PCE1 (DSM[3] 10344)	No	< 1	<1
12	*D. metallireducens* 853-15A[T](DSM[3] 15288)	No	ND	ND
13	*D. dichloroeliminans* DCA1[T4]	No	ND	ND
14	*D. youngia* (DSM[3] 17734[T])	Yes (2508507521)[5]	>10<25	>5<10
15	*D. orientis* (DSM[3] 765[T])	No	<1	<1
16	*D. meridiei* (DSM[3] 13257[T])	No	<1	<1
17	*D. acidiphilus* (DSM[3] 22704[T])	No	<1	<1

Table 1. *Desulfitobacterium* spp. (top) and *Desulfosporosinus* spp. (bottom) strains used in this study, Accession numbers are given in parentheses.[1] Received as a kind gift from Masatoshi Goto, Department of Bioscience and Biotechnology, Faculty of Agriculture, Kyushu University, Japan [2] Received as a kind gift from Jan Dirk van Elsas, Department of Microbial Ecology, University of Groningen, The Netherlands [3] DSM, Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany [4] Received as a kind gift from Nico Boon, Laboratory of Microbial Ecology and Technology, Ghent University, Belgium.[5] JGI gene ID

We constructed a phylogenetic tree using the nucleotide sequences of 512-bp fragments corresponding to part of the D-Ala-D-Lac ligase genes obtained in this study. This tree showed essentially the same clustering as when using the sequences of the entire *vanHAX* and homologuous gene clusters (data not shown), indicating that the sequence fragments obtained in this study are sufficient for robust phylogenetic analysis. We then included in the phylogenetic tree the 512-bp sequence fragments obtained from *Desulfitobacterium* spp., closely related *Desulfosporosinus* spp, and sequences from the newly described *vanM* cluster [213, 214], resulting in a tree containing representative sequences from all currently described vancomycin-resistance clusters of the D-Ala-D-Lac type (Fig. 3). The phylogenetic tree shows two main and deep branching clusters. Cluster I consists of *vanA*; *vanM**; *vanF*; *vanB* and *vanD,* while cluster II consists of *vanM*^, *vanN,* and a separate sub-cluster comprising all D-Ala-D-Lac ligase-encoding sequences obtained from *Desulfitobacterium* spp. and *Desulfosporosinus* spp. (Fig. 3). Although the sequences obtained from glycopeptide producing *Actinobacteria*, *Desulfitobacterium* spp. and *Desulfosporosinus* spp belong to cluster II, it can be assumed that desulfitobacteria do not produce vancomycin, as no vancomycin synthesis pathway was identified in the two published genomes of *Desulfitobacterium hafniense* DCB-2 and Y51 [19, 26], nor in the available draft genome sequence of strain DP7 (AFZX00000000.1).

Sequences obtained from *Desulfitobacterium* spp. and *Desulfosporosinus* spp. were found to be more similar to those of the D-Ala-D-Lac ligases belonging to cluster II than to any of the sequences from cluster I (Supplementary Table 1). Based on these results, it is clear that the D-Ala-D-Lac ligase gene found in *Desulfitobacterium* spp. and *Desulfosporosinus* spp. does not belong to the *vanA* cluster, but instead is part of a distinct vancomycin resistance cluster, that we propose to name *vanI* after its unique D-Ala-D-Lac ligase.

Presence of *vanI* outside the *Desulfitobacterium* genus

BlastP analysis against all sequenced bacterial genomes identified vanI homologues in the genomes of both Desulfosporosinus youngiae JW/Y JL-B18 and Desulfosporosinus sp. OT (Fig. 3), but not in Desulfosporosinus orientis Singapore I; D. meridiei S10 or D. acidiphilus SJ4 [213, 215, 216]. We did not detect vanI-like genes in any of the metagenomes publicly available in the JGI IMG/M repository, (as of 4-2-2013), demonstrating that vanI genes are not widespread in the environment or the human gut.

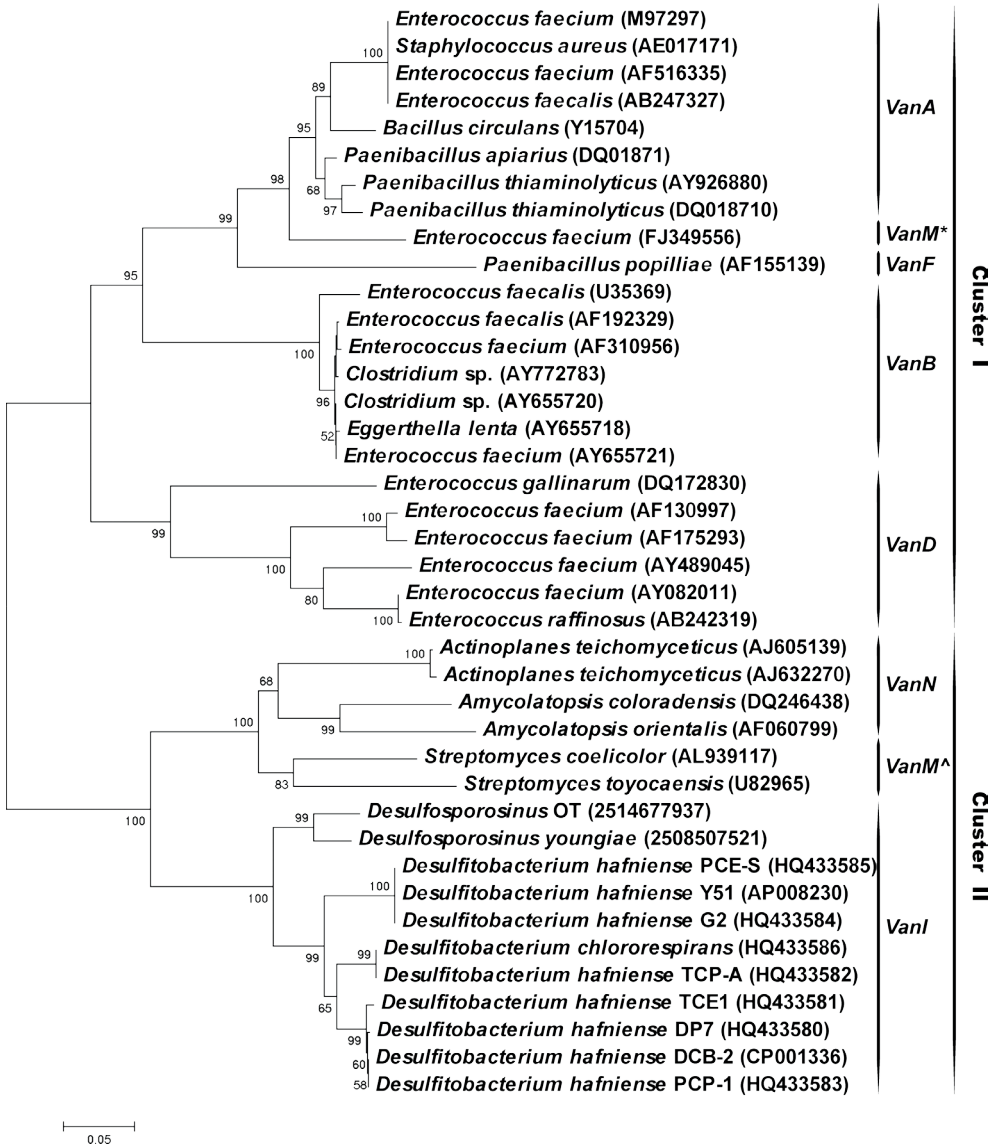


Figure 3. Phylogenetic tree based on a 512bp fragment of D-Ala-D-Lac ligase encoding genes representing the currently known *Van* gene clusters of the D-Ala-D-Lac type. Corresponding reference sequences [207, 212], *vanI* sequences obtained in this study and sequences retrieved from genome datasets [214-216] were aligned with MUSCLE [36], and a neighbour joining tree was constructed and validated with 1000 bootstraps analysis using the MEGA5 software package [35]. * *vanM* according to [207] ^ *vanM* according to [212]. Bootstrap values higher than 50% are given at corresponding nodes in the tree. The reference bar indicates 5% base substitutions per site. GenBank accession numbers or JGI gene IDs of sequences used in this analysis are given in parentheses. Definition of *van* gene clusters was adapted from [212].

Architecture of the *vanI* gene cluster

Inspection of the genomic sequences of *Desulfitobacterium* spp. *and Desulfosporosinus* spp. revealed the *vanX$_I$murF$_I$vanK$_I$W$_I$IR$_I$S$_I$* (*vanXmurFvanKWARS* in[209]) gene cluster to be conserved (Fig. 1). However, it should be noted, that *vanK$_I$*, encoding the enzyme adding the branched amino acid(s) to the stem pentapeptide of peptidoglycan precursors [217], was only found in *Desulfitobacterium hafniense* Y51 and *Desulfosporosinus* sp. OT. An interesting feature of the *vanI* gene cluster is the location of the gene encoding the D-Lac dehydrogenase (*vanH*) outside the vancomycin-resistance gene cluster. Usually *vanH* genes are part of the three genes *vanHAX* or homologues that form the core of vancomycin-resistance gene clusters [211]. Previously, four genes have been proposed as potential *vanH*-encoding genes (DSY0996, DSY1673, DSY3442, and DSY4020) in the genome of *D. hafniense* Y51 [209]. Homologs of these four genes were also found in the genomes of *D. hafniense* DCB-2 (DHAF1344, DHAF1980, DHAF2079 and DHAF2820) and *D. hafniense* DP7 (data not shown). The complete genomes of *D. hafniense* Y51 and DCB-2 [19, 26] were aligned and inspected for signs of horizontal transfer of the vancomycin-resistance gene cluster. No genes encoding integrases or transposases, nor large deviations in dinucleotide frequency or GC content, were found in the vicinity of the vancomycin-resistance gene clusters suggesting that they are not located on mobile genomic islands.

Vancomycin and teicoplanin susceptibility

In total 15 strains belonging to 3 different species of *Desulfitobacterium* and four species of *Desulfosporosinus* were tested for growth on plates containing different concentrations of vancomycin or teicoplanin. All strains for which a *vanI* gene was detected are resistant to vancomycin and teicoplanin (Table 1). Strains of *Desulfitobacterium* spp. showed high level of resistance with MIC values >150 µg/ml vancomycin and > 25 ug/ml Teicoplanin (Table 1). The two strains that tested negative for *vanI* genes, namely *D. dehalogenans* and the closely related strain PCE1, are highly susceptible to both vancomycin and teicoplanin with MIC values <1 µg/ml. The same was the case for the tested members of the *Desulfosporosinus* genus. *Desulfosporosinus youngia* whose genome contain a *vanI* gene cluster, is resistant to vancomyin and teicoplanin although their MICs were approximately ten-fold lower than those observed for *Desulfitobacterium* spp. The three tested *Desulfosporosinus* spp. strains lacking a *vanI* gene are all susceptible to both vancomycin and teicoplanin (Table 1).

Vancomycin as selective agent

Previous experiments with cultures of the PCE-dehalogenating strain *D. hafniense* TCE1, received from collaborators as well as from DSMZ, led us to suspect that the cultures contained a microbial contaminant. Partial 16S rRNA gene sequencing confirmed that the *D. hafniense* TCE1 culture was contaminated with a *Sedimentibacter* sp. (97% 16S rRNA gene sequence identity to *Sedimentibacter* sp. C7, AY766466.1) (data not shown). *Sedimentibacter* spp. are gram positive, members of the *Firmicutes*, and as a consequence likely susceptible to vancomycin [218]. We therefore decided to test vancomycin as a potential selective agent for the elimination of the *Sedimentibacter* contaminant.

A contaminated *D. hafniense* TCE1 culture was first grown in liquid medium with 20 mM lactate and 10 mM PCE as electron donor and acceptor, respectively, in order to pre-select for the dechlorinating *Desulfitobacterium*. Subsequently, an early exponential growth phase culture was plated on medium containing 40 mM pyruvate and supplemented with 50 μg/ml vancomycin. After one week, only one colony type was observed. Two colonies were picked and grown in liquid medium containing 20 mM lactate and 10 mM PCE, supplemented with 50 μg/ml vancomycin and subsequently transferred to liquid medium containing 40 mM pyruvate and no vancomycin. This resulted in a pure culture of *D. hafniense* TCE1 cultures without the presence of any *Sedimentibacter* as confirmed by screening of a general bacterial 16S rRNA gene clone library with primers specific for *Desulfitobacterium* spp.

Discussion

In this contribution the distribution of vancomycin-resistance among *Desulfitobacterium* spp. and *Desulfosporosinus* spp. is described. Nine of the thirteen tested strains of *Desulfitobacterium* spp. display high levels of vancomycin resistance, whereas one out of four strains of *Desulfosporosinus* spp. was found moderately vancomycin resistant. (Table 1). Previously a vancomycin MIC value of 64 μg/ml was reported for *D. hafniense* Y51 [209], which is considerably lower than the values obtained in this study (Table 1). A possible explanation for the higher MIC values could be differences in the plating medium and protocol between this study and the study of Kalan and colleagues [209]. It was, for instance, demonstrated that the growth on plates of *Desulfitobacterium dehalogenans* was much improved when 0.8% of highly purified agar was used instead of 1.5% agar of standard purity[119].

Phylogenetic analysis of predicted D-Ala-D-Lac ligase encoding genes showed that those obtained from *Desulfitobacterium* spp. together with *Desulfosporosinus* spp. belong to cluster II where they form a distinct branch, next to genes retrieved from glycopeptide producing members of the phylum *Actinobacteria* (Fig. 3). It has often been speculated that antibiotic resistance, including, vancomycin resistance, was developed as a mechanism to prevent suicide by antibiotic producing bacteria. According to this hypothesis, resistance genes then spread to soil bacteria before, finally, via an intermediate host bacterium, ending up in enterobacteria and other gut bacteria [212, 219-221]. The phylogenetic analysis of Aminov and Mackie (2007) suggested an early diversification between cluster I (soil and gut bacteria) and cluster II (glycopeptide producing *Actinobacteria*), with no later exchange of genes between these [212]. The vancomycin-resistance genes found in *Desulfitobacterium* spp and *Desulfosporosinus* spp. form a distinct group in cluster II, indicating that they have been acquired from vancomycin producers at a later stage than those belonging to cluster I (Fig. 3). Therefore, we propose that the vancomycin-resistance gene cluster found in *Desulfitobacterium* spp. and *Desulfosporosinus* spp. should be considered as a novel cluster, designated *vanI* rather than a variant of the *vanA* type as initially suggested for *D. hafniense* Y51[209].

The basic structure of the $vanX_I murF_I vanK_I W_I IR_I S_I$ gene cluster was found to be conserved in the published genomes of *D. hafniense* strain Y51, DCB-2, *Desulfosporosinus* sp. OT and *Desulfosporosinus youngiae* (Fig. 1) [19, 26, 213, 215]. The only exceptions to this conserved structure found so far are that $murF_I$ was not found in *Desulfosporosinus youngiae* and $vanK_I$ is only present in *D. hafniense* Y51 and *Desulfosporosinus* sp. OT (Fig. 1). A characteristic feature of the *VanI* vancomycin-resistance gene cluster is the absence of a gene encoding the D-Lac dehydrogenase (VanH). This is the first vancomycin-resistance gene cluster, where the *vanH* gene is located on the chromosome outside the resistance cluster.

Desulfitobacteria have been found both in the environment and in the human gut [16]. Hence, it is tempting to speculate that they potentially could act as a shuttle for vancomycin resistance between the environment and the gut microbiota. However, no indications were found in the available *Desulfitobacterium* spp. genomes that the vancomycin-resistance genes are or have been mobile. The location of *vanH* homologues elsewhere in the genome, rather than in the vancomycin resistance cluster, likely act as a barrier against horizontal transfer of this phenotype.

The genera *Desulfitobacterium* and *Desulforosporosinus* are phylogenetically closely related, both belonging to the family *Peptococcaceae* (Fig. 2) [222]. It seems likely that the *vanI* gene cluster was acquired by a common ancestor of *Desulfitobacterium* spp. and *Desulforosporosinus* spp. based on the very high sequence similarity between their *vanI* sequences, and the characteristic location of *vanH* homologues outside the resistance gene cluster (Fig. 1 and 3). It is interesting to note that *Streptomyces coelicolor vanM^* is the D-Ala-D-Lac ligase encoding homologue showing highest identity to *vanI* (Supplemental Table 1). This is also the only other vancomycin-resistance gene cluster in which *vanK* has been found [211, 223]. A *Streptomyces coelicolor* Δ*vanK* mutant was shown to be unable to grow in the presence of vancomycin [217]. In contrast, *vanK* is not essential for vancomycin resistance in *Desulfitobacterium* spp., as both *Desulfitobacterium hafniense* DP7 and DCB-2 are resistant to vancomycin, despite not encoding a $vanK_I$. Inspecting the genomes of *Desulfitobacterium hafniense* DCB-2 and DP7 confirmed that *vanK* unlike *vanH* is not located elsewhere in the genomes. *Streptomyces* spp. do not share a recent common ancestor with *Desulfitobacterium* spp. or *Desulfosporosinus* spp., thus it is plausibly that the *vanI* gene cluster was acquired horizontally by a common ancestor of *Desulfitobacterium* spp. and *Desulfosporosinus* spp from *Streptomyces coelicolor* or its ancestor. This would also offer a plausible explanation for the presence of $vanK_I$ in *Desulfitobacterium hafniense* Y51 and *Desulfosporosinus* sp. OT as a redundant reminiscence that has been lost from most *vanI*-containing strains due to lack of selection pressure.

It was investigated whether vancomycin could be used as a selective agent for purifying Desulfitobacterium spp. from enrichment cultures containing other, susceptible Gram-positive populations such as Sedimentibacter spp., which have been repeatedly co-enriched with organohalide respiring species of Desulfitobacterium and the closely related genus Dehalobacter [53, 224-226]. A D. hafniense TCE1 culture, contaminated with a Sedimentibacter sp., was successfully purified using vancomycin as a selection agent. Searching public available meta/genome databases showed that vanI could not be found in any of the screened metagenomes but only in the genomes of members of both D. hafniense or *chlororespirans* and the closely related Desulfosporosinus spp. supporting the hypothesis that vanI, unlike other vancomycin resistance clusters found in the gut microbiota, is not subject to frequent horizontal gene transfers. The use of vancomycin or teicoplanin may thus help in future isolation of new Desulfitobacterium or Desulfosporosinus spp. strains and may also be used as a convenient selection marker for genetic manipulations.

In conclusion we have shown that *D. hafniense*, *D. chlororespirans* and *Desulfosporosinus* spp. contain a novel chromosomal gene cluster, designated *vanI*, conferring medium to high levels of vancomycin resistance, and medium level teicoplanin resistance. This feature can be both an useful intrinsic selection marker, or a potential risk if mobilised into clinically relevant strains.

Nucleotide sequence accession numbers.

The obtained partial *vanI* D-Ala-D-Lac ligase gene sequences are available in the NCBI database under accession numbers HQ433580 to HQ433586

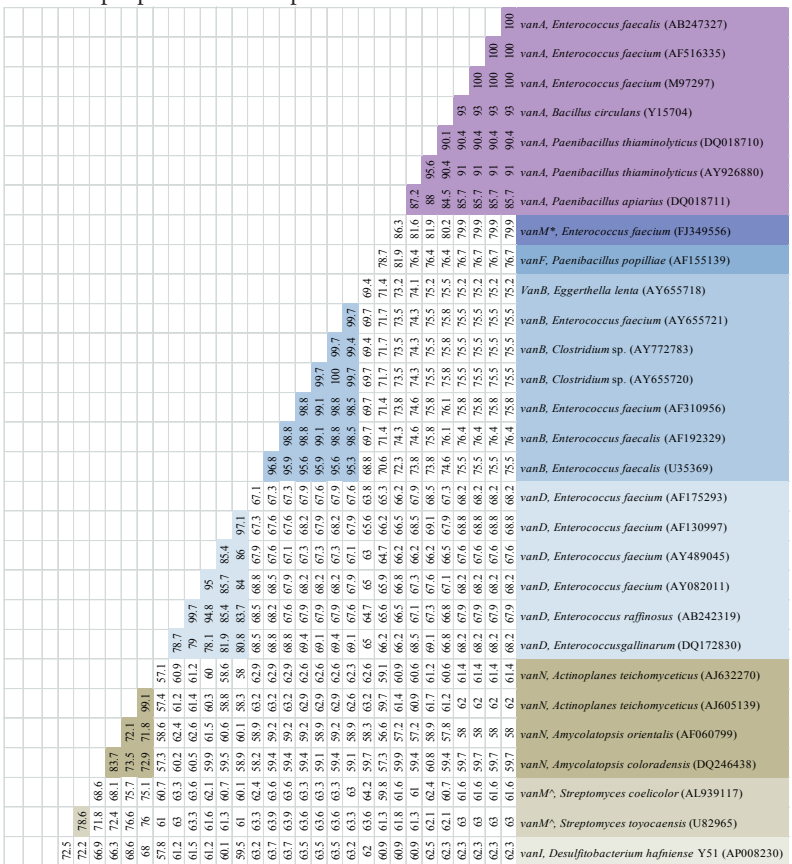
Acknowledgements

The Netherlands Genomics Initiative as well as the European Community's Seventh Framework Programme (FP7/ 2007-2013) are acknowledged for support through the Ecogenomics program, and the BACSIN and METAEXPLORE projects (grant agreements No. 211684 and 222625), respectively.

Supplementary Table 2. Primers used in this study

Name	Sequence 5'-3'	Target	Reference
27F	AGAGTTTGATCMTGGCTCAG	Bacterial 16s rRNA	[229]
1492R	GGTTACCTTGTTACGACTT		[230]
DSB 406F	GTACGACGAAGGCCTTCGGT	16s rRNA of *desulfitobacteria*	[231]
DSB619R	CCCAGGGTTGAGCCCTAGGT		
T7	TAATACGAACTCACTATAGG	T7 promoter	Promega
SP6	GATTTAGGTGACACTATAG	SP6 promoter	
K46F	AAGGCCAACGAGACAAGCC	*vanI*	This study
K47R	GCTGTTCCGAAGAACATCCCG		

Supplementary Table 1. Identity matrix showing percentage amino acid identity of D-Ala-D-Lac ligases representing the currently known van gene clusters of the D-Ala-D-Lac type. The sequences used here are translated, using the standard bacterial genetic code, from the full length genes used in figure 3. Accession numbers are given in parentheses. The identity matrix were generated using MatGat [113] as described under experimental procedures. Colour codes highlight the different vancomycin resistance clusters. Shades of brown indicates representatives of cluster II, whereas blue and purple indicates representatives of cluster I.



Chapter 6

Genomic, proteomic and biochemical analysis of the organohalide respiratory pathway in *Desulfitobacterium dehalogenans*

Thomas Kruse, Bram A. van de Pas, Ariane Atteia, Klaas Krab, Wilfred R. Hagen, Lynne Goodwin, Patrick Chain, Sjef Boeren, Farai Maphosa, Gosse Schraa, Willem M. de Vos, John van der Oost, Hauke Smidt and Alfons J.M. Stams

Supplementary materials are available through the online version of this manuscript

Published as: Kruse T, Van de Pas BA, Atteia A, Krab K, Hagen WR, Goodwin L, Chain P, Boeren S, Maphosa F, Schraa G, de Vos WM, Van der Oost J, Smidt H, Stams AJ (2015) Genomic, proteomic and biochemical analysis of the organohalide respiratory pathway in *Desulfitobacterium dehalogenans* J Bacteriol 195 (5):12. doi:10.1128/JB.02370-14

Abstract

Desulfitobacterium dehalogenans is able to grow by organohalide respiration using 3-chloro-4-hydroxyphenyl acetate (Cl-OHPA) as an electron acceptor. We used a combination of genome sequencing, biochemical analysis of redox active components and shotgun proteomics to study elements of the organohalide respiratory electron transport chain. The genome of *Desulfitobacterium dehalogenans* JW/IU-DC1^{T} consists of a single circular chromosome of 4321753 bp with a GC content of 44.97%. The genome encodes 4252 genes including six rRNA operons and six predicted reductive dehalogenases. One of the reductive dehalogenases, CprA, is encoded by a well characterised *cprTKZEBACD* gene cluster. Redox active components were identified in concentrated suspensions of cells grown on formate and Cl-OHPA or formate and fumarate, using electron-paramagnetic-resonance [232], visible spectroscopy, and high-performance–liquid-chromatography (HPLC) analysis of membrane extracts. In cell suspensions, these components were reduced upon addition of formate and oxidized after addition of Cl-OHPA indicating involvement in organohalide respiration. Genome analysis revealed genes that likely encode the identified components of the electron transport chain from formate to fumarate or Cl-OHPA. Data presented here suggest that the first part of the electron transport chain from formate to fumarate or Cl-OHPA is shared. Electrons are channelled from an outward facing formate dehydrogenase via menaquinones to a fumarate reductase located at the cytoplasmic face of the membrane. When Cl-OHPA is the terminal electron acceptor, electrons are transferred from menaquinones to outward facing CprA, via an yet unidentified membrane complex, and potentially an extracellular flavoprotein acting as an electron shuttle between the quinol dehydrogenase membrane complex and CprA.

Introduction

Microorganisms in anoxic environments have the ability to reductively dechlorinate highly chlorinated compounds such as tetrachloroethene (PCE), pentachlorophenol (PCP) and polychlorinated biphenyls (PCBs). Anaerobic bacteria from a phylogenetically broad range of genera, including *Desulfitobacterium*, *Dehalobacter*, *Geobacter*, *Sulfurospirillum* and *Dehalococcoides*, couple reductive dechlorination of chlorinated organic compounds to energy conservation in a respiratory manner, previously referred to as (de)halorespiration and currently as organohalide respiration [3, 6, 8]. Formate and hydrogen are common electron donors in this process. Since these compounds do not allow substrate-level phosphorylation, energy is most likely conserved by the formation of a proton gradient across the cytoplasmic membrane [6].

The localisation of the reductive dehalogenase (RD) and the topology of the electron transport pathway from electron donor to electron acceptor have been a matter of debate. In early models it was suggested that the electron donating reaction takes place outside the cell, whereafter electrons would be transported across the cell membrane to an intracellular RD, leading to formation of a proton gradient without vectorial proton translocation [74-76]. Contrasting this suggestion, all RDs of the PceA/CprA type (tetrachloroethene/ortho-chlorophenol reductive dehalogenase) contain a twin arginine translocation (TAT) signal sequence, which is recognised by the TAT export system, leading to the export of folded proteins with incorporated co-factors across the cell membrane [3, 233]. The current view is that the catalytic subunit of RDs is located at the outside of the cytoplasmic membrane [3, 16, 18, 19, 234]. If the electron-donating and accepting reactions both occur at the same side of the cell membrane, a proton-translocating mechanism is necessary to establish a proton gradient that can drive ATP synthesis [235, 236]. The electron transport chain from electron donor to RD has been proposed to involve both menaquinones and an intermediate electron carrier, as menaquinones are unable to directly reduce RDs [6, 82]. These electron-transferring elements have, however, not yet been identified.

The low GC Gram-positive *Desulfitobacterium dehalogenans* is able to use a range of terminal electron acceptors, including ortho-chlorophenols, sulfite, thiosulfate, sulfur, nitrate and fumarate [44, 237]. The catalytic subunit of the key enzyme in 3-chloro-4-hydroxyphenyl acetate (Cl-OHPA) dehalogenation, CprA, was purified and characterised as a corrinoid-containing Fe-S protein [64]. Molecular analysis of the *cpr* gene cluster in *D. dehalogenans* led to the identification of genes encoding putative regulatory proteins and protein-folding catalysts, whose transcription

was specifically induced under organohalide-respiring conditions [73]. Research on the electron transport chain from formate to Cl-OHPA has been hampered thus far by the lack of *D. dehalogenans* whole genome information. Previously the genomes of *Desulfitobacterium hafniense* Y51 and DCB-2 have been published [19, 26]. The number of published genomes from *Desulfitobacterium* spp. is expected to increase as the result of ongoing sequencing projects.

In this study, the genome of *D. dehalogenans* strain JW/IU-DC1^T was sequenced. The proteomes of cells grown with formate as electron donor and fumarate or Cl-OHPA as electron acceptor were compared with that of cells grown fermentatively on pyruvate. Involvement of redox active components in the transfer of electrons from formate to either Cl-OHPA or fumarate was investigated. The redox state of these components was analysed during their reduction by formate and subsequent oxidation by Cl-OHPA or fumarate.

Materials and methods

Organism and growth conditions

D. dehalogenans strain JW/IU-DC1^T was cultivated under anoxic conditions at 37°C under a 100% N_2 gas phase. Medium was prepared as described previously [64]. Cultures for the enzyme assays were grown in 3-L bottles containing 2 L medium supplemented with either 20 mM formate and 20 mM Cl-OHPA [238] or 20 mM formate and 20 mM fumarate (FF). Cultures for the proteomic analyses were grown in 250-ml bottles containing 100 ml medium supplemented with 20 mM formate and 20 mM Cl-OHPA or fumarate as electron donors/acceptors or with 40 mM pyruvate (P) for fermentative growth. Cultures used for genome sequencing were grown in 1-L bottles containing 0.5 L medium supplemented with 40 mM pyruvate.

Genome sequencing and analysis

For genome sequencing, cells grown fermentatively on pyruvate were harvested from an early stationary phase culture by centrifugation and resuspended in TE buffer (10 mM Tris, 20 mM EDTA, pH 8), before DNA was extracted following the cetyltrimethylammonium bromide (CTAB) protocol recommended by the DOE Joint Genome Institute (JGI) [239]. The genome of *D. dehalogenans*, gold stamp Gc02348, was generated by the JGI using a combination of Illumina and 454 sequencing [97, 98].

Genes were identified using Prodigal [105] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [106]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) non-redundant database, UniProt, TIGRFam, Pfam, PRIAM, KEGG, COG, and InterPro databases. These data sources were combined to assert a product description for each predicted protein. Non-coding genes and miscellaneous features were predicted using tRNAscan-SE [107], RNAMMer [108], Rfam [109], TMHMM [110], signalP [111], Pred-tat [240] and SMART [241]. The Artemis v14 genome browser was used for visualising the annotated genome [145].

The annotated genome sequence has been deposited under GenBank accession number CP003348.

Cell fractionation and enzyme assays

Cells were harvested in an anoxic glove box with an N_2/H_2 gas phase in a 95/5% ratio. Late-exponential phase cultures were harvested by 10 min centrifugation at 16,000 × g and resuspended to a total volume of 2 ml in buffer A, containing 100 mM potassium phosphate (KPi) pH 7.5 and 1 mM dithiothreitol (DTT). A portion, set aside on ice, was used as whole cell suspension. Whole cells were permeabilized by incubation in the presence of 0.1 % CTAB for 10 min at 4°C or 0.04% toluene when indicated. A few crystals of DNase I were added to a fraction of the whole cell suspension, and cells were disrupted by 6 cycles of 30 sec sonication and 30 sec cooling on ice. Undisrupted cells were removed by 5 min centrifugation at 20,000 × g. An aliquot of the cell-free extract was kept for enzyme assays. The remaining part was centrifuged for 90 min at 140,000 × g. The supernatant containing the soluble proteins was transferred to a glass vial and stored at 4°C. The pellet containing the membranes was resuspended in buffer A. Soluble and membrane fractions were stored at 4°C under a 100 % N_2 gas phase.

Protein concentration of the extracts was determined with the micro biuret method with bovine serum albumin as standard [242]. Formate dehydrogenase (FDH), chlorophenol reductive dehalogenase (CPR), and fumarate reductase (FRD) activities in cell extracts were determined by spectroscopic recording of changes in the reduced methyl viologen concentration at 578 nm as described previously [243]. Cl-OHPA and OHPA were measured by high performance liquid chromatography (HPLC) as described elsewhere [243].

Quinone-oxidation/reduction and extraction

To investigate the possible involvement of menaquinones (MK) in Cl-OHPA and fumarate respiration, concentrated cell suspensions (FC or FF) were incubated for 10 minutes in the presence of 10 mM formate. Dithionite was added to cell suspensions prior to addition of fumarate or Cl-OPHPA, to ensure total reduction of quinones at the start of the experiment. The reduced MKs became oxidized by addition of Cl-OHPA (10 mM) to formate and dithionite-reduced formate-Cl-OHPA grown cells, or when fumarate (10 mM) was added to formate and dithionite reduced FF grown cells. Quinones were extracted in methanol-petroleum ether and separated by HPLC as described previously [244]. The detection of quinones was followed spectrophotometrically at 245nm. The absorption spectrum of Vitamin K was used as control. Instability of (chemically prepared) menaquinol prevented reliable quantification of the quinol extinction coefficient. Therefore, the ratio between the peak areas representing, reduced (MKH_2) or oxidised (MK) menaquinones in the HPLC chromatogram was used as indicator of the degree of quinone reduction.

UV-visible spectroscopic analysis of cell suspensions

Cultures (2 L) of *D. dehalogenans* were grown with formate as electron donor and either fumarate or Cl-OHPA as electron acceptor. Cells were harvested and washed with 8 ml buffer containing 50 mM Tris-HCl pH 8.0 and 0.5 mM DTT resuspended in 50 mM Tris-HCl pH 8.0 to a final concentration of 0.36 g wet cell weight ml^{-1} in 1-ml N_2-flushed cuvettes. All handling was done under anoxic conditions. 90 µl of the concentrated cell suspension was resuspended in 0.9 ml 50 mM Tris-HCl (pH 8.0.) After incubation with reducing or oxidizing substance(s), the suspension was frozen in liquid N_2 and the spectrum (320-700 nm) was recorded at 77 K, on a DW-2a™ spectrophotometer (American Instrument Co.) as described in [245]. After recording, the base lines were corrected for light scattering by subtracting a straight line between 540 and 580 nm. The corrected spectra were deconvoluted into two Gaussian peaks of variable width, one signal (at 552 nm) representing cytochrome c and another signal (at 561 nm) representing cytochrome b. The spectra were analyzed as described in [246]. Peak areas were taken to be proportional to the reduction level of the cytochromes. Involvement of cytochromes in electron transfer from formate to fumarate or Cl-OHPA was investigated by recording the changes in their redox state after reduction of the cells with 1 mM formate and re-oxidation by addition of 10 mM Cl-OHPA or fumarate.

Protein extraction and proteomic analysis

Triplicate samples were prepared for proteomic analysis as described before [115], with the modification that the protein samples were run only over a short distance, approximately 1 cm, in the SDS-PAGE prior to in gel trypsin digestion of proteins. This was done to be able to cut and process entire lanes as single samples. Detection, identification and relative quantification of proteins were done as described before [115]. In brief, trypsin digested peptides were run on a Proxeon LTQ-Orbitrap XL nLC–MS and all spectra data obtained were analysed with MAXQUANT v. 1.3.0.5 [143] using a protein database generated from the *D. dehalogenans* genome. Filtering and further bioinformatic analysis of the MAXQUANT/Andromeda workflow output and the analysis of the abundances of the identified proteins were performed with the PERSEUS v. 1.3 module (available at the MAXQUANT suite). Only proteins showing a false discovery rate of less than 1% were accepted. Reversed hits were deleted from the MAXQUANT result table as well as all proteins showing less than 2 peptides or less than 1 unique peptide or with label free quantification (LFQ) values of 0 for both sample and control. Zero values for one of the Log LFQ columns were replaced by a value of 5.5 (just below the lowest measured value) to make sensible ratio calculations possible. Relative protein quantification of sample to control was conducted with PERSEUS by applying a two sample t-test using the 'Log LFQ intensity' columns obtained with FDR threshold set to 0.05 and S0=1. S0 = 1 indicates that both the average proteins abundance ratio as well as the t-tests p-value have equal weight to decide whether the protein is significantly different between the two different growth conditions. Whether a protein concentration changed significantly was also determined with PERSEUS using t-tests p-value and the average protein abundance ratio between the different growth conditions.

Results

Genome analysis of the ortho-chlorophenol respiring *D. dehalogenans*

The genome of *D. dehalogenans* is 4.3 Mbp with a G+C content of 45%, and encodes six predicted reductive dehalogenases. The general genome characteristics were determined and compared to the two published *Desulfitobacterium hafniense* genomes (Table 1) [19, 26].

Table 1. General genome features of *D. dehalogenans, D. hafniense* DCB-2 and Y51. # Genes, total number of predicted genes; # CDS, total number of predicted coding sequences; # 16S rRNA, number of 16S ribosomal RNA encoding genes; # Rdh, number of predicted reductive dehalogenase encoding genes.

Strain	Size (Mbp)	GC (%)	# Genes	# CDS	# 16S rRNA	# Rdh
D. dehalogenans JW/IU-DC1^T	4.3	45	4252	4142	6	6
D. hafniense DCB-2^T	5.3	48	5042	4953	5	7
D. hafniense Y51	5.7	47	5208	5060	6	1

We focused our initial genomic analysis on genes coding for enzymes of metabolic pathways expected to be active under the growth conditions tested, including respiration with formate as electron donor and fumarate or Cl-OHPA as electron acceptor (Table 2).

The genome of *D. dehalogenans* contains two predicted formate dehydrogenase (FDH) encoding gene clusters, *fdh$_1$ABCE* and *fdh$_2$AFEH* (Table 2). *Fdh$_1$ABCE* encodes a three subunit membrane-bound formate dehydrogenase, consisting of a molybdenum-containing catalytic subunit, an Fe-S protein, and a membrane bound cytochrome b. Both *fdh$_1$A* and *fdh$_1$B* encode TAT signal sequences indicating that this complex is located at the outer face of the membrane. The second formate dehydrogenase, *Fdh$_2$AFEH,* encodes an cytoplasmic NADH-dependent formate dehydrogenase, consisting of a catalytic subunit, two (21 and 51 kD) subunits involved in oxidation/reduction of NADH/NAD$^+$ and a 5-formyltetrahydrofolate cyclo-ligase.

The gene cluster *frdCAB* encodes a membrane bound fumarate reductase/succinate dehydrogenase, consisting of an FAD-containing catalytic subunit, an Fe-S protein and a membrane bound cytochrome b. The genes of that complex do not encode any signal sequences and is therefore predicted to be located at the cytoplasmic face of the membrane.

We identified four gene clusters predicted to encode membrane-bound hydrogenases, including three NiFe uptake hydrogenases and an energy-conserving hydrogenase (Hyc, Table 2). Two of these hydrogenases have previously been identified as potentially involved in organohalide respiration [119]. The first is an uptake hydrogenase, encoded by *hydABC* (desde_2201-2203), consisting of a NiFe hydrogenase catalytic subunit, an Fe-S protein and a membrane-bound cytochrome b (Table 2). The second, a putative energy-conserving hydrogenase (Hyc), is encoded by a six-gene cluster (3649-3654), of which four code for membrane proteins (Table 2). This cluster has some similarity (22-36% amino acid identity) with the formate hydrogen lyase-4 complex encoding gene cluster (*hyfABCDEFGHIR-focB*) described in *E. coli* [248]. In fact, the *D. dehalogenans hyc* operon reported here has previously been described as a *hyf* operon, based on sequence analysis of transposon insertion sites in organohalide respiration-deficient mutants of this strain [119]. Unlike the *E. coli* encoded *hyf* operon the *D. dehalogenans hyc* operon does not encode a formate dehydrogenase.

The *cprTKZEBACD* gene cluster encodes the well-characterized Cl-OHPA reductive dehalogenase (CprA, Table 2). The catalytic subunit, CprA, contains a TAT signal peptide, two Fe-S binding domains, and the holoenzyme contains a corrinoid as co-factor. This protein is thought to be anchored to the membrane by CprB, a small protein consisting of 103 amino acids, predicted to contain three transmembrane helices. The remaining genes in this gene cluster encode accessory proteins involved in maturation and regulation. The *cprTKZEBACD* gene cluster is also present in *Desulfitobacterium hafniense* DCB-2, and has been characterised in detail in both organisms [64, 71, 73]. The genome encodes a complete menaquinone synthesis pathway (*menFDHCEBIAG*) [249]. We found nine flavoproteins resembling FAD binding proteins present in a wide range of bacteria, often associated with electron transfer or reduction of terminal electron acceptors. Among these are the FAD binding c terminus of the Fe^{3+} induced flavocytochrome c fumarate reductase, Ifc3, from *Shewanella frigidimarina* NCIMB400 (Table 2), or a cytochrome c dependent methacrylate reductase from *Geobacter sulfurreducens* AM-1. However, these predicted flavoproteins from *D. dehalogenans*, unlike Ifc3, do not contain any N-terminal heme binding, CxxCH, motifs (data not shown) [250-253].

Two of these flavoproteins are annotated as flavocytochrome c, and the remaining seven as succinate dehydrogenase/fumarate reductases. All nine flavoproteins contain signal sequences, suggesting that they are exported outside the cell. Four of the flavoproteins contain an FMN binding domain in addition to the FAD-containing fumarate reductase-like domain (Table 2).

Table 2. Selected proteins potentially involved in the respiration of *D. dehalogenans* using formate as electron donor and either fumarate or Cl-OHPA as electron acceptors. Not shown prefixes to locus tags are desde, dsy and Dhaf for *D. dehalogenans*; *D. hafniense* Y51 or DCB-2, respectively. TMH, transmembrane helixes as predicted with TMHMM v 2.0 [110], S. Sq, predicted signal sequence, * signal sequence only detected after analysis with signal P in sensitive mode [247], $ only detected after analysis with signal P in sensitive mode and using alternative start codon. # overlaps with TAT signal sequence. ND not determined

Gene	Locus tag			Annotation	TMH	S. Sq.
	D. deh.	*D. h.* Y51	*D. h.* DCB-2			
fdh$_1$A	3637	3101	4271	Formate dehydrogenase, catalytic	no	yes*
fdh$_1$B	3636	3100	4270	Formate dehydrogenase, Fe-S	yes	yes
fdh$_1$C	3635	3099	4269	Formate dehydrogenase, cytochrome b	yes	no
fdh$_1$E	3634	3098	4268	Uncharacterized protein involved in formate dehydrogenase formation	no	no
fdh$_2$H	1315	3972	1395	5-formyltetrahydrofolate cyclo-ligase	no	no
fdh$_2$E	1316	3971	1396	NADH:ubiquinone oxidoreductase 24 kD subunit	no	no
fdh$_2$F	1317	3970	1397	NADH:ubiquinone oxidoreductase, 51 kD subunit	no	no
fdh$_2$A	1318	3969-68	1398	NAD-dependent formate dehydrogenase catalytic subunit	no	no
frdC	616	735	743	Fumarate reductase cytochrome b	yes	no
frdA	617	736	744	Menaquinone-dependent fumarate reductase, flavoprotein subunit	no	no
frdB	618	737	745	Fumarate reductase iron-sulfur protein	no	no
hydD	2200	1599	2740	Hydrogenase maturation protease	no	no
hydC	2201	1598	2741	Ni/Fe-hydrogenase, b-type cytochrome	yes	no
hydB	2202	1597	2742	Ni,Fe-hydrogenase I large subunit	no	no
hydA	2203	1596	2743	Ni,Fe-hydrogenase I small subunit	no	no$
-	3649	3114	4283	Ni,Fe-hydrogenase III small subunit	no	no
-	3650	3115	4284	Ni,Fe-hydrogenase III large subunit	no	no
-	3651	3116	4285	Formate hydrogenlyase subunit 3/multisubunit Na+/H+ antiporter, MnhD subunit	yes	no
-	3652	3117	4286	Hydrogenase4 membrane component (E)	yes	no
-	3653	3118	4287	Formate hydrogenlyase subunit 4	yes	no
-	3654	3119	4288	Formate hydrogenlyase subunit 3/multisubunit Na+/H+ antiporter, MnhD subunit	yes	no
cprT	602	-	732	FKBP-type peptidyl-prolyl cis-trans isomerase (trigger factor)	no	no
cprK	603	-	733	cAMP-binding protein	no	no
cprZ	604	-	734	Protein of unknown function (DUF2869)	no	no
cprE	605	-	735	Chaperonin GroEL	no	no
cprB	606	-	736	Putative membrane anchor	yes	no
cprA	607	-	737	Reductive dehalogenase	no	yes
cprC	608	-	738	FMN-binding protein	yes	yes
cprD	610	-	740	Chaperonin GroEL	no	no
-	212	285	232	Succinate dehydrogenase/fumarate reductase flavoprotein subunit (FAD)	no	yes*
-	309	-	-	Flavocytochrome c (FAD-FMN)	no	yes
-	595	-	-	Succinate dehydrogenase/fumarate reductase flavoprotein subunit (FAD)	no	yes*

-	643	-	-	Succinate dehydrogenase/fumarate reductase flavoprotein subunit (FMN-FAD)	no	yes
-	646	764	773	Succinate dehydrogenase/fumarate reductase flavoprotein subunit (FAD)	no	no
-	650	768	779	Succinate dehydrogenase/fumarate reductase flavoprotein subunit (FAD)	no	yes
-	2005	1391	2507	Succinate dehydrogenase/fumarate reductase flavoprotein subunit (FAD)	no	yes*
-	3368	2816	3960	Succinate dehydrogenase/fumarate reductase flavoprotein subunit (FMN-FAD)	yes[#]	yes
-	3673	3139	4309	Flavocytochrome c (FAD-FMN)	no	yes
nrfA	3613	3065	4234	Formate-dependent nitrite reductase, periplasmic cytochrome c552 subunit	no	yes*
nrfH	3614	3066	4235	Respiratory nitrite reductase specific menaquinol-cytochrome-c reductase	yes	no
nrfA	3046	2472	3631	Formate-dependent nitrite reductase, periplasmic cytochrome c552 subunit	no	yes
nrfH	3045	2471	3630	Nitrate/TMAO reductase, membrane-bound tetraheme cytochrome c subunit	yes	no
fixX	2231	1626	2773	Ferredoxin-like protein	no	no
fixC	2232	1627	2774	Flavin-dependent dehydrogenase	no	no
fixB	2233	1628	2775	Electron transfer flavoprotein, α subunit	no	no
fixA	2234	1629	2776	Electron transfer flavoprotein, β subunit	no	no
menF	1960	517	469	Isochorismate synthases	ND	ND
menD	1961	518	470	2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic-acid synthase	ND	ND
menH	1531	519	471	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase	ND	ND
menC	1964	522	474	O-succinylbenzoic acid (OSB) synthetase	ND	ND
menE	1963	521	473	O-succinylbenzoate-CoA ligase	ND	ND
menB	1962	520	472	Naphthoate synthase (dihydroxynaphthoic acid synthetase)	ND	ND
menI	2535	1906	3066	1,4-dihydroxy-2-naphthoyl-CoA thioesterase	ND	ND
menA	3397	2871	4028	1,4-dihydroxy-2-naphthoate octaprenyltransferase	ND	ND
menG	2536	1907	3067	Menaquinone biosynthesis methyltransferases	ND	ND

In addition, the genome encodes two predicted cytochrome c nitrite reductases, NrfAH, thought to link the menaquinones with periplasmic nitrite reductase [254]. Finally, the genome encodes a cluster of predicted membrane-associated electron-transferring flavodoxin and ferredoxin proteins, resembling FixABCX (Table 2). These proteins are involved in generating highly reduced ferredoxin that act as electron donor to the dinitrogenase in the diazotrophic *Rhodospirillum rubrum* [255, 256]. It has been speculated that the products of this gene cluster also may provide low redox potential electrons for organohalide respiration [130].

Enzyme activity

Formate alone does not allow fermentative growth of *D. dehalogenans*. Addition of fumarate or Cl-OHPA as electron acceptor led to exponential growth, with doubling times of 8 to 10 h. In the late-exponential phase, the FC and FF grown cultures reached an OD600 of 0.17 and 0.25, respectively.

Subcellular localization of key enzyme activities

Localization of formate dehydrogenase, chlorophenol reductive dehalogenase and fumarate reductase activity was determined in FC and FF grown cells by *in-vitro* enzyme assays using the membrane impermeable artificial electron donor methyl viologen (Table 3). The membrane and soluble fractions contained 90 to 100% of the activity that was measured in the cell extracts, except for CPR for which only 48% of the activity could be recovered, possibly due to partial inactivation or loss of co-factors. Furthermore, 20% and 64% of the total protein from the cell extract was recovered from the membrane and soluble fraction of FC and FF grown cells, respectively. The activities measured in the membrane and soluble fractions were used to predict the localization of enzymes.

In concentrated cell suspensions of formate-Cl-OHPA grown cells, an FDH activity of 14 nmol formate min^{-1} mg $protein^{-1}$ was measured, which increased 7.7-fold after permeabilization of the cells. In FF grown cells, the specific FDH activity in concentrated cells was 42% of that observed in cell extract and increased 2.8-fold upon permeabilization of the cells with CTAB. We found for both growth conditions that the specific activity of FDH was higher in the soluble fraction than in the membrane fraction, whereas over 90% of the total activity was recovered from the membrane fraction (Table 3).

Table 3. Specific activity of formate dehydrogenase (FDH), chlorophenol reductive dehalogenase (CPR) and fumarate reductase (FRD) in different fractions of *Desulfitobacterium dehalogenans* cells, grown with formate as electron donor and Cl-OHPA or fumarate as electron acceptor. The ratio of the total activity of the cell extract that is recovered from the membrane and cytoplasm fraction is given between brackets.

		Specific activity[a]				
Enzyme[g]	e-acceptor	Concentrated cells	Permeabilized cells[b]	CFE[c]	Membrane Fraction (%)[f]	soluble fraction (%)[f]
FDH	Cl-OHPA[d]	14	108	59	189 (60)	973 (39)
	Fumarate	47	133	110	167 (90)	654 (32)
CPR	Cl-OHPA	51	87[c]	89	234 (48)	7 (0)
FRD	Fumarate	37	151	294	457 (92)	0 (0)
Total protein	FC/FF mg	174/94	174/94	180/96	32/56	4/5

[a]: Specific activity is expressed as nmol of substrate converted per minute per milligram of protein, [b]: Permeabilization with 0.1% CTAB, [c]: Cells were permealized with 0.04% toluene to avoid inactivation of CPR by CTAB., [d]: data from [243], [e]: Cell free extract [f]:Values in parentheses are the ratios of the total activities of the cell extracts recovered from the membrane and cytoplasm fractions., [g]: FDH, formate dehydrogenase; CPR, chlorophenol reductive dehalogenase; FRD, fumarate reductase.

FRD was found to be membrane bound and the activity increased 4-fold when cells were permeabilized with CTAB. CPR activity increased 1.7-fold upon permeabilization with toluene, which was used because CPR was found to be inhibited by CTAB. It should be noted that FRD and FDH activities determined in toluene-permeabilized cells were 4 to 5 times lower than those determined in CTAB-treated cells, indicating that permeabilization with toluene was less efficient or may have an inhibitory effect of these (data not shown). CPR activity was mainly found in the membrane fraction.

These results suggest that FDH is membrane associated. The fact that activity was observed on both sides of the cytoplasmic membrane is in agreement with the finding that the genome encodes two FDHs, including a predicted membrane bound protein orientated towards the outside and a soluble cytoplasmic protein. Both CPR and FRD activity were found to be membrane bound. FRD was determined to be orientated at the cytoplasmic face of the cell membrane.

Menaquinone

Quinones were extracted from FC and FF grown cells, and their involvement in fumarate and organohalide respiration was investigated. Extracted quinones exhibited an UV absorbance spectrum characteristic for menaquinones (MK, data not shown). In their oxidized form, the isolated quinones showed two peaks at 245 and 270 nm and a broad absorption centered around 325 nm, whereas the dithionite-reduced samples exhibited a single peak at 245 nm and a broad absorption at 330 nm.

The reduced MK became oxidized by addition of Cl-OHPA (10 mM) to formate and dithionite-reduced formate-Cl-OHPA grown cells, or when fumarate (10 mM) was added to formate and dithionite reduced FF grown cells (Table 4). These findings indicate that MK is involved in electron transport from formate to both Cl-OHPA or fumarate.

Table 4. The ratio of peak areas of reduced and oxidized menaquinone extracted under different conditions from cells of *D. dehalogenans* grown with formate as electron donor and fumarate or Cl-OHPA as electron acceptor. Cells were incubated for 10 minutes with 10 mM formate, Cl-OHPA or fumarate prior to quinone extraction.

	MKH_2/MK ratio for cells grown with:	
Conditions	Formate/fumarate	Formate/Cl-OHPA
Formate	0.33	1.79
Cl-OHPA	nd[a]	0.09
Fumarate	0.01	nd[a]
No addition	0.12	0.30

[a] nd: not determined.

UV-visible and EPR spectroscopic analysis

The absorption spectra were recorded from concentrated cell suspensions of FF and FC grown cultures, to determine the involvement of cytochromes in fumarate and Cl-OHPA respiration. Changes in the redox state of cytochromes were followed by visible spectrometry at 552 and 561 nm (Fig. 1). The low concentration of cytochrome b in formate-Cl-OHPA cells did not allow us to determine changes in its redox state, whereas the concentration of both cytochrome b and c combined were sufficient to study their involvement in respiration during growth on formate and fumarate (Fig. 1). The maxima between 550 and 560 nm, seen in the absorption spectrum of concentrated cell suspensions of FF and FC grown cultures, indicated that the cytochromes were partially reduced (Fig. 1, trace A and F). Upon addition of formate to concentrated cell suspensions, a rapid reduction of the cytochromes could be seen (Fig. 1, trace B and G), which is similar to the level of reduction that was obtained upon addition of dithionite (Fig. 1, trace C and H). Incubation of concentrated formate-Cl-OHPA cells with Cl-OHPA or concentrated FF cells with fumarate resulted in a decrease of the absorption peaks, indicating that the cytochromes became oxidized (Fig. 1, trace D and I). Upon addition of 10 mM Cl-OHPA to formate-reduced formate-Cl-OHPA cells the cytochrome c became oxidized (Fig 1, trace E), although a complete re-oxidation of cytochrome c was neither observed at low nor at high Cl-OHPA concentrations. A complete re-oxidation of both cytochrome B and C was observed when 10 mM fumarate was added to formate-reduced FF cells (Fig. 1, trace J).

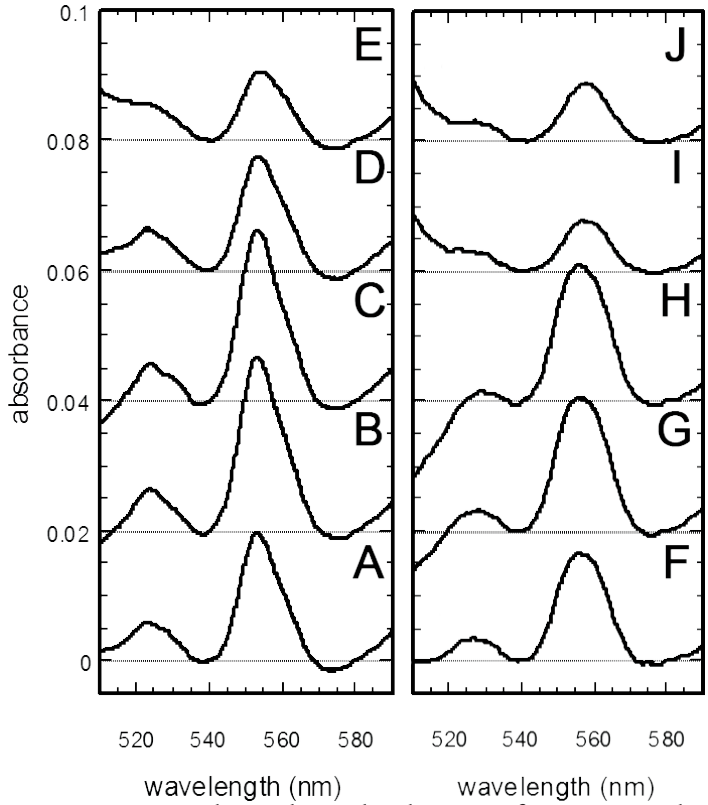


Figure 1. Low temperature cytochrome b, c α-band spectra of a concentrated cell suspension of *Desulfitobacterium dehalogenans* grown with formate as electron donor and either Cl-OHPA (trace **A**-**E**) or fumarate (trace **F**-**J**) as electron acceptor. Spectra were shifted upwards in steps of 0.02 absorbance units to clarify the Figure. Traces **A** and **F** show the absorption spectra from cells without additions. Traces **B** and **G** were recorded after incubation of the cells with 10 mM formate. Traces **C** and **H** were recorded after reduction of the sample with dithionite for 8 minutes. Traces **D** and **I** were recorded after incubation of the cells with 1 mM formate and 10 mM of Cl-OHPA or fumarate. Traces **E** and **J** were recorded after incubation of the cells with 10 mM Cl-OHPA or fumarate.

Electron paramagnetic resonance [232] spectra of concentrated formate-Cl-OHPA cells after addition of formate, and after addition of formate and Cl-OHPA, were recorded to obtain information on the involvement of Fe-S clusters and molybdenum. A detailed description of these experiments can be found in Supplementary File 1.

Comparison of the proteome of FC, FF and P grown cells

Shotgun proteomics was used to identify proteins produced at different levels in cells grown by respiration or fermentation. To this end, we compared the proteomes of FC and FF vs P grown cells. Furthermore, we compared proteomes of cells grown by organohalide vs fumarate respiration (FC vs. FF). We identified in total 578 different proteins out of a theoretical maximum of 4142 (Supplementary Table S1). In total 104 proteins were present at significantly ($P<0.05$) different levels between at least two growth conditions (Supplementary Table S1). Largest variations were seen between respiratory and fermentative growth, with 79 and 70 proteins varying significantly ($P<0.05$), between FC or FF vs P grown cells whereas only 20 proteins were found to differ significantly between FC and FF grown cells (Supplementary Table S1). For many enzyme complexes, we only detected the non-membrane embedded components. As extracting and detecting membrane proteins is still a challenging task[257], we assume that the values for the detected components are representative for the entire complex.

The catalytic subunit of the membrane-bound formate dehydrogenase complex, FdhA (desde_3637), increased 59 and 9 fold in FF or FC cells compared to P grown cells, respectively. However, only the first value was found to be statistically significant ($P<0.05$, Supplementary Table S1 Fig. 2). This upregulation of FdhA for both respiratory growth conditions compared to fermentative growth indicates that the same formate dehydrogenase is used when Cl-OHPA or fumarate are used as electron acceptor. In contrast, the cytoplasmic NADH-dependent formate dehydrogenase encoded by *fdh$_2$AFEH* was present in equal amounts both during respiratory and fermentative growth, indicating that it is part of other cellular processes, rather than energy conservation by respiration with formate as electron donor (Supplementary Table S1).

Two elements of the membrane-bound fumarate reductase encoded by *frdAB* (desde_617-618) were upregulated (10-14 fold) under both respiratory growth conditions, suggesting that their synthesis is regulated by redox and energy state, rather than presence of the substrate.

All elements encoded by the *cprTKZEBACD* gene cluster were identified, except the transcriptional regulator CprK and the highly hydrophobic CprB. This gene cluster was, as expected, highly (>270 fold) upregulated when *D. dehalogenans* was grown with Cl-OHPA (Fig. 2A and C, Supplementary Table S1).

Two putative extracellular electron-transferring flavoproteins were present in the proteome. A flavoprotein (Desde_3368) predicted to harbour two flavin-binding domains, an N-terminal FMN and a C-terminal FAD, increased strongly in abundance (>1000 fold) when *D. dehalogenans* was grown with Cl-OHPA. This increase is of the same magnitude as observed for elements of the *cpr* gene cluster. This suggests that this protein is involved in organohalide respiration, potentially as an electron shuttle between the MKs and CprA. The second identified flavoprotein (Desde_3673), also predicted to harbour two flavin binding domains, an N-terminal FAD and a C-terminal FMN, was found in high amounts at all growth conditions (Fig. 2, Supplementary Table S1).

Interestingly, we identified a protein annotated as cytochrome c nitrite reductase NrfA (Desde_3613), which is the outward facing component of the membrane bound NrfAH quinol dehydrogenase/nitrite reductase. Unfortunately, NrfA was not identified in all replica samples, preventing firm conclusions on eventual changes in abundance between growth conditions to be drawn. (Fig. 2, Supplementary Table S1).

The electron-transferring flavoproteins encoded by *fixABCX* (Desde_2231-34) were detected at all growth conditions. Their levels, however, did not vary significantly between growth conditions (Fig. 2, Supplementary Table S1).

An uptake hydrogenase (HydABC, Desde_2200-2203), and an energy-conserving hydrogenase (Hyc, Desde_3649-3654) have previously been suggested to play a role in organohalide respiration [119]. Subunits of the uptake hydrogenase HydABC were upregulated 10-18 fold during respiratory growth (FC and FF) compared to growth by fermentation (P). This suggests that expression is regulated by redox and energy state, rather than presence of the substrate. In contrast, we did not detect any subunits of the two alternative uptake hydrogenases in the proteome (Fig. 2, Supplementary Table S1).

Elements of the energy-conserving hydrogenase (Hyc) were not detected. Four of the six subunits of this complex contains several transmembrane helices, and are therefore tightly embedded in the membrane (Table 2). We cannot completely rule out that this complex is present in *D. dehalogenans* under the tested growth conditions, but has not been extracted from the membrane, and thus was not observed in the proteome analysis.

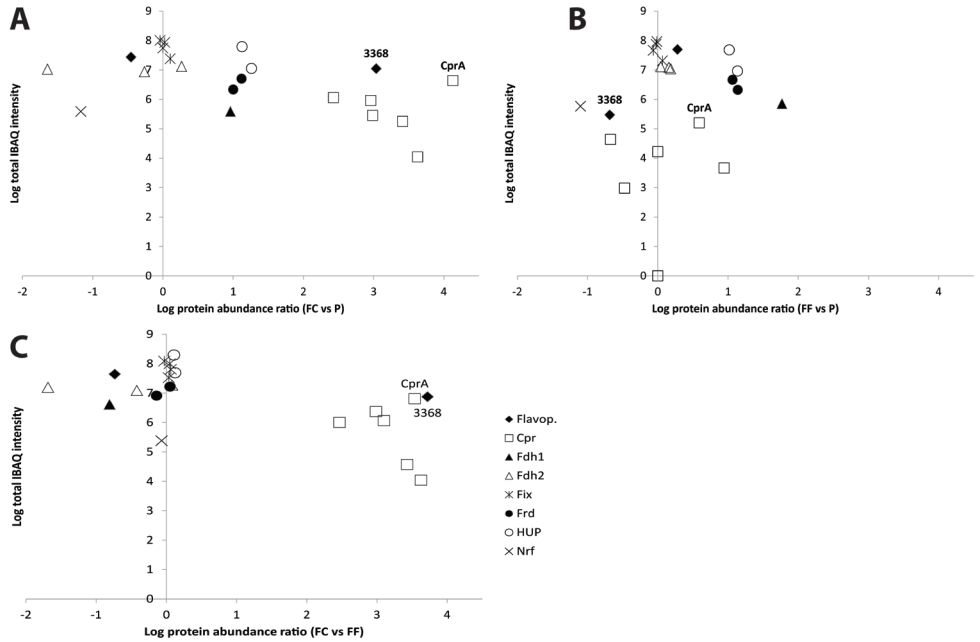


Figure 2. Changes in relative protein abundance between growth conditions where Formate and either Cl-OHPA or Fumarate (20 mM) was used as electron donors or accepters or growth by fermentation of Pyruvate (40 mM) **A**) Formate/Cl-OHPA vs Pyruvate **B**) Formate/Fumarate vs Pyruvate and **C**) Formate/Cl-OHPA vs Formate/fumarate. The logarithm of the total intensity based absolute quantification intensity (IBAQ) is plotted against the logarithm of the total protein abundance ratio, for selected proteins, (see also Fig. 3) The data points for CprA and the flavoprotein encoded by Desde_3368 are highlighted in the graphs.

Discussion

Formate-fumarate respiration

The genome of *D. dehalogenans* encodes two FDHs and one FRD. FDH_1 and FRD were found to be upregulated during both respiratory (FC and FF) growth conditions. FDH_1 and FRD from *D. dehalogenans* resemble the well characterised FDH and FRD of *Wolinella succinogenes* (Fig. 2 and 3A, Supplementary Table S1) [258] suggesting a similar topology of the electron transport chain from formate to fumarate in both bacteria. Electrons are channelled from an outward-facing FDH complex containing an Fe-S protein and a cytochrome b integral membrane protein via menaquinones to an FRD complex facing the cytoplasm, thereby creating a proton motive force [258, 259]. These findings are supported by both the genome and localization studies.

The genome predicts the presence of both an outward facing and a cytoplasmic formate dehydrogenase (FDH_1 and FDH_2, Table 2). In agreement with this we found FDH activity at both sides of the membrane, whereas FRD activity was localized at the cytoplasmic face of the cell membrane (Table 3, Fig. 3A). The involvement of menaquinones as electron carriers was demonstrated by the observation that they could be reduced with formate and re-oxidized with fumarate or Cl-OHPA (Table 4).

The second formate dehydrogenase, FDH_2, an NAD dependent formate dehydrogenase, does not contain any signal sequences or transmembrane domains, indicating a cytoplasmic localisation. Proteomic data showed that FDH_2 was present in the cells in equal amounts in all three tested growth conditions (Fig. 2, Supplementary Table 1). In the same operon we identified a gene predicted to encode a 5-formyltetrahydrofolate cyclo-ligase (Desde_1315). Analysis by tBLASTn revealed that this gene organisation is commonly found in sequenced genomes of bacteria beyond the *Desulfitobacterium* genus (data not shown). Taken together, these data suggest that FDH_2 is not part of the electron transport chain from formate to fumarate or Cl-OHPA, but rather may be part of the methyl branch of the Wood-Ljungdahl pathway, catalysing the reduction of CO_2 to formate (for review see [260]). It has previously been demonstrated that *D. dehalogenans* is capable of CO_2 fixation [243].

Formate-Cl-OHPA respiration

The electron transport pathway from electron donor to acceptor during organohalide respiration is still poorly understood. Previously, a simple model analogous to formate-fumarate respiration has been proposed [74, 76, 235]. According to this, the electron-donating and electron-accepting reactions are orientated towards the periplasmic and cytoplamic face, respectively, thereby creating a proton motive force driving ATP synthesis [74, 76, 235]. The *cprA* gene, however, like other reductive dehalogenase encoding genes, codes for a pre-protein with a TAT signal sequence, strongly suggesting that CprA is exported across the cytoplasmic membrane 24, [3, 57]. In *Desulfitobacterium hafniense* strain Y51 and *Sulfurospirillum multivorans PceA* is constitutively expressed. In both strains PceA was found to be present on both sides of the cell membrane when grown with PCE, but PceA was found solely inside the cells when grown without PCE [56, 77]. It seems likely that the tightly regulated CprA in *Desulfitobacterium dehalogenans* also is exported across the cell membrane. The increase in CprA activity after permeabilization would then be due to the presence of folded and active CprA that has not yet been exported.

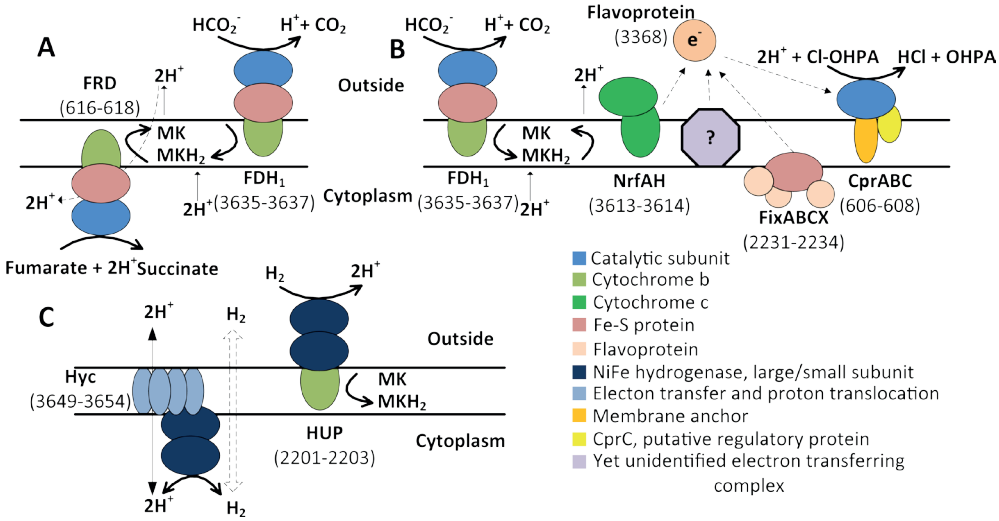


Figure 3. Overview of enzyme complexes shown to be, or predicted to be involved in energy conservation in *D. dehalogenans* growing with formate as electron donor and **A**) fumarate or **B**) Cl-OHPA as electron acceptor. **C**) putative interaction between the energy conserving hydrogenase Hyc and uptake hydrogenase HUP. Numbers in brackets indicate locus tags with the Desde_ prefix removed. Bold arrows show direction of reactions or movements, dashed arrows show potential direction of reactions or movements. FDH: formate dehydrogenase; FRD: fumarate dehydrogenase; NrfAH: cytochrome c quinol dehydrogenase/nitrite reductase; ?: unidentified membrane bound electron transfer complex; Fix: *fix* gene cluster products; CprABC; chloropehenol reductase; Hyc: energy conserving hydrogenase; HUP: uptake hydrogenase.

Our findings suggest that the electron transport chain from FDH_1 to Cl-OHPA consist of menaquinones, likely an extracellular flavoprotein and a membrane complex linking the menaquinones with extracellular respiratory processes (Fig. 3B). This is supported by the finding that menaquinones are oxidised by both Cl-OHPA and fumarate (Table 4). The involvement of menaquinones in the electron transport chain during organohalide respiration has also been shown for the closely related *Dehalobacter restrictus* [76]. The standard redox potential of menaquinones is -74 mV, whereas it has been demonstrated that PceA from *Sulfurospirillum multivorans* (initially named *Dehalospirillum multivorans*) requires an electron donor with a standard redox potential below -360 mV [82, 261]. Menaquinones are thus unlikely to deliver the electrons directly to the reductive dehalogenase. It has been suggested that strong reducing equivalents, necessary for reducing reductive dehalogenases, are generated by reverse electron flow or electron bifurcation [77, 82, 262]). Combining genome and proteomic data, we were able to identify some candidates for the link between the menaquinones and the periplasmic orientated CprA. (Fig. 3B).

One of these is NrfAH, a putative membrane bound cytochrome c-containing quinol dehydrogenase/nitrite reductase. The genome of *D. dehalogenans* encodes two NrfAH homologues (Table 2), of which one, Desde_3313-14, was identified in the proteomes from all tested growth conditions (Fig. 2 and 3B, supplementary Table S1). Some studies have found NrfH capable of donating electrons to other periplasmic electron donors than NrfA [254, 263], whereas other studies did not find evidence for such a mechanism [264]. This may suggest a role for NrfH as an electron hub between the menaquinones and extracellular electron acceptors, analoguous to the function of CymA in *Shewanella* spp. ([252] and references therein). Similar to this, it was recently suggested that an NapGH like quinol dehydrogenase acts as the link between the menaquinones and the tetrachloroethene reductase PceA in *Sulfurospirillum multivorans* [29]. The genome of *D. dehalogenans* does not encode any CymA homolog. Another candidate electron transport complex is encoded by the *fix*ABCX gene cluster (Desde_2231-2234), the products of which are involved in generation of reductants for nitrogen fixation by reversed electron flow in *Rhodospirillum rubrum* [255, 256]. As the genome of *D. dehalogenans* does not encode the minimum machinery necessary for nitrogen fixation [265], the FixABCX complex most likely is linked to other cellular processes in this bacterium. Recently, comparative transcriptomics of *D. hafniense* Y51 [131] and proteomics of *D. hafniense* TCE1 [130] showed expression of *fixABCX* to be induced under organohalide-respiring conditions, leading to the speculation that the encoded complex may provide low redox potential electrons for organohalide respiration [130]. In our proteomic dataset we saw no differences in the abundance of the *fixABCX* gene products between the tested growth conditions ($p<0.05$, Fig. 2 supplementary Table S1). The lack of differential expression indicates that if FixABCX would be part of the organohalide respiratory electron transport chain to CprA, it is also involved in other cellular processes. Furthermore FixABCX is located in the cytoplasm and loosely associated with the membrane [255]. If the FixABCX complex creates low potential electron donors needed for CprA activity, an as yet unidentified electron carrier must transport these across the membrane to the outward facing CprA (Fig. 3B).

Cytochromes were analysed in whole cells grown with Cl-OHPA or fumarate. The shifts in the absorption spectra of reduced cytochromes after addition of either Cl-OHPA or fumarate showed that cytochrome c could be completely oxidised by fumarate and partly by Cl-OHPA (Fig. 1). These findings would be in agreement with the activity of a membrane bound cytochrome c quinol dehydrogenase such as NrfAH. The oxidation of menaquinones by fumarate indirectly

leads to oxidation of NrfAH, although this complex is not part of the electron transport pathway from formate to fumarate (Fig. 3A). The seemingly less efficient oxidation of cytochrome c with Cl-OHPA might be due to partial loss of extracytoplasmic components of the electron transport chain from cytochrome c to Cl-OHPA, caused by handling of the samples (Fig. 3B).

The genomes of *D. dehalogenans*, *D. hafniense* strain Y51 and DCB-2 encode, unlike other metal-oxidising bacteria, only a limited number of cytochrome c's [19, 26], leading to the speculation that *Desulfitobacterium* spp. utilize other electron carriers than cytochrome c for organohalide respiration [19]. Two predicted extracytoplasmic electron-transferring flavoproteins (Desde_3368 and Desde_3673) were identified in the proteome in the present study. These resemble flavocytochome c's, but both lack known consensus heme binding domains (CxxCH) and are not found next to cytochrome c encoding genes in the genome. One of these flavoproteins (Desde_3368) increased in abundance in the presence of Cl-OHPA by three orders of magnitude, a similar increase as elements encoded by the *cprTKZEBACD* gene cluster (Fig. 2, Supplementary Table S1). This was a surprising finding as this type of flavoproteins has not previously been associated with organohalide respiration. Flavoproteins similar to those from *D. dehalogenans* have been found in a range of bacteria [266]. Usually these are linked with a tetraheme cytochrome c, either directly as a flavocytochrome c fusion protein or as a flavoprotein - cytochrome c enzyme complex [251-253]. In *Shewanella frigidimarina* a flavocytocrome c, Ifc_3, is induced in the presence of Fe^{3+}. It was speculated that Ifc3 acts as an electron shuttle between a membrane-bound quinol dehydrogenase, CymA, and extracellular electron acceptors [251]. A similar role has been suggested for other flavocytocrome c's [251]. A recent study on selenite reduction by *Shewanella oneidensis* MR-1 indicated that a deletion of *cymA* results in an inability to reduce selenite. Deleting the flavocytochrome c encoding gene *fccA* led to lower selenite reduction rates, which seems to be in agreement with a role for flavocytochrome c as an electron shuttle, although it cannot be ruled out that FccA acts as the terminal selenite reductase [267]. Methacrylate reduction in *Geobacter sulfurreducens* AM_1 was shown to be carried out by a periplasmic flavoprotein tetraheme cytochrome c complex [253]. The strong increase in abundance of Desde_3368 when the cells were grown in the presence of Cl-OHPA makes this flavoprotein (Desde_3368) a promising candidate for the yet unidentified electron shuttle between a membrane bound electron-donating complex and CprA (Fig. 2 and 3B). The genomes of *D. dehalogenans* and *D. hafniense* strains DCB-2 and Y51 encode 9, 26 and 30 homologues of these putative electron-transferring flavoproteins, respectively (Table 2) [19, 26]. Some of these have been annotated as

flavocytochrome c's, but all lack heme binding domains. The majority are, however, annotated as fumarate reductase/succinate dehydrogenase or flavoproteins (data not shown), showing that careful curation of automatically generated predictions of protein function is needed after more detailed functional studies.

One (Desde_2200-2203) out of the three uptake hydrogenases [148] encoded in the genome of *D. dehalogenans* increased 10 to 18 fold in abundance during respiratory growth compared to fermentative growth (Fig. 2, Supplementary Table S1). This is in line with previous findings of the involvement of hydrogenases in respiratory growth of *D. dehalogenans* [119]. Interestingly, a similar increase in expression of genes encoding an uptake hydrogenase corresponding to *hydABC* has been reported for *Desulfitobacterium hafniense* Y51 under respiratory growth compared to fermentative growth [131]. This points towards a more general role in respiratory metabolism rather than a specific role in organohalide respiration, potentially as part of a hydrogen recycling mechanism as described for diazotrophs and sulfate reducers [268-270]. No elements of the Hyc energy-conserving hydrogenase could be detected in the proteome. This may be due to the presence of transmembrane helixes in four of the six subunits, embedding it tightly in the cell membrane (Table 2). We cannot rule out that the Hyc complex is present in *D. dehalogenans* under the growth conditions tested. Based on the current data the exact role of HUP and Hyc cannot be established. It is tempting to speculate that they interact as speculatively shown in Fig. 3C, but whether they play a direct role in the respiratory electron transport or are involved in other metabolic processes as suggested for the organohalide-respiring *Dehalococcoides* spp. still requires further studies [120].

In this study we report the full genome sequence of *Desulfitobacterium dehalogenans* strain JW/IU-DC1^T, which combined with comparative proteomics and biochemical analysis allowed us to obtain an improved insight into FF and FC respiration in *D. dehalogenans*. FF respiration occurs in a way analogous to the well described mechanism for *Wolinella succinogenes* [258]. Our data show that the first part of the electron transport pathway from FDH to MKs is shared between FF and FC respiration. Although the exact link between the menaquinones and the outward facing CprA could not be determined, proteome analysis pointed to a possible role of gene products of *nrfAH* or alternatively *fixABCX*, but this link could also be carried out by other membrane complexes like Hyc, or the products of yet unidentified gene clusters. Finally, we identified an extracytoplasmic flavoprotein as a candidate for an electron shuttle between a membrane-bound electron-donating complex and the outward facing CprA.

Future investigations such as comparative as well as functional genomics of organohalide-respiring bacteria, including the characterization of knockout mutants, can be expected to further propel our understanding of the energy metabolism of desulfitobacteria and other organohalide-respiring bacteria.

Acknowledgements

This work was supported by the Ecogenomics Project of the Netherlands Genomics Initiative, as well as the IOP Environmental Biotechnology programme. We furthermore thank the European Community program FP7 (grants KBBE-211684; BACSIN, and KBBE-222625; METAEXPLORE) for financial support.

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Chapter 7

The *Desulfitobacterium* genome

Thomas Kruse, Willem M. de Vos and Hauke Smidt

Supplementary materials are available upon request

Manuscript in preparation

Abstract

The *Desulfitobacterium* genus constitutes anaerobic Gram-positive bacteria, the majority of which has been shown to be facultative organohalide respiring bacteria. Four genomes from members of this genus have previously been published, including three from *D. hafniense* strains and *D. dehalogenans* JW/IU-DC1^T. We here present the genomes of eight additional strains of *Desulfitobacterium* spp., including four strains of *D. hafniense*, one strain each from *D. dichloroeliminans* and *D. metallireducens*, and two strains that had not been assigned to any species prior to this study. The newly sequenced genomes were compared with the four previously published desulfitbacterial genomes. The average genome sizes are 5.5, 4.3 and 3.4 Mbp for *D. hafniense, D. dehalogenans and D. dichloroeliminans/metallireducens*, respectively. The genomes of ten out of the 12 available desulfitobacterial strains encoded one to seven reductive dehalogenases, whereas the genomes of both *D. hafniense* DP7 and *D. metallireducens* 853-15A^T did not encode any reductive dehalogenases. The latter result was a surprise as *D. metallireducens* 853-15A^T had previously been reported to utilize several chlorinated aromatic and aliphatic compounds as terminal electron acceptors. In general, reductive dehalogenase encoding genes are located in eight highly conserved gene clusters (dehaloclusters A-G), with each cluster being shared by two to four strains. The dehaloclusters do not show any signs of having being mobile or having been acquired recently by horizontal gene transfer, with the composite transposon encoding *pceABCT* as a notable exception. Dehalocluster C encodes five reductive dehalogenases, predicted to be involved in sequential dehalogenation of highly chlorinated phenols. In two of the three strains encoding CprA, enabling respiration with 3-chloro-4-hydroxyphenyl acetate (Cl-OHPA), an additional gene encoding a methyl-accepting chemotaxis protein was found to be part of the *cprTKZE-$_1$BACE$_2$O* gene cluster. Two of the reductive dehalogenases do not have the cognate membrane anchor-encoding gene *rdhB.* One of these was found in a gene cluster, *rdhSPAABMN*, encoding a two component regulator system and a membrane-bound quinol dehydrogenase, an organisation similar to what was recently reported from *Sulfurospirillum multivorans.* The genomes of *D. hafniense*, *D. dehalogenans* and *D. dichloroeliminans/ metallireducens* encode 44-65, 39-42 and 12-19 molybdopterin oxidoreductases, respectively. All sequenced desulfitobacterial genomes encode a complete cobalamin synthesis pathway. A menaquinone synthesis pathway was found in all strains except *D. dichloroeliminans* DCA1^T, explaining its requirement of exogenous menaquinones supplied in the growth medium. The availability of the genome sequence of twelve desulfitobacteria from four different species, confirmed the general view that this genus has an

extremely large metabolic repertoire. Future studies based on the insight gained from analysing these genomes will provide the framework for an improved understanding and exploitation of this exciting genus.

Introduction

The genus *Desulfitobacterium* belongs to the phylum *Firmicutes*, class *Clostridia*, order *Clostridiales*, family *Peptococcaceae*. These are Gram-positive anaerobic bacteria with a versatile metabolism, and all but two of the characterized isolates have been described as facultative organohalide-respiring bacteria (OHRB) [16]. The two exceptions are *D. hafniense* DP7 and *D. aromaticivorans* UKTL^T. The former was isolated from a human fecal sample, and has not been shown to degrade any chlorinated substrate [49]. The latter was isolated from a former coal gasification site, using toluene and ferric citrate as electron donor and acceptor, respectively, whereas the ability to carry out organohalide respiration (OHR) was not tested [50]. It should be noted, however, that indirect evidence for OHR stems from a study where 16S ribosomal RNA (rRNA) sequences with high similarity to that of *Desulfitobacterium aromaticivorans* UKTL^T were reported to be dominating in sediments where OHR of PCE took place [51].

The first member of the genus *Desulfitobacterium* was isolated in 1992 and named strain DCB-2 due to its similarity to *Desulfomonile tiedjei* DCB-1^T [43]. The publication of *D. dehalogenans* JW/IU-DC1^T introduced the *Desulfitobacterium* genus [44]. In 1996 strain DCB-2 was classified as the type strain of *Desulfitobacterium hafniense* [45]. In the following years a large number of additional strains of *Desulfitobacterium* have been isolated. *D. hafniense* strains PCP-1, TCE-1, DP7, TCP-A and G2 were originally published as members of a distinct species, *D. frappieri*. However, it was later shown that *D. hafniense* and *D. frappieri* both belong to the same species and have subsequently been merged under the name of the first described species, *D. hafniense* [16, 46]. The genus *Desulfitobacterium* contains at least six different species (Table 1, Figure 1). Currently five *Desulfitobacterium* species are listed in the "List of Prokaryotic names with Standing in Nomenclature" (LPSN), namely *D. hafniense*, *D. dehalogenans*, *D. metallireducens*, *D. aromaticivorans* and *D. chlororespirans* [47]. *D. dicloroeliminans* has not yet been included in LPSN, although the species name is generally recognised [16, 271].

Desulfitobacteria are characterised as capable of reducing sulfite but not sulfate [16, 44], although D. metallireducens 853-15AT does not reduce sulfite [272] and D. hafniense Y51 has been reported to reduce both sulfate and sulfite [273]. Furthermore, Desulfitobacterium spp. have been reported to use a wide range of organic and inorganic compounds as electron donors and terminal electron acceptors. For reviews see [16, 50].

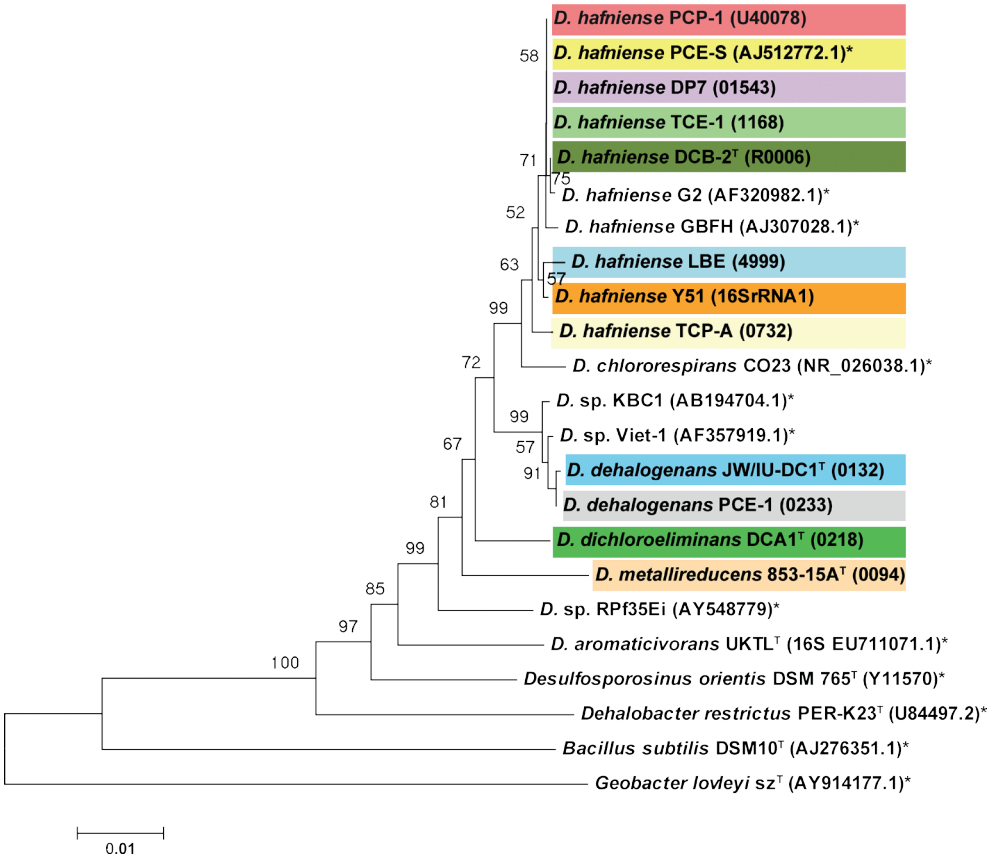


Figure 1. Phylogenetic tree based on 16S rRNA genes from *Desulfitobacterium* spp. and outgroup reference sequences. Sequences were aligned with Clustal W, [274], and a neighbour-joining tree was constructed and validated with 1000 bootstraps analysis using the MEGA5 software package [35]. Bootstrap values higher than 50% are given at corresponding nodes in the tree. The reference bar indicates 1% base substitutions per site. *Desulfitobacterium* is abbreviated *D.*; locus tags, with the prefix removed are given in parentheses; GenBank accession numbers of sequences when marked with an *. For colour coding and prefixes for locustags see (Table 1).

Table 1 Isolated representatives of the six *Desulfitobacterium* species currently described. *Desulfitobacterium* is abbreviated *D.* Locus tag prefixes are given for strains for which a full genome sequence is available. Locus tag prefixes are substituted with a colour as indicated in the remaining part of this chapter. DSM: Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. IMG: Integrated Microbial Genomes [34]. * BCCM/LMG, Belgian Culture Collection - Department of Biochemistry and Microbiology, Faculty of Sciences of Ghent University Belgium. $ Received as a kind gift from Jan Dirk van Elsas, Department of Microbial Ecology, University of Groningen, The Netherlands. [#] GenBank accession number, [€] European Molecular Biology Laboratory Database accession number. Δ, J. Maillard unpublished.

Name	Environment	Locus tag	DSM	IMG genome ID	Reference
D. hafniense DCB-2[T]	Sewage sludge, Denmark	Dhaf_	10664	CP001336[#]	[43, 45]
D. hafniense Y51	Soil contaminated with tetrachloroethene, Japan	DSY		AP008230[#]	[19, 273]
D. hafniense TCP-A	River sediment, Saale river Germany	DeshafDRAFT_	13557	2516653041	[275]
D. hafniense PCP-1	Methanogenic consortium from mixture of anaerobic sewage sludge and soil, Canada	A37YDRAFT_	12420	2513237149	[276]
D. hafniense TCE-1	Chloroethene-polluted soil, The Netherlands	DeshaDRAFT_	$	2513020008	[277]
D. hafniense DP7	Human feces of 28 year old female, The Netherlands	HMPREF0322_	13498	2522572064	[49]
D. hafniense LBE		DesLBEDRAFT_	Δ	2516493017	Δ
D. hafniense PCE-S	Soil from landfill site, Germany	DPCES_	14645	LK996017-LK996040[€]	[27, 170]
D. hafniense GBFH	River delta sediment, U.S.A.				[46]
D. hafniense G2	Subsurface clay bedding, U.S.A.		16228		[278]
D. dehalogenans JW/IU-DC1[T]	Freshwater pond sediment, U.S.A.	Desde_	9161	CP003348[#]	[44]
D. dehalogenans PCE-1	Chloroethene-polluted soil, The Netherlands	DesPCE1DRAFT_	10344	2512875014	[279]
D. dichloroeliminans DCA1[T]	Soil polluted with 1,2-DCA	Desdi_	LMG P-21439*	2507149019	[271]
D. metallireducens 853-15A[T]	Uranium-contaminated aquifer, U.S.A	Desme_	15288	2507149024	[272]
D. aromaticivorans UKTL[T]	Soil of a former coal gasification site, Poland		19510		[50]
D. chlororespirans CO23[T]	Compost soil, U.S.A.		11544		[280]

The genome of the tetrachloroethene (or perchloroethene; PCE)-degrading *D. hafniense* Y51 was published in 2006, only one year after the elucidation of the first two genomes of OHRB, namely those of *Dehalococcoides mccartyi* strains195^T (previously *Dehalococcoides ethenogenes* 195^T) and CBDB1 [17-19]. The *D. hafniense* Y51 genome has a size of 5.7 Mbp, which is more than 4 Mbp larger than those of *Dehalococcoides mccartyi* 195^T and CBDB1. Contrary to the 17-32 reductive dehalogenase found in currently available *Dehalococcoides mccartyi* genomes, the genome of *D. hafniense* Y51 only encodes a single reductive dehalogenase. The gene encoding this PCE-reductive dehalogenase are part of the *pceABCT* cluster that is flanked by two transposases, in what has been demonstrated to be a functional transposon [118]. Other notable findings from the *D. hafniense* Y51 genome were the presence of a complete corrinoid synthesis pathway, large numbers of hydrogenases including three HUP type NiFe uptake hydrogenases, 31 genes annotated to encode fumarate reductase flavoproteins and 57 molybdopterin oxidoreductases (pfam01568), the latter being the highest numbers reported for any sequenced bacterial genomes [19, 26]. This first desulfitobacterial genome was followed in 2012 and 2015 by the genomes of *D. hafniense* DCB-2^T, PCE-S and *D. dehalogenans* JW/IU-DC1^T [25-27]. These encode seven, two and six reductive dehalogenases respectively, whereof the majority are full length and thus likely to be functional. The genomes of both *D. hafniense* PCE-S and Y51 encode a copy of the *pceABCT* gene cluster, although the organisation of the transposon structure differs slightly between the two strains [27]. The well-characterised 3-chloro-4-hydroxyphenyl acetate (Cl-OHPA) reductase CprA encoded by the *cprTKZE$_1$BACE$_2$* gene cluster is encoded on the genome of both *D. hafniense* DCB-2^T and *D. dehalogenans* JW/IU-DC1^T [25, 26]. While the availability of additional *Desulfitobacterium* spp. genomes confirmed the picture of desulfitobacteria as bacteria with an extremely versatile metabolism, some interesting differences between the species could be observed. The size of the genomes of the three strains of *D. hafniense* ranges from 5.3-5.7 Mbp, which is considerably larger than the 4.3 Mbp *D. dehalogenans* JW/IU-DC1^T genome, suggesting that *D. dehalogenans* has a less versatile metabolism than *D. hafniense*. In line with this, it was found that all four desulfitobacterial genomes encode three HUP type NiFe uptake hydrogenases, and 27-31 and 10 genes were annotated as fumarate reductase flavoproteins and 53-57 or 39 molybdopterin oxidoreductase genes (pfam01568) were found in the genomes of *D. hafniense* and *D. dehalogenans* respectively.

The availability of full genome sequences has greatly aided in understanding the metabolism of *Desulfitobacterium* spp. For example analysis of the genome of *D. hafniense* DCB-2^T led to

identification of genes for nitrogen fixation, a trait that was later verified experimentally [26]. A microarray targeting all predicted genes from *D. hafniense* DCB-2^T has been designed and used for studying gene expression patterns under different growth conditions [26], and the proteomes of *D. hafniense* Y51 and *D. dehalogenans* JW/IU-DC1^T were obtained from cells grown with different electron donors and acceptors. For *D. dehalogenans* this led to identification of one of the annotated fumarate reductase flavoproteins as a possible component of the respiratory electron transport chain during organohalide respiration OHR [25, 131]. Similar comparative proteomic experiments have been done for *D. hafniense* TCE-1. In that study, a protein database generated from the genome of *D. hafniense* Y51 was used for identifying proteins, due to lack of a full genome sequence of strain TCE-1 at the time of the study [130].

We here present the genome sequence of eight additional members of the *Desulfitobacterium* genus, including *D. hafniense* TCE-1, bringing the total number of published genome sequences from this genus up to 12 (Tables 1 and 2), and covering four species, namely *D. dehalogenans*, *D. hafniense*, *D. dichloroeliminans* and *D. metallireducens*. We analysed and compared all 12 currently available genomes of *Desulfitobacterium* spp. with special emphasis on linking current knowledge on physiology with insights gained from genome analysis.

Table 2. Genome properties of all currently available desulfitobacterial genomes. *Desulfitobacterium* is abbreviated *D.*, * not counting the non canonical reductive dehalogenase, crdA, which is found in all desulfitobacteria except *D. metallireducens* 853-15AT

Strain	Size Mbp	GC %	scaffold #	Number of genes, % of total in ()					Reference
				16S	Total	With signal peptides	With trans membrane helices	*rdhA**	
D. hafniense DCB-2^T	5.3	48	1	5	5042	212 (4.2)	337 (6.7)	7	[26]
D. hafniense Y51	5.7	47	1	6	5208	304 (5.8)	818 (15.7)	1	[19]
D. hafniense TCP-A	5	47	2	3	4877	194 (40)	1221 (25.0)	5	This study
D. hafniense PCP-1	5.6	48	45	3	5358	347(6.5)	1395 (26.0)	7	This study
D. hafniense TCE-1	5.7	47	6	4	5570	232 (4.2)	1427 (25.6)	1	This study
D. hafniense DP7	5.2	48	151	1	5456	177 (3.2)	1370 (25.1)	0	This study
D. hafniense LBE	5.5	47	1	6	5341	232 (4.3)	1387 (26.0)	2	This study
D. hafniense PCE-S	5.7	47	24	2	5494	296 (5.5)	1624 (30.0)	2	[27]
D. dehalogenans JW/IU-DC1^T	4.3	45	1	6	4252	143 (3.4)	1129 (26.6)	6	[25]
D. dehalogenans PCE-1	4.2	45	1	3	4123	139 (3.4)	1086 (26.3)	6	This study
D. dichloroeliminans DCA1^T	3.6	44	1	6	3537	125 (3.5)	933 (26.4)	1	This study
D. metallireducens 853-15A^T	3.2	42	1	8	3152	77 (2.4)	841 (26.7)	0	This study

Materials and Methods

In this study, we obtained the genome sequence of eight strains of *Desulfitobacterium* spp. Genomic DNA of *Desulfitobacterium hafniense* strains TCP-A and LBE were provided, respectively, by Dr. Ute Lechner, Martin-Luther Universität, Halle-Wittenberg, Germany and Dr. Julien Maillard, Laboratory for Environmental Biotechnology, ENAC-IIE-LBE, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. Genomic DNA of *Desulfitobacterium hafniense* strains PCP-1, TCE-1, DP7, *Desulfitobacterium dichloroeliminans* DCA1^T, *Desulfitobacterium metallireducens* 853-15A^T and *Desulfitobacterium* sp. PCE-1 were obtained as described below.

Growth of strains

Desulfitobacterium spp. were obtained from different sources as listed in (Table 1), cultivated under anoxic conditions in 500 ml bottles containing 250 ml culture medium under an atmosphere of 100% N_2. *D. dichloroeliminans* DCA1^T was grown as described previously (De Wildeman, et al., 2003). *Desulfitobacterium metallireducens* 853-15A^T was grown in DSM 838 medium as recommended by the DSMZ. All other strains were cultivated in basal medium as described previously (Neumann, et al., 1994), supplemented with 1 g/l yeast extract, and with additional modifications from the original medium as follows. After autoclaving, the medium was supplemented with vitamins and trace elements from anaerobic filter sterilised (0.22 µm pore size) stocks giving final concentrations per litre medium of 1.34 µM EDTA; 10.06 µM $FeCl_2 \cdot 4H_2O$; 0.51 µM $MnCl_2 \cdot 4H_2O$; 0.8 µM $CoCl_2 \cdot 6H_2O$; 0.51 µM $ZnCl_2$; 0.02 µM $CuCl_2$; 0.04 µM $AlCl_3 \cdot 6H_2O$; H_3BO_3; 0.15 µM Na_2MoO_4; 0.1 µM $NiCl_2 \cdot 6H_2O$, 0.75 mM $CaCl_2 \cdot 2H_2O$; 0.5 mM $MgCl_2 \cdot 6H_2O$; 3 µg Na_2SeO_3; 9 µg $Na_2WO_4 \cdot 2H_2O$; 0.04 µl concentrated HCl; 50 µg biotin; 250 µg p-aminobenzoate; 50 µg panthothenate; 20 µg folic acid; 50 µg lipoic acid; 100 µg pyridoxine; 550 µg nicotinamide; 100 µg thiamine HCl; 50 µg riboflavine; 50 µg cyanocobalamine. Finally, reducing and buffer solutions were added to a final concentration of 1mM $Na_2S \cdot 9H_2O$, 5.6 mM NH_4HCO_3 and 44.4 mM $NaHCO_3$. Cultures were routinely grown fermentatively on 40 mM pyruvate at the temperatures recommended by DSMZ. Pyruvate was added from sterile anoxic stocks.

DNA extraction

For genome sequencing, cells grown fermentatively on pyruvate were harvested from early-stationary-phase cultures by centrifugation, resuspended in TE buffer (10mM Tris, 20 mM EDTA, pH 8), and stored at -80 °C, until DNA was extracted according to the cetyltrimethylammonium bromide (CTAB) protocol recommended by the DOE Joint Genome Institute (JGI, Walnut Creek, CA) [239]. DNA was stored in H_2O at -80 °C, prior to sequencing. Integrity of the extracted DNA was validated by gel electrophoresis of aliquots of the extracts. Purity was validated by PCR amplification, cloning and sequence analysis of 16S rRNA genes as described previously [281]. Genome sequencing and annotation was done by the Joint Genome Institute (JGI) following in house standard procedures as described before [24], except for *Desulfitobacterium hafniense* DP7, that was sequenced and annotated by the Genome Center, Washington University School of Medicine (St. Louis, MO).

Tools used for genome analysis and visualization

Genomes were analysed using the tools available through IMG, the JGI integrated microbial genomes comparative analysis system [34], and the EMBOSS software suite [282]. Predicted function of genes was evaluated using the Simple Modular Architecture Research Tool (SMART) [241] and Interpro [283]. Genomes were visualised and aligned using the Artemis genome browser version 16 and the Artemis comparison tool version 13 [145, 187]. Phylogenetic analyses were done with JSpecies [284], and pan and core genomes were determined with the CMG-biotools software package, using the methods described previously [285, 286]. Protein identities were calculated using Matrix Global Alignment Tool (MatGAT) with default settings [113]. Sequence alignments and phylogenetic trees were constructed with Clustal-X [138] and MEGA5 [35]. The Interactive Tree of Life (ITOL) was used for visualising circular phylogenetic trees [287]

Results & Discussion

Genome features and strain classification

Here we report the genome sequences of eight members of the genus *Desulfitobacterium,* bringing the total number of published genome sequences from this genus up to 12 (Tables 1 and 2), now covering four species, namely *D. dehalogenans*, *D. hafniense*, *D. dichloroeliminans* and *D. metallireducens*. Four of the newly sequenced genomes have been closed whereas the other four are in the form of permanent draft genomes consisting of 2-151 scaffolds (Table 2). There are pronounced differences in the genome size between the different *Desulfitobacterium* species. The average genome sizes are 5.5, 4.3 and 3.4 Mbp for *D. hafniense, D. dehalogenans and D. dichloroeliminans/metallireducens*, respectively. Similarly, the GC content per genome slightly differed per species with average values of 47.4, 45 and 43 %, respectively. The genomes were predicted to encode up to seven reductive dehalogenases (Table 2), which is in line with the general view of *Desulfitobacterium* spp. as having a versatile metabolism, with OHR being just one colour on the palette. Two of the strains sequenced in this study had not previously been assigned to any species. *D. dehalogenans* PCE-1 was initially published as *Desulfitobacterium* sp. PCE-1. Although the authors noted that the 16S rRNA gene of strain PCE-1 showed less than 1% sequence divergence from that of *D. dehalogenans* JW/IU-DCl^{T}, they initially decided to not assign strain PCE-1 to the species *D. dehalogenans*, as they did not have DNA-DNA hybridisation data [279]. The average nucleotide identity (ANI) has been proposed as an easily reproducible in silico alternative to wet lab DNA-DNA hybridizations for strains for which the complete genome sequence is available [288]. It has been shown that ANI values of 94-96% correspond to the 70% threshold used as the species boundary for DNA-DNA hybridizations [284, 288, 289].

ANI analysis confirmed the species assignment of all sequenced desulfitobacteria and clearly showed that *Desulfitobacterium* sp. PCE-1 is a member of *D. dehalogenans* (Table 3), and therefore will be referred to as *D. dehalogenans* PCE-1 from hereon. In the same manner the newly isolated strain LBE was found to belong to *D. hafniense* and will therefore be referred to as *D. hafniense* LBE in the following. The strains with closed genomes harbour between three and eight 16S rRNA gene copies (Table 2). *Desulfitobacterium* spp. have unusually long 16S rRNA sequences due to the presence of variable insertions in the 5' region. Previously, eight different types of insertions have been identified in the 16S rRNA genes of *Desulfitobacterium* spp. [290].

This pattern was confirmed by aligning and analysing the 16S rRNA gene sequences obtained in this study (data not shown). The presence of multiple 16S rRNA gene sequences containing different 5' insertions within the same strain also implies that the 5' end should not be used for phylogenetic analysis of *Desulfitobacterium* spp. strains.

Table 3. Average nucleotide identity (ANI), calculated as described in [288] using the JSpecies V 1.2.1 software [284]. The DNA sequences of type strains were used as query, and the DNA sequences of all other sequenced desulfitobacteria were used as subject. *Desulfitobacterium* is abbreviated *D.*, Results are given in columns with strain designation of the query sequence given on top. Strains showing an ANI > 95 % to the type strain are considered to belong to the same species.

	DCB-2^T	JW/IU-DC1^T	DCA1^T	853-15A^T
D. hafniense DCB-2^T	---	85.5	75.9	69.2
D. hafniense Y51	97.8	85.5	76.1	69.2
D. hafniense TCP-A	96.3	85.4	75.7	68.7
D. hafniense PCP-1	97.9	85.2	75.5	68.5
D. hafniense TCE-1	97.7	85.5	76.2	69.2
D. hafniense DP7	97.8	85.5	75.7	68.9
D. hafniense LBE	97.3	86.0	76.3	69
D. hafniense PCE-S	96.7	84.9	75.7	69.2
D. dehalogenans JW/IU-DC1^T	85.0	---	76.3	69.3
D. dehalogenans PCE-1	84.9	99.0	76.3	69.4
D. dichloroeliminans DCA1^T	75.4	76.0	---	69.5
D. metallireducens 853-15A^T	68.7	69.1	69.3	---

The pan and core genome consists of all genes found in at least one or all members of a defined group, respectively [291]. The core genome of a species or genus represents basic functions often related with basic traits of the group of interest. In contrast, the pan genome is considered a flexible pool of genes associated with adaptation to specific niches occupied by different members within a group [292]. The pan and core genome of a single genome will have the same size and consist of all genes present. Adding more genomes will in/decrease the size of the pan/core genome towards a theoretical asymptote representing the true pan/core genome of the investigated group. We determined the pan and core genome of the twelve sequenced desulfitobacteria, using the CMG-Biotools software package [285]. The pan and core genome consist of 10341 and 1410 genes, respectively, and from the slope of the graphs it can be concluded that these

numbers do not represent the true values for the entire *Desulfitobacterium* genus (Figure 2). Both graphs bends at the species boundaries in agreement with the expectation that the interspecies genetic variation exceeds the intraspecies variation. When *D. hafniense* DP7 was included in the analysis the pan genome increased stronger than with inclusion of other strains of *D. hafniense*, whereas no dramatic change in the core genome was observed. A more detailed analysis showed that the genome of *D. hafniense* DP7 encodes 643 protein families not found in other members of *D. hafniense*, and 629 of these were not found in any other of the sequenced desulfitobacterial genomes. In a preliminary analysis we found that many of these were annotated as proteins involved in transport or degradation of polysaccharides. This finding is probably a reflection of the natural habitat of *D. hafniense* DP7, that unlike other desulfitobacteria has not been isolated from soil or sediment but from a human fecal sample, an environment rich in polysaccharides (Table 1). Work on a more detailed analysis of the core and pan genome of the *Desulfitobacterium genus* is currently in progress.

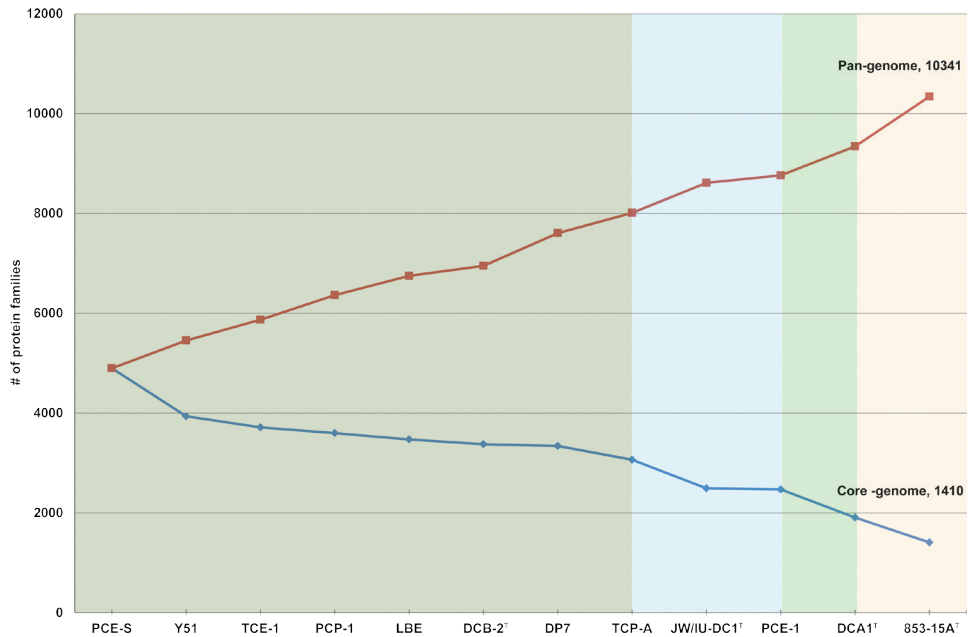


Figure 2. Pan and core genome calculated for all sequenced *Desulfitobacterium* spp. Genomes were added one by one in the order *hafniense*, *dehalogenans*, *dichloroeliminans*, *metallireducens*. Strain designations are indicated on the X axis. A BLAST cutoff of 50% identity and 50% coverage of the longest gene was used. If two proteins within a genome matched according to the 50/50% cutoff, they were clustered into one protein family [285].

Reductive dehalogenases

The majority of the currently described desulfitobacteria are able to couple reductive dehalogenation of one or more chlorinated or other halogenated organic compounds to energy conservation in a respiratory manner, a process known as organohalide respiration [5, 6, 16]. The key enzymes in this processes are the reductive dehalogenase RDase, catalysing the removal of one or more halogens from the organohalide substrate. Reductive dehalogenases are characterised by the presence of two iron sulfur cluster motifs, a twin arginine translocation (TAT) signal sequence and the requirement for a corrinoid co-factor in the active site. The catalytic subunit (RdhA) is translocated across the cellular membrane, where it is believed to be attached to the membrane via a small hydrophobic protein RdhB acting as a membrane anchor [3, 56, 64, 65, 77]. This minimum *rdhAB* gene set can be accompanied by a flexible pool of accessory genes, but only for a few of these the function has been established experimentally [293]. The *rdhK* gene encodes a transcriptional activator of the CRP/FNR type. The CprK protein from *Desulfitobacterium hafniense* DCB-2^T was shown to react with 3-chloro-4-hydroxyphenylacetic acid (Cl-OHPA), before binding to the promoter sequence of *cprA* encoding the Cl-OHPA reductase [67, 71]. The other characterised accessory gene, *rdhT*, encodes a trigger factor like protein, predicted to be involved in folding and maturation of RdhA [73]. Trigger factors contain three domains, an *N*-terminal ribosome-binding domain; a peptidyl-prolyl *cis*/*trans* isomerase domain; and a *C*-terminal domain. The trigger factor binds to the large ribosomal subunits near the exit channel, and chaperones the folding of newly synthesized proteins [294]. The *rdhT* gene does, however, not encode an N-terminal ribosome-binding domain, showing that it most likely does not function as a classical trigger factor. It was shown that PceT from *D. hafniense* Y51 and TCE-1 binds specifically to the TAT signal sequence, which may delay the translocation of PceA, thereby increasing the chance of correct folding and incorporation of all co-factors [68, 70]. Furthermore, the peptidyl-prolyl *cis*/*trans* isomerase activity of PceT was confirmed, supporting the suggestion that PceT, and likely RdhTs in general, act as dedicated chaperones for the correct folding of RdhAs [68, 70]. In agreement with this, co-expression of *pceA* and *pceT* from *D. hafniense* Y51 or $rdhA_3$ and $rdhT_3$ from *D. hafniense* DCB-2^T in *Shimwellia blattae* led to increased amounts of soluble RdhA compared to when *pceA* or $rdhA_3$ were expressed alone [69]. In general, accessory genes are traditionally named with an *rdh* prefix, or a specific prefix for the associated reductive dehalogenase, and a one letter suffix specific for genes with the same predicted function. As an example the CRP/FNR type regulators are named *cprK* when associated with *cprA* and *rdhK*

when associated with uncharacterised reductive dehalogenases [26, 73]. For the naming of genes associated with reductive dehalogenases we will continue this approach with the modifications proposed in Chapter 8 of this thesis [293].

The four previously published genomes of *Desulfitobacterium* spp. were found to encode one to seven reductive dehalogenases [19, 25-27], in stark contrast to the large numbers found in the genomes of the obligate organohalide respiring *Dehalococcoides mccartyi* and *Dehalobacter* spp. where numbers up to 36 or 25, respectively, have been reported [24, 31].

This trend was confirmed by our analysis of the eight additional *Desulfitobacterium* spp. genomes that were predicted to encode up to seven reductive dehalogenases (Table 2). The genomes of *D. hafniense* DP7 and *D. metallireducens* 853-15A^T both do not encode any putative reductive dehalogenases (Table 4). For strain DP7 this result came as no surprise as it has been shown unable to grow by OHR on any tested compound [49]. However, for *D. metallireducens* 853-15A^T this was an unexpected result as Cl-OHPA, PCE and TCE have been reported to support growth with formate or lactate as electron donor [272]. In the initial characterization of this species, increase in optical density over several transfers was interpreted as ability to utilize Cl-OHPA, PCE and TCE as electron acceptors [272]. Two possible explanations could explain these observations. An obvious one is that the coding capacity for reductive dehalogenase(s) has been lost after the strain was characterized, which has been observed for *pceA* as will be discussed below in detail for dehalocluster A. Alternatively, *D. metallireducens* 853-15A^T did derive energy in a manner not requiring the action of reductive dehalogenases. It would therefore be of significant interest to measure the degradation products of *D. metallireducens* 853-15A^T grown in the presence of Cl-OHPA, PCE and TCE. The reductive dehalogenase encoding genes are located in seven gene clusters, in the following termed "dehalocluster A-G" (Figure 3, Table 4). Each of these encodes one to five reductive dehalogenases, and the different clusters are highly conserved between strains (Figure 3). In total we found thirteen different *rdhA* homologues to be encoded on the desulfitobacterial genomes. At least one of these homologues has evolved to have specificity for different substrates in different strains. Dehalocluster A is found in four *D. hafniense* strains where it encodes the reductive dehalogenase PceA whereas in *D. dichloroeliminans* DCA1^T it encodes DcaA (Figure 3). PceA converts PCE to TCE and further to cis-1,2-dichloroethene [78], whereas DcaA dechlorinates 1,2-dichlororethane to ethene by dichloroelimination, i.e. the simultaneous removal of two chlorides [159, 271]. In addition to PceA, four other reductive dehalogenases from *Desulfitobacterium* spp. have been partially purified and biochemically characterized (Table 4).

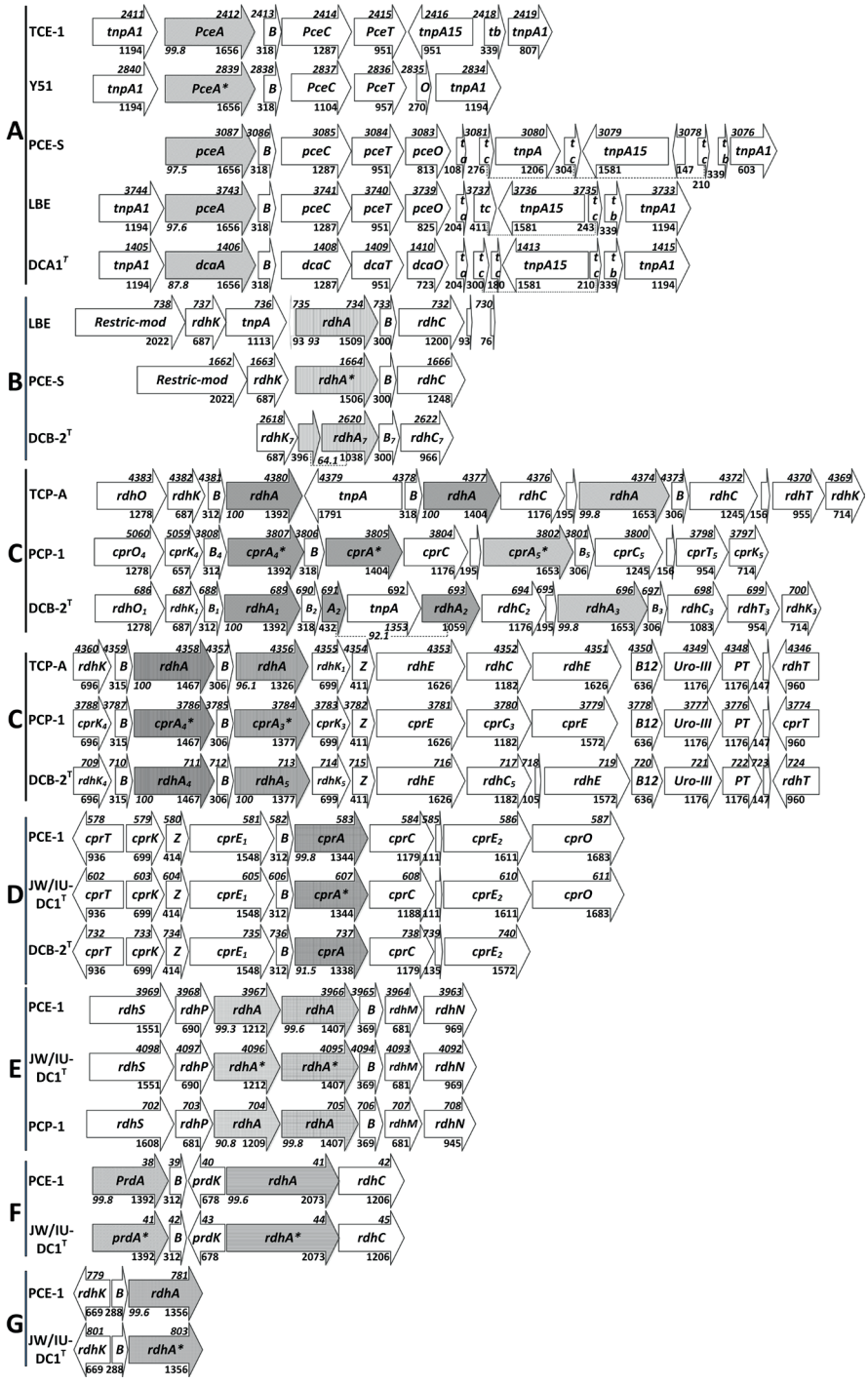

A
TCE-1
tnpA1 PceA B PceC PceT tnpA15 tb tnpA1
Y51
tnpA1 PceA* B PceC PceT O tnpA1
PCE-S
pceA B pceC pceT pceO tnpA tnpA15 tnpA1
LBE
tnpA1 pceA B pceC pceT pceO tnpA15 tnpA1
DCA1T
tnpA1 dcaA B dcaC dcaT dcaO tnpA15 tnpA1
B
LBE
Restric-mod rdhK tnpA rdhA B rdhC
PCE-S
Restric-mod rdhK rdhA* B rdhC
DCB-2T
rdhK7 rdhA7 B7 rdhC7
C
TCP-A
rdhO rdhK B rdhA tnpA B rdhA rdhC rdhA B rdhC rdhT rdhK
PCP-1
cprO4 cprK4 B4 cprA4* B cprA* cprC cprA5* B5 cprC5 cprT5 cprK5
DCB-2T
rdhO1 rdhK1 B1 rdhA1 B2 A2 tnpA rdhA2 rdhC2 rdhA3 B3 rdhC3 rdhT3 rdhK3
C
TCP-A
rdhK B rdhA B rdhA rdhK1 Z rdhE rdhC rdhE B12 Uro-III PT rdhT
PCP-1
cprK4 B cprA4* B cprA3* cprK3 Z cprE cprC3 cprE B12 Uro-III PT cprT
DCB-2T
rdhK4 B rdhA4 B rdhA5 rdhK5 Z rdhE rdhC5 rdhE B12 Uro-III PT rdhT
D
PCE-1
cprT cprK Z cprE1 B cprA cprC cprE2 cprO
JW/IU-DC1T
cprT cprK Z cprE1 B cprA* cprC cprE2 cprO
DCB-2T
cprT cprK Z cprE1 B cprA cprC cprE2
E
PCE-1
rdhS rdhP rdhA rdhA B rdhM rdhN
JW/IU-DC1T
rdhS rdhP rdhA* rdhA* B rdhM rdhN
PCP-1
rdhS rdhP rdhA rdhA B rdhM rdhN
F
PCE-1
PrdA B prdK rdhA rdhC
JW/IU-DC1T
prdA* B prdK rdhA* rdhC
G
PCE-1
rdhK B rdhA
JW/IU-DC1T
rdhK B rdhA*

Figure 3. (Opposite page) Genetic organisation of the reductive dehalogenase encoding gene clusters (dehaloclusters) from the *Desulfitobacterium* spp.. Bottom bars indicate dehaloclusters **A-G**, and strain names without species designation are given below the corresponding gene clusters. Locus tags without prefix are given above arrows (see also Table 1), some locus tags has ben omitted for practical reasons, gene sizes in bp are given under arrows to the right. Percentage amino acid similarity between homologues reductive dehalogenases are shown under arrows to the left, * indicates reference *rdhA* homologue. Gene names are given inside arrows, reductive dehalogenases and associated genes are given an *rdh* prefix, unless other names have been used in the literature, see text for details. The three letter prefix has been left out for some smaller genes due to space limitations. For predicted functions of reductive dehalogenase associated genes see Chapter 8, Table 2). Other genes are *tnpA*: transposase; *ta*, *tb* and *tc*: *tatA*, *B* and *C*; *Restric-mod*: restriction modification; *B12*: vitamin B12 binding enzyme; *Uro-III*: Uroporphyrinogen-III decarboxylase; *PT*: pterrin binding enzyme; Empty arrows indicate hypothetical protein encoding genes.

Table 4. Distribution of dehalocluster A-G in the genomes of sequenced desulfitobacteria, *Desulfitobacterium* are abbreviated *D.*; Locus tags are given without prefix (see Table 1). Name of reductive dehalogenases that are functionally characterised, or in the case for *dcaA*, where a substrate has been indirectly deduced, are given followed by the substrate (see text for details) *pceA*: tetrachloroethene [78], *dcaA*: 1,2-dichlororethane [159], $cprA_2/rdh_2$: 3,5- and 2,4-dichlorophenols [69, 171], $cprA_3$: pentachlorophenol [295], *cprA*: 3-chloro-4-hydroxyphenylacetate [64], *prdA*: tetrachloroethene [296].

	A	B	C					D	E		F		G
	pceA, *dcaA				***$cprA_5/rdh_3$***		***$cprA_3$***	***cprA***			***prdA***		
D. dehalogenans JW/IU-DC1^T								607	4096	4095	41	44	803
D. dehalogenans PCE-1								583	3967	3966	38	41	781
D. hafniense PCP-1			3807	3805	3802	3786	3784		704	705			
D. hafniense TCP-A			4380	4377	4374	4358	4356						
D. hafniense DCB-2 ^{T}		2620	689	693	696	711	713	737					
D. hafniense LBE	3743	734											
D. hafniense PCE-S	3087	1664											
D. hafniense TCE-1	2412												
D. hafniense Y51	2839												
D. hafniense DP7													
D. dichloroeliminans DCA1^T	1406*												
D. metallireducens 853-15A^T													

Dehalocluster A includes the *pce/dcaABCT* genes and a truncated methyl accepting chemotaxis sensory protein *pce/dcaO* gene (Figure 3). These genes are sandwiched by two identical transposases belonging to the IS256 family, where the predicted promoter of the *pceA* gene partly overlaps with the integrating site of the flanking transposase [116, 118]. This cluster has been shown to be a functional transposon, able to excise itself from the genome of *D. hafniense* TCE1 and Y51, whereafter it forms a circular intermediate [116, 118]. This gene cluster does not encode any predicted transcriptional regulator and has been shown to be constitutively expressed in *D. hafniense* Y51 and TCE1 [56, 130]. It is interesting to note, however, that PceA is not translocated across the cell membrane in the absence of PCE [56], suggesting some form of post-translational regulation.

Dehalocluster A is rapidly lost when *D. hafniense* Y51 or TCE1 are cultivated in the absence of PCE, likely reflecting the burden of the constitutively expressed *pceABCT* [55, 56]. Interestingly, we found dehalocluster A to be integrated in different genomic regions in all five strains, suggesting that the transposon has reintegrated after excision from the genome.

Dehalocluster B is present in *D. hafniense* strains DCB-2^T, LBE and PCE-S and contains the *rdhKABC* genes. The reductive dehalogenase in DCB-2^T is probably non-functional due to a nonsense mutation in *rdhA* (Figure 3) [26, 27]. The substrate of the reductive dehalogenase is not known, although it has been speculated to be involved in the degradation of pentachlorophenol and 2,4,5-trichlorophenol in *D. hafniense* PCE-S [27]. Dehalocluster B is located in the same area of the syntenous genomes in the three strains. A large tRNA array is found 15 to 39 kbp upstream, and a tRNA-ala is located directly downstream dehalocluster B. It is well known that tRNA´s often serves as integration sites for integrases [297]. It is therefore tempting to speculate that dehalocluster B was acquired by the integration of a mobile element into the upstream tRNA array.

Dehalocluster C was found in *D. hafniense* DCB-2^T, PCP-1and TCP-A, and consists of two arrays with three and two reductive dehalogenases and several accessory genes, respectively. Among these are four *rdhK* genes encoding transcriptional regulators of the CRP/FNR type and a single *rdhO* coding for a membrane bound methyl accepting chemotaxis sensory protein (for details see Figure 3, Table 4). In *D hafniense* PCP-1 dehalocluster C is divided into two different contigs, where *rdhO* and its cognate *rdhK* (5060-5059) are located at the edge of one contig while *rdhB* (3308) and the remaining part of dehalocluster C are located at the edge of another contig (Figure 3). The region upstream the *D. hafniense* PCP-1 (5060-5059) genes matches perfectly with the corresponding region in the genome of *D. hafniense* DCB-2^T, we therefore propose that the two contigs should be joined as shown in Figure 3.

The two reductive dehalogenase arrays are separated by a 9.5-kbp region conserved among all strains encoding dehalocluster C. This region contains eight genes that are predicted to encode a CysG^A-HemD, fusion protein, involved in synthesis of corrinoids (see corrinoid synthesis paragraph for details), two CRP/FNR family transcriptional regulators, a cupin domain protein, an efflux permease, two predicted FAD/FMN containing dehydrogenases and an Fe/S oxidoreductase. None of the latter three gene products are predicted to contain signal sequences or transmembrane helices. The possible involvement of these proteins in reductive dehalogenation

has not been investigated. Components of dehalocluster C have been studied in detail for *D. hafniense* DCB-2^T and PCP-1. Unfortunately, different names have been used for homologues genes in the two strains. For clarity we will here use the names as given in the cited literature, followed by locus tags without prefix in brackets (Figure 3).

Two of the five reductive dehalogenases encoded by dehalocluster C have been characterized. The product of *cprA*$_3$ (3784) was partially purified from *D. hafniense* PCP-1 and shown to degrade pentachlorophenol (PCP) to 2,3,4,5-tetrachlorophenol (TeCP) and further to 3,4,5-trichloropenol (TCP). In addition, 2,3,4-TCP was also degraded to 3,4-dichlorophenol (DCP) by this dehalogenase [295]. The product of *cprA*$_5$ (3802) was partially purified from *D. hafniense* PCP-1 and shown to degrade 3,5- and 2,4-DCP to 4-chlorophenol [171, 298]. The latter finding has recently been confirmed by heterologous expression and characte-rization of the *cpra*$_5$ homologue *rdh*$_3$ (696) (99.8% amino acid identity) from *D. hafniense* DCB-2^T (Figure 3) in *Shimwellia blattae* [69]. Two transcriptional regulators *cprK*$_2$ (709) and *cprK*$_4$ (687) from *D. hafniense* DCB-2^T were cloned and expressed in *E. coli*. CprK$_2$ induced expression of its cognate *rdhA*$_4$ (711) in the presence of Cl-OHPA [66]. Interestingly CprK$_4$ induced expression of *rdhO* (686), encoding a membrane bound methyl accepting chemotaxis protein. This regulator was shown to react with a range of chlorophenols such as 2,4,6-TCP, 2,3-DCP and 2,5-DCP but not Cl-OHPA [66]. In *D. hafniense* PCP-1 quantitative real time PCR has been used to quantify the expression of the reductive dehalogenases encoded by dehalocluster C [160]. Expression of *cprA*$_3$ (3784) was strongly induced in the presence of its substrate PCP. Furthermore, expression of *cprA*$_3$ (3784) and *cprA*$_5$ (3802) increased in a dose-dependent manner in the presence of 2,4,6-TCP and 3,5-DCP. Finally, it was shown that when *D. hafniense* PCP-1 was grown in the presence of 2,4,6-TCP, expression of *cprA*$_2$ (3787), *cprA*$_3$ (3784) and *cprA*$_5$ (3802) was induced in a sequential manner, probably as a response to the production of intermediates from the dehalogenation of 2,4,6-TCP. Taken together, these data suggest that dehalocluster C functions as one tightly regulated unit enabling *D. hafniense* DCB-2^T, PCP-1 and TCP-A to utilize a large variety of chlorophenols as terminal electron acceptors. The finding that *rdhK*$_4$ after exposure to a range of chlorophenols induced expression of a methyl accepting chemotaxis protein encoding gene, *rdhO*, is especially intriguing, as it strongly suggests that desulfitobacteria are able to sense and response in a chemotactic manner to the presence of chlorophenols, although this still needs to be confirmed experimentally.

Full genome alignment of *D. hafniense* DCB-2^T, PCP-1 and TCP-A revealed that dehalocluster C is located in the same genomic context in all three strains. Comparison with the genomes of *D. hafniense* LBE and Y51 that do not encode cluster C, showed that the genomes are highly similar up and downstream of dehalocluster C. We did not find any strong deviations in GC content or dinucleotide frequencies or other indications of recent horizontal gene transfer [299]. By aligning and comparing strains with and without dehalocluster C, we identified five additional genes as being part of dehalocluster C (Dhaf_0720-724, DeshafDRAFT_4350-4346 and A37YDRAFT_03778-03774 for *D. hafniense* strains DCB-2^T, TCP-A and PCP-1, respectively). These additional five genes encode a predicted vitamin B_{12}-binding protein, a uroporphyrinogen-III decarboxylase, a pterrin binding enzyme, a small hypothetical protein and an RdhT trigger factor like protein (Figure 3). These may have a role in insertion or modification of corriniod co-factors. We performed a tblastn search against the NCBI database, excluding uncultured organisms, the best hits to organisms outside the *Desulfitobacterium* genus were to the genome of *Dehalobacter restrictus* PER-K23^T. We found four copies of an operon consisting of three genes with high similarity to the predicted vitamin B_{12}-binding protein, Uroporphyrinogen-III decarboxylase, and the small hypothetical protein (but not the pterrin binding enzyme) that were predicted to be organised in the same order as in dehalocluster C. All four copies are located within a region, *rdh* cluster B, encoding a large number of reductive dehalogenases, strengthening the speculation that these genes have a role in OHR [24].

Dehalocluster D was found in *D. hafniense* strain DCB-2^T, and in *D. dehalogenans* strains JW/IU-DC1^T and PCE-1 (Figure 3, Table 4). It harbours a single reductive dehalogenase, CprA, that has been shown to remove halides from the ortho position from Cl-OHPA, 2, 3-DCP, and 2-Br-4-CP [64]. This gene cluster has been characterised in detail for both *D. hafniense* DCB-2^T and *D. dehalogenans* strain JW/IU-DC1^T. In addition to the reductive dehalogenase CprA, and its membrane anchor CprB, it encodes several proteins potentially involved in folding and maturation of CprA. The transcriptional regulator CprK has been shown to induce expression of genes found in dehalocluster D. CprK reacts with Cl-OHPA but not the dehalogenated OHPA, whereafter it interacts with a dehalo-box motif upstream of the -10 and -35 promoter regions, inducing expression from 3 promoters within dehalocluster D. [64, 71, 73].

This gene cluster has previously been described as *cprTKZE$_1$BACE$_2$* [25, 26, 73]. However, careful inspection of the genomes revealed that both strains of *D. dehalogenans* encode an additional

gene *cprO,* located downstream from the *cprTKZE$_1$BACE$_2$* cluster (Figure 3). The product of this gene is predicted to be a membrane bound methyl-accepting chemotaxis sensory protein. Aligning and comparing the full genome sequences of all sequenced desulfitobacteria showed that dehalocluster D is located in the same genomic context in all three strains.

D. hafniense strain DCB-2^T encodes both dehalocluster C and D, where they are located adjacent to each other. It is therefore tempting to speculate that these should be seen as one large 54 kbp genomic region dedicated to utilization of a wide range of chlorophenols as terminal electron acceptors.

Dehalocluster E was found in *D. hafniense* PCP-1, and in *D. dehalogenans* strains JW/IU-DC1^T and PCE-1 (Figure 3, Table 4). This cluster encodes two predicted reductive dehalogenases, for which the substrate has not yet been determined. The organisation of this gene cluster is unique among *Desulfitobacterium* spp. (Figure 3). Instead of a RdhK transcriptional regulator, we found genes encoding a two component transcriptional regulator system, consisting of a sensory histidine kinase, RdhS and a DNA binding response regulator, RdhP. These are followed by genes for two predicted reductive dehalogenases, one of which does not have the cognate membrane anchor-encoding gene *rdhB*. This reductive dehalogenase does also not contain a TAT signal sequence, suggesting a cytoplasmic location. Alternatively, the two reductive dehalogenases could form a complex and be co-translocated across the cell membrane, in a manner analogous to what has been described previously for hydrogenases [300]. In addition to this, we found two genes, *rdhMN*, encoding a NapGH like membrane bound menaquinone (MK) dehydrogenase electron transferring complex. The RdhMN complex likely facilitates electron transport from the membrane embedded MKs to the outward facing reductive dehalogenase as previously speculated [25, 29]. In *Sulfurospirillum multivorans* a similar *rdhMN* gene pair is located together with *pceA,* encoding a PCE reductive dehalogenase. Interestingly, expression of these genes was found to be induced in the presence of PCE, supporting the speculation that RdhMN are part of the electon transport chain from menaquinones to the reductive dehalogenase. It is worth noting that the *pceA* gene cluster in *S. multivorans* encodes a two component transcriptional regulators system instead of an RdhK homologue [29].

In all strains encoding dehalocluster E, this is found in the same genomic context. Comparing the genomes of *D. hafniense* DCB-2^T and TCP-A that do not encode dehalocluster E with those of *D. hafniense* PCE-1 and *D. dehalogenans* JW/IU-DC1^T showed that the synteny of the ge-

nomes is conserved on both sides of dehalocluster E. suggesting that this gene cluster was present in the last common ancestor of both species and later has been lost from most of the currently sequenced *D. hafniense* strains.

Dehalocluster F was only found in *D. dehalogenans* strains JW/IU-DC1^T and PCE-1. It encodes two predicted reductive dehalogenases (Figure 3, Table 4). One of these, PrdA, has been partially purified from *D. dehalogenans* strain PCE-1 and shown to convert PCE to TCE [296]. The other putative reductive dehalogenase-encoding gene has some unusual features, as it does not code for a cognate membrane anchor, does not encode a TAT signal sequence, and is unusual large with a size of 2073 bp. The large size is due to the unusual architecture of the encoded protein it consist of a predicted N terminal dehalogenase domain consisting of approximately 200 amino acids followed by an iron sulfur cluster and then a shorter dehalogenase domain consisting of approximately 100 amino acids followed by a second iron sulfur cluster. The amino acid sequence of the encoded protein is 99.6% identical between the two strains, suggesting that either the gene is functional or the rearrangements has happened too recently to allow the gene to decay. Comparing the genomes of *D. hafniense* strains DCB-2^T and TCP-A that do not encode dehalocluster F with those of *D. dehalogenans* strains JW/IU-DC1^T and PCE-1 showed that the synteny of the genomes is conserved on either side of dehalocluster F.

Dehalocluster G was only found in *D. dehalogenans* strains JW/IU-DC1^T and PCE-1. It harbours one predicted reductive dehalogenase (Figure 3, Table 4). The substrate of this reductive dehalogenase has not yet been determined. Comparative genomics showed that the genomic context of dehalocluster G is conserved between both strains of *D. dehalogenans*, whereas interspecies comparisons revealed that there have been several species-specific insertions or deletions in this region in other *Desulfitobacterium* spp. as there was variation between genomes of *D. dehalogenans*, *D. hafniense* and *D. dichloroeliminans* but high conservation between strains belonging to the same species.

General metabolism

In the following sections we discuss a series of selected features linked to the general metabolism that can be deduced from the genomes.

Molybdopterin oxidoreductases

Currently available genomes of *Desulfitobacterium* spp. encode unusually large numbers of molybdopterin oxidoreductases (pfam01568), with the largest numbers found in members of *D. hafniense* (44-65) followed by *D. dehalogenans* (39-42), and *D. dichloroeleminans* and *D. metallireducens* with 19 and 12, respectively (Table 5, Figure 4, Table S1). The function of the majority of these is still unknown, while for some a specific function has been suggested, as will be discussed together with other aspects of the energy metabolism throughout the following sections.

Table 5. Selected metabolic traits discussed in detail in the text. M. OX.: molybdopterin oxidoreductases (see also Figure 4); FRD: fumarate reductase; FDH: formate dehydrogenase; Sel: selenocysteine; Pyr: pyrrolysine. P: presence of pathway; brackets indicate that pathway is disrupted or that some genes are missing; N: not present. [$] Encodes the minimum gene set required for nitrogen fixation. * Encodes both a cytoplasmic and an extracellular Fe only hydrogenase. # three of these are disrupted by transposases. Numbers indicate numbers of genes. *D. haf*: *Desulfitobacterium hafniense*; *D. deh.*: *Desulfitobacterium dehalogenans*; *D. dic.*; *Desulfitobacterium dichloroeliminans*; *D. met.*; *Desulfitobacterium metallireducens*.

	Synthesis of		*nif* cluster						Hydrogenases			membrane bound			M.OX.	Synthesis of					
	Corrinoid	MK	1[$]	2	3	4	5	6	HUP	Hyc	Fe	FRD	FDH	*odmA*	Pfam 01568	Sel	Pyr	*arrA*	*vanI*	*arsC* I-II	TAT/sec
D. haf. DCB-2[T]	P	P	P	P	P	P	N	N	3	P	P	P	P	23	52	P	P	P	P	I	P/P
D. haf. Y51	P	P	P	P	P	P	P	P	3	P	P	P	P	15	57	P	P	P,(P)	P	I-(II)	P/P
D. haf. TCP-A	P	P	P	P	P	P	N	N	3	P	P	P	P	14	44	P	P	N	P	N	P/P
D. haf. PCP-1	P	P	P	P	P	P	N	P	3	P	P	P	P	21	54	P	P	P	P	I	P/P
D. haf. TCE-1	P	P	P	P	P	P	P	P	3	P	P	P	P	16	65	P	P	P	P	I	P/P
D. haf. DP7	P	P	P	P	P	P	P	P	3	P	P	P	P	20	52	P	P	N	P	N	P/P
D. haf. LBE	P	P	P	P	P	P	N	N	3	P	P	P	P	18	56	P	P	P	P	II	P/P
D. haf. PCE-S	P	P	P	P	P	P	P	P	3	P	P	P	P	19	52	P	P	P	P	I-(II)	P/P
D. deh. JW/IU-DC1[T]	P	P	N	N	(P)	N	N	N	3	P	P*	P	P	10	39	P	P	N	N	(II)	P/P
D. deh. PCE-1	P	P	N	N	(P)	N	N	N	3	P	P*	P	P	10	42	P	P	N	N	(II)	P/P
D. dic. DCA1[T]	P	N	N	N	N	N	N	N	5[#]	P	P	N	P	4	19	P	N	P	P	(I-II)	P/P
D. met. 853-15A[T]	P	P	N	N	N	N	N	N	0	P	N	N	P	0	12	P	N	N	N	(I)	P/P

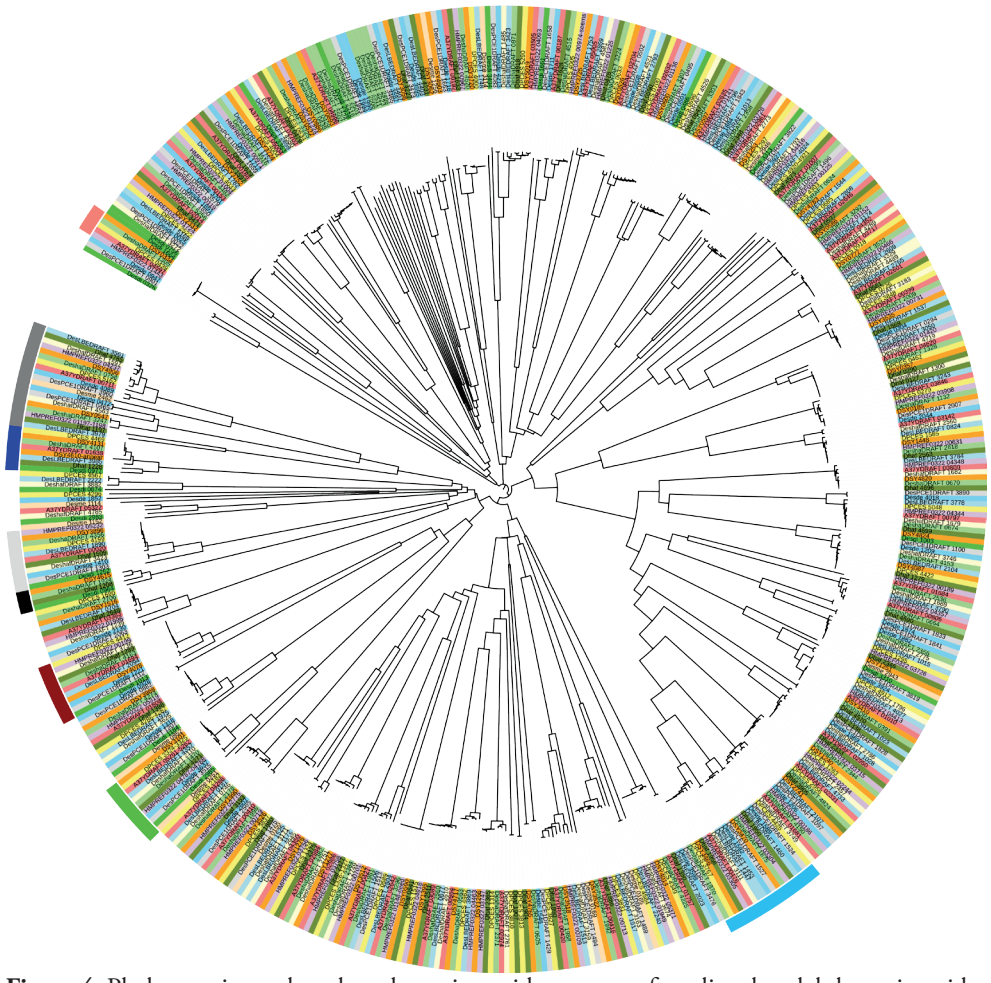


Figure 4. Phylogenetic tree based on the amino acid sequence of predicted molybdopterin oxidoreductases (pfam 01568), from all sequenced genomes of *desulfitobacteria* [301]. Sequences were aligned with Clustal W, and a neighbour-joining tree was constructed and validated with 1000 bootstraps analysis [274]. Tree was visualised using Interactive Tree Of Life (ITOL) [287]. Colours of leaf labels indicate strains, see Table 1. Colour strips indicates genes with a predicted function, *narG,* *dmsA,* *fdhA*, *napA*, *phsA*/*prsA*, *phsA*/*prsA* /FeIII reductase, *arrA* *ttrA.*

Sulfur metabolism

Desulfitobacteria are generally described as capable of utilizing sulfite, thiosulfate and sulphur, but not sulfate, as terminal electron acceptors. Being exceptions to this rule, *D. hafniense* Y51 has been reported to reduce sulfate, whereas *D. metallireducens* does not reduce sulfite [16, 44, 302]. The activity of three gene products, namely ATP sulfurylase (Sat), and APS reductase (AprAB) are required for sulfate reduction [303]. We used the amino acid sequences of Sat (DesyoDRAFT_1712) and AprAB (DesyoDRAFT_1713-1714) from the closely related *Desulfosporosinus youngiae* DSM17734^T as a query for a blastp analysis against all available genomes of *Desulfitobacterium* spp. [215]. We did not find *sat* or *aprAB* homologues in any of the genomes, reinforcing the notion that *Desulfitobacterium* spp. do not possess the genetic machinery for utilizing sulfate as electron acceptor.

Table 6. *Desulfitobacterium* are abbreviated *D*. Locus tags, without prefix (see Table 1), of the genes encoding the catalytic subunit of tetrathionate reductase (*ttrA*) [304, 305] and thiosulfate/polysulfide reductase (*phsA*/*psrA*) [305]; dissimilatory sulfite reductase (*dsrA*) [306]. (C): cytoplasmic location, (P): orientated towards the outside of the cell. *: Annotated as two genes due to a frameshift mutation. When there are two putative catalytic genes, brackets are used to indicate homologues genes.

	***ttrA* (P)**	***phsA*/*psrA* (P)**	***dsrA* (C)**
Catalysed reactions =>	$S_4O_6^{2-}$ => $S_2O_3^{2-}$	$S_2O_3^{2-}$ => S^{2-}+SO_3^{2-}	SO_3^{2-} => S^{2-}
	$S_3O_6^{2-}$ => $S_2O_3^{2-}$ + SO_3^{2-}	S_{n+1}^{2-} => S_n^{2-} + S^{2-}	
D. hafniense DCB-2 ^{T}	1197 (4785)	1509 (1208)	252
D. hafniense Y51	547 (4906)	3896 (4615)	309
D. hafniense TCP-A	4563 (1597)	3423	4756
D. hafniense PCP-1	(711)	20	4573
D. hafniense TCE-1	3247 (758)	4356 (1924)	1278
D. hafniense DP7	1197-1198*(4237)	5232	1468
D. hafniense LBE	3979 (3561)	1890	3088
D. hafniense PCE-S	(5162)	4199	414
D. dehalogenans JW/IU-DC1^T	472 (4089)	1410	266
D. dehalogenans PCE-1	465 (3960)	1303	256
D. dichloroeliminans DCA1^T		1252 (964)	242
D. metallireducens 853-15A^T	(2462)		

The presence of two tetrathionate reductase homologues (*ttrA*) [304, 305], Dhaf_1197 and Dhaf_4785, has previously been reported for *Desulfitobacterium hafniense* DCB-2 [26]. We found a Dhaf_4785 homologue in the genomes of all sequenced desulfitobacteria except for *D. dichloroeliminans* DCA1^{T} that does not encode any *ttrA* homologues. In contrast, homologues of Dhaf_1197 were absent from the genomes of *D. dichloroeliminans* DCA1^{T}, *D. metallireducens* 853-15A^{T}, and *D. hafniense* strains PCP-1 and PCE-S (Table 6, Figure 4, Table S1). To the best of our knowledge it still remains to be tested if desulfitobacteria can use tetrathionate as electron acceptor.

Thiosulfate is used as terminal electron acceptor by all desulfitobacteria that have been tested for this trait [16]. Two different genes have previously been suggested as possible candidates to encode thiosulfate/polysulfide reductase PhsA/PsrA (Dhaf_1208 and Dhaf_1509/DSY3896) in *D. hafniense* DBC-2 and *D. dehalogenans* JW/IU-DC1^{T} [26, 139, 305, 307]. We used these and the sequences of PhsA and PsrA from *Salmonella enterica* or *Wolinella succinogenes* [305] as query for a blastP analysis against all *Desulfitobacterium* spp. genomes. We identified a number of operons containing genes belonging to two adjacent deeply rooted branches of the pfam 1568 tree (Figure 4, Table 6). Each of the available *Desulfitobacterium* spp. genomes encodes one representative of the Dhaf_1509 branch, whereas the Dhaf_1208 branch only contains sequences from *D. hafniense* strains DCB-2^{T}, Y51 and TCE-1 and *D. dichloroeliminans* DCA1^{T} (Figure 4, Table 6). As an exception, the genome of *D. metallireducens* 853-15A^{T} does not encode any sequences clustering with the proposed *phsA*/*psrA* genes (Figure 4, Table 6). This strain does, however, use thiosulfate as terminal electron acceptor [272], pointing towards the presence of an alternative *phsA* gene in *D. metallireducens* 853-15A^{T}.

All described desulfitobacteria, with the exception of *D. metallireducens* 853-15A^{T}, are able to use sulfite as terminal electron acceptor [16]. In agreement with this, we found genes encoding the dissimilatory sulfite reductase, *dsrAB*, in the genomes of all desulfitobacteria, except *D. metallireducens*. The *dsrAB* genes are part of a larger gene cluster conserved among all sulfite-reducing desulfitobacteria (Figure 5, Table 6, Table S1) [306, 308]. This gene cluster consists of seventeen genes, and homologues to these are found in the closely related *Desulfosporosinus youngiae* DSM17734^{T}. The organisation of these genes is conserved between *Desulfitobacterium* spp. and *Desulfosporosinus* the only difference being that seven genes are located separately in the genome of the latter (Figure 5) [215].

The organisation of the dissimilatory sulfite reduction genes is also almost identical to the *dsr* operon from the sulfur-oxidising bacterium *Allochromatium vinosum* (Figure 5) [308, 309]. The *dsr* gene cluster from desulfitobacteria and *Desulfosporosinus youngiae* DSM17734^T does, however, unlike *Allochromatium vinosum*, not encode any *dsrEFHL* homologues, which is in line with previous observations that these genes are solely found in sulfur oxidisers and are thus not present in the genomes of sulfite-reducers [308].

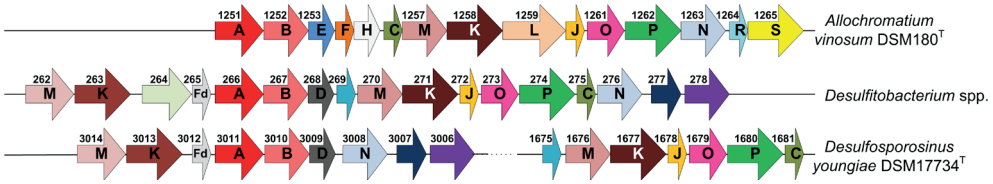


Figure 5. organisation of *dsr* gene clusters in the sulfur oxidizing bacterium *Allochromatium vinosum* DSM180^T (locus tag prefix: Alvin_) [309], and the sulfur reducing bacteria *Desulfitobacterium dehalogenans* JW/IU-DC1^T. (locus tag prefix: Desde_) [293] and *Desulfosporosinus youngiae* DSM17734^T (locus tag prefix: DesyoDRAFT_) [215]. Numbers above arrows are locus tags without prefix. Homologous genes are indicated with same colour, and letters inside arrows are gene names without *dsr* prefix. Predicted function for conserved genes without a *dsr* prefix ■ NAD(FAD) dependent dehydrogenase, ■ Fd: ferrodoxin, ■ hypothetical protein, ■ transcriptional regulator, ■ putative amidohydrolase. See Table S1 for remaining locus tags for *dsr* genes from *Desulfitbacterium* spp. Figure was adapted from [306].

Nitrogen metabolism

Nitrogen Fixation

Three different subtypes of nitrogenases are currently known, Nif, Vnf and Anf, which contain either MoFe, VFe or FeFe metal clusters in their active site, respectively. Nitrogenases consist of two subunits, including a monomeric component I encoded by *nifH* and a di or trimeric component II encoded by *nifDK* or *anfI vnfDKG* [265]. Dos Santos and co-workers recently analysed the genomes of known nitrogen fixing prokaryotes and identified the catalytic *nifHDK* and the accessory *nifENB* as the minimum gene sets required for nitrogen fixation [265]. Genes predicted to be involved in nitrogen fixation have previously been identified in *D. hafniense* Y51 and DCB-2^T. Nitrogen fixation was demonstrated in *D. hafniense* DCB-2^T, DP7 PCP-1 and *D. chlororespirans*, and for the former it was suggested that it possesses both a standard and an alternative vanadium-dependent nitrogenase [19, 26]. The genome of *D. hafniense* DCB-2^T encodes four gene clusters encoding predicted *nifH* homologues, which are called *nif*1-4, according to the

order in which they occur in the genome [26]. We found a total of six different *nif* clusters in the genomes of *Desulfitobacterium* spp. (Table 5). Clusters 1-4 were found in all *D. hafniense* strains, a partial decayed cluster 3 was also present in both *D. dehalogenans* genomes, whereas cluster 5 and 6 were only found in four and five members of *D. hafniense*, respectively. No *nif* gene clusters were identified in *D. dichloroeliminans* DCA1^{T}or *metallireducens* 853-15A^T (Table 5, Table S1). *Nif*1 encodes *nifHDKENB*, the minimum gene set for nitrogen fixation, as suggested by Dos Santos and co-workers [265]. It has been speculated that *Nif*2 encodes a vanadium-dependent nitrogenase [26]. This gene cluster does, however, only contain genes with homology to *vnfHEN*, but no genes encoding the trimeric component II VnfDKG. We used *vnfG* and *anfG* sequences from well-known diazotrophs [265] as query for a tblastn search against all sequenced desulfitobacterial genomes. We were unable to identify any putative *vnfG* or *anfG* genes, which are essential for the vanadium and iron-dependent nitrogenases. These findings render it unlikely that *nif*2 encodes a functional vanadium-dependent nitrogenase. However, *D. hafniense* DCB-2^T, DP7, PCP-1 and *D. chlororespirans* CO23^T have been reported to fix nitrogen in medium where molybdenum was substituted by vanadium [26]. A possible explanation for these seemingly contradicting observations could be the presence of trace amounts of molybdenum in the medium originating either from the inoculum or impurities in the chemicals used for medium preparation. Alternatively, the desulfitobacterial genomes may encode a currently unknown molybdenum independent nitrogen fixation system. *Nif*2-5 are all found together with ABC transporters, in line with previous observations in other bacteria where it was speculated that this couples metal transporters with the assembly of metal co-factors [265]. Finally *nif*6 only contains *nifH* and *nifD/K* like genes but no associated ABC transporters. This cluster is located on the opposite strand as the surrounding genes, directly upstream of two toxin antitoxin genes of the *relE/stbE* family. In conclusion, this suggests that all *D. hafniense* strains are able to fix nitrogen utilizing the MoFe nitrogenase encoded by *Nif*1, whereas *D. dehalogenans, dichloroeliminans* DCA1^T and *metallireducens* 853-15A^T are unlikely to be diazotrophs.

Nitrogen respiration

In addition to the *nif* genes employed for nitrogen fixation we also found several genes with a predicted function in utilization of nitrate, nitrite, nitric oxide and nitrous oxide as terminal electron acceptors (Table 7, Figure 4, Table S1).

Table 7. Shows locus tags, without prefix, (see Table 1), of genes encoding the catalytic subunit of nitrate reductase (*napA*, *narG*), ammonifying nitrite reductase (*nrfA*), nitric oxide reductase (*norB*) and nitrous oxide reductase (*nosZ*) [254] (C): cytoplasmic location, (P): orientated towards the outside of the cell. When there are putative two catalytic genes () are used to indicate homologues genes. *Desulfitobacterium* are abbreviated *D*.

	***napA* (P)**	***narG* (C)**	***nrfA* (P)**	***norB* (P)**	***nosZ* (P)**
Catalysed reactions =>	NO_3^- => NO_2^-	NO_3^- => NO_2^-	NO_2^- => NH_4^+	NO => N_2O	N_2O => N_2
D. hafniense DCB-2^T	1289		3631 (4235)	2253	209
D. hafniense Y51	4078	334	2472 (3066)	1165	261
D. hafniense TCP-A	3736		242 (2148)	2869	4827
D. hafniense PCP-1	1694	4597	1860 (4010)	4287	4835
D. hafniense TCE-1	4163	1305	1432 (5105)	2930	1221
D. hafniense DP7	199	1443	1558 (4748)	3631	243
D. hafniense LBE	2094		4400 (5017)	1116	3150
D. hafniense PCE-S	4412	439	2678 (3289)	1201	359
D. dehalogenans JW/IU-DC1^T	1221		3046 (3613)	1787	186
D. dehalogenans PCE-1	1112		3017 (3491)	1745	177
D. dichloroeliminans DCA1^T	1013		1519 (2945)	1536	167
D. *metallireducens* 853-15A^T				1736	

All sequenced desulfitobacterial genomes except for the one of *D. metallireducens* encode a Nap-type outward orientated nitrate reductase. The organisation of the genes in the *nap* operon is generally conserved between all desulfitobacteria, and has previously been described as *napDF-BAG* for *Desulfitobacterium hafniense* DCB-2 [26]. A comparison of the *nap* operon found in desulfitobacteria with the *napAGHBFLD* genes found in *Wolinella succinogenes* [310] revealed that the gene previously referred to as *napG* encodes a fusion protein of which the N- terminus corresponds to NapH and the C-terminus to NapG, a feature that seems to be unique for desulfitobacteria [311]. Therefore, the *nap* operon from desulfitobacteria should be referred to as *napDFBAH:G* [311]. For an excellent review of the function and location of the components of different *nap* systems see [254]. In brief both NapA and NapB are translocated across the cell membrane where they form a complex, NapA catalyses the reduction of nitrate to nitrite. NapH, a membrane bound MK dehydrogenase, passes electrons through NapG via NapB to NapA.

NapD and NapF are cytoplasmic proteins likely involved in maturation of NapA. In both *D. dehalogenans* strains we found two additional genes to be part of the *nap* operon. These are *ccsB* and *resB*, with a predicted function in synthesis of cytochrome C. These genes are located directly upstream of the *nap* genes (Table S1). Genome analysis revealed that the two available genomes of *D. dehalogenans* encode two copies of *ccsB* and *resB*. Homologues of the *ccsB* and *resB* pair not associated with *nap* genes are present in all sequenced desulfitobacteria and are located directly upstream of corrinoid synthesis cluster I (data not shown).

In addition to the periplasmic Nap operon, five out of the eight sequenced *D. hafniense* genomes also encode the cytoplasm facing Nar-type nitrate reductase (Table 7, Figure 4, Table S1). The *nar* operon consists of two genes encoding nitrate/nitrite transporters and *narGHJI*. The NarGHI complex is attached to the cytoplasmic side of the cytoplasmic membrane. Electrons are channelled from MKs to the membrane embedded NarI and via NarH to NarG catalysing the reduction of nitrate to nitrite [254]. NarJ is a chaperone involved in assembly of the nitrite reductase. Aligning the genomes of *D. hafniense* strains with or without *nar* genes showed that the *nar* genes reside on an approximately 12 kbp region that in addition to the six *nar* genes encodes three genes with predicted function as transcriptional regulators, potentially regulating expression of the *nar* genes. The synteny of all *D. hafniense* genomes is conserved up and downstream of the *nar* region. We did not find any indications of horizontal gene transfer, such as genes encoding recombinases, transposases or deviations in the tetranucleotide frequency associated with the *nar* region [299]. This could indicate that either the acquisition of the *nar* region is an ancient event, or alternatively, all *D. hafniense* in the past encoded a *nar* operon, which then was lost from some but not all genomes of *D. hafniense.* The presence of both *nap* and *nar* operons is commonly observed for members of the *Proteobacteria*, where expression of the two is linked to presence of external factors such as nitrate concentration [254]. This could very well be the case also for desulfitobacteria, although they belong to the *Firmicutes*.

We did not find any *nir* genes encoding for nitrite reductases, catalyzing the reduction of nitrite to nitric oxide, in any of the sequenced desulfitobacteria genomes. Which is in line with previous observations that *nir* genes are rarely found in *firmicutes* [312]

Instead we found two predicted membrane-bound outward facing, nitrite reductases, NrfAH, to be encoded on the genomes of all desulfitobacteria except *D. metallireducens* 853-15A^{T}. NrfH is embedded in the membrane where it receives electrons from the MK pool, and subsequently

transfers these to the catalytic NrhA that reduces nitrite to ammonia (Table 7, Table S1). The conversion of nitrite to ammonium can both act as a respiratory process for energy conservation and a detoxification mechanism [254]. Reduction of other compounds such as hydroxylamine, hydrogen peroxide and nitric oxide are also commonly observed for NrfAH complexes [254] It is therefore possible that the two NrfAH complexes have roles beyond nitrite ammonification in *Desulfitobacterium* spp.

In addition to this, all twelve sequenced desulfitobacterial genomes contain *norBC* genes that encode nitric oxide reductase. The NorBC complex is attached to the cell membrane where it catalyses the extracellular reduction of nitric oxide to nitrous oxide [254] (Table 7, Table S1).

Lastly, all currently available desulfitobacterial genomes except that of *D. metallireducens* 853-15S^{T} encode an extracellular protein, NosZ, catalyzing the reduction of nitrous oxide to dinitrogen (Table 7). Five *nos* genes have been identified in desulfitobacteria, namely *nosZDLYF*. The order of these is conserved and is, in the direction of transcription, *nosZ*-hypothetical gene-*nosDLYF*. The catalytic subunit NosZ contains a signal sequence and is translocated across the cell membrane. NosDLYF forms a membrane bound complex involved in maturation of NosZ [313]. The role of the hypothetical protein encoded by the gene located between *nosZ* and *nosD* has not yet been determined. The encoded protein consists of around 180 amino acids and has four predicted transmembrane helices. Interestingly, using this gene as query for a tblastn against the genome of *Wolinella succinogenes,* the best hit (28% identity, 75% query coverage) was to a similar hypothetical gene also located between *nosZ* and *nosD* [310]. The desulfitobacterial *nos* genes are very similar to the *nos* genes from the gram positive bacterium *Geobacillus thermodenitrificans* NG80-2. the *nosZ* gene from bacterium *Geobacillus thermodenitrificans* NG80-2 has been heterologously expressed and determined to be a functional nitrous oxide reductase [314]. The *nos* genes are located in the downstream region of a conserved gene cluster consisting of nine additional genes located directly upstream of *nosZ*. A homologue of the gene located directly upstream of *nosZ*, encoding a cytochrome c, was also found in the *nos* operon of *Geobacillus thermodenitrificans* NG80-2 where the name nosC was proposed [314]. *D. dehalogenans* genomes have two (JW/IU-DC1^{T}) or three (PCE-1) additional genes but share the conserved nine gene cluster with *D. hafniense* and *D. dichloroeliminans* DCA1^{T}. The majority of the corresponding gene products have one or more predicted transmembrane regions and are predicted to be Fe-S cluster-, cytochrome B or C domain proteins. This fact, in combination with the conservation in all *nos* encoding desulfitobacteria, makes it tempting to speculate that this gene cluster, or part of

it, encodes a membrane-bound complex involved in electron transfer to NosZ.

Arsenate reduction

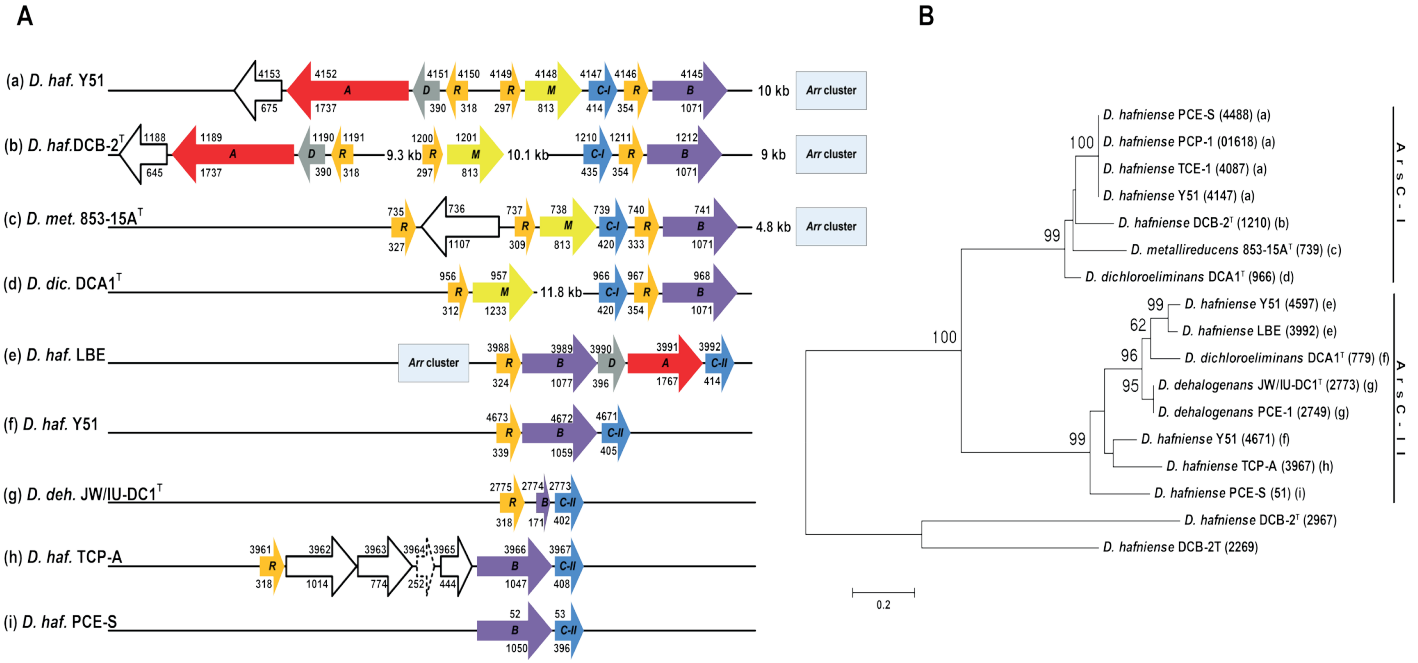
D. hafniense strains DCB-2^T, TCE-1, PCP-1 and GBF1 have been reported to couple the oxidation of formate and reduction of arsenate As(V) to arsenite As with growth [46, 315, 316]. We found an *arrA* gene encoding arsenate reductase in seven of the twelve sequenced desulfitobacterial genomes (Table 5, Figure 4, Table S1), including *D. hafniense* strains DCB-2^T, TCE-1 and PCP-1. Three of the sequenced strains, *D. dehalogenans* JW/IU-DC1^T, PCE-1 and *D. hafniense* DP7, have been shown not to utilize arsenate as terminal electron acceptor [46, 316]. In line with this, the genome of these strains does not contain an *arrA* gene, supporting the claim that the reported ArrA encoding genes indeed are responsible for arsenate reduction in desulfitobactera. Aligning the genomes of all ArrA encoding desulfitobacteria allowed us to identify a conserved gene cluster, *arrTSRCABD*. The *arrSR* genes encode a two- component transcription regulator system, consisting of a membrane bound sensory histidine kinase ArrS and a response regulator ArrR. The role of ArrT is not clear, however, the encoding gene is predicted to be a phosphonate binding protein, attached to the outward face of the cytoplasmic membrane. It seems plausible that the substrate in desulfitobateria is arsenate, as speculated previously [317]. The *arrB* gene encodes an Fe-S cluster protein predicted to form a complex with ArrA, facilitating electron transfer to the latter. Electrons may be transferred from the MK pool to the outward facing ArrAB complex via ArrC, a membrane embedded electron transferring NrfD family protein, although the exact electron transport pathway is not yet known [317, 318]. Finally, *arrD* encodes a chaperone, involved in maturation of the Arr proteins [318]. The genome of *D hafniense* Y51 encodes an additional copy of *arrABD* where the predicted *arrA* gene product is N terminally truncated. A gene located directly upstream of these genes encodes a TraG like protein, which is involved horizontal gene transfer via conjugation [319].

Arsenate resistance

Three arsenate-resistance associated *arsC* genes were reported to be located on the genome of *D. hafniense* DCB-2 [26]. One of these *arsC* genes (Dhaf_1210) did not have a homologue in the genome of *D. dehalogenans* JW/IU-DC1^T, whereas homologues of the other two were found in both strains. Growth of *D. dehalogenans* JW/IU-DC1^T has been shown to be inhibited in the presence of arsenate [46], and it therefore seems likely that the ArsC homologue encoded by (Dhaf_1210) confers resistance to arsenate, whereas the two other predicted ArsC homolouges may have other functions. Further support for this stems from the absence of other *ars* genes in close proximity to these putative *arsC* genes. We identified one to three *arsC* genes in the genomes of all sequenced *Desulfitobacterium* spp. except *D. hafniense* DP7, belonging to two distinct phylogenetic groups *arsC*-I and II (Table 6, Figure 6). The products of *arsCB* are essential components of arsenate resistance [320]. The absence of *arsAB* from the genome of *D. dehalogenans* JW/IU-DC1^T likely explains why this strain is inhibited by arsenate despite the presence of *arsC*-II.

The order of the *ars* genes differs between the two clusters, namely *arsADRRMCRB* and *arsRB-DAC* for clusters I and II, respectively (Figure 6). The function of the product of the *ars* genes is in brief as follows: ArsA, ATPase; ArsB, efflux pump; ArsC, arsenate reductase; ArsRD transcriptional regulators; ArsM, As[315] S-adenosylmethionine methyltransferase [320-322]. Resistance is achieved by cytoplasmic reduction of arsenate to arsenite by ArsA, whereafter it is pumped out of the cell by ArsB [320]. ArsM methylates arsenite to mono-, di- and tri-methylarsine, all of which are more toxic than both arsenate and arsenite but also volatile, potentially leading to complete removal of As from the site by outgassing [321-323]. The presence of *arsM* in one but not the other of the two *ars* gene clusters clearly shows that Ars-I and II represent two different detoxication mechanisms. The *ars*-I gene cluster is generally located upstream of the *arrTSRCABD* gene cluster, whereas *ars*-II in *D. hafniense* LBE is located directly downstream of *arrTSRCABD* (Figure 6). The *ars*-II cluster is partially deleted in all other strains, without any associated *arr* genes, although *D. hafniense* Y51 encodes a second truncated *arr* cluster upstream of one of two partial *ars*-II clusters (Figure 6).

Figure 6. A) Examples of organisation of *ars* cluster I [39] and *ars* cluster II (e-i). Numbers above arrows are locus tags without prefix, numbers below arrows are gene sizes in bp. Homologous genes are indicated with the same colour, and letters inside arrows are gene names without *ars* prefix. **B**) Phylogenetic tree based on the amino acid sequence of arsenate reductase ArsC from all sequenced desulfitobacteria. Sequences were aligned with Clustal W [274], a neighbour-joining tree was constructed and validated with 1000 bootstraps analysis using the MEGA5 software package [35]. Bootstrap values higher than 50% are given at corresponding nodes in the tree. The reference bar indicates 20% base substitutions per site. The two outgroup sequences shown at the lower branch have previously been proposed to be ArsC homologues in *D. hafniense* DCB-2^T [26]. numbers in brackets are locus tags without prefix, see also Table 1, letters in brackets refers to the organisation of the surrounding *ars* clusters as depicted in Figure 6A.



O-demethylases

The presence of several *O*-demethylase corrinoid proteins (OdmA) has been reported for *Desulfitobacterium hafniense* DCB-2 and Y51 [19, 26, 301] One of these (Dhaf_4611), from *Desulfitobacterium hafniense* DCB-2, has been heterologously expressed and functionally characterised as a functional *O*-demethylase showing highest specific activity towards guaiacol and methyl chloride (Figure 7) [324]. Another *odmA* paralogue (DSY3155) was found to be highly upregulated when *Desulfitobacterium hafniense* Y51 was grown with vanillate instead of pyruvate as electron donor (Figure 7) [131]. Previously, the presence of 4 to 23 putative OdmA encoding genes in *D. hafniense*, *D. dehalogenans* and *D. dichloroeliminans* DCA1^{T}, but none in *D. metallireducens* 853-15A^{T}, has been reported, with the highest numbers found in members of *D. hafniense* and the lowest in *D. dichloroeliminans* [301] (Figure 7, Table 5). Ten different *Desulfitobacterium* spp. strains covering the species *D. hafniense*, *D. dehalogenans*, *D. chlororespirans* CO23^{T} and *D. metallireducens* were tested for their ability to grow with a range of phenyl methyl ethers as electron donor. All but *D. metallireducens* 853-15A^{T} were able to utilize phenyl ethers for growth. This is in line with the absence of *odmA* genes in the genome of *D. metallireducens* 853-15A^{T} [301].

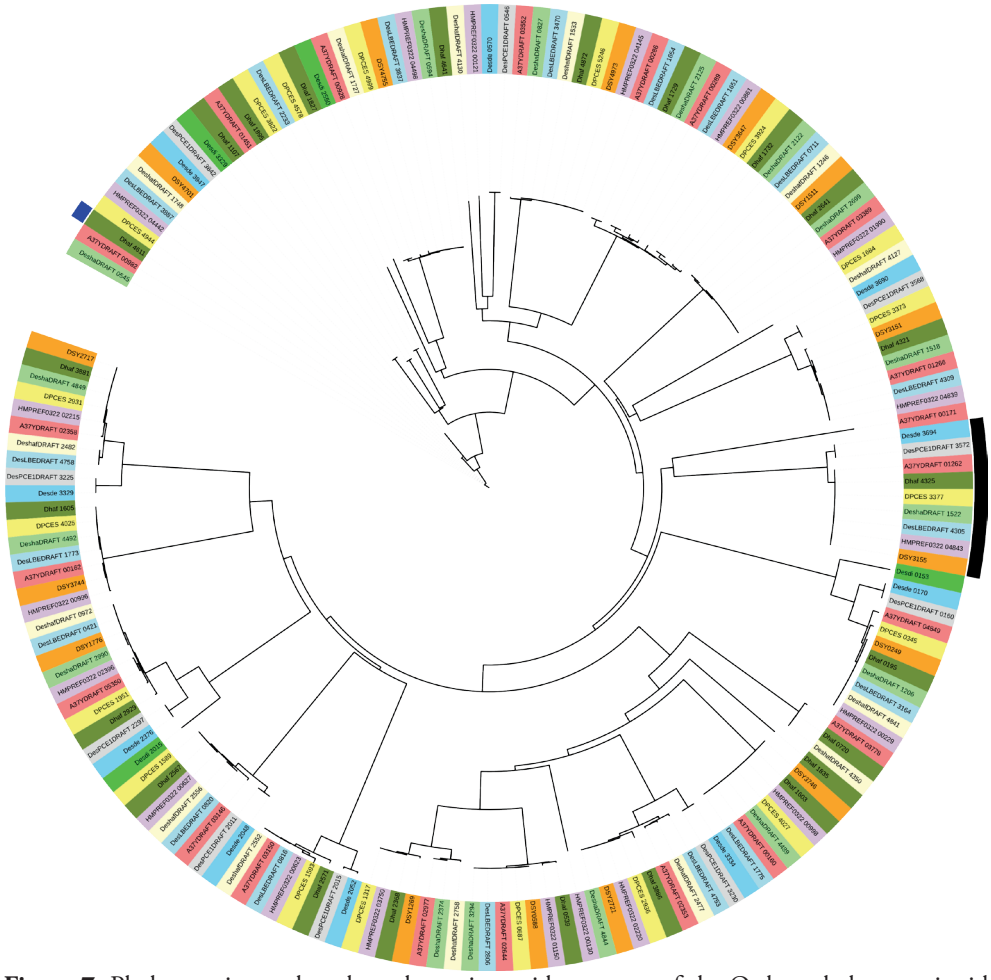


Figure 7. Phylogenetic tree based on the amino acid sequence of the O-demethylase corrinoid protein (OdmA), from all sequenced desulfitobacteria [301]. Sequences were aligned with Clustal W, and a neighbour-joining tree was constructed and validated with 1000 bootstraps analysis [274]. Tree was visualised using Interactive Tree Of Life (ITOL) [287]. Colours of leaf labels indicate strains, see Table 1. Colour strips indicate ■ Characterised guaiacol and methyl chloride demethylase [324] or ■ indirectly predicted function as vanillate demethylase [131]. One sequence from *D. hafniense* DP 7 was excluded as the gene is split into two (HMPREF0322_00864-865).

Hydrogenases

Most desulfitobacteria have been shown to utilize H_2 as electron donor. In agreement with this, all sequenced desulfitobacterial genomes encode one or more different types of hydrogenases (Table 5, Table S1) [16]. One of these, a six-subunit membrane-bound NiFe hydrogenase (Hyc), is encoded in the genomes of all *Desulfitobacterium* spp. The corresponding gene cluster resembles the *hyfBCEFGI* components of the formate hydrogen lyase 4 complex-encoding gene cluster (*hyfABCDEFGHIR-focB*) described for *E. coli* [25, 119, 248]. Furthermore, the genomes of all but *D. metallireducens* 853-15A^T encode three membrane bound outward orientated uptake NiFe hydrogenases (HUP_{1-3}) and a membrane associated Fe only hydrogenase. In addition *D. dehalogenans* genomes encode a cytoplasmic NADH dependent Fe only hydrogenase. The exact role of the various hydrogenases in the metabolism of *Desulfitobacterium* spp. has not yet been studied in detail. Disruption of either *hyc* or hup_1 surprisingly lead to lack of growth with formate or lactate as electron donor and Cl-OHPA or nitrate as electron acceptors in *D. dehalogenans* JW/IU-DC1^T [119], pointing towards a central role of hydrogenases in the energy metabolism of *Desulfitobacterium* spp.

The *hyc* complex was previously described as a formate hydrogenlyase [19, 119], however, it does not encode a formate dehydrogenase, and can thus only act as a formate hydrogenlyase if it forms a complex with a formate dehydrogenase encoded elsewhere in the genome. Interestingly all *Desulfitobacterium* spp. encode a cytoplasmic NADH dependent formate dehydrogenase, predicted to be part of the Wood-Ljungdahl CO_2 fixation pathway (Table S1) as reviewed previously [197]. Alternatively the *hyc* complex may use reverse electron flow to generate low potential electrons used for biosynthesis or reducing Co present in the active site of reductive dehalogenases as suggested for a similar complex found in *Dehalococcoides mccartyi* [18, 120]. Phylogenetic analysis of the large subunits of the HUP hydrogenases, encoded by $hydB_{1-3}$, showed that they form three distinct clusters that within the clusters form a pattern similar to a 16S rRNA-based tree (Figure 1 and 8), clearly showing that the presence of three different HUP is a trait originating from the last common ancestor and not the result of recent gene duplications or horizontal gene transfer. Sequences within clusters are highly conserved at 88-100% amino acid identity whereas the between clusters identity is limited to 30-36% (data not shown). The fact that these three HUP have been conserved in nearly all *Desulfitobacterium* spp. may suggest that they have a role in the metabolism beyond respiratory growth with H_2 as primary electron donor. Differential expression of HUP hydrogenase encoding genes has previously been associated with e.g. symbiosis or

as a way to recycle H_2 produced as a by-product of N_2 fixation [325, 326]. HUP_1 was found to be upregulated in both *D. dehalogenans* JW/IU-DC1^T and *D. hafniense* Y51 during growth by respiration compared to fermentation [25, 131]. Similar to the situation in desulfitobacteria, also the genome of the closely related obligate OHRB *Dehalobacter restrictus* PER-K23^T encodes three HUP hydrogenases that cluster together with HUP 1-3 from *Desulfitobacterium* spp. (Figure 8)[24]. Interestingly, HUP_1 was the only HUP type hydrogenase detected in the proteome of *Dehalobacter restrictus* PER-K23^T grown with H_2 as electron donor and PCE as electron acceptor.

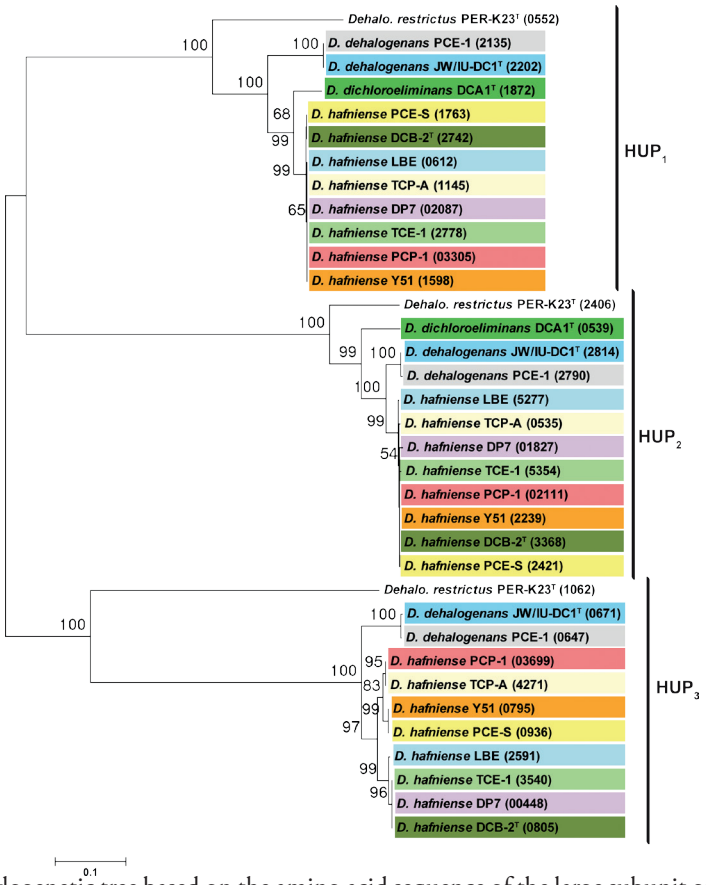


Figure 8. Phylogenetic tree based on the amino acid sequence of the large subunit of NiFe uptake hydrogenases *hydB*. Sequences were aligned with Clustal W, [274] and a neighbour-joining tree was constructed and validated with 1000 bootstraps using the MEGA5 software package [35]. Bootstrap values higher than 50% are given at corresponding nodes in the tree. The reference bar indicates 10% aminoacid substitutions per site. The *D. dichloroeliminans* DCA1^T Hup_3 sequences are left out of this analysis as they are disrupted by transposase activity. *Desulfitobacterium* is abbreviated *D.*, and *Dehalobacter* is abbreviated *Dehalo*. Locus tags, with the prefix removed are given in parentheses, GenBank accession numbers of sequences when marked with an *. For colour coding see Table 1.

This knowledge, combined with the lack of growth by respiration after disrupting HUP_1 may suggest an yet unidentified role of hydrogen in the respiration of *Desulfitobacterium* spp [119]. *D. dichloroeliminans* DCA1^T encodes three copies of HUP_3 that all have been disrupted by transposases. Despite this the strain is still able to utilize H_2 as electron donor [271].

All sequenced desulfitobacterial genomes, with the exception of *D. metallireducens*, encode an extracellular NiFe hydrogenase, with the corresponding genes found closely together with a gene (*ssc*) encoding a periplasmic split soret cytochrome c, a gene organisation that seems to be unique for desulfitobacteria [26]. Recently it was reported that a split soret cytochrome c receives electrons from a periplasmic NiFe hydrogenase in *Desulfovibrio desulfuricans* [327].

In conclusion, the large number of hydrogenases and the finding of a putative role of hydrogenases in respiration without H_2 in the headspace, makes it tempting to speculate that hydrogen could play a more general role in the respiration of desulfitobacteria, as has previously been shown for *Desulfuvibrio* spp. [328].

Fumarate reductase paralogs

The presence of 31 and 19 putative fumarate reductase paralogous (FRP) flavoprotein encoding genes (cog1053) has been reported for *D. hafniense* Y51 and DCB-2, respectively [19, 26]. In line with this, we identified 24 to 31 of these genes in available genomes of *D. hafniense* including eight additional putative FRPs in *D. hafniense* DCB-2^T, 8 to 12 in *D. dehalogenans* and *dichloroeliminans* DCA1^T and none in *D. metallireducens* 853-15A^T (Figure 9, Table 8, Table S1). Most probably, some of these have functions beyond fumarate reduction or succinate oxidation, especially since *D. dichloroeliminans* DCA1^T does not grow with fumarate as electron acceptor, despite the presence of genes predicted to encode seven full length putative fumarate reductases. One operon, *frdABC,* is conserved in all genomes of *D. dehalogenans* and *D. hafniense* but absent from the genome of *D. dichloroeliminans* DCA1^T and *D. metallireducens* 853-15A^T. This operon encodes a fumarate reductase, an Fe-S protein and a membrane bound cytochrome b, with high similarity to the well characterised, membrane bound fumarate reductase encoded by *frdABC* in *Wolinella succinogenes* [258]. Transposon insertional inactivation, as well as comparative proteomics of *D. dehalogenans* JW/IU-DCl^T grown in the presence or absence of fumarate confirmed this gene cluster as the most likely candidate for the fumarate reductase in *Desulfitobacterium dehalogenans* [25, 329]. The absence of this operon from the genomes of *D. dichloro-*

eliminans DCA1^T and *metallireducens* 853-15A^T probably explains the inability of these strains to utilize fumarate as electron acceptor (Figure 9, Table 5)[271, 272].

Table 8. Number of genes encoding fumarate reductase paralogs (FRP, cog-1053), from all sequenced desulfitobacteria. Top row: products of genes putatively co-transcribed with the FRP, Cyt.: Cytochrome, OYE: old yellow enzyme family, Hyp.:hypothetical protein. Second row: location of co-transcribed gene compared to the FRP, US: upstream, DS: downstream. For multiple domains, FRP type domains are indicated in the order from N to C terminus. The truncated desdi_2762 sequence has been excluded. See also Figures 9 and 10. *Desulfitobacterium* are abbreviated *D.*

	Cyt. C		FMN		Fe-S		OYE	Hyp.	single			Total
							FMN-FMN-FAD					
	US	DS	US	DS	US	DS	DS	DS	FMN-FAD	FAD-FMN	FAD	#
D. hafniense DCB-2^T	3	3	0	3	1	3	1	0	3	3	7	27
D. hafniense Y51	3	3	0	6	1	3	1	0	4	3	7	31
D. hafniense TCP-A	2	1	1	5	0	2	1	0	4	1	7	24
D. hafniense PCP-1	2	2	1	6	1	3	1	0	4	3	7	30
D. hafniense TCE-1	3	3	0	4	1	3	1	0	3	3	7	28
D. hafniense DP7	2	2	0	6	1	3	1	0	4	3	7	29
D. hafniense LBE	2	2	1	6	1	3	1	0	5	3	7	31
D. hafniense PCE-S	2	2	0	5	1	4	1	0	3	3	7	28
D. dehalogenans JW/IU-DC1^T	1	0	0	1	0	1	0	1	3	1	2	10
D. dehalogenans PCE-1	1	1	0	1	0	1	0	1	4	1	2	12
D. dichloroeliminans DCA1^T	2	1	0	0	2	1	0	0	0	0	2	8
D. *metallireducens* 853-15A^T	0	0	0	0	0	0	0	0	0	0	0	0

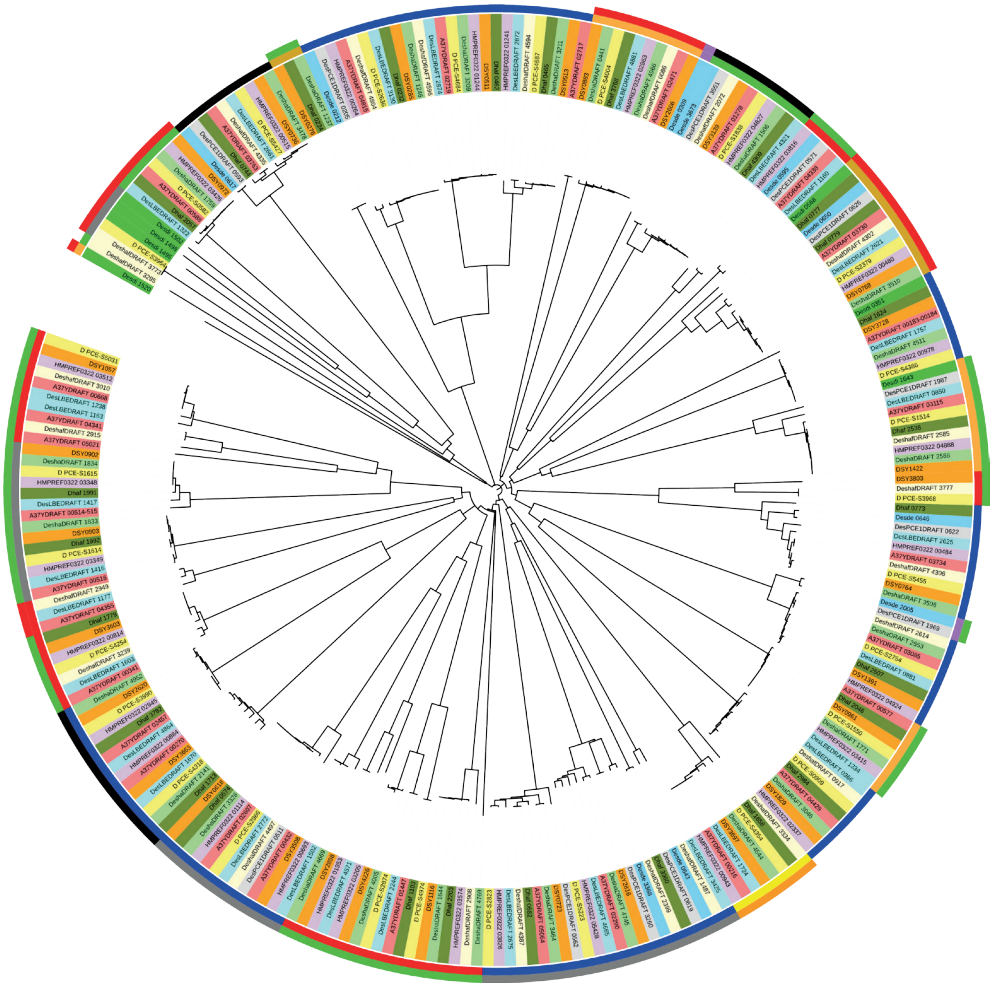


Figure 9. Phylogenetic tree based on the amino acid sequence of the fumarate reductase paralogs (FRP, cog-1053), from all sequenced desulfitobacteria, excluding the truncated desdi_2762 sequence. Sequences were aligned with Clustal W, and a neighbour-joining tree was constructed and validated with 1000 bootstraps [274]. Tree was visualised using Interactive Tree Of Life (ITOL) [287]. Colours of leaf labels indicate strain, see Table 1. **Inner ring**: Types of genes found associated with the FRP, colour codes are ■ No associated gene, ■ Protein with a single heme, CxxCH, binding domain, ■ Multiheme cytochrome C, ■ Fe-S cluster protein, ■ Fumarate reductase, FrdA ■ FMN binding domain protein, ■ Part of a larger operon/gene cluster, ■ Protein of unknown function, ■ NADH flavin oxidoreductase with high similarity to Old yellow enzyme family proteins. **Outer ring**: ■ Associated protein located upstream of the FRP, ■ Associated protein located downstream of the FRP, ■ Multiple domain FRP organised as FMN-FAD binding domains, associated protein located downstream of the FRP, ■ Multiple domain FRP organised as FAD-FMN binding domains, ■ Multiple domain FRP organised as FMN-FAD binding domains, ■ Multiple domain FRP organised as FMN-FMN-FAD binding domains, associated protein located downstream of the FRP

The function of the remaining putative FRPs has not yet been determined experimentally. A similar flavoprotein, FccA, was found in *Wolinella succinogenes*. The *fccA* gene was found co-transcribed with two cytochrome C encoding genes *fccB* and *fccC*, with FccB and FccA being believed to form a complex [330]. FccC is a predicted membrane bound NapC/NirT type protein, mediating electron transfer from MKs in the membranes to the FccAB complex (Figure 10). It has been suggested that FccA is involved in methacrylate reduction in *Wolinella succinogenes* [331]. Both FccA and FRPs from *Desulfitobacterium* spp. displays high similarity to the C-terminus of a fumarate reductase flavocytochrome c_3 encoded by *ifcA*, found in *Shewanella frigidimarina* NCIMB400 and *Shewanella putrefaciens* [250, 251, 330]. Unlike IfcA, however, neither FccA nor the FRPs from *Desulfitobacterium* spp. contain any N-terminal heme binding CxxCH motifs (Figure 10)[250, 251]. It has been suggested that *ifcA* resulted from a fusion of a small cytochrome c and a larger flavoprotein encoding gene, like the aforementioned *fccAB* [332, 333]. This type of flavoprotein - cytochrome C complexes or flavocytochrome C fusion proteins has been suggested to be involved in periplasmic electron transport, for examples see [252, 334-336].

Careful analysis of the FRPs from *Desulfitobacterium* spp. and their genomic context showed that they can be divided into five main groups. Genes putatively co-transcribed with either 1) cytochrome Cs, 2) FMN binding proteins 3) Fe-S cluster proteins 4) or single genes, containing an FAD binding domain or 5) dual flavin proteins containing both FMN and FAD binding domains (Figures 9, 10, Table 8, Table S1). Cytochrome Cs and flavoproteins play important roles as electron carriers in extracellular oxidation/reduction processes both in the form of periplasmic electron carriers or as membrane bound MK dehydrogenase linking the MK pool with extracellular redox processes [252, 337]. Multiheme proteins of the NapC/NrfH family are attached to the membrane by a single N-terminal transmembrane helix. They transfer electrons from reduced MKs to extracellular proteins. Some interact with a specific electron acceptor, i.e. NapC donates electrons to the nitrate reductase NapAB [254], whereas the NapC/NrfH family protein CymA from *Shewanella* spp. has been shown to act as a versatile electron transfer hub, linking many different electron acceptors to the MK pool. This ability to transfer electrons to different alternative periplasmic electron carriers provides a robust and flexible electron -ransferring network in *Shewanella* spp. [338, 339]. We found two *napC/nrfH* type tetraheme cytochrome C encoding genes located upstream of the FRPs. One (Desdi_1520) was solely found in *D. dichloroeliminans* DCA1^T, whereas the other was found in all members of *D. hafniense* (Table 5, Figure 9).

The former was found together with a gene coding for another tetraheme cytochrome C, and blast analysis revealed that these two proteins resemble the NrfAH complex, mediating electron transfer from MKs to the periplasmic nitrite reductase [254, 340]. In contrast the latter either functions as a dedicated electron donor to the associated FRP or maybe has a more general role in electron transfer to extracellular electron acceptors, similar to CymA in *Shewanella* spp. [252, 339]. All sequenced desulfitobacterial genomes, except *D. metallireducens,* encode FRPs associated with both mono- and hexaheme cytochrome Cs. These all contain TAT signal sequences, so they are likely translocated across the cytoplasmic membrane. In addition to this we found several additional FRPs in connection to cytochrome C-encoding genes in the desulfitobacterial genomes.

The Fe-S cluster associated FRPs include the previously described fumarate reductase encoded by *frdABC* [25, 329] and are generally part of larger gene clusters. It could be speculated that these mainly are catalytic enzymes like the aforementioned FrdA.

A large number of the FRPs are either diflavins i.e. containing both FAD and FMN binding domains or encoded in an operon also containing a gene coding for a small FMN binding protein (Table 8, Figures 9, 10). Interestingly all these accompanying FMN binding proteins contain a signal sequence, showing that they like the FRPs are exported across the membrane. It is believed that diflavins are the result of a fusion of FAD and FMN binding proteins [341, 342]. It is tempting to speculate that the FRPs co-transcribed with FMN binding proteins form a complex with their accompanying FMN binding proteins after being exported from the cell, and thus function as diflavins. Diflavins are involved in electron transport to terminal electron accepting reactions [342]. A diflavin reductase (Desde_3368) was recently reported in *D. dehalogenans* to strongly increase in abundance when cells were grown with Cl-OHPA as terminal electron acceptor, suggesting a role in the electron transport chain to the Cl-OHPA reductase CprA [25].

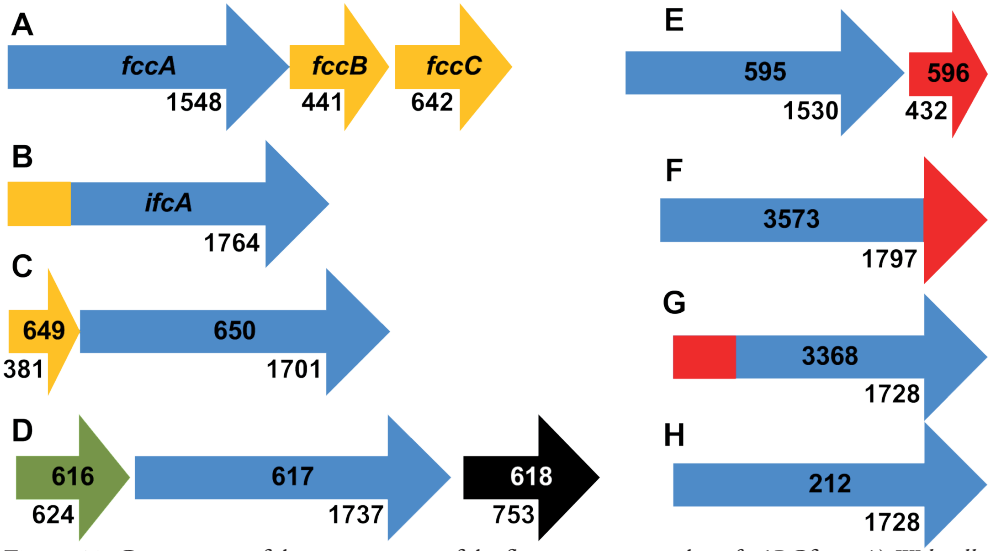


Figure 10. Comparison of the organisation of the flavoprotein encoding *fccABC* from **A**) *Wolinella succinogenes* [330]; **B**) flavocytochrome encoding *ifcA* from *Shewanellla frigidimarina* NCIMB400 [251] with representative examples of fumarate reductase paralogs found in *Desulfitobacterium* spp. **C**) flavoprotein with an upstream cytochrome c; **D**)fumarate reductase *frdCAB*; **E**) flavoprotein with a downstream FMN binding protein; **F-G**) Diflavins organised as FAD-FMN or FMN-FAD; **H**) flavoprotein. ■ FAD binding flavoprotein; ■ Cytochrome c; ■ FMN binding protein, ■ cytochrome b; n Fe-S cluster protein. Numbers under arrows indicate size in bp, and numbers inside arrows for C-H indicate locus tags without the *Desulfitobacterium dehalogenans* JW/IU-DC1^{T} desde_ prefix

Synthesis of cobalamins

Reductive dehalogenases require cobalamin, also termed vitamin B_{12}, as an essential co-factor [177]. In recent years it has been shown that not all OHRB utilize the same type of cobalamin and that many, if not all, are able to modify available cobalamins to suit their specific requirements (for a recent review see [293]). As the cobalt centre in the corrinoid backbone is always conserved we will from hereon refer to these collectively as cobalamins, unless we specifically deal with a certain type of cobalamin.

The synthesis of cobalamins is a highly complex process, involving around 30 enzymatic steps. Two different cobalamin synthesis pathways have been described, an anaerobic and an aerobic one, which differ mainly by the timing of cobalt insertion and are therefore also referred to as the early or late cobalt insertion pathway, respectively. The aerobic pathway has been well characterised since long whereas all steps of the anaerobic pathway have been characterised only recently [161, 343, 344].

The presence of an intact and functional anaerobic cobalamin synthesis pathway has been reported for the genomes of *Desulfitobacterium hafniense* Y51, TCE1 and DCB-2^T [19, 55, 56]. Our analysis showed that all sequenced desulfitobacterial genomes encode intact cobalamin synthesis pathways, indicating that this is a general trait of this genus (Table 5). The synthesis pathway is organised in a similar fashion in all sequenced desulfitobacteria. It consists of three clusters (I-III) located separately in the genomes (Figure 11). Cluster I encodes the proteins needed for the synthesis of the basic ring structure from glutamyl-tRNA to sirohydrochlorin (also called factor II), cluster II encodes the steps starting from cobalt insertion until adenosyl cobinamide, and finally cluster III encodes all proteins needed for the insertion of the lower ligand leading to cobalamin, with the exception of *cobC* that is located elsewhere in the genome (Figure 11 Table S1).

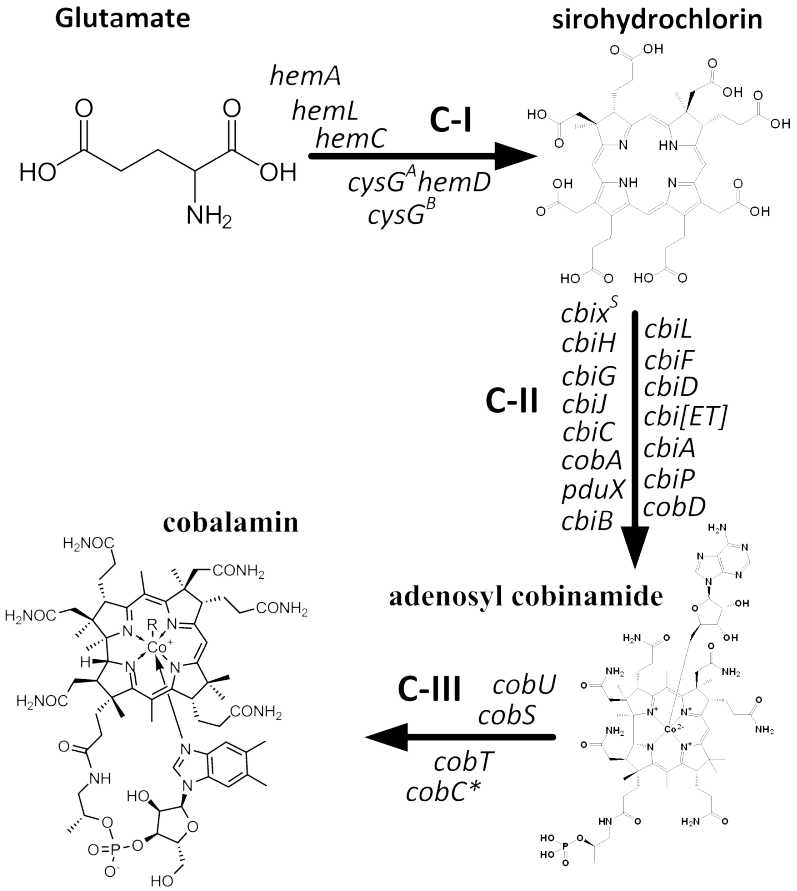


Figure 11. The cobalamin synthesis pathway of *Desulfitobacterium* spp. Substrates and end products of cluster I-III are shown. Order of gene names indicates order of the enzymatic steps catalyzed by the gene products. * Located outside cluster III. Figure modified from [55, 161].

In some bacteria e.g.. *E. coli* and *Salmonella enterica,* sirohaem synthesis from uroporphyrinogen III is mediated by a single multifunctional protein CysG, where the C-terminus contains the uroporphyrinogen III methyl transferase activity and the N-terminus the dehydrogenase and chelatase activities [345, 346]. In other bacteria, including *Desulfitobacterium* spp., the C terminal $Cysg^{A}$ has fused with HemD, leaving a short version of CysG termed $CysG^{B}$. The $CysG^{A}$-HemD, fusion protein catalyses the conversion of hydroxymethylbilane to precorrin 2, and $CysG^{B}$ catalyzes the final step to sirohydrochlorin, as described in detail for *Desulfovibrio vulgaris Hildenborough* [346]. The genomes of *D. hafniense* DCB-2^{T}, PCP-1 and TCP-A encode an additional *cysG*/*hemD* homologue. Interestingly this homologue is found in dehalocluster C (Table S1 and Figure 3). The genome of *D. dichloroeliminans* $DCA1^{T}$ encodes two copies of *cbiD* encoding a protein that catalyses the conversion of cobalt precorrin-5B to cobalt precorrin-6A. The duplication of *cbiD* seems to be the result of movement of a transposase leading to duplication of a 3.6 kb fragment encoding a transposase, *cbiD*, and a *marR* type transcriptional regulator. The transposase-encoding gene is found in 14 additional copies scattered throughout the genome of *D. dichloroeliminans* $DCA1^{T}$ (data not shown). An additional *cbiA* homologue located outside the cobalamin synthesis clusters has previously been reported for *D. hafniense* DCB-2^{T} and Y51 [55]. This homologue is present in all sequenced *Desulfitobacterium* spp. except *D. metallireducens* 853-$15A^{T}$. A closer inspection of the genomes revealed that this homologue encodes DsrN, a putative siro(haem)amidase as part of a dissimilatory sulfite reductase gene cluster [347]. The presence of multiple *cobA* homologues was reported for *D. hafniense* DCB-2; Y51 and TCE1 [55]. The *cobA* gene encodes cob(I)yrinic acid a,c-diamide adenosyltransferase catalyzing the conversion of cob(I)yrinate *a,c*-diamide to adenosyl-cobyrinate *a,c*-diamide [348].

We identified one to three additional *cobA* homologues outside the cobalamin synthesis clusters in the genomes of *Desulfitobacterium* spp. (Table S1). The product of *cobA* has, however, also been associated with uptake of corrinoids from the environment [349] and might therefore be involved in other processes than *de novo* corrinoid synthesis in desulfitobacteria.

One homologue, $cobA_1$, is present in all twelve *Desulfitobacterium* spp. in a conserved gene cluster consisting of $cobA_1$, a hypothetical protein encoding gene, and a gene coding for a tetratricopeptide repeat (TPR) protein. TPR proteins are believed to mediate protein-protein interactions [350]. The $cobA_2$ gene is present in *D. hafniense* Y51; TCE1; PCP-1, PCE-S and DP7, and is located as a single gene without any obvious functional connection with the surrounding genes.

Finally *cobA*$_3$ is found in all *D. hafniense* and *D. dehalogenans* strains together with a gene encoding methyl-malonyl CoA mutase, a vitamin B_{12} dependent enzyme [351].

Synthesis of Menaquinones (MKs)

MKs, also called vitamin K2, are lipophilic electron carriers, acting as electron shuttles between membrane bound oxidoreductases [236, 352, 353]. All sequenced genomes of *Desulfitobacterium* spp. with the exception of *D. dichloroeliminans* DCA1^T contain one copy of the *menFDHCEBIAG* MK synthesis pathway [249, 354] (Table 5, Table S1). The genes *menFDHCEB* are organized in a single operon, which is absent from the genome of *D. dichloroeliminans*. This is in line with previous findings that *D. dichloroeliminans* DCA1^T requires vitamin K2 to grow in pure culture [271]. The *menFDHCEB* operon structure is conserved in all other desulfitobacteria, with the exception that *menH* is located elsewhere in the genome in both *D. dehalogenans* strains. The enzyme catalysing the step from 1,4-dihydroxy-2-naphthoyl-CoA to 1,4-dihydroxy-2-naphthoate had until recently not been identified. Chen and co-workers identified a 1,4-dihydroxy-2-naphthoyl-CoA thioesterase as responsible for this step in *E. coli*, and proposed to designate this gene *menI* (previously *ydiI*) [249]. We identified a putative thioesterase encoding gene located directly downstream of *menG* in all sequenced desulfitobacteria. This gene likely encodes MenI in desulfitobacteria (Table S1). An alignment of the genomes of *D. metallireducens* and *D. dichloroeliminans* showed a high degree of conservation in the area up- and downstream of *menFDHCEB*. As mentioned above, the *menFDHCEB* genes and one gene at each side of these are absent in *D. dichloroeliminans.* Instead we found a 18.5 kbp region at this location. This region has a slightly higher GC deviation (G-C/G+C) compared to the direct up- and downstream part of the genome, which is an indicator of horizontal gene transfer (see [299] and references therein). Furthermore the encoded genes are oriented in the opposite direction as compared to the surrounding genomic area. These findings may suggest the 18.5 kbp region has been acquired horizontally by *D. dichloroeliminans* DCA1^T leading to loss of the *menFDHCEB* genes.

Alternative amino acids

All sequenced *Desulfitobacterium* spp. encode several genes predicted to contain selenocysteine, among which is the membrane bound formate dehydrogenase FdhABCE (Table 5, Figure 4, Table S1).The catalytic subunit FdhA contains a selenocysteine [19, 25, 26]. Selenocysteine is also referred to as the 21st amino acid and is encoded by TGA which usually read as a (opal)

stopcodon [355, 356]. In agreement with this we found the necessary gene set for synthesis and incorporation of selenocysteine in all sequenced desulfitobacteria (Table 5, Table S1), comprising *selA*, encoding a selenocysteine synthase, *selB* that codes for an elongation factor, *selC* $tRNA^{sec}$, and *selD* coding for selenophosphate synthetase [357, 358].

In addition to this it has previously been reported that *Desulfitobacterium hafniense* strains DCB-2^T, DP7, Y51 and *D. dehalogenans* JW/IU-$DC1^T$ are able to use pyrrolysine, also known as the 22^{nd} amino acid encoded by TAG that else acts as a (amber) stopcodon [359-362]. We found the pyrrolysine synthesis operon *pylScBCDSn* [363] in all sequenced members of *D. dehalogenans* and *hafniense* (Table 5, Table S1). Both *D. dehalogenans* strains JW/IU-$DC1^T$ and PCE-1 contain a 5.4 kb insert between *pylSc* and *pylB*, showing that this insertion must have happened before the split between these two strains. Interestingly, neither *D. dichloroeleminans* $DCA1^T$ nor *D. metallireducens* 853-$15A^T$ encode a *pyl* operon.

Protein secretion systems

All twelve sequenced desulfitobacteria encodes the general secretory system (SEC) (Table 5, Table S1), enabling translocation of proteins in the unfolded state across the cell membrane [364, 365]. The sec system is conserved between sequenced desulfitobacteria with the exception that in *D. hafniense* and *D. dichloroeliminans* $DCA1^T$, *secD* and *secF* are fused to encode a protein with a N-terminal SecD and a C-terminal SecF. In *D. dehalogenans* and *D. metallireducens* 853-$15A^T$, these are separate genes. The SecDF complex is located in the membrane and believed to assist in protein translocation by preventing backwards movements of partly translocated proteins [364].

In addition, all sequenced desulfitobacterial genomes encode a twin arginine translocation (TAT) pathway. The TAT system facilitates translocation of proteins, such as reductive dehalogenases, in their folded state with incorporated co-factors across the cell membrane. The TAT machinery is composed of TatB and TatC, two membrane proteins involved in recognition of the TAT signal sequence and the membrane pore consisting of a multimer of TatA proteins [59]. The TAT genes are organised in a similar way in the genomes of all sequenced desulfitobacteria. Two genes, *tatBC*, are found together in a small operon, whereas one to four copies of *tatA* are found scattered throughout the genomes (Table S1). The genomes of *D. hafniense* TCE-1, LBE and *D. dichloroeliminans* $DCA1^T$ encode an additional *tatBC* set, which is, however, disrupted by transposons and thus not functional. This additional *tatBC* gene set is located in close proximity

to genes encoding the PCE or 1,2-DCA dechlorinating reductive dehalogenases PceA or DcaA. Both *D. hafniense* PCS-S and Y51 encode *pceA*, but no additonal *tatBC* genes.

Distribution and organisation of v*anI* clusters

A novel gene cluster, *vanI*, conferring a high level (MIC > 150 µg/ml) of resistance to the glycopeptide vancomycin has previously been identified in members of the genus *Desulfitobacterium* (Table 5)[209, 281]. This gene cluster generally consists of either *vanXmurFvanKWIRS* or *vanXmurFvanWIRS* (Figure 12) Vancomycin resistance results from the production of modified peptidoglycan precursors terminating in D-alanyl-D-lactate instead of D-alanyl-D-alanine. Vancomycin exhibits low binding affinity to D-alanyl-D-lactate compared to D-alanyl-D-alanine. The function of the proteins encoded by the *vanI* gene cluster are in brief: VanX is a dipeptidase cleaving D-alanyl-D-alanine, thereby removing the target for vancomycin; MurF adds the D-alanyl-D-lactate depsipeptide to the growing peptidoglycan chain; VanK catalyzes cross linking of peptidoglycan; the function of VanW is not yet known; VanI is an ATP-dependent depsipeptide ligase involved in the synthesis of D-Ala-D-Lac; VanRS constitutes a two component transcriptional regulator [205, 217]. The products of three genes are considered essential for vancomycin resistance, namely *vanX*, *vanI* and homologues and *vanH*. The latter encodes a D-lactate dehydrogenase converting pyruvate to D-lactate [205]. The *vanI* gene cluster is the first vancomycin resistance gene cluster described that is not encoding a *vanH* gene [209, 281]. We did not find a *vanI* gene cluster in *D. dehalogenans* or *D. metallireducens* 853-15A^{T}, whereas all *D. hafniense* strains encodes *vanI*. Unexpectedly, we also found a *vanI* gene cluster in *D. dichloroeliminans* DCA1^{T}, contradicting previous findings [281]. The sequence difference between *vanI* from *D. dichloroeliminans* DCA1^{T} and *D. hafniense* Y51 probably explains why the vanI primers designed from the D. hafniense Y51 sequence failed to amplify *VanI* from *D. dichloroeliminans* DCA1^{T} previously (Figure 11)

The genomes of two strains, *D. hafniense* Y51 and PCE-S, contain *vanXmurFvanKWIRS*, whereas the genomes of five other *D. hafniense* strains harbour *vanXmurFvanWIRS* (Table 5, Figure 12). *D. hafniense* TCP-A encodes a truncated and differently organised *vanI* cluster consisting of *vanWIX*. In *D. dichloroeliminans* DCA1^{T} the *vanI* gene cluster is also organised as *vanWIX* with the genes encoding the two component transcriptional regulator VanRS located approximately 8.6 kbp upstream. In *D. hafniense* strain TCP-A *vanRS* has been disrupted by a hypothetical gene leaving a small pseudogene corresponding to the 3' end of *vanR* as the only remnants. Directly

upstream of the hypothetical gene we found two genes encoding a toxin/antitoxin addiction system. In both *D. dichloroeliminans* DCA1^T and *D. hafniense* strain TCP-A a gene, *vanZ*, encoding a protein conferring low levels of teicoplanin resistance are encoded on the opposite strand directly downstream of *vanRS* or the truncated *vanR*, respectively (Figure 12) [366]. The *vanZ* gene was not found elsewhere on the genome of any other sequenced desulfitobacterial genome.

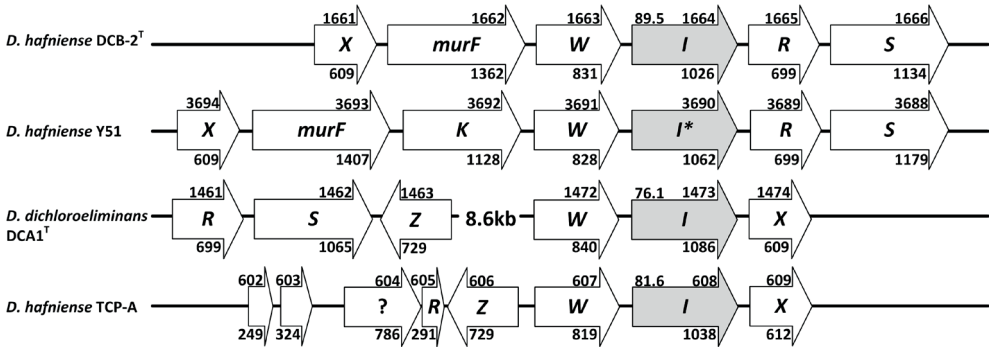


Figure 12. Alignment of vancomycin resistance gene clusters representing the different organisations found in the sequenced genomes of *Desulfitobacterium* spp. Arrows show the direction of transcription, while arrow size indicates gene size. Letters inside arrows denote gene names with the '*van*' prefix removed. ?: hypothetical gene. Light grey arrows represent the D-Ala-D-Lac ligases giving name to the gene clusters. The two genes (602-3) shown at the 5'end of the *D. hafniense* TCP-A clusters are predicted to encode a toxin/anti toxin system, see text for details. Numbers under arrows indicate gene size in bp, numbers at the 3' end above arrows indicate locus tags without their prefix (see Table 1). Numbers above the *vanI* genes at the 5' end show percentage amino acid identity to the first characterised VanI from *D. hafniense* Y51 [209]. Amino acid identities were calculated using MatGAT with default settings [113]. The figure was adapted from [281]

Chemotaxis and Motility

Desulfotobacteria are generally described as motile, and genes encoding the *che* chemotaxis signalling pathway have previously been reported as present in the genome of *D. hafniense* DCB-2^T [16, 66]. We found genes encoding *cheAYWRCD* in all available desulftobacterial genomes (Table 9), with the exceptions that *cheWY* has been disrupted by a transposase in *D. metallireducens* 853-15A^T and that the genome of *D. hafniense* strain TCE-1 apparently does not encode any *cheAWY* genes. However, in strain TCE-1, the *cheR* gene is located at the end of a scaffold and the *cheCD* genes are located close to the edge of another contig. It is therefore possible that the missing genes are present in strain TCE-1, but located in a gap between *cheR* and *cheCD* in the draft genome (Table 2 and 9). The *che* genes are located in a highly conserved region on the genomes, encoding large numbers of genes involved in motility such as synthesis of flagella and flagella motor proteins.

The function of the Che proteins are, in brief, as follows: CheRWA interact with the membrane bound methyl accepting chemotaxis protein (MCP). When the MCP reacts to an external stimulus, CheA gets phosphorylated and then phosphorylates the response regulator CheY. The phosphorylated CheY binds to the flagellar motor causing an alteration in the direction of swimming. The exact role of CheCD is not known. [66, 367].

Table 9. Locus tags, without prefix see (Table 1), of the *che* chemotaxis signalling pathway [367]. * disrupted by a transposase.

	cheA	***cheW***	***cheR***	***cheY***	***cheC***	***cheD***
D. hafniense DCB-2 ^{T}	4155	4154	4153	4203	4130	4129
D. hafniense Y51	2994	2993	2992	3036	2970	2969
D. hafniense TCP-A	2208	2209	2210	2174	2232	2233
D. hafniense PCP-1	4077	4078	4079	4040	4101	4102
D. hafniense TCE-1			1382		4593	4594
D. hafniense DP7	4681	4680	4679	4720	4658	4657
D. hafniense LBE	4473	4474	4475	4436	4497	4498
D. hafniense PCE-S	3233	3232	3231	3266	3208	3209
D. dehalogenans JW/IU-DC1 ^{T}	3530	3529	3528	3562	3506	3505
D. dehalogenans PCE-1	3414	3413	3412	3440	3390	3389
D. dichloroeliminans DCA1 ^{T}	2877	2876	2875	2916	2853	2852
D. metallireducens 853-15A ^{T}	2528	2526-2527*	2524*	2554	2502	2501

Conclusions and future perspectives

In the present contribution, we present the full genome sequence of eight members of the *Desulfitobacterium* genus, including the first genomes of *D. dichloroeliminans* and *D. metallireducens*. When the genomes of *D. hafniense* Y51 and *D. hafniense* DCB-2^T were published, it was noticed that they only encode a limited number of reductive dehalogenases. Both genomes were found to encode the full cobalamin synthesis pathways and to provide the genetic machinery for utilizing a broad range of organic and inorganic electron donors and acceptors [19, 26]. In fact, both genomes were reported to encode the largest number of molydopterin oxidoreductases of any fully sequenced bacterial genome at the time of publication [26]. The sequencing of ten additional desulfitobacterial genomes (Table 2) confirmed this general trend, with 65 predicted molydopterin oxidoreductases encoded on the genome of *D. hafniense* TCE-1 as an impressive maximum (Table 5). Furthermore, all but one of the genomes encode the genes necessary for utilizing both of the two alternative aminoacids selenocysteine and pyrrolysin (Table 5). A striking observation made here was the fact that the four sequenced *Desulfitobacterium* species. have large differences in genome size (Table 2). The data presented here shows, not too surprising, that the decrease in genome size is reflected in a smaller metabolic potential (Table 5, 8). For many of the traits not found in the species with smaller genomes we found truncated pathways or genes, suggesting that these traits have been present and subsequently lost, as exemplified by the discussion on arsenate resistance (Figure 6). Thus, gene loss at least partially explains the genome size differences between the *Desulfitobacterium* species, although gene gain by *D. hafniense* also may have played a role. More detailed analysis may give a better understanding of the role of gene gain and gene loss in the evolution of desulfitobacteria.

Main focus on desulfitobacteria has been on their possible application for bioremediation due to their ability to carry out OHR. The genomes of obligate OHRB such as *Dehalocoocoides mccartyi* and *Dehalobacter restrictus* PER-K23^T encode large numbers of reductive dehalogenases, ranging from 11 to 36, located in two genomic clusters. The content of these clusters is highly variable between strains, it has been speculated that this is due to frequent horizontal gene transfers taking place in these regions [293]. This is not the case in desulfitobacteria, where we found a limited number of reductive dehalogenases, that do not seem to be the result of recent horizontal genetransfers. Two dehaloclusters, D and E, (Table 4, Figure 3) were found in members of both *D. hafniense* and *D. dehalogenans.*

The fact that both dehaloclusters as well as the surrounding areas of the genomes are conserved between the species indicated that these must have been present in the last common ancestor of *D. hafniense* and *D. dehalogenans*. A notable exception is the gene cluster *pceA*/*dcaA* that is located on a functional transposon (Figure 3) [116, 118]. This transposon was found integrated in a different region of the genome of all five strains where it is found, likely due to several events of excision and reintegration in the chromosome (data not shown).

One very interesting finding is the identification of genes encoding methyl accepting chemotaxis proteins, RdhO, in dehalocluster C and D, and a partial *rdhO* gene in dehalocluster A. Expression of *rdhO* from dehalocluster C has previously been shown to be induced in the presence of chlorophenols [66]. A chemotactic response towards organohalides has not been demonstrated for any desulfitobacteria, or to the best of our knowledge any other OHRB. Using bacteria that actively move towards and dehalogenate organohalides would be a very attractive solution for bioaugmentation [368], as they potentially could be applied at a few spots and then would be able to move around in the sediment leading to efficient removal of the contamination. Very recently three papers reported for the first time heterologous expression of functional reductive dehalogenases from *Desulfitobacterium*, *Dehalococcoides mccartyi* and *Nitratireductor pacificus* [69, 79, 80]. It would be of great interest to express and functional characterise the previously uncharacterized reductive dehalogenases from *Desulfitobacterium* spp. and thereby unlock their full potential for bioremediation of environments contaminated with organohalides.

We have in the present contribution sought to link insight from genome analysis with current knowledge on the metabolism of desulfitobacteria. It is obvious from the presented data that our current knowledge is still far from complete. For example, for the vast majority of the predicted *O* demethylases and molydopterin oxidoreductases the substrate is not yet known although some advances especially on the *O* demethylases have been made in recent years [301, 324]. We found indications that desulfitobacteria may be used for bioremediation of arsenate contamination (Table 5 Figure 6), suggesting that desulfitobacteria may have a role in bioremediation beyond OHR.

One of the remaining questions is the composition of the electron transport chain from the MK pool to the reductive dehalogenase, as this as yet is not known [80]. Involvement of a NapGH like membrane bound quinol dehydrogenase encoded by *rdhMN* was recently repor-ted from *Sulfurospirillum multivorans* [29]. Comparative proteomics revealed that a diflavin encoded by an FRP increased strongly in abundance in *D. dehalogenans* JW/IU-DC1^T during growth by OHR, leading to the suggestion of possible role as an electron carrier between a membrane bound electron donating complex and the reductive dehalogenase [25]. Dehalocluster E encodes an *rdhMN* gene set, but as this cluster is only present in three strains (Figure 3), other links between the MK pool and the reductive dehalogenases must exist. Genes encoding likely candidates for this link could be identified by genome analysis combined with transcriptomics or proteomics, and verified by directed gene knockouts as described for the *frdABC* operon in *D. dehalogenans* JW/IU-DC1^T [329].

Finally, comparative genomics both within the *Desulfitobacterium* genus and across genera, with for example members of the closely related genus *Dehalobacter*, will help unravel genetic traits associated with adaptation to a specific niche as exemplified by the identification of three genes from dehalocluster C as also being present in the *rdh* B genomic region in *Dehalobacter restrictus* PER-23.

The availability of complete genome sequence information from twelve desulfitobacteria has given us an invaluable road map to the journey towards understanding the metabolic capacity of this truly fascinating genus.

Acknowledgements

This work was supported by the Ecogenomics Project of the Netherlands Genomics Initiative, as well as the IOP Environmental Biotechnology program.

We furthermore thank the European Community program FP7 (grants KBBE-211684, BACSIN, and KBBE-222625, METAEXPLORE) for financial support.

We are greatly indebted to the U.S. Department of Energy Joint Genome Institute that is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231 for sequencing and annotation of seven of the here described desulfitobacterial genomes.

We are greatly indebted to George M. Weinstock, Erica Sodergren and Elizabeth A. Lobos from the Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park, St. Louis, MO63108, USA for sequencing and annotation of the *Desulfitobacterium hafniense* DP7 genome.

Chapter 8

Comparative genomics and transcriptomics of organohalide respiring bacteria and regulation of *rdh* gene transcription

Thomas Kruse, Hauke Smidt and Ute Lechner

Accepted for publication as: Kruse T, Smidt H, Lechner U (2015) Comparative genomics and transcriptomics of organohalide respiring bacteria and regulation of *rdh* gene transcription. In: Adrian, L Löffler FR (eds) Organohalide Respiring Bacteria. Springer

Abstract

Comparison of the genomes of organohalide-respiring bacteria has improved our understanding of the genetic background of the organohalide respiration process. In this chapter the remarkable differences between obligate and facultative organohalide-respiring bacteria in the number of reductive dehalogenase encoding genes and the numbers and types of accessory genes are discussed in relation to different lifestyles and evolutionary aspects. Furthermore, the putative function of accessory genes is discussed and a unifying nomenclature is proposed. The genomes also reflect distinct mechanisms for the synthesis or acquisition of the corrinoid cofactors of reductive dehalogenases, which are well in accord with the observed growth requirements of the respective organohalide-respiring bacteria. The value of microarray-based comparative genomics, transcriptomics and quantitative transcription analyses for understanding the physiology and environmental significance of organohalide respiration is discussed. The reductive dehalogenase genes are in general associated with genes encoding transcriptional regulators, which are likely involved in sensing the halogenated electron acceptors. The role of two types of regulators in transcriptional regulation of organohalide respiration has been investigated. A multiple antibiotic resistance regulator (MarR)-type regulator was shown to regulate negatively the transcription of reductive dehalogenase genes in *Dehalococcoides mccartyi*. In *Desulfitobacterium hafniense*, the cAMP receptor protein/fumarate and nitrate reduction (CRP/FNR) regulator, CprK, activates transcription of reductive dehalogenase genes. The molecular mechanism of how *ortho*-chlorophenols act as effectors has been elucidated and how, through the induction of structural changes, they lead to DNA binding of the regulator.

The full genome sequences of two organohalide-respiring bacteria (OHRB) were published in 2005, just 10 years after the publication of the first bacterial genome [17, 18, 369]. These two genome sequences were obtained for *Dehalococcoides mccartyi* strains CBDB1 and 195^{T} (previously *D. ethenogenes* 195^{T}, here and throughout this chapter ^{T} indicates a type strain) both of which belong to the *Dehalococcoides* genus, whose members are known as obligate OHRB using hydrogen as sole electron donor [11]. In 2006 the first genome of a facultative OHRB, *Desulfitobacterium hafniense* Y51, was published [19]. Currently (March 2015) 23 genomes of OHRB have been formally published, and many more are listed in the Genomes Online database (GOLD; https://gold.jgi-psf.org/) as planned or on-going projects (Table 1) [20]. With the rapid advances in sequencing technology and the trend towards publishing the full genome sequence, together with the physiological characterisation of new strains, it is expected that these numbers will increase in the future.

The availability of full genome sequences has boosted our understanding of the physiology and evolution of OHRB. The genome size gives a crude indication of an organism's metabolic capacity. Organisms with a diverse metabolism therefore generally have larger genomes than organisms inhabiting more restricted ecological niches. In line with this assumption, the obligate organohalide-respiring *Dehalococcoides mccartyi* and *Dehalogenimonas lykanthroporepellens* have very small genomes with an average size of 1.4 Mbp, which is among the smallest among free-living bacteria (Table 1) [17], whereas *Dehalobacter* spp. have an intermediate genome size of 3 Mbp, which is rather large for an organism with such a restricted metabolism (Table 1). It has been shown that at least some *Dehalobacter* strains are able to ferment dichloromethane, which suggests this genus has a more extended metabolic repertoire than previously assumed [41, 42]. In contrast, facultative OHRB in general have larger genomes ranging from 3.2 to 6.5 Mbp (Table 1). For example, the genomes of *Desulfitobacterium* spp. encode the largest numbers (53 to 57) of molybdopterin oxidoreductases (protein family (pfam) 01568) of any organism with a sequenced genome, highlighting their versatile metabolism. The function of these proteins is to a large extent unknown. However, an example of a well-characterized member of this group is the catalytic subunit FdhA of the membrane-bound, outward-facing formate dehydrogenase FdhABC (for details see chapter 10, of the book wherein this chaper will be published) [19, 25, 26]. Furthermore, genome sequencing of obligate OHRB such as *Dehalococcoides mccartyi*, *Dehalogenimonas lykanthroporepellens* and *Dehalobacter* spp. has revealed large numbers of genes encoding reductive dehalogenase homologues (Rdh), ranging from 11 to 36 in number (Table 1, and references

therein). Generally, facultative OHRB carry a lower number of reductive dehalogenase-encoding genes, ranging from one to seven for currently available OHRB genomes. These *rdhAB* genes are scattered throughout the genomes (Table 1, and references therein). Notably, some strains of the genus *Desulfitobacterium* don´t carry *rdhAB* genes in their genomes. A possible explanation could be that the genes encoding reductive dehalogenases have been lost from the genome prior to isolation or genome sequencing. For example, genes encoding reductive dehalogenases were absent in the genome of *Desulfitobacterium metallireducens*^{T} although this strain was previously reported to be an OHRB [272].

Dehalococcoides mccartyi strains have highly conserved genomes and the core genome, i.e. genes present in all members of the group (strains 195^{T}; CBDB1; BAV1 and VS), has been reported to comprise 1029 to 1118 genes, corresponding to 63-77 % of the genome [31, 121]. The inter-strain variability is mainly found in two genomic regions located at either side of the origin of replication, often referred to as 'high plasticity regions' [17, 31]. Most of the *rdh* genes are located in these regions, leading to the speculation that the 'high plasticity regions' serve as evolutionary hotspots with frequent horizontal gene transfers, allowing rapid evolution of *Dehalococcoides mccartyi* strains while at the same time preserving core metabolic functions [31]. Interestingly, 21 out of 25 *rdh* genes encoded in the genome of *Dehalobacter restrictus* PER-K23^{T} are also located in two genomic regions. The majority of the *rdh* genes from these regions were not found in the genome of the closely related *Dehalobacter* sp. E1, although the genomes showed an otherwise high degree of similarity [22, 24]. *Dehalococcoides* and *Dehalobacter* belong to the Chloroflexi and Firmicutes phyla, respectively, indicating that evolution favours this genomic organisation for metabolically restricted OHRB, and it is generally accepted that horizontal gene transfer was crucial in the spreading of *rdh* genes [19, 24, 31, 370].

Table 1. Basic genome characteristics of organohalide-respiring bacteria for which the genome sequences are available. [1] Plasmid, [2] not counting the non-canonical 2.4.6-trichlorophenol reductase, CrdA [371], [3] available through the European Molecular Biology Laboratory (EMBL), [4] Genome online database (Gold) ID. [T] Type strain.

Organism	Genome size (Mbp)	GC %	Genes #	RDH #	Closed	Accession #	Published	Reference
Dehalococcoides mccartyi 195[T]	1.5	49	1647	17	Yes	NC_002936	2005	[18]
Dehalococcoides mccartyi CBDB1	1.4	47	1517	32	Yes	NC_007356	2005	[17]
Dehalococcoides mccartyi VS	1.4	47	1489	36	Yes	NC_013552	2009	[31]
Dehalococcoides mccartyi BAV1	1.3	47	1443	11	Yes	NC_009455	2009	[31]
Dehalococcoides mccartyi BTF08	1.5	47	1580	20	Yes	CP004080	2013	[372]
Dehalococcoides mccartyi DCMB5	1.4	47	1526	23	Yes	CP004079	2013	[372]
Dehalococcoides mccartyi CG1	1.5	47	1640	35	Yes	CP006949	2014	[373]
Dehalococcoides mccartyi CG4	1.4	49	1510	15	Yes	CP006950	2014	[374]
Dehalococcoides mccartyi CG5	1.4	47.	1490	26	Yes	CP006951	2014	[374]
Dehalococcoides mccartyi SG1	1.4 (23 kbp)[1]	47	1486	28	Yes	JPRE01000000.	2014	[374]
Dehalococcoides mccartyi GT	1.4	47	1483	20	Yes	NC_013890		Unpublished, JGI
Dehalococcoides mccartyi GY50	1.4	47	1591	25	Yes	NC_022964.1		Unpublished, National University of Singapore
Dehalogenimonas lykanthroporepellens	1.7	55	1771	25	Yes	Nc_014314	2012	[21]
Dehalobacter sp E1	2.6	45	2587	10	No	CANE01000001 - CANE01000102	2012	[22]
Dehalobacter sp CF	3.1	44	3040	17	Yes	CP003870	2012	[23]
Dehalobacter sp 11DCA	3.1	45	3038	17	Yes	CP003869	2012	[23]
Dehalobacter restrictus PER-K23[T]	2.9	45	2908	25	Yes	CP007033	2013	[24]
Dehalobacter sp UNSWDHB	3.2	45	3041	17	No	AUUR00000000	2013	[52]
Desulfitobacterium hafniense Y51	5.7	47	5208	1[2]	Yes	NC_007907	2006	[19]
Desulfitobacterium hafniense DCB-2[T]	5.3	48	5042	7[2]	Yes	CP001336	2012	[26]
Desulfitobacterium dehalogenans JW/IU-DC1[T]	4.3	45	4252	6[2]	Yes	NC_018017	2015	[25]
Desulfitobacterium hafniense PCE-S	5.7	47	5456	2[2]	No	LK996017-LK996040[3]	2015	[27]
Desulfitobacterium hafniense TCP-A	5	47	4877	5[2]	No	KB900390:KB900391		Kruse et al. in preparation
Desulfitobacterium hafniense PCP-1	5.6	48	5358	7[2]	No	KB902317:KB902361		Kruse et al. in preparation
Desulfitobacterium hafniense TCE-1	5.7	47	5570	1[2]	No	Gp0007974		Kruse et al. in preparation
Desulfitobacterium hafniense DP7	5.2	48	5456	0[2]	No	JH414432:JH414494		Kruse et al. in preparation
Desulfitobacterium hafniense LBE	5.5	47	5341	2[2]	Yes	Gp0007973		Kruse et al. in preparation
Desulfitobacterium sp. PCE-1	4.2	45	4123	6[2]	Yes	Gp0006289[4]		Kruse et al. in preparation
Desulfitobacterium dichloroeliminans LMG[T]	3.6	44	3537	1[2]	Yes	NC_019903		Kruse et al. in preparation
Desulfitobacterium metallireducens[T]	3.2	42	3152	0	Yes	NZ_AGJB00000000		Kruse et al. in preparation
Geobacter lovley SZ[T]	3.9 (77kbp)[1]	55	3777	2	no	NC_010815	2012	[28]
Shewanella sediminis HAW-EB3[T]	5.5	46	4666	5	Yes	NC_009831	2007	Unpublished, JGI
Sulfurospirillum multivorans[T]	3.2	41	3301	2	Yes	CP007201		[29]
Anaeromyxobacter dehalogenans 2CP-C	5	75	4421	2	Yes	CP000251.1	2008	[30]
Anaeromyxobacter dehalogenans 2CP-1[T]	5	75	4540	2	yes	CP001359.1		Unpublished, JGI
Desulfomonile tiedjei DCB-1[T]	6.5 (27kbp)[1]	50	5628	1	Yes	CP003360.1		Unpublished, JGI

Comparative genomics and functionality of corrinoid biosynthesis

The availability of full genome sequences has paved the way for further research on OHRB. Microarrays have been designed for both *Dehalococcoides mccartyi* and *Desulfitobacterium hafniense* DCB-2^T for comparative genomics and transcriptomics [26, 31, 32, 174, 375-377]. The first OHRB microarray was designed based on the genome of *Dehalococcoides mccartyi* 195^T. This microarray was used to analyse gDNA obtained from a microbial consortium, ANAS, containing various *Dehalococcoides mccartyi* strains able to degrade PCE completely to ethene [174]. DNA from the ANAS culture hybridized with 87% of the probes. The majority of the probes without a match were for genes located in predicted insertion elements or the previously mentioned 'high plasticity regions', revealing for the first time that the presence of 'high plasticity regions' is a general characteristic of *Dehalococcoides mccartyi* and not merely a feature of the 195^T and CBDB1 strains [174]. The detection range of *Dehalococcoides mccartyi* microarrays was later expanded by adding probes from additional strains as more genomes became available [376] or by designing probes targeting consensus sequences conserved between strains of *Dehalococcoides mccartyi* in order to construct a pan-genome microarray targeting both characterized and uncharacterized *Dehalococcoides mccartyi* strains [377]. Finally a shotgun metagenome microarray has been constructed, based on sequences from the KB-1 enrichment culture containing several *Dehalococcoides mccartyi* strains [378]. Microarrays have been used to gain insight into the global response of OHRB to external factors such as oxygen stress, the presence of symbionts, and variations in the levels of nutrients or corrinoids [26, 32, 184, 199, 379]. Another interesting finding using these microarrays was increased transcription of prophage genes in the late growth phase [133] or after starving the culture by omitting chlorinated compounds from the medium [378]. Although no *rdhA* genes were found within the prophage region, the *tceA* gene, encoding an RDase that dechlorinates trichlorethene to dichloroethene, was shown to be located in close proximity to a prophage in the KB-1 culture and in strain BTF08 [372]. Thus, it can not be excluded that mispackaging of phage particles occasionally leads to transfer of *tceA* together with the phage [378]. It has been speculated that phages might act as shuttles for horizontal transfer of *rdh* genes in both *Dehalococcoides mccartyi* [372, 378, 380] and *Sulfurospirillum multivorans* [29], although experimental evidence for this hypothesis is still lacking.

In some cases the presence or absence of specific genes has led to directly testable hypotheses. For example, genes encoding a protein with high similarity to PceA, catalysing the degradation of tetrachloroethene and trichloroethene to dichloroethene, was discovered in the genome of

Dehalobacter sp. E1 and *Shewanella sediminis* [54, 57, 381]. Similarly, the predicted minimum gene set for nitrogen fixation was identified in the genome of *Desulfitobacterium hafniense* DCB-2^T [26, 265]. The functionality of both was confirmed experimentally, reinforcing the design of experiments based on genome analysis [26, 54, 381]. On the other hand, analysis of the genome of *Dehalococcoides mccartyi* 195^T suggested incomplete synthesis pathways for some amino acids [18]. However, later experimental findings demonstrated that complete amino acid synthesis machineries are present in *Dehalococcoides mccartyi* 195^T, indicating the existence of new biosynthetic pathways [122].

Another important result of comparative genomics of OHRB was the identification of differences in the acquisition of the corrinoid cofactor. The RdhA catalytic subunit of reductive dehalogenases requires incorporation of a corrinoid cofactor in the active site [177]. The facultative OHRB, *Geobacter lovleyi*, *Desulfitobacterium* spp. and *Sulfurospirillum* spp. all encode the full gene set for *de novo* synthesis of corrinoids [19, 25, 26, 28, 29]. The latter has been shown to synthesise norpseudo B_{12}, a novel type of corrinoid [60]. In contrast, the obligate OHRB, *Dehalococcoides mccartyi, Dehalogenimonas lykanthroporepellens*, *Dehalobacter restrictus* and *Dehalobacter* sp. E1, all encode incomplete *de novo* corrinoid synthesis pathways [21, 22, 24, 32, 33, 63]. The presence of a seemingly intact corrinoid synthesis pathway has been observed in the genome of four *Dehalobacter* spp. strains [382]. It has not yet been determined, however, whether these strains are capable of *de novo* corrinoid synthesis.

This may seem counterintuitive considering the essential role of corrinoids for the metabolism of obligate OHRB. A possible explanation could be the energy cost associated with *de novo* corrinoid synthesis, a process involving more than 30 enzymatic steps [344]. Organohalides are present in limited quantities in pristine natural environments, probably favouring a salvaging strategy rather than *de novo* synthesis of corrinoids for obligate OHRB [2]. Two studies investigated the effect of cultivating the corrinoid prototrophic *Desulfitobacterium hafniense* TCE1 or Y51 in the presence or absence of corrinoids. It was found that the absence of corrinoids led to either a prolonged lag phase or faster loss of the transposon encoding the corrinoid-containing RDase, PceA, [55, 56]. Both *Dehalococcoides mccartyi* and *Dehalobacter* spp. encode the prokaryotic *btuBFCD* corrinoid uptake system belonging to the ABC transporter family [22, 33, 115, 383]. Inter-species transfer of corrinoids between different symbionts and *Dehalococcoides mccartyi* has been demonstrated [63, 182, 184].

The nature of the lower ligand is essential for corrinoids utilized by *Dehalococcoides mccartyi* as 5',6'-dimethylbenzimidazole (DMB) and some other ligands promoted growth, whereas corrinoids containing e.g. phenolic lower ligands did not [61]. *Dehalococcoides mccartyi* can substitute the lower ligands to yield the suitable form of corrinoids if the correct lower ligand is present in the medium [61, 184]. In fact it has been shown that remodelling of corrinoids is intimately linked to salvage [61]. The *btuFCD* genes encoding the corrinoid transporter, are located next to *cbiZ*, which codes for the amidohydrolase responsible for cleaving the lower ligand in both *Dehalococcoides mccartyi* and *Dehalobacter* spp. [61, 149, 184, 382]. A recent study indicated remodelling of the phenolic ligand-containing corrinoids provided by the co-cultivated *Pelosinus fermentans* R7 to *Dehalococcoides mccartyi* 195^{T} in the presence of added DMB. Based on a microarray analysis, up-regulation of *btuFCD* expression was confirmed, whereas an uncharacterised operon encoding a putative Fe^{3+}/cobalamin transporter lacking an associated c*biZ* gene was strongly down-regulated in co-cultures supplemented with DMB compared to co-cultures treated with B_{12} [184]. Based on these observations, it is tempting to speculate on the presence of two salvage systems, one of which would be dedicated to corrinoid remodelling and a second system would be responsible for uptake without adjustment of the lower ligand. To what extent remodelling of corrinoids takes place in *Dehalogenimonas* spp. and *Dehalobacter* spp. has still not been resolved. It is noteworthy, however, that the genomes of *Dehalococcoides mccartyi* encode several c*biZ* homologues associated with BtuFCD transporters and reductive dehalogenases (Table 2). This will be discussed in more detail together with other genes associated with *rdhA* later in this chapter.

Organisation of *rdh* gene clusters

With few exceptions, the *rdhA* genes encoding the catalytic subunits of reductive dehalogenases are adjacent to a small gene, *rdhB*, coding for a hydrophobic protein believed to act as a membrane anchor for the RdhA [58, 64, 65]. This minimal *rdhAB* gene cluster is frequently accompanied by a variable set of accessory genes, for most of which the exact function is still unkown. Some of these have been shown to encode proteins regulating expression of the *rdhAB* genes [66, 67, 72] or are chaperones probably assisting the folding of RdhA [68-70]. Traditionally, the nomenclature of these accessory genes has followed the name of the reductive dehalogenase subunit, followed by a letter specific for genes predicted to be homologues of previously described genes. Rather confusingly in some cases, however, different letters have been assigned to accessory genes predicted to encode proteins with similar functions, both within the same *rdh* gene cluster or between different organisms. Similarly, the same letter has been used for genes for which the

gene product seems to have different functions (Table 2) [17, 31, 67, 73]. As more genomes of OHRB become available it will be necessary to develop an uniform nomenclature for these accessory *rdh* genes.

In the following paragraphs we will give a short overview of accessory genes identified in *Dehalococcoides mccartyi, Dehalobacter* spp., *Desulfitobacterium* spp. and *Sulfurospirillum* spp. comprising the four most thoroughly studied genera of OHRB. We also propose names for *rdh*-associated genes that have not yet been named as such. The known or proposed function of Rdh proteins is schematically shown in Figure 1. Figure 2 shows exemplarily the organization of *rdh* gene clusters in Firmicutes and Chloroflexi.

Table 2. Accessory genes found to be associated with *rdhAB* genes from four well-characterised genera of organohalide-respiring bacteria and the elucidated or predicted function of the encoded proteins. Columns from left to right: 1) currently used nomenclature; proposed changes of previously published names are indicated between brackets; #indicates that this gene had not previously been assigned an *rdh* prefix; 2) first traceable use of suffix for genes encoding reductive dehalogenase or associated genes; 3) verified or predicted function; 4-7). Representative examples from *Desulfitobacterium*, *Dehalococcoides*, *Dehalobacter* and *Sulfurospirillum* are shown. Gene names are given when the substrate of the encoded proteins is known, followed by locus tag and NCBI gene identity (GI) number, *unclear if associated with organohalide respiration.

Nomenclature	First traceable use of name	Verified or predicted function	*Desulfitobacterium*	*Dehalococcoides*	*Dehalobacter*	*Sulfurospirillum*
rdhA	[65]	Reductive dehalogenase, catalytic subunit	*cprA*, Dhaf_0737, 219537062	*vcrA*, DhcVS_1291, 270154553	*pceA*, DEHRE_12145, 570739860	*pceA*, SMUL_1531, 584602139
rdhB	[65]	Reductive dehalogenase, membrane anchor	*cprB*, Dhaf_0736, 219537061	*vcrB*, DhcVS_1290, 270154552	*pceB*, DEHRE_12140, 570739859	*pceB*, SMUL_1532, 584602140
rdhC	[73]	Putative function in regulation or electron transport	*cprC*, Dhaf_0738, 219537063	*vcrC*, DhcVS_1289, 270154551	*pceC*, DEHRE_12135, 570739858	Not observed
rdhC (Use discouraged, changed to *rdhS*)	[17]	Sensory histidine kinase, part of two component regulatory system	see *rdhS* for examples			
rdhD (Use discouraged, changed to *rdhP*)	[17]	DNA binding response regulator, part of two component regulatory system	see *rdhP* for examples			
rdhD (Use discouraged, changed to *rdhE*)	[73]	Chaperonine 60 protein	*cprD*, Dhaf_0735, 219537060	Not observed	DEHRE_03960, 570738474	Not observed
rdhE	[73]	Chaperonine 60 protein	*cprE*, Dhaf_0740, 219537065	Not observed	DEHRE_03955, 570738473	Not observed
rdhF	[17]	CbiZ like protein, potentially involved in modification or insertion of corrinoids	Not observed	cbdb_A87, 73748025	Not observed	Not observed

Table 2. Continiued from opposite page

rdhG	[17]	hypothetical protein, putatively involved in insertion of metal co-factors	Not observed	cbdb_A97, 73748036	Not observed	Not observed
rdhH	[17]	hypothetical protein, containing a DUF71 domain, putatively involved in corrinoid modification or insertion	not observed	cbdbA89, 3623418	Not observed	Not observed
rdhI	[17]	radical SAM domain protein, putatively involved in co-factor modification or insertion	Not observed	cbdb_A93, 73748032	Not observed	Not observed
rdhJ (Previously named *rdhI*)	[31]	Vitamin B12 binding radical SAM domain protein	not observed	cbdb_A92, 73748031	Not observed	Not observed
rdhK	[73]	CRP/FNR type regulator	*cprK*, Dhaf_0733, 219537058	cbdbA262, 3624107*	DEHRE_03985, 570740417	Not observed
rdhM[#]		NapG like membrane bound menaquinone dehydrogenase	Desde_4093, 390526623	Not observed	Not observed	SMUL_1541, 584602149
rdhN[#]		NapH like membrane bound menaquinone dehydrogenase	Desde_4092, 390526622	Not observed	Not observed	SMUL_1542, 84602150
rdhO[#]		Methyl accepting chemotaxis protein	Dhaf_0686, 219537011		Not observed	Not observed
rdhP		DNA binding response regulator, part of two component regulatory system	Desde_4097, 13116968	DET0170, 3230558	Not observed	SMUL_1539, 584602147
rdhR	[17]	MarR type transcriptional regulator	Not observed	cbdb_A1625, 73749322	Not observed	Not observed
rdhS		Sensory histidine kinase, part of two component regulatory system	Desde_4098, 13119763	DET0171, 3230557	Not observed	SMUL_1538, 584602146
rdhT	[73]	Trigger factor	*cprT*, Dhaf_0732, 219537057	Not observed	DEHRE_12130, 570739857	Not observed
rdhZ	[73]	Putative chaperone, pfam 11068	*cprZ*, Dhaf_0734, 219537059	Not observed	DEHRE_03965, 570738475	Not observed

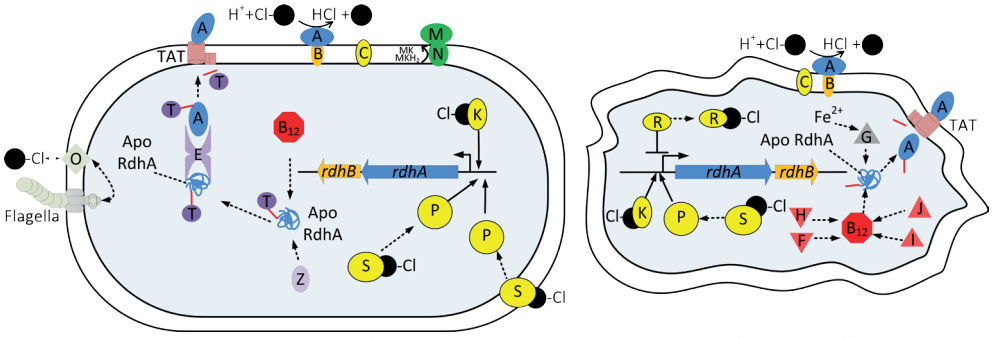


Figure 1. Schematic representation of proteins encoded by accessory genes found in *rdh* gene clusters. Left, consensus cell representing *Desulfitobacterium* spp., *Dehalobacter* spp. and *Sulfurospirillum* spp., right *Dehalococcoides mccartyi*. Note not all *rdhAB* accessory genes are found in one single organism (see Table 2). Letters inside symbols indicate the protein name without the Rdh prefix. Curved arrows, chemical reactions; dashed arrows, direction of movement or interaction; bendt arrow, transcriptional start site; straight solid lines, interaction of transcriptional regulators with *rdhA* promoters (arrow indicates activation, an inverse T repression); blue, reductive dehalogenase; orange, membrane anchor; red bar, TAT signal sequence; purple, chaperones; red, corrinoids; light red, corrinoid interacting protein, grey: metal co-factor binding protein; green: electron transferring complex; yellow, transcriptional regulators; black circles, chlorinated compounds; light green, chemotaxis and motility. See main text or Table 2 for details on function of individual proteins.

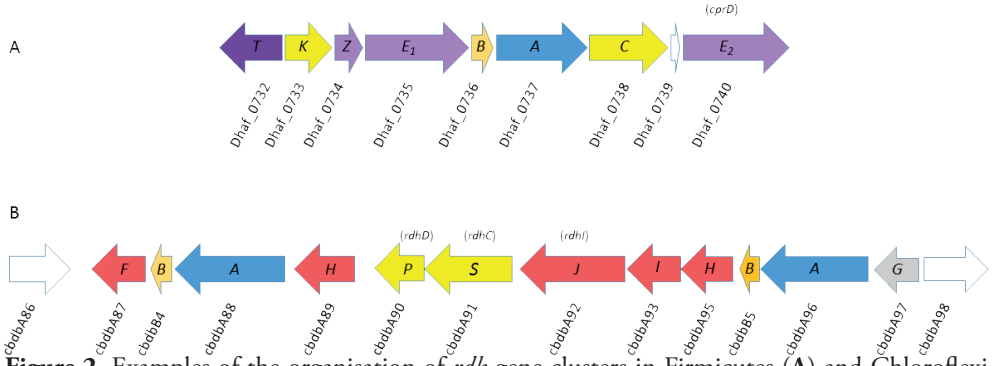


Figure 2. Examples of the organisation of *rdh* gene clusters in Firmicutes (**A**) and Chloroflexi (**B**). **A**: The *rdh* gene cluster of *Desulfitobacterium hafniense* strain DCB-2^T and *Desulfitobacterium dehalogenans* is depicted, which encodes the chlorophenol reductive dehalogenase CprA [26, 73]. **B**: A detail of 'high plasticity region' 1 of the genome of *Dehalococcoides mccartyi* strain CBDB1 is shown. The *rdhAB* and accessory genes are given without the *rdh* prefix and colored corresponding to the respective proteins shown in Figure 1. When a new gene designation is proposed the originally reported name is given in brackets (for details see the text). Empty arrows, unknown function. The locus tags are indicated under arrows.

The ***rdhA*** gene encodes the catalytic subunit of reductive dehalogenases, catalysing the removal of one or more halogens from the organohalide substrate. Reductive dehalogenases contain a twin arginine translocation (TAT) signal sequence, two Fe/S clusters and a corrinoid co-factor, although the type of the incorporated corrinoid may vary (for details see chapter 17) [58-63].

The ***rdhB*** gene encodes a small membrane protein believed to function as a membrane anchor for RdhA [58, 64, 65]. As mentioned already, the *rdhA* gene is usually accompanied by an *rdhB* gene; however, *Dehalogenimonas lykanthroporepellens* BL-DC-9T is a recently described exception where only six of twenty-five *rdhAs* have a cognate *rdhB* [21, 58]. Furthermore, the order of *rdhA* and *rdhB* (i.e. *rdhBA* vs. *rdhAB)* can differ for different gene clusters.

The ***rdhC*** gene encodes a membrane-bound flavin mononucleotide-binding protein resembling NosR- and NirI-type proteins [73, 234]. These proteins have previously been described as membrane-bound proteins controlling the expression of gene clusters encoding nitrous oxide and nitrite reductases [384, 385]. The first described *rdhC* gene was *cprC*, which is part of the *cpr* gene cluster in *Desulfitobacterium dehalogenans* JW/IU-DC1^T [73]. Based on amino acid similarity, it was speculated that RdhC might have a similar function to NosR- and NirI-type proteins [73, 116, 234]. However, it was shown that NosR likely plays a role in electron transfer [386], for a recent review see [254]. It has been recently proposed that RdhC has a function in electron transport to RdhA [387, 388], although there is still no experimental evidence to support this proposition.

rdhD see *rdhE* below

The ***rdhE*** gene encodes a protein with a high level of amino acid sequence similarity to the 60 kD subunit of the well-characterized *E. coli* GroEL/GroES two-component chaperonins [73]. This complex assists in protein folding and the mechanism has been described in detail in excellent reviews [389, 390]. A nomenclature where the large subunit is named *cpn60* and the small subunit *cpn10*, has been proposed, with multiple *cpn60*s being numbered by their order of appearance in the genome [391, 392]. The *cpr* gene cluster of *Desulfitobacterium* spp. encoding a chlorophenol RDase has been reported to harbour two *cpn60* homologues, designated *cprD* and *cprE* [71, 73]. We suggest to rename these to $cprE_1$ and $cprE_2$, following the nomenclature suggested by [391]. Consequently *cpn60* homologues found in *rdh* gene clusters should be named *rdhE* with numbers in subscript if there are multiple *rdhE* homologues, unless the substrate for the associated *rdhA* is known.

The ***rdhF*** gene encodes a protein resembling CbiZ, an amidohydrolase involved in salvaging cobinamide from the environment [17, 61]. In bacteria, *cbiZ* genes are often found in the proximity of genes encoding corrinoid transporters (BtuFCD) [393]. It has been suggested that *cbiZ* homologues located distant from genes related to synthesis or salvage of cobinamide play a role in processes other than cobalamin salvage [149, 393, 394]. A role in remodelling the lower ligand of cobalamin has been demonstrated in *Rhodobacter sphaeroides* [395]. Although no experimental evidence for the function of *rdhF* exists, it is possible that it plays a role in ensuring incorporation of the correct form of cobalamin into RDases.

The ***rdhG*** gene encodes a hypothetical protein without any signal sequence or transmembrane helices, suggesting a cytoplasmic location. This protein belongs to pfam03692. The function of this protein is not yet known but it encodes a conserved CxxCxxCC domain, which has been speculated to be involved in zinc or iron chelation [396]. Reductive dehalogenases contain two N-terminal Fe/S clusters [57, 58], hence it is possible that this protein plays a role in maturation of RDases. One part of RdhG contains a zinc-dependent carboxypeptidase-like domain, leading to the proposal that it might have protease activity [17].

The ***rdhH*** gene encodes a hypothetical protein without any signal sequence or transmembrane helices, strongly suggesting a cytoplasmic location. This protein contains a Domain of Unknown Function 71, but a recent comparative genomics study using the published genome sequences of bacteria and archaea suggested that it originates from archaea and may have a function similar to CbiZ in modifying the lower ligand in cobalamins [394].

The ***rdhI*** gene encodes a hypothetical protein without any signal sequences or transmembrane helices, strongly suggesting a cytoplasmic location. This protein contains a predicted radical SAM domain, and therefore might have a role in incorporation or modification of cofactors [17].

RdhJ. In the literature this gene has also been referred to as *rdhI* [31].We here propose to rename this gene *rdhJ* to distinguish these two non-homologous proteins. The *rdhJ* gene encodes a hypothetical protein without any signal sequence or transmembrane helices, strongly suggesting a cytoplasmic location. This protein contains both a predicted *N*-terminal vitamin B12-binding domain and a *C*-terminal radical SAM domain, again suggesting a role in the modification or incorporation of cofactors [17, 31].

The ***rdhK*** gene encodes a CRP/FNR (cAMP receptor protein-fumarate and nitrate reduction)-type transcriptional regulator. Among these, CprK from *Desulfitobacterium hafniense* DCB-2^T has been characterized in detail (see 15.5), and shown to act as a transcriptional activator of *cprA*, which encodes a chlorophenol RDase found in several strains of *Desulfitobacterium* spp. [16, 26, 73, 397]. CprA catalyses the degradation of 3-chloro-4-hydroxyphenylacetic acid to 4-hydroxyphenylacetic acid [64, 398]

The ***rdhM*** and ***rdhN*** genes encode two proteins resembling NapG and NapH, respectively. NapH is a membrane-bound Fe/S protein that receives electrons from the menaquinol pool. Electrons are channelled from NapH to the periplasmic NapG, from which the electrons are transferred to the periplasmic NapA, catalysing the reduction of nitrate to nitrite [254]. RdhN is an Fe/S protein predicted to contain several transmembrane regions, whereas RdhM contains both a signal sequence and sequences with homology to Fe/S clusters suggesting that, like NapG, it is a periplasmic protein involved in electron transfer. In the newly published genome of *Sulfurospirillum multivorans* a pair of *rdhMN* genes was found in the same gene cluster as the genes encoding the PCE dehalogenase PceA and another RDase for which the substrate is not known [29]. Expression of these genes was only detected when PCE was used as electron acceptor, but not when fumarate was used [29]. These findings, and the similarity to NapGH, makes it plausible that RdhMN facilitate electron transport from membrane-embedded menaquinones to RDases attached to the outside of the cell membrane.

The ***rdhO*** gene encodes a protein predicted to function as a membrane-bound methyl-accepting chemotaxis protein. These act as membrane-bound receptors able to detect extracellular factors like the presence of suitable electron acceptors. Information is channelled from membrane-bound methyl-accepting chemotaxis protein via the Che signalling pathway to the flagellar motor enabling the cell to alter the direction of swimming (for reviews see [367, 399, 400].

The ***rdhP*** gene was originally named *rdhD* [17]. In order to distinguish it from the unrelated aforementioned *rdhD,* we propose to rename it *rdhP* according to the function of the gene product as the phosphoryl group-accepting DNA-binding response regulator of a two-component signal transduction system. Probably regulating gene expression upon phosphorylation or dephosphorylation by the associated sensory histidine kinase RdhS.

The ***rdhR*** gene encodes a transcriptional regulator of the MarR-type. It was recently demonstrated that expression of some reductive dehalogenases in *Dehalococcoides mccartyi* is indeed regulated by RdhR proteins, as will be discussed in detail later in this chapter [72].

The rdhS gene was originally named *rdhC* [17]. To distinguish it from the previously mentioned *rdhC* we here propose to rename it *rdhS*. The *rdhS* gene encodes a **s**ensory histidine kinase that is part of a two-component signal transduction system. The histidine kinases possess diverse signal input domains often located outside the cell allowing sensing of extracellular signals, although it should be noted that, in *Dehalococcoides*, the RdhS proteins are cytoplasmic. In two-component signal transduction systems, a conserved autophosphorylation and phosphotransfer pathway links the input signals to output responses mediated by the phosphorylated response regulator [401]. RdhS is proposed to sense specific organohalides, but it may also recognise other types of stimuli such as the redox status of the cell [17].

The ***rdhT*** gene encodes a trigger factor-like protein, predicted to be involved in folding and maturation of RdhA [73]. Full-length trigger factors contain three domains: an *N*-terminal ribosome-binding domain; a peptidyl-prolyl *cis*/*trans* isomerase domain; and a *C*-terminal domain. The trigger factor binds to the large ribosomal subunits near the exit channel and chaperones the folding of newly synthesized proteins (reviewed in [294]. However, the *rdhT* gene does not encode an *N*-terminal ribosome-binding domain, showing that it most likely does not function as a classical trigger factor. It was shown that PceT from *Desulfitobacterium hafniense* Y51 and TCE1 binds specifically to the TAT signal sequence, which may delay the translocation of PceA thereby increasing the chance of correct folding and incorporation of all cofactors [68, 70]. Furthermore, the peptidyl-prolyl *cis*/*trans* isomerase activity of PceT was confirmed, which supports the suggestion that PceT and likely RdhTs in general act as dedicated chaperones for the correct folding of RdhAs [68, 70]. In agreement with this, co-expression of *pceA* and *pceT* from *Desulfitobacterium hafniense* Y51 or $rdhA_3$ and $rdhT_3$ from *Desulfitobacterium hafniense* DCB-2^T in *Shimwellia blattae* led to increased amounts of soluble reductive dehalogenase compared to when *pceA* or $rdhA_3$ were expressed alone [69]. The PceA encoding gene cluster in *Desulfitobacterium hafniense* Y51 and TCE1 is constitutively expressed [56, 130]. Cultivation of strain Y51 in medium devoid of PCE led to the formation of intracellular protein aggregates, consisting of apo-PceA with bound PceT and CobT, a protein involved in the later stage of corrinoid synthesis [56, 402, 403]. This may suggest post-translational regulation, but the role, if any, of RdhT in this process still warrants further studies.

The ***rdhZ*** gene encodes a small protein with a size of approximately 135 amino acids belonging to pfam11068 [396]. This protein does not contain any signal sequence or membrane helices suggesting a cytoplasmic localisation. Based on analysis of the crystal structure of a member of the same protein family, obtained from the non-organohalide-respiring bacterium *Synechococcus* sp. strain WH8102, a function as a molecular chaperone has been suggested (http://pfam.xfam.org/structure/4dci). The location of *rdhZ* next to the chaperone-encoding *rdhE* gene in both *Desulfitobacterium* spp. and *Dehalobacter restrictus* lends support to this speculation.

Transcription of *rdh* genes

The transcription of *rdh* genes has been studied to elucidate the molecular mechanisms underlying the perception of, and response to, organohalides. This work has also aimed to identify the RdhA enzyme catalyzing the degradation of a specific organohalide compound and to develop RNA-based tools for the prediction of dehalogenation activity. Here, the current knowledge in the field of *rdh* gene transcription and regulation is summarized.

As already mentioned most known OHRBs possess several *rdhA* genes in their genome. The function of some of these has been elucidated using traditional biochemistry (for a review see Chapter 17). Until recently, the lack of a system for functional heterologous expression of RDases [69] has prevented detailed biochemical analysis of the majority of RdhAs. Instead, transcription analyses can provide first hints to the substrate range of an RdhA when gene expression is induced by a specific substrate. Transcriptional analyses included reverse transcription of total RNA, the amplification of *rdhA* fragments using degenerate primers, followed by cloning and sequencing. This approach identified BvcA, the vinyl chloride (VC) RDase of *Dehalococcoides mccartyi* strain BAV1 [404], DcpA, the 1,2-dichloropropene-dichloro-eliminating RDase of *Dehalogenimonas lykanthroporepellens* and of non-identified *Dehalococcoides mccartyi* strains in two enrichment cultures [405], and RdhA1, responsible for dichloro-elimination of 1,2-dichloroethane by a *Dehalobacter* strain [127]. The involvement of DceA6 (MbrA) in the dechlorination of PCE to *trans*-DCE by *Dehalococcoides mccartyi* MB was also supported by transcription analyses [406]. Transcriptional analyses also gave hints to an extended substrate range of the PCE reductive dehalogenase PceA of *Dehalococcoides mccartyi* 195^T. The *pceA* gene was one of the most highly up-regulated *rdhA* genes in strain 195^T, not only during growth with PCE but also during growth with 2,3-dichlorophenol, suggesting that both compounds are substrates of PceA [132].

Simultaneous transcription of multiple *rdhA* genes has also been observed, such as in one of the first transcriptional studies conducted on the *Dehalococcoides*-containing enrichment culture KB1 [407]. This appears logical, if reductive dechlorination of a compound involves a series of dehalogenation steps. Indeed, dissection of the transcriptional response of all four *rdhA* genes *cprA*$_2$ to *cprA*$_5$ in *Desulfitobacterium hafniense* PCP-1 during the course of 2,4,6-trichlorophenol (TCP) dechlorination revealed that transcription of *cprA*$_3$, which encodes an *ortho*-dechlorinating RDase [295] (and of *cprA*$_2$ with an unknown function) preceded the induction of *cprA*$_5$ [160]. This probably reflected the turnover of 2,4,6-TCP to the transiently formed intermediate 2,4-dichlorophenol, which is a substrate of CprA$_5$ [171]. This expression pattern points to a transcriptional regulation by the successive substrates and dechlorination intermediates in *Desulfitobacterium*. Remarkably, *rdhA* transcription in obligate OHRB such as *Dehalobacter*, *Dehalococcoides* and *Dehalogenimonas* seems to be controlled by gene- and substrate-specific mechanisms as well as higher levels of regulation. Although the catabolic genes were transcribed to the highest levels in the presence of their putative substrate, transcripts of many additional *rdhA* genes were also detected. In *Dehalobacter restrictus*, transcripts of almost all 25 rdhA genes were identified during growth on PCE [115], although the *rdhA24* transcript encoding PceA was most abundant. In three representatives of the Chloroflexi, *Dehalococcoides mccartyi* strains 195^T and CBDB1 and *Dehalogenimonas lykanthroporepellens* BL-DC-9^T, a similar general up-regulation of almost all *rdhA* genes was observed. In strain 195^T, a total of 17, 13 and 9 *rdhA* transcripts were detected in the presence of PCE, TCE or 2,3-DCP, respectively, however, in levels varying over five orders of magnitude [132]. Among those, the *rdhA* gene DET0162 exhibited one of the highest transcript levels despite containing a point mutation probably leading to a non-functional protein [132, 136]. For strain CBDB1, transcription of all 32 *rdhA* genes was induced during growth on 1,2,3- or 1,2,4-trichlorobenzene [173], and a similar response was observed during growth with 2,3-dichlorodibenzo-*p*-dioxin [72], however, with a distinct up-regulation of the chlorobenzene RDase-encoding *cbrA* to the highest number of transcripts observed. Recently, the transcription of all 25 *rdhA* genes present in the genome of strain BL-DC-9^T was analysed during dichloro-elimination of 1,2-dichloroethane, 1,2-dichloropropane and 1,2,3-trichloropropane and indicated transcription of 19 *rdhA* genes [408]. Among these were 13 full-length and, similar to DET0162 in *Dehalococcoides mccartyi* 195^T, six incomplete *rdhA* genes lacking a predicted *N*-terminal TAT leader sequence or lacking iron-sulfur-cluster-binding motifs in the *C*-terminus [408]. A comparative analysis of microarray–based transcriptomic studies of *Dehalococcoides mccartyi* 195^T and *Dehalococcoides* strains in the KB-1 enrichment culture also indicated

the multitude of transcribed *rdhA* genes under a variety of conditions [409]. This apparently gratuitous expression might reflect a fundamental strategy in the obligate OHRBs for the adaptation to newly emerging halogenated compounds in their environment by ensuring the continuous presence of low levels of various *rdhA* transcripts. It is conceivable that post-transcriptional or –translational regulation mechanisms of RDase activity exist in OHRB, which is strongly supported by a recent comparison of mRNA and protein abundances per cell [410].

The expression of *rdhA* genes showed considerable differences depending on the growth phase and variations in other environmental parameters. Initial transcriptional studies quantified the up-regulation of specific functional *rdhA* genes (e.g. *tceA* and *vcrA*) in the *Dehalococcoides mccartyi*-containing mixed cultures ANAS and D2 [136, 411, 412]. Within the first hours of exposure, a several-fold increase in gene expression was observed, reaching a plateau lasting beyond the decline of active dechlorination and followed by a slow mRNA decay, which leveled off at low basal transcript numbers [412]. A similar initial response and long-lasting stability of the functional gene transcript *cbrA* to trichlorobenzene exposure was reported for *Dehalococcoides mccartyi* CBDB1 [173]. To explore the suitability of functional *rdhA* transcripts as bio-indicators for predicting *in situ* dehalogenation activity, the relation of transcript copy numbers with growth or dehalogenation activity was analysed. The pseudo steady-state respiration rates in mixed cultures containing *Dehalococcoides mccartyi* 195^{T} continuously fed with PCE or TCE were positively correlated with the *tceA* transcript abundance [413, 414]. Further, microarray-derived transcriptomic profiles suggested a close inter-relationship between the global transcriptional regulation and the organohalide respiration status of the cells [375]. DNA microarray studies of *Dehalococcoides mccartyi* 195^{T} growing on TCE demonstrated a major shift in global gene expression during transition into the stationary phase with a large number of genes being up- or down-regulated [133]. Several *rdhA* transcripts showed differential expression dynamics across the time-course of the experiment. The *tceA* gene was highly expressed throughout the experiment. Another group of *rdhA*s (DET0173, DET0180, DET1534 and DET1545) showed strong up-regulation during the transition from the exponential to the stationary phase. Interestingly, DET0180 is part of a conserved syntenic region present in all sequenced *Dehalococcoides mccartyi* genomes [57]. The delayed transcriptional induction of DET0173 and DET1545 orthologs in a mixed culture was repeatedly observed [415]. DET1545 possesses orthologs in most *Dehalococcoides mccartyi* genomes and is localised within a 12 kb DNA region with high gene synteny close to the 3′end of 'high plasticity region' 2 [31].

DET1545 and its orthologs lack a directly associated regulatory gene, however, its transcriptional up-regulation was positively correlated with starvation-related stress such as low respiration rates [375, 414], consumption of the electron acceptor [136], absence of a chlorinated compound [378, 409] or fixed nitrogen limitation [134]. This suggests an important role of the RDase in the survival strategy of *Dehalococcoides mccartyi*.

Transcripts of functional *rdhA* genes can be used as indicators of ongoing *in situ* reductive dechlorination. For instance, during bio-stimulation and bio-augmentation studies in a TCE-contaminated groundwater site [416] and in microcosms [417] the expression of *vcrA* and *bvcA* was correlated with the observed dechlorination of VC to ethene. In a PCE-dechlorinating flow column [418] the transcript abundance of *bvcA*, *tceA* and *vcrA* correlated with the zone of most active dechlorination. Also, in PCE-amended microcosms enriched from Yangtze River sediment the dynamics of transcript copy numbers of *tceA* and *vcrA* correlated with the observed turnover of TCE and VC [419]. However, transcription analyses also revealed some inconsistencies. Although a complete conversion of TCE [416, 417] or PCE [418, 419] to ethene occurred, transcripts of *pceA* or *tceA* encoding proteins catalyzing the conversion of PCE or TCE to DCE were absent or inconsistently detected. This might suggest the presence of other OHRB such as *Desulfitobacterium* or *Dehalobacter* [419] or *Eubacterium* [418], which compete successfully for the primary steps of PCE or TCE dechlorination, while their genes were probably not targeted by the primers used. Another possible reason for the discrepancy in the results might be due to the expression of *rdhA*s in *Dehalococcoides mccartyi* with an as yet unrecognized function. For instance, PCE-to-*trans*-DCE dechlorination by an enrichment culture [420] or *Dehalococcoides mccartyi* strain MB [421], containing neither *pceA* nor *tceA* pointed to the existence of other RDases with this specific function. Furthermore, transcript abundance did not always correlate with dechlorination activity. For instance, oxygen-amended, dechlorination-inactive *Dehalococcoides mccartyi* cells still formed *tceA* and *vcrA* transcripts [422]. An increase in *rdhA* gene expression was also observed at elevated temperatures in cultures with an impaired dechlorination activity suggesting that this up-regulation was a general stress response [423]. These observations highlight the necessity for a deeper understanding of composition and function of the dehalogenating community to allow an appropriate biomarker selection.

Regulation of *rdh* gene transcription

Despite the multitude of observed *rdh* transcripts from a broad range of reported OHRBs, relatively little is known about the regulatory mechanisms controlling their synthesis. From the genome sequences available it appears that these mechanisms might differ according to OHRB phylogeny rather than the organism´s status as facultative or obligate OHRB. In members of the Firmicutes, i.e. *Desulfitobacterium* and *Dehalobacter,* the canonical *rdh* operon structure (Figure 2) is well conserved and often associated upstream or downstream with *rdhK,* which encodes a member of the CRP/FNR family of transcriptional regulators [66, 115, 423] (Table 2, Figure 2). Its function in the *ortho*-chlorophenol-dependent regulation of organohalide respiration has been elucidated in detail for CprK (see below). In the Epsilonproteobacterium *Sulfurospirillum multivorans*, the genes encoding PceAB are closely associated with two *rdhPS* pairs encoding two component signal transduction system response regulators and histidine kinases (Table 2) [29]. The RdhS homologs found in *Sulfurospirillum multivorans* are predicted to contain several transmembrane helices and thus are most likely located in the cell membrane [29]. For *Sulfurospirillum multivorans*, induction of dechlorination in response to PCE or TCE concomitant with the formation of active PceA was demonstrated [424], suggesting an involvement of these two-component signal transduction systems in the regulation of organohalide respiration. Contrary to this observation, the *Sulfurospirillum multivorans pceA* mRNA and the PceA protein were present over many generations in the absence of PCE, suggesting that an additional, still unknown signal might contribute to this kind of long-term 'memory' [424]. In the *Dehalococcoidia* class of the Chloroflexi, the *rdhA* genes are mostly associated with genes encoding either two-component signal transduction system or MarR regulators. Analysis of the deduced amino acid sequences of the two-component signal transduction systems suggest a cytoplasmic localization for the histidine kinase components, as no transmembrane helices are predicted, raising the question whether the chlorinated compound itself or another signal such as the redox status of the cell is the recognized signal [17]. The MarR-type regulator family mediates cellular responses to changing environmental conditions, such as antibiotic or peroxide stress or adaptation to the catabolism of aromatic compounds [425]. It is characterized by a winged helix-turn-helix (HTH) motif mediating contact between the dimeric regulator and short palindromic sites in the target DNA. A high number of MarR-type regulators is encoded in *rdh* clusters in *Dehalococcoides mccartyi* genomes [72]. Recently, one representative was also detected in an *rdh* cluster in *Dehalobacter* sp. E1 [22]. Two further types of putative transcriptional regulators have been

predicted in *Dehalococcoides mccartyi.* In most *Dehalococcoides mccartyi* genomes a single RdhK is encoded, however, this is not closely associated with *rdhA* genes. The *vcrC* gene encoding a member of the NirI/NosR family (RdhC) is part of the *vcrABC* operon, which is essential for VC respiration in strain VS. Notably, members of the NosI/NosR family are also encoded in the canonical *rdhTKZE$_1$BACE$_2$* gene cluster of the Firmicutes (CprC; Table 2, Figure 2) [73].

The role of one MarR-type regulator, CbdbA1625, in the transcriptional control of the *rdhA* gene *cbdbA1624* in *Dehalococcoides mccartyi* CBDB1 has been studied in detail. The *rdhA* gene *cbdbA1624* is divergently oriented to the MarR-encoding gene and was specifically up-regulated in the presence of 1,2,4-trichlorobenzene [72, 173]. To determine whether CbdbA1625 exerts a regulatory function, it was heterologously produced and the interaction with the PCR-amplified intergenic region was studied *in vitro*. Electrophoretic mobility shift assays pointed to two binding sites within a perfect 40 bp palindrome, which contained short inverted repeats in its half sites typical of MarR binding motifs. This region also contained the overlapping transcriptional start sites and the -10 promoter regions of the *rdhA* and the MarR-encoding gene (*marR*). Transcriptional promoter-*lacZ* fusions introduced as single copy into the *E. coli* chromosome were used to study the *in vivo* interaction of CbdbA1625 (produced heterologously from a pBAD vector) with promoters of the *rdhA* genes cbdbA1624 (and cbdbA1453 as a negative control) and the *marR* gene. The resulting β-galactosidase activities of the *marR* [72] and *rdhA* (cbdbA1624) (Lydia Krasper, pers. communication) promoters were significantly reduced upon synthesis of CbdbA1625, whereas the activity of the control promoter was not impaired, indicating specificity of interaction and negative autoregulation (Figure 1). MarR regulators are known to be released from their target promoters by allosteric binding of a ligand [425]; however, the signals acting on MarR regulators in *Dehalococcoides mccartyi* are still unknown.

In summary, the MarR-dependent repression of *rdhA* gene transcription may be the rule rather than the exception in *Dehalococcoides mccartyi,* ensuring specific *rdhA* transcription when halogenated compounds become available.

Several studies have provided molecular insight into *rdhA* gene regulation by CprK, a CRP-FNR-family regulator from *Desulfitobacterium* spp. This family of regulators is characterized by a length of 230-250 amino acids, an *N*-terminally located nucleotide (effector)-binding domain and a *C*-terminal HTH DNA-binding motif. The family responds to a broad spectrum of signals leading to activation of gene transcription.

To accomplish the regulatory function the *N*-terminal domain binds allosteric effector molecules, after which a signal is transmitted to the DNA-binding domain [426]. CprK is encoded in the *cprTKZE$_1$BACE$_2$* gene cluster (Figure 2, Table 2) in *Desulfitobacterium dehalogenans* JW/IU-DC1^T and *Desulfitobacterium hafniense* DCB-2^T including the gene encoding the *ortho*-chlorophenol reductive dehalogenase CprA [26, 64, 73]. Transcription of this gene cluster was initiated from three promoters in the presence of 3-chloro-4-hydroxyphenylacetic acid, whereas *cprK* was constitutively expressed at a low level. Electrophoretic mobility shift assays, DNase footprinting studies and promoter-*lacZ* fusion experiments using heterologously expressed CprK from the two *Desulfitobacterium* spp. strains mentioned above indicated that CprK acts as a transcriptional activator of the *cpr* gene cluster in the presence of 3-chloro-4-hydroxyphenylacetic acid [67, 71]. 3-Chloro-4-hydroxyphenylacetic acid binds with high affinity to CprK, which promotes its specific interaction with a dehalo-box motif upstream of the -10 and -35 promoter regions of several *cpr* genes. CprK is inactivated by oxygen *in vitro* [67] due to the formation of an intermolecular disulphide bridge (Cys11 and Cys200); however, a physiological role in redox-sensing seems unlikely [71]. The availability of the genome sequence of *Desulfitobacterium hafniense* DCB-2^T allowed a comparative analysis of a number of CprK proteins [66]. Genes encoding five CprK-type proteins encoded in the genome are clustered with *cpr* genes. The c*prK1* gene is part of the cluster encoding the biochemically characterized *ortho*-chlorophenol RDase CprA1 of *Desulfitobacterium hafniense* [398]. Eleven dehalo-box motifs were identified within the five *cpr* gene clusters. They consisted of a 5 bp imperfect inverted repeat with 4 nt spacing and exhibited TTAGT-N$_4$-ACTAA as consensus sequence. Interaction studies of three heterologously produced CprK proteins with the corresponding dehalo-box motifs in the presence and absence of different chlorinated compounds revealed partly overlapping functions. Two CprK proteins were specifically activated by 3-chloro-4-hydroxyphenylacetic acid and other *ortho*-halogenated phenols, and another one by *meta*-chlorinated phenols. Thus, a range of chlorophenols can be sensed by *Desulfitobacterium hafniense* DCB-2^T in a cooperative manner. [66].

The crystal structures of oxidized *Desulfitobacterium hafniense* CprK1 with bound 3-chloro-4-hydroxyphenylacetic acid and of reduced, ligand-free CprK from *Desulfitobacterium dehalogenans* JW/IU-DC1^T allowed the first mechanistic insight and identified allosteric conformational changes induced by ligand binding [155]. Both the phenolic OH-group and the *ortho*-chlorine were required for tight binding, in which a conserved lysine residue (Lys133) in the central α-helix interacts with the OH group.

Reduced and oxidized protein dimerize without forming a disulphide bond. Based on results of DNA-binding assays and macromolecular native mass spectrometry, in combination with limited proteolysis, a dynamic model for the activation process was proposed [157]. This was refined by a complete structural description of the redox-dependent and allosteric molecular arrangements of CprK1 [156]. In the absence of the ligand, the dimeric CprK is flexible. Binding of the ligand leads to a concerted reorganization of both monomers that moves the *C*-terminal DNA-binding domains into positions compatible with binding to the palindromic dehalo-box motif sequence [156]. This motion is induced by interactions of the phenolate with Lys133 as well as docking of the chloride atom into a hydrophobic pocket provided by the central α-helix connecting the *N*- and *C*-terminal domains. As a consequence, three amino acids of the *N*-terminal effector domain interact with the DNA-binding domain and thus stabilize the DNA-binding conformation. Nucleotide-specific contacts are formed between the HTH motif recognition helix and the DNA that lead to the observed specificity for the AT-rich dehalo-box motif. The physical presence of the bulky halogen atom as well as its inherent electronegativity lowering the pK_a of the phenol are necessary for the formation of the Lys133-phenolate salt bridge, resulting in high affinity binding [427]. Therefore, neither the end product of dechlorination, 4-hydroxyphenylacetic acid, nor 3-chlorophenylacetic acid lacking either the chlorine substituent or the hydroxyl group are recognized by CprK. Lys133 is strictly conserved within the CprK family suggesting that its function might be limited to sensing phenolic compounds in the environment. Indeed, *pceA* genes in *Desulfitobacterium* are often part of the canonical transposon-encoded *pceABCT* gene cluster lacking a *cprK* ortholog and are constitutively expressed [78, 117, 131], perhaps reflecting an evolutionary adaptation to high anthropogenic concentrations of PCE and TCE in the environment.

These first studies of regulators in OHRB revealed the existence of completely different mechanisms of transcriptional regulation in two phylogenetically distant bacteria. It can be expected that the investigation of other types of regulators such as two-component-system members will further expand our understanding of how bacteria sense and respond to halogenated compounds as electron acceptors for organohalide respiration.

Acknowledgement

This work was funded by the German Research Foundation (DFG research unit FOR 1530). Work of HS and TK was financially supported by the EcoLinc Project of the Netherlands Genomics Initiative, the BE-BASIC foundation through projects F07.001.05 and F08.004.01, as well as the European Community program FP7 (grants KBBE-211684; BACSIN, and KBBE-222625; METAEXPLORE). We are grateful to Gary Sawers for comments on the manuscript.

Chapter 9

General discussion

Introduction

It took exactly thirty years from the isolation in pure culture of the first representative of organohalide-respiring bacteria (ORHB), *Desulfomonile tiedjei* DCB1^T, to the first heterologous production of functional reductive dehalogenases and resolving of its crystal structure [10, 69, 81]. In between these two hall marks numerous other exciting discoveries have been done, and many more will undoubtedly follow in the coming years (Figure 1).

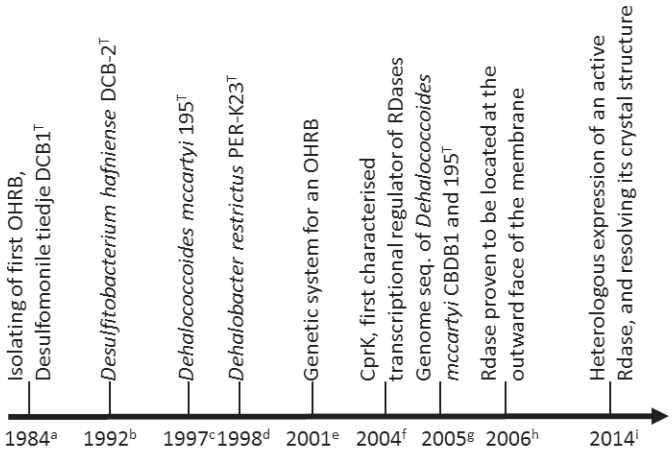


Figure 1. Time line showing an non exhaustive selection of hall marks in the study of OHRB a:[10]; b: [43]; c: [15]; d: [12]; e: [329]; f: [67]; g: [17, 18]; h: [77]; i:[69] [81]

Genome sequencing of OHRB

Ten years ago the publication of the genomes of *Dehalococcoides mccartyi* 195^T (Formerly *Dehalococcoides ethenogenes* 195^T) and *Dehalococccoides mccartyi* CBDB1 marked the entrance to a new era in the study of OHRB [17, 18]. In the following ten years, many more genomes of OHRB have been obtained as discussed in **chapter 8**. Much focus has been on the members of the species *Dehalococcoides mccartyi,* due to their capacity for degrading a very large number of organohalides, including the complete dehalogenation of PCE to ethene [11]. All *Dehalococcoides mccartyi* strains analyzed so far have small genomes, with an average size of 1.4 Mbp, and

these are highly conserved except in two regions located close to but at either side of the origin of replication, the so called high plasticity regions. Most of the *rdh* genes are located in these latter regions, leading to the speculation that the high plasticity regions serve as evolutionary hotspots with frequent horizontal gene transfers, allowing rapid evolution of *Dehalococcoides mccartyi* strains while at the same time preserving core metabolic functions [31]. This type of comparative genomic studies have until the present thesis work not been possible for OHRB from other genera, due to the unavailability of full genome sequences.

Insights from genomic analysis

When the genome of *Desulfitobacterium hafniense* Y51 was published in 2006, it revealed the presence of genes encoding an unprecedented number of molybdopterin oxidoreductases (pfam01568) and only a single reductive dehalogenase, *pceA*, located on a composite transposon [19]. This provided a picture of a bacterium with an extremely versatile metabolism, but not particularly specialised in organohalide respiration (OHR).

The availability of twelve genomes from the facultative organohalide respiring *Desulfitobacterium* genus (Chapters 6-7)[19, 26, 27], and three closed genomes from the obligate OHR *Dehalobacter* spp. [23], representing closely related genera with contrasting lifestyles, has changed this view.

The reductive dehalogenases found in desulfitobacteria do not seem to have been acquired by recent horizontal gene transfer as discussed in Chapter 7. On the contrary, it seems likely that gene loss has played an important role in the evolution of desulfitobacteria. The presence of dehaloclusters D and E in a conserved region of the genome in *D. dehalogenans* and *D. hafniense* PCP-1 and DCB-2^{T} , respectively, strongly suggests that the last common ancestor of *D. hafniense* and *D. dehalogenans* possessed both dehaloclusters, which then, subsequently, have been lost from some but not all members of *D. hafniense*.

Whether the dehaloclusters only found in one species were also present in the last common ancestor can not be deduced with certainty from the currently available data. Nevertheless, the lack of indications of horizontal gene transfer and the presence of *rdhK* transcriptional regulator encoding genes, typical for *Firmicutes*, points towards the fact that these genes could have been present in the last common ancestor, or have been horizontally acquired from other *Firmicutes*. This taken together with the tight regulation of expression, the presence of genes linked to chemotaxis, and the presence of genes for sequential degradation of highly chlorinated chlorophe-

nols, points towards the possibility that desulfitobacteria evolved as bacteria with both a fine tuned machinery for OHR and a versatile metabolism.

In contrast to this, the genome of *Dehalobacter restrictus* PER-K23^{T} was found to encode only three genes annotated as coding for molybdopterin oxidoreductases (pfam01568), but an unexpectedly large number of reductive dehalogenases, suggesting a much larger potential for OHR than previously anticipated. Interestingly, 21 out of a total of 25 *rdhA* genes were located in two genomic regions that we named *rdh* cluster A and B (**chapter 2**). A similar organisation of the *rdhA* genes was observed in the genomes of the only two other *Dehalobacter* strains for which closed genome sequence were published, namely *Dehalobacter* sp. 11DCA and *Dehalobacter* sp. CF (Figure 2) [23]. All four *rdhA* genes located outside *rdh* cluster A and B are conserved between *Dehalobacter restrictus* PER-K23^{T}, the draft genome of *Dehalobacter* sp. E1 and the closed genomes of *Dehalobacter* sp. 11DCA and *Dehalobacter* sp. CF (**Chapter 2** and data not shown), whereas there are pronounced differences between *rdh* clusters A and B (Figure 2).

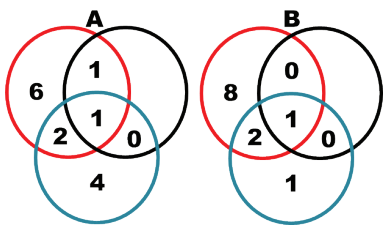


Figure 2 Venn diagram showing numbers of shared *rdhA* genes between ■: *Dehalobacter restrictus* PER-K23^{T}; ■: *Dehalobacter* sp. E1 [22] and ■: The almost identical genomes of *Dehalobacter* sp. 11DCA and *Dehalobacter* sp. CF [23]. A and B refers to *rdh* cluster A and B as introduced in **chapter 2.** Since the genome of *Dehalobacter* sp. E1. is not closed we did not attempt to estimate the total number of *rdhA* genes in *rdh* cluster A and B in this strain.

Dehalobacter restrictus PER-K23^{T} and *Dehalobacter* sp. E1 were obtained from European sediment samples [53, 83] whereas *Dehalobacter* sp. 11DCA and *Dehalobacter* sp. CF were obtained from samples collected in North America [428]. The presence of *rdh* cluster A and B in all four investigated *Dehalobacter* genomes suggests that this is a common feature of *Dehalobacter* spp. and not the result of recent horizontal gene transfer.

Drawing a direct parallel to the high plasticity regions in the likewise obligate OHRB *Dehalococcoides mccartyi* seems straightforward. However, more detailed comparative genomic analyses are needed before any conclusions on the cause of the heterogeneity between *rdh* cluster A and B in the different *Dehalobacter* spp. strains can be drawn.

Possible explanations could be that they are targeted by frequent horizontal gene transfers, as speculated for the high plasticity regions of *Dehalococoides mccartyi* [31, 380]. The multitude of

rdhA genes could also be the result of gene duplications or simply the remnants of even larger *rdh* A and B clusters.

Cobalamins are essential co-factors in reductive dehalogenases [3]. Therefore, OHRB capable of *de novo* cobalamin synthesis are well suited for bioremediation as they do not depend on external cobalamin sources.

In **chapter** 7 we reported the presence of an intact cobalamin synthesis pathway in all twelve currently available desulfitobacterial genomes. In contrast, the cobalamin synthesis pathway in *Dehalobacter restrictus* PER-K23^{T} is non-functional due to 101 bp deletion in the *cbiH* gene (**Chapter 4)**. Similarly, several genes from the cobalamin synthesis pathway were found missing from the draft genome of *Dehalobacter* sp E1 [22]. An intact cobalamin synthesis pathway was, however, encoded by the genomes of *Dehalobacter* sp. 11DCA and *Dehalobacter* sp. CF [23].

Dehalobacter restrictus PER-K23^{T} has routinely been maintained in medium supplemented with cobalamin [12], and hence it could be speculated that the deletion in *cbiH* occurred during or after its isolation. If *Dehalobacter* spp., like *Desulfitobacterium* spp., in general are capable of synthesising cobalamins *de novo* combined with the presence of a very large number of *rdhA* genes, that would mean that they represent a largely unexplored potential for bioremediation of organohalides.

Respiration

The details of the electron transport chain from electron donor to the reductive dehalogenase are to a large extent still not known. The study of electron transport chains is notoriously difficult, due to the fact that many electron shuttles can transfer electrons between several different electron donors or acceptors, and often more than one electron shuttle can perform the same task, and therefore they can substitute for each other, as reviewed by others [254, 339, 429, 430]. In **chapter 6** we demonstrate that the first part of the electron transport chain to the chlorophenol reductase CprA in *Desulfitobacterium dehalogenans* JW/IU-DC1^{T} involves menaquinones (MKs). The involvement of MKs in OHR has previously been observed for other OHRB, such as *Dehalobacter restrictus* PER-K23^{T}, *Sulfurospirillum multivorans* and *Desulfomonile tiedjei* DCB-1 [12, 29, 74], but not for *Dehalococcoides mccartyi* [195]. How the electrons are transferred from the membrane-embedded MKs to the outward facing RdhA has not yet been elucidated. This could happen either via A) a membrane protein or protein complex transferring electrons directly from

MKs to the RdhA, or B) indirectly via a soluble electron shuttle receiving electrons from a membrane bound menaquinone dehydrogenase (Figure 3).

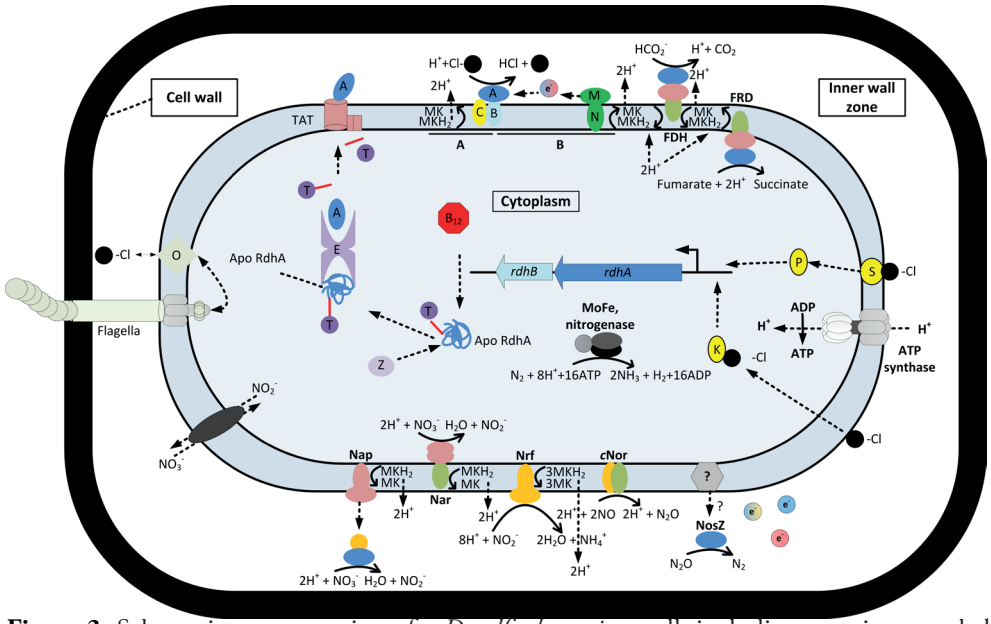


Figure 3. Schematic representation of a *Desulfitobacterium* cell, including proteins encoded by genes found in dehaloclusters. Letters inside symbols indicate the protein name without the Rdh prefix. See **chapter 8** for detailed review on function. Electron transport chain A or B are indicated with a horizontal line, (see text for details). FDH: formate dehydrogenase; FRD: fumarate reductase. Colour codes for elements of FDH; FRD; Nap; Nar; Nrf; cNor; NosZ: ■ cytochrome b; ■ iron sulfur cluster protein; ■ catalytic subunit; ■ cytochrome c. Fumarate reductase paralogs are depicted as filled circles with e^- colour of circles indicating the different domains present ■ FAD binding; ■ FMN binding; ■ cytochrome c. See **chapter 7** for detailed review.

In **chapter 6** we identified an extracytoplasmic diflavin as being more abundant in *Desulfitobacterium dehalogenans* JW/IU-DC1^T when the strain utilized Cl-OHPA instead of fumarate as terminal electron acceptor. The fumarate reductase complex receives electrons directly from the MKs, and is therefore not dependent on other electron carriers. The increase in abundance of a putative soluble electron carrier under OHR conditions does therefore lend support to electron transport model B). Further support for model B) may be obtained from the findings that a gene encoding a cytochrome c with a predicted signal sequence and no membrane helices was shown to be induced in *Desulfomonile tiedjei* DCB-1 under OHR conditions [431]. And in the genome of *Sulfurospirillum multivorans* two genes, *rdhMN,* encoding a NapGH like membrane-bound menaquinone dehydrogenase were found to be located in close proximity of two *rdhA* genes, whereof one encodes the PCE reductase PceA. Expression of *rdhMN* was induced in the

presence of PCE [29]. Interestingly, a similar *rdhMN* gene set was found to be located directly downstream of an *rdhAB* gene pair in the genome of *Desulfomonile tiedjei* DCB-1 [29] and in three *Desulfitobacterium* strains. **(Chapter 7)**. Determining if the electron transport chain follows model A or B or an yet unknown model C requires further studies. It is also plausible that there exist deviations in the electron transport chain between different reductive dehalogenases. Moreover, the important question still stands on how a proton gradient necessary for energy conservation is generated. This could both be produced by simple proton pumping by the MKs, as speculatively shown in Figure 3, or involve an yet unidentified proton pump.

The aforementioned extracytoplasmic diflavin belongs to cog-1053 annotated as fumarate reductases or succinate dehydrogenases, and referred to as fumarate reductase paralogs (FRP) [19, 26]. FRP encoding genes are present in large numbers (varying from 8 to 31) in *D. hafniense*, *D. dehalogenans* and *D. dichloroeliminans* but not in *D. metallireducens*. In **chapter 7** we analysed these FRP encoding genes and found that they all encode signal sequences but no transmembrane helices, and at least one FAD binding domain. Many of the FRP-encoding genes are predicted to be co-transcribed with genes that code for cytochrome c or FMN binding proteins. This suggest that they could play a role in electron transport at the outside of the plasma membrane (Figure 3). Similar enzymes have been studied intensively in members of the Gram-negative genera *Wolinenella* and *Shewanella*, were they have been shown to be active as catalytic enzymes and function as electron shuttles in the periplasmic space [330, 334, 335].

This observation is particular noteworthy as both *Desulfitobacterium* and *Dehalobacter* are Gram-positive bacteria [12, 44]. We did, however, find the coding capacity for many typical Gram-negative enzymes in the desulfitobacterial genomes, such as the gene clusters for the Nap and Nos type nitrate and nitrous oxide reductases, which both are located in the periplasm as discussed in **chapter 7** (Figure 3). The canonical view is that Gram-negative bacteria have two cell membranes with a periplasmic space in between that contains the peptidoglycan, whereas Gram-positive bacteria have a single membrane surrounded by a thick peptidoglycan wall. It has, however, been shown that some Gram-positive bacteria belonging to the *Firmicutes* contain a layer between the peptidoglycan wall and the cell membrane, termed the inner wall zone or the Gram-positive periplasm [432-434]. Seen in the light of the data presented in this thesis it would be highly interesting to investigate if members of the *Desulfitobacterium* and *Dehalobacter* genera possess a Gram positive periplasm and its role in their energy metabolism.

Setting sails for new horizons

In recent years it has become increasingly clear that the type of the lower ligand of cobalamins plays and essential role for function of reductive dehalogenases. In both *Sulfurospirillum multivorans* and *Dehalococcoides mccartyi*, experiments with cobalamin containing different types of lower ligands have shown that some lower ligands led to the production of functional RDases whereas others did not support growth by OHR, as discussed in **Chapter 8** [61, 62, 183].

The presence of a full cobalamin synthesis pathway in the genomes of *Desulfitobacterium* spp. and some *Dehalobacter* spp. makes them independent of exogenous cobalamins supplies. Although it still reamins to be tested if *Dehalobacter* spp. synthesize cobalamins *de novo*. Nevertheless, growth was improved when cobalamin was added to the medium [55, 56], probably reflecting the energy cost of *de novo* synthesis. It should furthermore be noted that we did not yet identify the genes coding for synthesis of the lower ligand in *Desulfitobacterium* spp. *Dehalobacter* spp. Such analyses have now for the first time become feasible, as this pathway was elucidated only very recently [435].

The finding that "one size fits all" does not apply for the lower ligand of cobalamin should be taken into account when culturing or isolating OHRB for which the capacity of *de novo* cobalamin synthesis is not known. In our and many other laboratories working with isolation and cultivation of OHRB, cobalamins are routinely given in the form of cyanocobalamin with 5,6-dimethylbenzimidazole (DMB) as the lower ligand. In cyanocobalamin the upper ligand has been substituted with cyanide as this increases shelf life of the cobalamin. After uptake in the cell the cyanide is replaced with another upper ligand leading to the formation of a biologically active molecule.

Improved results in isolating new strains, or improved growth of already isolated strains, could probably be achieved by adding cobalamin with different types of lower ligands or cyanocobalamin together with a cocktail of different lower ligands, thereby providing the bricks needed for the bacteria to remodel the cobalamins to meet their specific requirements.

A second focus point should be obtaining an improved understanding of the function of the genes found associated with *rdhA* genes, which in **Chapter 8** are referred to as accessory genes. Main focus in the past has been on identifying RdhA encoding genes. Since very little was known

about the role of the accessory genes these have often just been given an *rdh* prefix and a letter from the alphabet without any attempts on further characterization. This has unfortunately led to some confusion in the terminology of these accessory genes. In **chapter 8** we proposed a unifying nomenclature for the accessory genes and summarised the current knowledge on their function. Understanding the role of these genes will help us making predictions related to the function and regulation of the *rdhA* that they accompany. As an example we can assume that the vast majority of the currently identified RdhAs are attached to the membrane because they are co transcribed with a gene, *rdhB,* encoding a membrane protein anchoring RdhA to the membrane [3, 64]. Some accessory genes are only found with a subset of *rdhA*s, and elucidating their function would enable us to make predictions specifically related to these *rdhAs*. As an example *rdhF* was only found associated with some but not all *rdhAs* in *Dehalococcoides mccartyi*. This gene encodes a CbiZ like protein predicted to be responsible for removing the lower ligand of cobalamins [61, 149, 184]. The Cbiz encoding gene was found together with *btuFCD* genes encoding cobalamin transporters on the chromosome of *Dehalococcoides mccartyi*, and it has been shown that remodelling of corrinoids is intimately linked to salvaging [61].

If remodelling of cobalamins would be linked to the salvaging process, what is then the role of the CbiZ homologues encoded by *rdhF*? If one speculates freely, a possible explanation for the presence of *rdhF* in some but not all *rdhA* gene clusters could be that not all RdhAs from the same organism require cobalamins with the same lower ligand and that *rdhF* removes unsuitable lower ligands.

Another example is *rdhC* encoding a membrane-bound FMN-binding, iron sulfur protein, originally suggested to be a transcriptional regulator [73]. However, this view has been abandoned, and more recently a possible role in electron transport has been suggested [387, 388]. One could also speculate on possible other roles such as an activating enzyme (AE) similar to what is known from *O* demethylases, where the AE reactivates the corrinoid enzymes by reducing Co^{2+} to Co^{+} [436].

Knowing the role of these and other *rdhA* associated gene products would greatly improve our ability to predict the function of *rdhA* encoding genes from genome analysis, and identify traits associated with specific groups of *rdhA*s.

A third focus point should be to create a model organism for studying OHR and processes related to this, such as elucidating different components of the electron transport chain. The three recent papers describing successful heterologous production of active reductive dehalogenases represent a true hall mark. Allowing expression, functional characterisation and obtaining the crystal structure of reductive dehalogenases has already led to a greatly improved understanding of the mechanism of OHR [69, 79-81]. Nevertheless, all three studies were based on expressing extracting, purifying and studying the enzyme outside the cells. It would be a valuable addition to have a model organism where *rdhA*s could be expressed together with their accessory genes so their interplay could be studied together in a cell.

The criteria for an ideal model organism would be that 1) it should not already encode any reductive dehalogenases or their accessory genes 2) it should be fast growing and easy to maintain 3) the full genome sequence information should be available 4) genetic tools should be available 5) it should be able to express functional reductive dehalogenases.

Desulfitobacterium hafniense DP7 might be a strong candidate for such a model organism, as it fulfils criteria 1) to 3). Protocols for transforming *Desulfitobacterium dehalogenans* JW/IU-DC1^{T} both by conjugation and electroporation have previously been published [119, 329]. These protocols could probably be adapted to other desulfitobacteria. *Desulfitobacterium hafniense* DP7 encodes a full cobalamin synthesis pathway and is very closely related to OHRB, and thus it seems likely that the strain would be able to express functional reductive dehalogenases. Finally the vancomycin resistance, encoded by the *vanI* gene cluster offers a convenient selection marker (**Chapter 5**).

Finally our analysis of the desulfitobacterial genomes as presented in **Chapter 7** confirmed the previous observations that this is an genus with an extremely versatile metabolism [16, 19, 26]. It would be very interesting to obtain a better understanding on the contribution of desulfitobacteria to the global cycling of for example nitrogen, sulfur and their potential for and mechanism of dissimilatory metal reduction, as this is currently poorly understood for Gram positive bacteria [312, 437, 438].

From the data presented in the present thesis, it is obvious that great progress has been made in the study of OHRB during recent years, and equally obvious that we just have begun our journey towards understanding the world of OHR, including the exciting metabolic repertoire of the *Desulfitobacterium* genus.

References

1. Van Pée, K.H., *Enzymatic chlorination and bromination*, in *Methods in Enzymology*. 2012. p. 237-257.
2. Gribble, G.W., *Naturally Occurring Organohalogen Compounds.* 1998. **31**(3): p. 141-152.
3. Smidt, H. and W.M. de Vos, *Anaerobic microbial dehalogenation.* Annual Review Of Microbiology, 2004. **58**: p. 43-73.
4. Fetzner, S., *Bacterial dehalogenation.* Applied Microbiology and Biotechnology, 1998. **50**(6): p. 633-657.
5. Leys, D., L. Adrian, and H. Smidt, *Organohalide respiration: microbes breathing chlorinated molecules.* Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 2013. **368**(1616): p. 20120316.
6. Holliger, C., G. Wohlfarth, and G. Diekert, *Reductive dechlorination in the energy metabolism of anaerobic bacteria.* Fems Microbiology Reviews, 1999. **22**(5): p. 383-398.
7. Meckenstock, R.U., et al., *Biodegradation: Updating the Concepts of Control for Microbial Cleanup in Contaminated Aquifers.* Environmental Science & Technology, 2015. **49**(12): p. 7073-7081.
8. Maphosa, F., W.M. de Vos, and H. Smidt, *Exploiting the ecogenomics toolbox for environmental diagnostics of organohalide-respiring bacteria.* Trends in Biotechnology, 2010. **28**(6): p. 308-316.
9. DeWeerd, K.A., et al., *Desulfomonile tiedjei gen. nov. and sp. nov., a novel anaerobic, dehalogenating, sulfate-reducing bacterium.* Archives of Microbiology, 1990. **154**(1): p. 23-30.
10. Shelton, D.R. and J.M. Tiedje, *Isolation and partial characterization of bacteria in an anaerobic consortium that mineralizes 3-chlorobenzoic acid.* Applied and Environmental Microbiology, 1984. **48**(4): p. 840-848.
11. Löffler, F.E., et al., *Dehalococcoides mccartyi gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, Dehalococcoidia classis nov., order Dehalococcoidales ord. nov. and family Dehalococcoidaceae fam. nov., within the phylum Chloroflexi.* International Journal of Systematic and Evolutionary Microbiology, 2013. **63**(PART2): p. 625-635.
12. Holliger, C., et al., *Dehalobacter restrictus gen. nov. and sp. nov., a strictly anaerobic bacterium that reductively dechlorinates tetra- and trichloroethene in an anaerobic respiration.* Archives Of Microbiology, 1998. **169**(4): p. 313-321.
13. Yan, J., et al., *Isolation of novel bacteria within the Chloroflexi capable of reductive dechlorination of 1,2,3-trichloropropane.* Environmental Microbiology, 2009. **11**(4): p. 833-843.
14. Taş, N., et al., *The little bacteria that can - Diversity, genomics and ecophysiology of 'Dehalococcoides' spp. in contaminated environments.* Microbial Biotechnology, 2010. **3**(4): p. 389-402.
15. Maymó-Gatell, X., et al., *Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene.* Science, 1997. **276**(5318): p. 1568-1571.

16. Villemur, R., et al., *The Desulfitobacterium genus.* Fems Microbiology Reviews, 2006. **30**(5): p. 706-733.

17. Kube, M., et al., *Genome sequence of the chlorinated compound respiring bacterium Dehalococcoides species strain CBDB1.* Nature Biotechnology, 2005. **23**(10): p. 1269-1273.

18. Seshadri, R., et al., *Genome sequence of the PCE-dechlorinating bacterium Dehalococcoides ethenogenes.* Science, 2005. **307**(5706): p. 105-108.

19. Nonaka, H., et al., *Complete genome sequence of the dehalorespiring bacterium Desulfitobacterium hafniense Y51 and comparison with Dehalococcoides ethenogenes 195.* Journal Of Bacteriology, 2006. **188**(6): p. 2262-2274.

20. Pagani, I., et al., *The Genomes OnLine Database (GOLD) v.4: Status of genomic and metagenomic projects and their associated metadata.* Nucleic acids research, 2012. **40**(D1): p. D571-D579.

21. Siddaramappa, S., et al., *Complete genome sequence of Dehalogenimonas lykanthroporepellens type strain (BL-DC-9 T) and comparison to "Dehalococcoides" strains.* Standards in Genomic Sciences, 2012. **6**(2): p. 251-264.

22. Maphosa, F., et al., *Metagenome analysis reveals yet unexplored reductive dechlorinating potential of Dehalobacter sp. E1 growing in co-culture with Sedimentibacter sp.* Environmental Microbiology Reports, 2012. **4**(6): p. 604-616.

23. Tang, S., Y. Gong, and E.A. Edwards, *Semi-Automatic In Silico Gap Closure Enabled De Novo Assembly of Two Dehalobacter Genomes from Metagenomic Data.* PLoS ONE, 2012. **7**(12).

24. Kruse, T., et al., *Complete genome sequence of Dehalobacter restrictus PER-K23.* Standards in Genomic Sciences, 2013.

25. Kruse, T., et al., *Genomic, proteomic and biochemical analysis of the organohalide respiratory pathway in Desulfitobacterium dehalogenans* Journal of Bacteriology, 2015. **195**(5): p. 12.

26. Kim, S.-H., et al., *Genome sequence of Desulfitobacterium hafniense DCB-2, a Gram-positive anaerobe capable of dehalogenation and metal reduction.* Bmc Microbiology, 2012. **12**(1): p. 21.

27. Goris, T., et al., *Draft genome sequence and characterization of Desulfitobacterium hafniense PCE-S.* Standards in Genomic Sciences, 2015. **10**(1): p. 15.

28. Wagner, D.D., et al., *Genomic determinants of organohalide-respiration in Geobacter lovleyi, an unusual member of the Geobacteraceae.* 2012. **13**(1).

29. Goris, T., et al., *Insights into organohalide respiration and the versatile catabolism of Sulfurospirillum multivorans gained from comparative genomics and physiological studies.* Environmental Microbiology, 2014: p. n/a-n/a.

30. Thomas, S.H., et al., *The mosaic genome of Anaeromyxobacter dehalogenans strain 2CP-C suggests an aerobic common ancestor to the delta-proteobacteria.* PLoS ONE, 2008. **3**(5).

31. McMurdie, P.J., et al., *Localized Plasticity in the Streamlined Genomes of Vinyl Chloride Respiring Dehalococcoides.* Plos Genetics, 2009. **5**(11).

32. Johnson, D.R., et al., *Transcriptomic microarray analysis of corrinoid responsive genes in dehalococcoides ethenogenes strain 195.* FEMS Microbiology Letters, 2009. **294**(2): p. 198-206.

33. Zhang, Y., et al., *Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization.* Bmc Genomics, 2009. **10**(1): p. 78.

34. Markowitz, V.M., et al., *IMG 4 version of the integrated microbial genomes comparative analysis system.* Nucleic acids research, 2013.

35. Tamura, K., et al., *MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods.* Molecular Biology and Evolution, 2011. **28**(10): p. 2731-2739.

36. Edgar, R.C., *MUSCLE: Multiple sequence alignment with high accuracy and high throughput.* Nucleic Acids Research, 2004. **32**(5): p. 1792-1797.

37. Saitou, N. and M. Nei, *The neighbor-joining method: a new method for reconstructing phylogenetic trees.* Molecular Biology and Evolution, 1987. **4**(4): p. 406-425.

38. Wang, S., et al., *Isolation and characterization of a novel Dehalobacter species strain TCP1 that reductively dechlorinates 2,4,6-trichlorophenol.* Biodegradation, 2014. **25**(2): p. 313-323.

39. Sun, B., et al., *Microbial Dehalorespiration with 1,1,1-Trichloroethane.* Science, 2002. **298**(5595): p. 1023-1025.

40. Wild, A., R. Hermann, and T. Leisinger, *Isolation of an anaerobic bacterium which reductively dechlorinates tetrachloroethene and trichloroethene.* Biodegradation, 1996. **7**(6): p. 507-511.

41. Lee, M., et al., *Complete chloroform dechlorination by organochlorine respiration and fermentation.* Environmental Microbiology, 2012. **14**(4): p. 883-894.

42. Justicia-Leon, S.D., et al., *Dichloromethane fermentation by a Dehalobacter sp. in an enrichment culture derived from pristine river sediment.* Applied and Environmental Microbiology, 2012. **78**(4): p. 1288-1291.

43. Madsen, T. and D. Licht, *Isolation and characterization of an anaerobic chlorophenol-transforming bacterium.* Applied and Environmental Microbiology, 1992. **58**(9): p. 2874-2878.

44. Utkin, I., C. Woese, and J. Wiegel, *Isolation and characterization of Desulfitobacterium dehalogenans Gen-Nov, Sp-Nov, an anaerobic bacterium which reductively dechlorinates chlorophenolic compounds.* International Journal Of Systematic Bacteriology, 1994. **44**(4): p. 612-619.

45. Christiansen, N. and B.K. Ahring, *Desulfitobacterium hafniense sp. nov., an anaerobic, reductively dechlorinating bacterium.* International Journal of Systematic Bacteriology, 1996. **46**(2): p. 442-448.

46. Niggemyer, A., et al., *Isolation and characterization of a novel As(V)-reducing bacterium: implications for arsenic mobilization and the genus Desulfitobacterium.* Appl Environ Microbiol, 2001. **67**(12): p. 5568-80.

47. Parte, A.C., *LPSN - List of prokaryotic names with standing in nomenclature.* Nucleic Acids Research, 2014. **42**(D1): p. D613-D616.

48. De Wildeman, S., et al., *Growth-substrate dependent dechlorination of 1,2-dichloroethane by a homoacetogenic bacterium.* Biodegradation, 2003. **14**(4): p. 241-247.

49. van de Pas, B.A., et al., *A Desulfitobacterium strain isolated from human feces that does not dechlorinate chloroethenes or chlorophenols.* Archives Of Microbiology, 2001. **175**(6): p. 389-394.

50. Kunapuli, U., et al., *Desulfitobacterium aromaticivorans sp. nov. and Geobacter toluenoxydans sp. nov., iron-reducing bacteria capable of anaerobic degradation of monoaromatic hydrocarbons.* International Journal of Systematic and Evolutionary Microbiology, 2010. **60**(3): p. 686-695.

51. Wiegert, C., et al., *Carbon and chlorine isotope fractionation during microbial degradation of tetra-and trichloroethene.* Environmental Science and Technology, 2013. **47**(12): p. 6449-6456.

52. Deshpande, N.P., et al., *Genome sequence of Dehalobacter UNSWDHB, a chloroform-dechlorinating bacterium.* 2013. **1**: p. e00720-e00713.

53. Van Doesburg, W., et al., *Reductive dechlorination of β-hexachlorocyclohexane (β-HCH) by a Dehalobacter species in coculture with a Sedimentibacter sp.* Fems Microbiology Ecology, 2005. **54**(1): p. 87-95.

54. Maphosa, F., *Chasing Organohalide Respirers: Ecogenomics Approaches to Assess the Bioremediation Capacity of Soils*, in *Laboratory of Microbiology*. 2010, Wageningen University: Wageningen. p. 178.

55. Choudhary, P.K., et al., *Diversity of Cobalamin Riboswitches in the Corrinoid-Producing Organohalide Respirer Desulfitobacterium hafniense.* Journal of Bacteriology, 2013. **195**(22): p. 5186-5195.

56. Reinhold, A., et al., *Impact of Vitamin B 12 on formation of the tetrachloroethene reductive dehalogenase in Desulfitobacterium hafniense strain Y51.* Applied and Environmental Microbiology, 2012. **78**(22): p. 8025-8032.

57. Hug, L.A., et al., *Overview of organohalide-respiring bacteria and a proposal for a classification system for reductive dehalogenases.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1616).

58. Magnuson, J.K., et al., *Trichloroethene reductive dehalogenase from Dehalococcoides ethenogenes: Sequence of tceA and substrate range characterization.* 2000. **66**(12): p. 5141-5147.

59. Palmer, T. and B.C. Berks, *The twin-arginine translocation (Tat) protein export pathway.* Nature Reviews Microbiology, 2012. **10**(7): p. 483-496.

60. Krautler, B., et al., *The cofactor of tetrachloroethene reductive dehalogenase of Dehalospirillum multivorans is norpseudo-B-12, a new type of a natural corrinoid.* Helvetica Chimica Acta, 2003. **86**(11): p. 3698-3716.

61. Yi, S., et al., *Versatility in corrinoid salvaging and remodeling pathways supports corrinoid-dependent metabolism in Dehalococcoides mccartyi.* Appl Environ Microbiol, 2012. **78**(21): p. 7745-52.

62. Keller, S., et al., *Exogenous 5,6-dimethylbenzimidazole caused production of a non-functional tetrachloroethene reductive dehalogenase in Sulfurospirillum multivorans.* Environmental Microbiology, 2013.

63. Yan, J., et al., *Unexpected specificity of interspecies cobamide transfer from geobacter spp. to organohalide-respiring dehalococcoides mccartyi strains.* Applied and Environmental Microbiology, 2012. **78**(18): p. 6630-6636.

64. Van de Pas, B.A., et al., *Purification and molecular characterization of ortho-chlorophenol reductive dehalogenase, a key enzyme of halorespiration in Desulfitobacterium dehalogenans.* Journal Of Biological Chemistry, 1999. **274**(29): p. 20287-20292.

65. Neumann, A., G. Wohlfarth, and G. Diekert, *Tetrachloroethene dehalogenase from Dehalospirillum multivorans: Cloning, sequencing of the encoding genes, and expression of the pceA gene in Escherichia coli.* Journal Of Bacteriology, 1998. **180**(16): p. 4140-4145.

66. Gábor, K., et al., *Divergent roles of CprK paralogues from Desulfitobacterium hafniense in activating gene expression.* Microbiology, 2008. **154**(12): p. 3686-3695.

67. Pop, S.M., R.J. Kolarik, and S.W. Ragsdale, *Regulation of Anaerobic Dehalorespiration by the Transcriptional Activator CprK.* Journal Of Biological Chemistry, 2004. **279**(48): p. 49910-49918.

68. Maillard, J., P. Genevaux, and C. Holliger, *Redundancy and specificity of multiple trigger factor chaperones in Desulfitobacteria.* Microbiology, 2011. **157**(8): p. 2410-2421.

69. Mac Nelly, A., et al., *Functional Heterologous Production of Reductive Dehalogenases from Desulfitobacterium hafniense Strains.* Applied and Environmental Microbiology, 2014. **80**(14): p. 4313-4322.

70. Morita, Y., et al., *Functional characterization of the trigger factor protein PceT of tetrachloroethene-dechlorinating desulfitobacterium hafniense Y51.* Applied Microbiology and Biotechnology, 2009. **83**(4): p. 775-781.

71. Gábor, K., et al., *Characterization of CprK1, a CRP/FNR-type transcriptional regulator of halorespiration from Desulfitobacterium hafniense.* Journal Of Bacteriology, 2006. **188**(7): p. 2604-2613.

72. Wagner, A., et al., *Regulation of reductive dehalogenase gene transcription in Dehalococcoides mccartyi.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1616).

73. Smidt, H., et al., *Transcriptional regulation of the cpr gene cluster in ortho-chlorophenol-respiring Desulfitobacterium dehalogenans.* Journal Of Bacteriology, 2000. **182**(20): p. 5683-5691.

74. Louie, T.M. and W.W. Mohn, *Evidence for a chemiosmotic model of dehalorespiration in Desulfomonile tiedjei DCB-1.* Journal Of Bacteriology, 1999. **181**(1): p. 40-46.

75. Mohn, W.W. and J.M. Tiedje, *Microbial reductive dehalogenation.* Microbiological Reviews, 1992. **56**(3): p. 482-507.

76. Schumacher, W. and C. Holliger, *The proton/electron ratio of the menaquinone-dependent electron transport from dihydrogen to tetrachloroethene in "Dehalobacter restrictus".* Journal Of Bacteriology, 1996. **178**(8): p. 2328-2333.

77. John, M., et al., *Growth substrate dependent localization of tetrachloroethene reductive dehalogenase in Sulfurospirillum multivorans.* Archives of Microbiology, 2006. **186**(2): p. 99-106.

78. Suyama, A., et al., *Molecular characterization of the PceA reductive dehalogenase of Desulfitobacterium sp. strain Y51.* Journal of Bacteriology, 2002. **184**(13): p. 3419-3425.

79. Payne, K.A.P., et al., *Reductive dehalogenase structure suggests a mechanism for B12-dependent dehalogenation.* Nature, 2015. **517**(7535): p. 513-516.

80. Parthasarathy, A., et al., *Biochemical and EPR-Spectroscopic Investigation into Heterologously Expressed Vinyl Chloride Reductive Dehalogenase (VcrA) from Dehalococcoides mccartyi Strain VS.* Journal of the American Chemical Society, 2015. **137**(10): p. 3525-3532.

81. Bommer, M., et al., *Structural basis for organohalide respiration.* Science, 2014.

82. Miller, E., G. Wohlfarth, and G. Diekert, *Studies on tetrachloroethene respiration in Dehalospirillum multivorans.* Archives of Microbiology, 1997. **166**(6): p. 379-387.

83. Holliger, C., et al., *A highly purified enrichment culture couples the reductive dechlorination of tetrachloroethene to growth.* Appl. Environ. Microbiol., 1993. **59**(9): p. 2991-2997.

84. Takayuki, E.H., C., *The firmicutes*, in *Bergey's Manual of Systematic Bacteriology*, P.G. Vos, G.; Jones, D.; Krieg, N.R.; Ludwig, W.; Rainey, F.A.; Schleifer, K.-H.; Whitman, W.B., Editor. 2009. p. 969-975.

85. Field, D., et al., *The minimum information about a genome sequence (MIGS) specification.* Nature Biotechnology, 2008. **26**(5): p. 541-547.

86. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya.* Proceedings of the National Academy of Sciences of the United States of America, 1990. **87**(12): p. 4576-4579.

87. Gibbons, N.E. and R.G.E. Murray, *Proposals concerning the higher taxa of bacteria.* International Journal of Systematic Bacteriology, 1978. **28**(1): p. 1-6.

88. Garrity, G.M. and J.G. Holt, *The Road Map to the Manual*, in *Bergey's Manual of Systematic Bacteriology*, G.M. Garrity, D.R. Boone, and R.W. Castenholz, Editors. 2001, Springer: New York. p. 119-169.

89. Murray, R.G.E., *The Higher Taxa, or a Place for Everything...?*, in *Bergey's Manual of Systematic Bacteriology*, J. Holt, Editor. 1984, The Williams and Wilkins Co.: Baltimore. p. 31-34.

90. Rainey, F.A., *Class II. Clostridia class nov.*, in *Bergey's Manual of Systematic Bacteriology*, P. de Vos, et al., Editors. 2009, Springer Verlag: New York.

91. *List of new names and new combinations previously effectively, but not validly published. List no. 132.* International Journal of Systematic and Evolutionary Microbiology, 2010. **60**(3): p. 469-472.

92. Prévot, A.R., *Dictionnaire des Bactéries Pathogenes* 2ed. 1953, Paris: Masson et Cie.

93. Skerman, V.B.D., V. McGowan, and P.H.A. Sneath, *Approved lists of bacterial names.* International Journal of Systematic Bacteriology, 1980. **30**(1): p. 225-420.

94. Rogosa, M., *Peptococcaceae, a New Family To Include the Gram-Positive, Anaerobic Cocci of the Genera Peptococcus, Peptostreptococcus, and Ruminococcus.* International Journal of Systematic Bacteriology, 1971. **21**(3): p. 234-237.

95. *Validation if publication of new names and new combinations previously effectively published outside the IJSB. List No. 66.* International Journal of Systematic Bacteriology, 1998. **48**(3): p. 631-632.

96. Ashburner, M., et al., *Gene ontology: Tool for the unification of biology.* Nature Genetics, 2000. **25**(1): p. 25-29.

97. Bennett, S., *Solexa Ltd.* Pharmacogenomics, 2004. **5**(4): p. 433-438.

98. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors.* Nature, 2005. **437**(7057): p. 376-380.

99. *http://www.jgi.doe.gov/*.

100. Zerbino, D.R. and E. Birney, *Velvet: Algorithms for de novo short read assembly using de Bruijn graphs.* Genome Research, 2008. **18**(5): p. 821-829.

101. Ewing, B. and P. Green, *Base-calling of automated sequencer traces using phred. II. Error probabilities.* Genome Research, 1998. **8**(3): p. 186-194.

102. Ewing, B., et al., *Base-calling of automated sequencer traces using phred. I. Accuracy assessment.* Genome Research, 1998. **8**(3): p. 175-185.

103. Gordon, D., C. Abajian, and P. Green, *Consed: A graphical tool for sequence finishing.* Genome Research, 1998. **8**(3): p. 195-202.

104. Han, C. and P. Chain. *Finishing Repetitive Regions Automatically with Dupfinisher*. Proceedings of the 2006 International Conference on Bioinformatics Computational Biology, BIOCOMP'06 2006.

105. Hyatt, D., et al., *Prodigal: Prokaryotic gene recognition and translation initiation site identification*, in *Bmc Bioinformatics*. 2010.

106. Pati, A., et al., *GenePRIMP: A gene prediction improvement pipeline for prokaryotic genomes.* Nature Methods, 2010. **7**(6): p. 455-457.

107. Lowe, T.M. and S.R. Eddy, *tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence.* Nucleic Acids Research, 1997. **25**(5): p. 955-964.

108. Lagesen, K., et al., *RNAmmer: Consistent and rapid annotation of ribosomal RNA genes.* Nucleic Acids Research, 2007. **35**(9): p. 3100-3108.

109. Griffiths-Jones, S., et al., *Rfam: An RNA family database.* Nucleic Acids Research, 2003. **31**(1): p. 439-441.

110. Krogh, A., et al., *Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes.* Journal Of Molecular Biology, 2001. **305**(3): p. 567-580.

111. Bendtsen, J.D., et al., *Improved prediction of signal peptides: SignalP 3.0.* Journal Of Molecular Biology, 2004. **340**(4): p. 783-795.

112. Kelley, B., et al., *Conserved pathways within bacteria and yeast as revealed by global protein network alignment.* PNAS, 2003. **100**(20): p. 11394 - 11399.

113. Campanella, J.J., L. Bitincka, and J. Smalley, *MatGAT: An application that generates similarity/identity matrices using protein or DNA sequences.* BMC Bioinformatics, 2003. **4**.

114. Maillard, J., et al., *Characterization of the corrinoid iron-sulfur protein tetrachloroethene reductive dehalogenase of Dehalobacter restrictus.* Applied and Environmental Microbiology, 2003. **69**(8): p. 4628-4638.

115. Rupakula, A., et al., *The restricted metabolism of the obligate organohalide respiring bacterium Dehalobacter restrictus: lessons from tiered functional genomics.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1616).

116. Maillard, J., C. Regeard, and C. Holliger, *Isolation and characterization of Tn-Dha1, a transposon containing the tetrachloroethene reductive dehalogenase of Desulfitobacterium hafniense strain TCE1.* Environmental Microbiology, 2005. **7**(1): p. 107-117.

117. Duret, A., C. Holliger, and J. Maillard, *The opportunistic physiology of Desulfitobacterium hafniense strain TCE1 towards organohalide respiration with tetrachloroethene.* Applied and Environmental Microbiology, 2012.

118. Futagami, T., et al., *Emergence of two types of nondechlorinating variants in the tetrachloroethene-halorespiring Desulfitobacterium sp strain Y51.* Applied Microbiology And Biotechnology, 2006. **70**(6): p. 720-728.

119. Smidt, H., et al., *Random transposition by Tn916 in Desulfitobacterium dehalogenans allows for isolation and characterization of halorespiration-deficient mutants.* Journal of Bacteriology, 1999. **181**(22): p. 6882-6888.

120. Morris, R.M., et al., *Transcription and mass-spectroscopic proteomic studies of electron transport oxidoreductases in Dehalococcoides ethenogenes.* Environmental Microbiology, 2006. **8**(9): p. 1499-1509.

121. Ahsanul Islam, M., E.A. Edwards, and R. Mahadevan, *Characterizing the metabolism of Dehalococcoides with a constraint-based model.* PLoS computational biology, 2010. **6**(8).

122. Tang, Y.J., et al., *Investigation of carbon metabolism in "Dehalococcoides ethenogenes" strain 195 by use of isotopomer and transcriptomic analyses.* Journal of Bacteriology, 2009. **191**(16): p. 5224-5231.

123. Childers, S.E., S. Ciufo, and D.R. Lovley, *Geobacter metallireducens accesses insoluble Fe(III) oxide by chemotaxis.* Nature, 2002. **416**(6882): p. 767-769.

124. Lovley, D.R., et al., *Geobacter. The Microbe Electric's Physiology, Ecology, and Practical Applications*, in *Advances in microbial physiology*. 2011. p. 1-100.

125. Tran, H.T., et al., *Comparative genomics of Geobacter chemotaxis genes reveals diverse signaling function.* Bmc Genomics, 2008. **9**.

126. Duhamel, M., et al., *Chloroform respiration to dichloromethane by a Dehalobacter population.* Environmental Microbiology, 2010. **12**(4): p. 1053-1060.

127. Grostern, A. and E.A. Edwards, *Characterization of a Dehalobacter coculture that dechlorinates 1,2-dichloroethane to ethene and identification of the putative reductive dehalogenase gene.* Applied And Environmental Microbiology, 2009. **75**(9): p. 2684-2693.

128. Nelson, J.L., et al., *A role for dehalobacter spp. in the reductive dehalogenation of dichlorobenzenes and monochlorobenzene.* Environmental Science and Technology, 2011. **45**(16): p. 6806-6813.

129. Yoshida, N., et al., *A novel Dehalobacter species is involved in extensive 4,5,6,7- tetrachlorophthalide dechlorination.* Applied and Environmental Microbiology, 2009. **75**(8): p. 2400-2405.

130. Prat, L., et al., *Physiological adaptation of Desulfitobacterium hafniense strain TCE-1 to tetrachloroethene respiration.* Applied And Environmental Microbiology, 2011. **77**(11): p. 3853-3859.

131. Peng, X., et al., *Global transcriptome analysis of the tetrachloroethene-dechlorinating bacterium Desulfitobacterium hafniense Y51 in the presence of various electron donors and terminal electron acceptors.* Journal of Industrial Microbiology and Biotechnology, 2011: p. 1-14.

132. Fung, J.M., et al., *Expression of reductive dehalogenase genes in Dehalococcoides ethenogenes strain 195 growing on tetrachloroethene, trichloroethene, or 2,3-dichlorophenol.* Applied and Environmental Microbiology, 2007. **73**(14): p. 4439-4445.

133. Johnson, D.R., et al., *Temporal transcriptomic microarray analysis of "Dehalococcoides ethenogenes" strain 195 during the transition into stationary phase.* Applied and Environmental Microbiology, 2008. **74**(9): p. 2864-2872.

134. Lee, P.K., et al., *Global transcriptomic and proteomic responses of dehalococcoides ethenogenes strain 195 to fixed nitrogen limitation.* Applied and Environmental Microbiology, 2012. **78**(5): p. 1424-1436.

135. Morris, R.M., et al., *Comparative Proteomics of Dehalococcoides spp. Reveals Strain-Specific Peptides Associated with Activity.* Appl. Environ. Microbiol., 2007. **73**(1): p. 320-326.

136. Rahm, B.G., R.M. Morris, and R.E. Richardson, *Temporal expression of respiratory genes in an enrichment culture containing Dehalococcoides ethenogenes.* Applied and Environmental Microbiology, 2006. **72**(8): p. 5486-5491.

137. Altschul, S.F., et al., *Basic Local Alignment Search Tool.* Journal of Molecular Biology, 1990. **215**(3): p. 403-410.

138. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-2948.

139. Prat, L., et al., *An unusual tandem-domain rhodanese harbouring two active sites identified in Desulfitobacterium hafniense.* FEBS journal, 2012. **279**(15): p. 2754-2767.

140. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding.* Analytical Biochemistry, 1976. **72**(1-2): p. 248-254.

141. Rappsilber, J., M. Mann, and Y. Ishihama, *Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips.* Nature Protocols, 2007. **2**(8): p. 1896-1906.

142. Lu, J., et al., *Filter-aided sample preparation with dimethyl labeling to identify and quantify milk fat globule membrane proteins.* Journal of proteomics, 2011. **75**(1): p. 34-43.

143. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* 2008. **26**(12): p. 1367-1372.

144. Cox, J., et al., *Andromeda: A peptide search engine integrated into the MaxQuant environment.* Journal of Proteome Research, 2011. **10**(4): p. 1794-1805.

145. Rutherford, K., et al., *Artemis: Sequence visualization and annotation.* Bioinformatics, 2000. **16**(10): p. 944-945.

146. Hubner, N.C., et al., *Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions.* Journal of Cell Biology, 2010. **189**(4): p. 739-754.

147. Nesvizhskii, A.I., *Computational and informatics strategies for identification of specific protein interaction partners in affinity purification mass spectrometry experiments.* Proteomics, 2012. **12**(10): p. 1639-1655.

148. Schmehl, M., et al., *Identification of a new class of nitrogen fixation genes in Rhodobacter capsulatus: A putative membrane complex involved in electron transport to nitrogenase.* Molecular and General Genetics, 1993. **241**(5-6): p. 602-615.

149. Woodson, J.D. and J.C. Escalante-Semerena, *CbiZ, an amidohydrolase enzyme required for salvaging the coenzyme B 12 precursor cobinamide in archaea.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(10): p. 3591-3596.

150. Kovács, A.T., et al., *Ubiquitous late competence genes in Bacillus species indicate the presence of functional DNA uptake machineries: Minireview.* Environmental Microbiology, 2009. **11**(8): p. 1911-1922.

151. Londono-Vallejo, J.A. and D. Dubnau, *comF, a Bacillus subtilis late competence locus, encodes a protein similar to ATP-dependent RNA/DNA helicases.* Molecular Microbiology, 1993. **9**(1): p. 119-131.

152. Engelberg-Kulka, H. and G. Glaser, *Addiction modules and programmed cell death and antideath in bacterial cultures*, in *Annual Review of Microbiology*. 1999. p. 43-70.

153. Lioy, V.S., et al., *Plasmid pSM19035, a model to study stable maintenance in Firmicutes.* Plasmid, 2010. **64**(1): p. 1-17.

154. Gupta, N. and S.W. Ragsdale, *Dual roles of an essential cysteine residue in activity of a redox-regulated bacterial transcriptional activator.* Journal of Biological Chemistry, 2008. **283**(42): p. 28721-28728.

155. Joyce, M.G., et al., *CprK Crystal Structures Reveal Mechanism for Transcriptional Control of Halorespiration.* Journal Of Biological Chemistry, 2006. **281**(38): p. 28318-28325.

156. Levy, C., et al., *Molecular basis of halorespiration control by CprK, a CRP-FNR type transcriptional regulator.* Molecular Microbiology, 2008. **70**(1): p. 151-167.

157. Mazon, H., et al., *Transcriptional Activation by CprK1 Is Regulated by Protein Structural Changes Induced by Effector Binding and Redox State.* Journal Of Biological Chemistry, 2007. **282**(15): p. 11281-11290.

158. Pop, S.M., et al., *Transcriptional activation of dehalorespiration: Identification of redox-active cysteines regulating dimerization and DNA binding.* Journal of Biological Chemistry, 2006. **281**(36): p. 26382-26390.

159. Marzorati, M., et al., *A Novel Reductive Dehalogenase, Identified in a Contaminated Groundwater Enrichment Culture and in Desulfitobacterium dichloroeliminans Strain DCA1, Is Linked to Dehalogenation of 1,2-Dichloroethane.* Appl. Environ. Microbiol., 2007. **73**(9): p. 2990-2999.

160. Bisaillon, A., et al., *Quantitative analysis of the relative transcript levels of four chlorophenol reductive Dehalogenase genes in Desulfitobacterium hafniense PCP-1 exposed to chlorophenols.* Applied and Environmental Microbiology, 2011. **77**(17): p. 6261-6264.

161. Roessner, C.A. and A.I. Scott, *Fine-tuning our knowledge of the anaerobic route to cobalamin (Vitamin B12).* Journal of Bacteriology, 2006. **188**(21): p. 7331-7334.

162. Woodson, J.D., C.L. Zayas, and J.C. Escalante-Semerena, *A New Pathway for Salvaging the Coenzyme B12 Precursor Cobinamide in Archaea Requires Cobinamide-Phosphate Synthase (CbiB) Enzyme Activity.* Journal of Bacteriology, 2003. **185**(24): p. 7193-7201.

163. Gray, M.J. and J.C. Escalante-Semerena, *In vivo analysis of cobinamide salvaging in Rhodobacter sphaeroides strain 2.4.1.* Journal of Bacteriology, 2009. **191**(12): p. 3842-3851.

164. Zayas, C.L. and J.C. Escalante-Semerena, *Reassessment of the late steps of coenzyme B12 synthesis in Salmonella enterica: Evidence that dephosphorylation of adenosylcobalamin- 5β-phosphate by the CobC phosphatase is the last step of the pathway.* 2007. **189**(6): p. 2210-2218.

165. Kreher, S., A. Schilhabel, and G. Diekert, *Enzymes involved in the anoxic utilization of phenyl methyl ethers by Desulfitobacterium hafniense DCB-2 and Desulfitobacterium hafniense PCE-S.* Archives of Microbiology, 2008. **190**(4): p. 489-495.

166. Marco-Urrea, E., et al., *Stable isotope peptide mass spectrometry to decipher amino acid metabolism in dehalococcoides strain CBDB1.* Journal of Bacteriology, 2012. **194**(16): p. 4169-4177.

167. Moparthi, V.K. and C. Hägerhäll, *The evolution of respiratory chain complex I from a smaller last common ancestor consisting of 11 protein subunits.* Journal Of Molecular Evolution, 2011: p. 1-14.

168. Claverys, J.P., M. Prudhomme, and B. Martin, *Induction of competence regulons as a general response to stress in gram-positive bacteria*, in *Annual Review of Microbiology*. 2006. p. 451-475.

169. Krasotkina, J., et al., *Characterization of the B12- and Iron-Sulfur-containing Reductive Dehalogenase from Desulfitobacterium chlororespirans.* Journal of Biological Chemistry, 2001. **276**(44): p. 40991-40997.

170. Miller, E., G. Wohlfarth, and G. Diekert, *Comparative studies on tetrachloroethene reductive dechlorination mediated by Desulfitobacterium sp. strain PCE-S.* Archives of Microbiology, 1997. **168**(6): p. 513-519.

171. Thibodeau, J., et al., *Purification, cloning, and sequencing of a 3,5-dichlorophenol reductive dehalogenase from Desulfitobacterium frappieri PCP-1.* Applied and Environmental Microbiology, 2004. **70**(8): p. 4532-4537.

172. Stackebrandt, E., et al., *Phylogenetic analysis of the genus Desulfotomaculum: Evidence for the misclassification of Desulfotomaculum guttoideum and description of Desulfotomaculum orientis as desulfosporosinus orientis gen. nov., comb. nov.* International Journal of Systematic Bacteriology, 1997. **47**(4): p. 1134-1139.

173. Wagner, A., et al., *Transcription analysis of genes encoding homologues of reductive dehalogenases in "Dehalococcoides" sp. Strain CBDB1 by using terminal restriction fragment length polymorphism and quantitative PCR.* Applied and Environmental Microbiology, 2009. **75**(7): p. 1876-1884.

174. West, K.A., et al., *Comparative Genomics of Dehalococcoides ethenogenes 195 and an Enrichment Culture Containing Unsequenced Dehalococcoides Strains.* Appl. Environ. Microbiol., 2008: p. AEM.01835-07.

175. Bolhuis, A., et al., *Signal peptide peptidase- and ClpP-like proteins of Bacillus subtilis required for efficient translocation and processing of secretory proteins.* Journal of Biological Chemistry, 1999. **274**(35): p. 24585-24592.

176. Wang, P., et al., *Escherichia coli signal peptide peptidase A is a serine-lysine protease with a lysine recruited to the nonconserved amino-terminal domain in the S49 protease family.* Biochemistry, 2008. **47**(24): p. 6361-6369.

177. Banerjee, R. and S.W. Ragsdale, *The many faces of vitamin B12: Catalysis by cobalamin-dependent enzymes*, in *Annual Review of Biochemistry*. 2003. p. 209-247.

178. Moore, S.J. and M.J. Warren, *The anaerobic biosynthesis of vitamin B12.* Biochem Soc Trans, 2012. **40**(3): p. 581-586.

179. Scott, A.I., *Discovering nature's diverse pathways to vitamin B12: A 35-year odyssey.* Journal of Organic Chemistry, 2003. **68**(7): p. 2529-2539.

180. Schumacher, W., et al., *Redox chemistry of cobalamin and iron-sulfur cofactors in the tetrachloroethene reductase of Dehalobacter restrictus.* FEBS Letters, 1997. **409**(3): p. 421-425.

181. Hug, L.A., et al., *Comparative metagenomics of three Dehalococcoides-containing enrichment cultures: the role of the non-dechlorinating community.* Bmc Genomics, 2012. **13**(1).

182. Yan, J., et al., *Guided cobalamin biosynthesis supports Dehalococcoides mccartyi reductive dechlorination activity.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1616).

183. Men, Y., et al., *Identification of specific corrinoids reveals corrinoid modification in dechlorinating microbial communities.* 2014.

184. Men, Y., et al., *Sustainable growth of Dehalococcoides mccartyi 195 by corrinoid salvaging and remodeling in defined lactate-fermenting consortia.* 2014. **80**(7): p. 2133-2141.

185. Maillard, J., et al., *Reductive dechlorination of tetrachloroethene by a stepwise catalysis of different organohalide respiring bacteria and reductive dehalogenases.* Biodegradation, 2011. **22**(5): p. 949-960.

186. Burge, S.W., et al., *Rfam 11.0: 10 years of RNA families.* Nucleic Acids Research, 2013. **41**(D1): p. D226-D232.

187. Carver, T.J., et al., *ACT: the Artemis comparison tool.* Bioinformatics, 2005. **21**(16): p. 3422-3423.

188. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular cloning: a laboratory manual*. 2 ed. 1989, New York: Cold Spring Harbor Laboratory.

189. Smaczniak, C., et al., *Proteomics-based identification of low-abundance signaling and regulatory protein complexes in native plant tissues.* Nature Protocols, 2012. **7**(12): p. 2144-2158.

190. Chenna, R., et al., *Multiple sequence alignment with the Clustal series of programs.* Nucleic Acids Research, 2003. **31**(13): p. 3497 - 3500.

191. Fan, C. and T.A. Bobik, *The PduX enzyme of Salmonella enterica is an L-threonine kinase used for coenzyme B12 synthesis.* Journal of Biological Chemistry, 2008. **283**(17): p. 11322-11329.

192. Poehlein, A., et al., *An ancient pathway combining carbon dioxide fixation with the generation and utilization of a sodium ion gradient for ATP synthesis.* PLoS ONE, 2012. **7**(3).

193. Nahvi, A., J.E. Barrick, and R.R. Breaker, *Coenzyme B12 riboswitches are widespread genetic control elements in prokaryotes.* Nucleic Acids Research, 2004. **32**(1): p. 143-150.

194. Grostern, A. and E.A. Edwards, *Growth of Dehalobacter and Dehalococcoides spp. during degradation of chlorinated ethanes.* Applied And Environmental Microbiology, 2006. **72**(1): p. 428-436.

195. Schipp, C.J., et al., *Organic cofactors in the metabolism of dehalococcoides mccartyi strains.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1616).

196. Stupperich, E., H.J. Eisinger, and B. Kräutler, *Diversity of corrinoids in acetogenic bacteria. P-cresolylcobamide from Sporomusa ovata, 5-methoxy-6-methylbenzimidazolylcobamide from Clostridium formicoaceticum and vitamin B12 from Acetobacterium woodii.* European Journal of Biochemistry, 1988. **172**(2): p. 459-464.

197. Ragsdale, S.W. and E. Pierce, *Acetogenesis and the Wood-Ljungdahl pathway of CO2 fixation.* Biochimica et Biophysica Acta - Proteins and Proteomics, 2008. **1784**(12): p. 1873-1898.

198. He, J.Z., et al., *Influence of vitamin B-12 and cocultures on the growth of Dehalococcoides isolates in defined medium.* Applied and Environmental Microbiology, 2007. **73**(9): p. 2847-2853.

199. Men, Y., et al., *Sustainable syntrophic growth of Dehalococcoides ethenogenes strain 195 with desulfovibrio vulgaris hildenborough and methanobacterium congolense: Global transcriptomic and proteomic analyses.* 2012. **6**(2): p. 410-421.

200. Löffler, F.E., R.A. Sanford, and K.M. Ritalahti, *Enrichment, cultivation, and detection of reductively dechlorinating bacteria*, in *Methods in Enzymology*. 2005. p. 77-111.

201. Schweizer, H.P., *Bacterial genetics: Past achievements, present state of the field, and future challenges.* Biotechniques, 2008. **44**(5): p. 633-641.

202. Davies, J. and D. Davies, *Origins and evolution of antibiotic resistance.* Microbiology And Molecular Biology Reviews, 2010. **74**(3): p. 417-433.

203. Courvalin, P., *Vancomycin resistance in gram-positive cocci.* Clinical Infectious Diseases, 2006. **42**(SUPPL. 1): p. S25-S34.

204. Fraimow, H., et al., *Putative VanRS-like two-component regulatory system associated with the inducible glycopeptide resistance cluster of Paenibacillus popilliae.* Antimicrobial Agents and Chemotherapy, 2005. **49**(7): p. 2625-2633.

205. Depardieu, F., et al., *Modes and modulations of antibiotic resistance gene expression.* Clinical Microbiology Reviews, 2007. **20**(1): p. 79-114.

206. Boyd, D.A., et al., *Molecular Characterization of Enterococcus faecalis N06-0364 with Low-Level Vancomycin Resistance Harboring a Novel D-Ala-D-Ser Gene Cluster, vanL.* Antimicrob. Agents Chemother., 2008. **52**(7): p. 2667-2672.

207. Xu, X., et al., *vanM, a New Glycopeptide Resistance Gene Cluster Found in Enterococcus faecium.* Antimicrob. Agents Chemother., 2010. **54**(11): p. 4643-4647.

208. Lebreton, F., et al., *D-Ala-D-Ser VanN-type transferable vancomycin resistance in Enterococcus faecium.* Antimicrobial Agents and Chemotherapy, 2011. **55**(10): p. 4606-4612.

209. Kalan, L., et al., *Noncanonical vancomycin resistance cluster from Desulfitobacterium hafniense Y51.* Antimicrobial Agents and Chemotherapy, 2009. **53**(7): p. 2841-2845.

210. Marshall, C.G., et al., *Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms.* Antimicrobial Agents and Chemotherapy, 1998. **42**(9): p. 2215-2220.

211. Hong, H.J., M.I. Hutchings, and M.J. Buttner, *Vancomycin resistance VanS/VanR two-component systems*, in *Advances in Experimental Medicine and Biology*. 2008. p. 200-213.

212. Aminov, R.I. and R.I. Mackie, *Evolution and ecology of antibiotic resistance genes.* Fems Microbiology Letters, 2007. **271**(2): p. 147-161.

213. Abicht, H.K., et al., *Genome Sequence of Desulfosporosinus sp. OT, an Acidophilic Sulfate-Reducing Bacterium from Copper Mining Waste in Norilsk, Northern Siberia.* Journal of Bacteriology, 2011. **193**(21): p. 6104-6105.

214. Teo, J.W.P., et al., *Detection of an unusual van genotype in a vancomycin-resistant Enterococcus faecium hospital isolate.* Journal Of Clinical Microbiology, 2011. **49**(12): p. 4297-4298.

215. Pester, M., et al., *Complete Genome Sequences of Desulfosporosinus orientis DSM765T, Desulfosporosinus youngiae DSM17734T, Desulfosporosinus meridiei DSM13257T, and Desulfosporosinus acidiphilus DSM22704T.* Journal of Bacteriology, 2012. **194**(22): p. 6300-6301.

216. Markowitz, V.M., et al., *IMG: The integrated microbial genomes database and comparative analysis system.* Nucleic acids research, 2012. **40**(D1): p. D115-D122.

217. Hong, H.J., et al., *The role of the novel fem protein VanK in vancomycin resistance in Streptomyces coelicolor.* Journal of Biological Chemistry, 2005. **280**(13): p. 13055-13061.

218. Breitenstein, A., et al., *Reclassification of Clostridium hydroxybenzoicum as Sedimentibacter hydroxybenzoicus gen. nov., comb. nov., and description of Sedimentibacter saalensis sp. nov.* International Journal of Systematic and Evolutionary Microbiology, 2002. **52**(3): p. 801-7.

219. Benveniste, R. and J. Davies, *Aminoglycoside antibiotic inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic resistant bacteria.* Proceedings of the National Academy of Sciences of the United States of America, 1973. **70**(8): p. 2276-2280.

220. D'Costa, V.M., et al., *Sampling the antibiotic resistome.* Science, 2006. **311**(5759): p. 374-377.

221. Forsberg, K.J., et al., *The shared antibiotic resistome of soil bacteria and human pathogens.* Science, 2012. **337**(6098): p. 1107-1111.

222. Spring, S. and F. rosenzweig, *The Genera Desulfitobacterium and Desulfosporosinus*, in *The Prokaryotes*, E. rosenberg, et al., Editors. 2006, Springer-Verlag Berlin Heidelberg. p. 771–786.

223. Hong, H.J., et al., *Characterization of an inducible vancomycin resistance system in Streptomyces coelicolor reveals a novel gene (vanK) required for drug resistance.* Molecular Microbiology, 2004. **52**(4): p. 1107-1121.

224. Bedard, D.L., et al., *Development and characterization of stable sediment-free anaerobic bacterial enrichment cultures that dechlorinate aroclor 1260.* Applied and Environmental Microbiology, 2006. **72**(4): p. 2460-2470.

225. Breitenstein, A., J.R. Andreesen, and U. Lechner, *Analysis of an anaerobic chemostat population stably dechlorinating 2,4,6-trichlorophenol.* Engineering in Life Sciences, 2007. **7**(4): p. 380-387.

226. Sánchez-Andrea, I., et al., *Screening of anaerobic activities in sediments of an acidic environment: Tinto River.* Extremophiles, 2012: p. 1-11.

227. Neumann, A., H. Scholz-Muramatsu, and G. Diekert, *Tetrachloroethene metabolism of Dehalospirillum multivorans.* Archives of Microbiology, 1994. **162**(4): p. 295-301.

228. Markowitz, V.M., et al., *IMG/M: The integrated metagenome data management and comparative analysis system.* Nucleic acids research, 2012. **40**(D1): p. D123-D129.

229. Lane, D.J., *16S/23S rRNA sequencing*, in *Nucleic acid techniques in bacterial systematics*, E. Stackebrandt and M. Goodfellow, Editors. 1991, John Wiley & Sons ltd: West Sussex. p. 115-175.

230. Turner, S., et al., *Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis.* Journal of Eukaryotic Microbiology, 1999. **46**(4): p. 327-338.

231. Smits, T.H.M., et al., *Development of a real-time PCR method for quantification of the three genera Dehalobacter, Dehalococcoides, and Desulfitobacterium in microbial communities.* Journal of Microbiological Methods, 2004. **57**(3): p. 369-378.

232. Bock, A., et al., Selenocysteine the 21st amino-acid. Molecular Microbiology, 1991. **5**(3): p. 515-520.

233. Robinson, C. and A. Bolhuis, *Protein targeting by the twin-arginine translocation pathway.* Nature reviews. Molecular cell biology, 2001. **2**(5): p. 350-356.

234. Futagami, T., M. Goto, and K. Furukawa, *Biochemical and genetic bases of dehalorespiration.* Chemical Record, 2008. **8**(1): p. 1-12.

235. Holliger, C., C. Regeard, and G. Diekert, *Dehalogenation by anaerobic bacteria*, in *Dehalogenation: microbial processes and environmental applications*. 2003, Kluwer. p. 115-157.

236. Simon, J., R.J.M. van Spanning, and D.J. Richardson, *The organisation of proton motive and non-proton motive redox loops in prokaryotic respiratory systems.* Biochimica et Biophysica Acta - Bioenergetics, 2008. **1777**(12): p. 1480-1490.

237. Utkin, I., D.D. Dalton, and J. Wiegel, *Specificity of reductive dehalogenation of substituted ortho- chlorophenols by Desulfitobacterium dehalogenans JW/IU-DC1.* Applied And Environmental Microbiology, 1995. **61**(1): p. 346-351.

238. Jozefczuk, S., et al., *Metabolomic and transcriptomic stress response of Escherichia coli.* Molecular Systems Biology, 2010. **6**.

239. Feil, W.S.H. *Bacterial genomic DNA isolation using CTAB*. 2004 [cited 2004 8-25]; 2:[Available from: http://my.jgi.doe.gov/general/protocols.html.

240. Bagos, P.G., et al., *Combined prediction of Tat and Sec signal peptides with hidden Markov models.* Bioinformatics, 2010. **26**(22): p. 2811-2817.

241. Letunic, I., T. Doerks, and P. Bork, *SMART 6: Recent updates and new developments.* Nucleic Acids Research, 2009. **37**(SUPPL. 1): p. D229-D232.

242. Goa, J., *A micro biuret method for protein determination; determination of total protein in cerebrospinal fluid.* Scandinavian journal of clinical and laboratory investigation, 1953. **5**(3): p. 218-222.

243. Van de Pas, B.A., et al., *Energy yield of respiration on chloroaromatic compounds in Desulfitobacterium dehalogenans.* Applied And Environmental Microbiology, 2001. **67**(9): p. 3958-3963.

244. Van den Bergen, C.W.M., et al., *The relationship between electron flux and a redox poise of the quinone pool in plant mitochondria. Interplay between quinol-oxidizing and quinone-reducing pathways.* European Journal of Biochemistry, 1994. **226**(3): p. 1071-1078.

245. Van Wielink, J.E., et al., *A method for in situ characterization of b- and c-type cytochromes in Escherichia coli and in Complex III from beef heart mitochondria by combined spectrum deconvolution and potentiometric analysis.* BBA - Bioenergetics, 1982. **681**(2): p. 177-190.

246. Krab, K., et al., *Identification of the site where the electron transfer chain of plant mitochondria is stimulated by electrostatic charge screening.* European Journal of Biochemistry, 2000. **267**(3): p. 869-876.

247. Petersen, T.N., et al., *SignalP 4.0: Discriminating signal peptides from transmembrane regions.* Nature Methods, 2011. **8**(10): p. 785-786.

248. Andrews, S.C., et al., *A 12-cistron Escherichia coli operon (hyf) encoding a putative proton-translocating formate hydrogenlyase system.* Microbiology, 1997. **143**(11): p. 3633-3647.

249. Chen, M., et al., *Identification of a Hotdog Fold Thioesterase Involved in the Biosynthesis of Menaquinone in Escherichia coli.* Journal of Bacteriology, 2013. **195**(12): p. 2768-2775.

250. Bamford, V., et al., *Open conformation of a flavocytochrome c3 fumarate reductase.* Nature Structural Biology, 1999. **6**(12): p. 1104-1107.

251. Dobbin, P.S., et al., *Characterization of a flavocytochrome that is induced during the anaerobic respiration of Fe^{3+} by Shewanella frigidimarina NCIMB400.* Biochemical Journal, 1999. **342**(2): p. 439-448.

252. Fonseca, B.M., et al., *Mind the gap: Cytochrome interactions reveal electron pathways across the periplasm of Shewanella oneidensis MR-1.* Biochemical Journal, 2013. **449**(1): p. 101-108.

253. Mikoulinskaia, O., et al., *Cytochrome c-dependent methacrylate reductase from Geobacter sulfurreducens AM-1.* Eur J Biochem, 1999. **263**(2): p. 346-52.

254. Simon, J. and M.G. Klotz, *Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations.* Biochimica et Biophysica Acta - Bioenergetics, 2013. **1827**(2): p. 114-135.

255. Edgren, T. and S. Nordlund, *The fixABCX genes in Rhodospirillum rubrum encode a putative membrane complex participating in electron transfer to nitrogenase.* Journal Of Bacteriology, 2004. **186**(7): p. 2052-2060.

256. Selao, T.T., et al., *Identification of chromatophore membrane protein complexes formed under different nitrogen availability conditions in Rhodospirillum rubrum.* Journal of Proteome Research, 2011. **10**(6): p. 2703-2714.

257. Gilmore, J.M. and M.P. Washburn, *Advances in shotgun proteomics and the analysis of membrane proteomes.* Journal of proteomics, 2010. **73**(11): p. 2078-2091.

258. Kröger, A., et al., *Fumarate respiration of Wolinella succinogenes: Enzymology, energetics and coupling mechanism.* Biochimica et Biophysica Acta - Bioenergetics, 2002. **1553**(1-2): p. 23-38.

259. Lancaster, C.R.D., et al., *Experimental support for the "E pathway hypothesis" of coupled transmembrane e- and H^+ transfer in dihemic quinol:fumarate reductase.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(52): p. 18860-18865.

260. Ragsdale, S.W., *Enzymology of the Wood–Ljungdahl Pathway of Acetogenesis.* Annals of the New York Academy of Sciences, 2008. **1125**(1): p. 129-136.

261. Thauer, R.K., K. Jungermann, and K. Decker, *Energy-Conservation in Chemotropic Anaerobic Bacteria.* Bacteriological Reviews, 1977. **41**(1): p. 100-180.

262. Buckel, W. and R.K. Thauer, *Energy conservation via electron bifurcating ferredoxin reduction and proton/Na+ translocating ferredoxin oxidation.* Biochimica et Biophysica Acta - Bioenergetics, 2013. **1827**(2): p. 94-113.

263. Pittman, M.S., et al., *Growth of Campylobacter jejuni on nitrate and nitrite: Electron transport to NapA and NrfA via NrfH and distinct roles for NrfA and the globin Cgb in protection against nitrosative stress.* Molecular Microbiology, 2007. **63**(2): p. 575-590.

264. Kern, M. and J. Simon, *Characterization of the NapGH quinol dehydrogenase complex involved in Wolinella succinogenes nitrate respiration.* Molecular Microbiology, 2008. **69**(5): p. 1137-1152.

265. Dos Santos, P.C., et al., *Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes.* Bmc Genomics, 2012. **13**(1).

266. Arkhipova, O.V., G.V. Mikulinskaya, and A.S. Galushko, *Comparative analysis of the N-terminal sequence of Geobacter sulfurreducens AM-1 methacrylate reductase.* 2012. **81**(5): p. 555-564.

267. Li, D.-B., et al., *Selenite reduction by Shewanella oneidensis MR-1 is mediated by fumarate reductase in periplasm.* 2014. **4**.

268. Bothe, H., et al., *Nitrogen fixation and hydrogen metabolism in cyanobacteria.* Microbiology and Molecular Biology Reviews, 2010. **74**(4): p. 529-551.

269. Odom, J.M. and H.D. Peck Jr, *Hydrogenase, electron-transfer proteins, and energy coupling in the sulfate-reducing bacteria Desulfovibrio.* Annual Review of Microbiology, 1984. **38**: p. 551-592.

270. Voordouw, G., *Carbon monoxide cycling by Desulfovibrio vulgaris Hildenborough.* Journal of Bacteriology, 2002. **184**(21): p. 5903-5911.

271. De Wildeman, S., et al., *Stereoselective microbial dehalorespiration with vicinal dichlorinated alkanes.* Applied and Environmental Microbiology, 2003. **69**(9): p. 5643-5647.

272. Finneran, K.T., et al., *Desulfitobacterium metallireducens sp. nov., an anaerobic bacterium that couples growth to the reduction of metals and humic acids as well as chlorinated compounds.* Int J Syst Evol Microbiol, 2002. **52**(6): p. 1929-1935.

273. Suyama, A., et al., *Isolation and characterization of Desulfitobacterium sp strain Y51 capable of efficient dehalogenation of tetrachloroethene and polychloroethanes.* Bioscience Biotechnology And Biochemistry, 2001. **65**(7): p. 1474-1481.

274. Thompson, J.D., D.G. Higgins, and T.J. Gibson, *CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.* Nucleic Acids Research, 1994. **22**(22): p. 4673-4680.

275. Breitenstein, A., et al., *Analysis of a 2,4,6-trichlorophenol-dehalogenating enrichment culture and isolation of the dehalogenating member Desulfitobacterium frappieri strain TCP-A.* Archives of Microbiology, 2001. **175**(2): p. 133-142.

276. Bouchard, B., et al., *Isolation and characterization of Desulfitobacterium frappieri sp. nov., an anaerobic bacterium which reductively dechlorinates pentachlorophenol to 3-chlorophenol.* International Journal of Systematic Bacteriology, 1996. **46**(4): p. 1010-1015.

277. Gerritse, J., et al., *Influence of Different Electron Donors and Acceptors on Dehalorespiration of Tetrachloroethene by Desulfitobacterium frappieri TCE1.* Appl. Environ. Microbiol., 1999. **65**(12): p. 5212-5221.

278. Shelobolina, E.S., C.G. Vanpraagh, and D.R. Lovley, *Use of ferric and ferrous iron containing minerals for respiration by Desulfitobacterium frappieri.* Geomicrobiology Journal, 2003. **20**(2): p. 143-156.

279. Gerritse, J., et al., *Desulfitobacterium sp strain PCE1, an anaerobic bacterium that can grow by reductive dechlorination of tetrachloroethene or ortho-chlorinated phenols.* Archives of Microbiology, 1996. **165**(2): p. 132-140.

280. Sanford, R.A., et al., *Characterization of Desulfitobacterium chlororespirans sp. nov., which grows by coupling the oxidation of lactate to the reductive dechlorination of 3-chloro-4-hydroxybenzoate.* Applied and Environmental Microbiology, 1996. **62**(10): p. 3800-8.

281. Kruse, T., et al., *vanI: a novel d-Ala-d-Lac vancomycin resistance gene cluster found in Desulfitobacterium hafniense.* Microbial Biotechnology, 2014: p. n/a-n/a.

282. Rice, P., L. Longden, and A. Bleasby, *EMBOSS: The European Molecular Biology Open Software Suite.* Trends in Genetics, 2000. **16**(6): p. 276-277.

283. Mitchell, A., et al., *The InterPro protein families database: the classification resource after 15 years.* Nucleic Acids Research, 2015. **43**(Database issue): p. D213-D221.

284. Richter, M. and R. Rosselló-Móra, *Shifting the genomic gold standard for the prokaryotic species definition.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(45): p. 19126-19131.

285. Vesth, T., et al., *CMG-Biotools, a Free Workbench for Basic Comparative Microbial Genomics.* PLoS ONE, 2013. **8**(4).

286. Snipen, L., T. Almøy, and D.W. Ussery, *Microbial comparative pan-genomics using binomial mixture models.* BMC Genomics, 2009. **10**.

287. Letunic, I. and P. Bork, *Interactive Tree of Life v2: Online annotation and display of phylogenetic trees made easy.* Nucleic acids research, 2011. **39**(SUPPL. 2): p. W475-W478.

288. Goris, J., et al., *DNA-DNA hybridization values and their relationship to whole-genome sequence similarities.* International Journal of Systematic and Evolutionary Microbiology, 2007. **57**(1): p. 81-91.

289. Kim, M., et al., *Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes.* International Journal of Systematic and Evolutionary Microbiology, 2014. **64**(PART 2): p. 346-351.

290. Villemur, R., et al., *Heterogeneity between 16S ribosomal RNA gene copies borne by one Desulfitobacterium strain is caused by different 100-200 bp insertions in the 5 ' region.* Canadian Journal of Microbiology, 2007. **53**(1): p. 116-128.

291. Tettelin, H., et al., *Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: Implications for the microbial "pan-genome".* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(39): p. 13950-13955.

292. Tettelin, H., et al., *Comparative genomics: the bacterial pan-genome.* Current Opinion in Microbiology, 2008. **11**(5): p. 472-477.

293. Kruse, T., H. Smidt, and U. Lechner, *Comparative genomics and transcriptomics of organohalide respiring bacteria and regulation of rdh gene transcription*, in *Organohalide Respiring Bacteria*, L. Adrian and F.E. Löffler, Editors. 2015, Springer

294. Kim, Y.E., et al., *Molecular chaperone functions in protein folding and proteostasis*, in *Annual Review of Biochemistry*. 2013. p. 323-355.

295. Bisaillon, A., et al., *Identification and characterization of a novel CprA reductive dehalogenase specific to highly chlorinated phenols from Desulfitobacterium hafniense strain PCP-1.* Applied and Environmental Microbiology, 2010. **76**(22): p. 7536-7540.

296. Van de Pas, B.A., et al., *Two distinct enzyme systems are responsible for tetrachloroethene and chlorophenol reductive dehalogenation in Desulfitobacterium strain PCE1.* Archives of Microbiology, 2001. **176**(3): p. 165-169.

297. Williams, K.P., *Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: Sublocation preference of integrase subfamilies.* Nucleic Acids Research, 2002. **30**(4): p. 866-875.

298. Letunic, I., et al., *SMART 4.0: towards genomic data integration.* Nucleic Acids Research, 2004. **32**: p. D142 - D144.

299. Karlin, S., *Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes.* Trends In Microbiology, 2001. **9**(7): p. 335-343.

300. Rodrigue, A., et al., *Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway.* Journal Of Biological Chemistry, 1999. **274**(19): p. 13223-13228.

301. Mingo, F.S., S. Studenik, and G. Diekert, *Conversion of phenyl methyl ethers by Desulfitobacterium spp. and screening for the genes involved.* FEMS Microbiology Ecology, 2014. **90**(3): p. 783-790.

302. Wiegel, B.L.J., *The firmicutes*, in *Bergey's Manual of Systematic Bacteriology*, P.G. Vos, G.; Jones, D.; Krieg, N.R.; Ludwig, W.; Rainey, F.A.; Schleifer, K.-H.; Whitman, W.B., Editor. 2009. p. 975-982.

303. Meyer, B. and J. Kuevert, *Phylogeny of the alpha and beta subunits of the dissimilatory adenosine-5β -phosphosulfate (APS) reductase from sulfate-reducing prokaryotes - Origin and evolution of the dissimilatory sulfate-reduction pathway.* Microbiology, 2007. **153**(7): p. 2026-2044.

304. Hensel, M., et al., *The genetic basis of tetrathionate respiration in Salmonella typhimurium.* Molecular Microbiology, 1999. **32**(2): p. 275-287.

305. Hinsley, A.P. and B.C. Berks, *Specificity of respiratory pathways involved in the reduction of sulfur compounds by Salmonella enterica.* Microbiology, 2002. **148**(11): p. 3631-3638.

306. Venceslau, S.S., et al., *The "bacterial heterodisulfide" DsrC is a key protein in dissimilatory sulfur metabolism.* Biochimica et Biophysica Acta - Bioenergetics, 2014. **1837**(7): p. 1148-1164.

307. Simon, J. and P.M.H. Kroneck, *Microbial Sulfite Respiration*, in *Advances in microbial physiology*. 2013. p. 45-117.

308. Dahl, C., et al., *Novel genes of the dsr gene cluster and evidence for close interaction of Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium Allochromatium vinosum.* Journal of Bacteriology, 2005. **187**(4): p. 1392-1404.

309. Weissgerber, T., et al., *Complete genome sequence of Allochromatium vinosum DSM 180 T.* Standards in Genomic Sciences, 2011. **5**(3): p. 311-330.

310. Baar, C., et al., *Complete genome sequence and analysis of Wolinella succinogenes.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(20): p. 11690-11695.

311. Sparacino-Watkins, C., J.F. Stolz, and P. Basu, *Nitrate and periplasmic nitrate reductases.* Chemical Society Reviews, 2014. **43**(2): p. 676-706.

312. Verbaendert, I., et al., *Denitrification in Gram-positive bacteria: An underexplored trait.* Biochemical Society Transactions, 2011. **39**(1): p. 254-258.

313. Wunsch, P., et al., *Requirements for CuA and Cu-S center assembly of nitrous oxide reductase deduced from complete periplasmic enzyme maturation in the nondenitrifier Pseudomonas putida.* Journal of Bacteriology, 2003. **185**(3): p. 887-896.

314. Liu, X., et al., *The nos gene cluster from gram-positive bacterium Geobacillus thermodenitrificans NG80-2 and functional characterization of the recombinant NosZ.* FEMS Microbiology Letters, 2008. **289**(1): p. 46-52.

315. Beliaev, A.S., et al., *Gene and protein expression profiles of Shewanella oneidensis during anaerobic growth with different electron acceptors.* OMICS A Journal of Integrative Biology, 2002. **6**(1): p. 39-60.

316. Luijten, M.L.G.C., et al., *Anaerobic reduction and oxidation of quinone moieties and the reduction of oxidized metals by halorespiring and related organisms*. Vol. 49. 2004. 145-150.

317. Saltikov, C.W., *Regulation of Arsenic Metabolic Pathways in Prokaryotes*, in *Microbial Metal and Metalloid Metabolism*. 2011, American Society of Microbiology.

318. Van Lis, R., et al., *Arsenics as bioenergetic substrates.* Biochimica et Biophysica Acta - Bioenergetics, 2013. **1827**(2): p. 176-188.

319. Schröder, G., et al., *TraG-like proteins of DNA transfer systems and of the Helicobacter pylori type IV secretion system: Inner membrane gate for exported substrates?* Journal of Bacteriology, 2002. **184**(10): p. 2767-2779.

320. Stolz, J.F., et al., *Arsenic and selenium in microbial metabolism*, in *Annual Review of Microbiology*. 2006. p. 107-130.

321. Zhang, J., et al., *Arsenic methylation and volatilization by arsenite s-adenosylmethionine methyltransferase in Pseudomonas alcaligenes NBRC14159.* Applied and Environmental Microbiology, 2015. **81**(8): p. 2852-2860.

322. Qin, J., et al., *Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(7): p. 2075-2080.

323. Bentley, R. and T.G. Chasteen, *Microbial Methylation of Metalloids: Arsenic, Antimony, and Bismuth.* Microbiology and Molecular Biology Reviews, 2002. **66**(2): p. 250-271.

324. Studenik, S., M. Vogel, and G. Diekert, *Characterization of an O-Demethylase of Desulfitobacterium hafniense DCB-2.* Journal of Bacteriology, 2012. **194**(13): p. 3317-3326.

325. Ruiz-Argüeso, T., J.M. Palacios, and J. Imperial, *Regulation of the hydrogenase system in Rhizobium leguminosarum.* Plant and Soil, 2001. **230**(1): p. 49-57.

326. Richau, K.H., et al., *Structural and gene expression analyses of uptake hydrogenases and other proteins involved in nitrogenase protection in Frankia.* Journal of Biosciences, 2013. **38**(4): p. 703-712.

327. Da Silva, S.M., I. Pacheco, and I.A.C. Pereira, *Electron transfer between periplasmic formate dehydrogenase and cytochromes C in Desulfovibrio desulfuricans ATCC 27774.* Journal of Biological Inorganic Chemistry, 2012. **17**(5): p. 831-838.

328. Odom, J.M. and H.D. Peck Jr, *Hydrogen cycling as a general mechanism for energy coupling in the sulfate-reducing bacteria. Desulfovibrio sp.* FEMS Microbiology Letters, 1981. **12**(1): p. 47-50.

329. Smidt, H., J. van der Oost, and W.M. de Vos, *Development of a gene cloning and inactivation system for halorespiring Desulfitobacterium dehalogenans.* Applied And Environmental Microbiology, 2001. **67**(2): p. 591-597.

330. Simon, J., et al., *A periplasmic flavoprotein in Wolinella succinogenes that resembles the fumarate reductase of Shewanella putrefaciens.* Archives of Microbiology, 1998. **169**(5): p. 424-433.

331. Gross, R., J. Simon, and A. Kröger, *Periplasmic methacrylate reductase activity in Wolinella succinogenes.* Archives of Microbiology, 2001. **176**(4): p. 310-313.

332. Pealing, S.L., et al., *Sequence of the gene encoding flavocytochrome c from Shewanella putrefaciens: A tetraheme flavoenzyme that is a soluble fumarate reductase related to the membrane-bound enzymes from other bacteria.* Biochemistry, 1992. **31**(48): p. 12132-12140.

333. Tsapin, A.I., et al., *Identification of a Small Tetraheme Cytochrome c and a Flavocytochrome c as Two of the Principal Soluble Cytochromes c in Shewanella oneidensis Strain MR1.* Applied and Environmental Microbiology, 2001. **67**(7): p. 3236-3244.

334. Schuetz, B., et al., *Periplasmic Electron Transfer via the c-Type Cytochromes MtrA and FccA of Shewanella oneidensis MR-1.* Applied and Environmental Microbiology, 2009. **75**(24): p. 7789-7796.

335. Brutinel, E.D. and J.A. Gralnick, *Shuttling happens: Soluble flavin mediators of extracellular electron transfer in Shewanella.* Applied Microbiology And Biotechnology, 2012. **93**(1): p. 41-48.

336. Gralnick, J.A., et al., *Extracellular respiration of dimethyl sulfoxide by Shewanella oneidensis strain MR-1.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(12): p. 4669-4674.

337. Clarke, T.A., et al., *Structure of a bacterial cell surface decaheme electron conduit.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(23): p. 9384-9389.

338. Marritt, S.J., et al., *The roles of CymA in support of the respiratory flexibility of Shewanella oneidensis MR-1.* Biochem Soc Trans, 2012. **40**(6): p. 1217-1221.

339. Marritt, S.J., et al., *A functional description of CymA, an electron-transfer hub supporting anaerobic respiratory flexibility in Shewanella.* Biochem J, 2012. **444**(3): p. 465-74.

340. Simon, J., et al., *A NapC/NirT-type cytochrome c (NrfH) is the mediator between the quinone pool and the cytochrome c nitrite reductase of Wolinella succinogenes.* Molecular Microbiology, 2000. **35**(3): p. 686-696.

341. Aigrain, L., et al., *Dynamic control of electron transfers in diflavin reductases.* 2012. **13**(11): p. 15012-15041.

342. Murataliev, M.B., R. Feyereisen, and F.A. Walker, *Electron transfer by diflavin reductases.* Biochimica et Biophysica Acta - Proteins and Proteomics, 2004. **1698**(1): p. 1-26.

343. Rodionov, D.A., et al., *Comparative Genomics of the Vitamin B12 Metabolism and Regulation in Prokaryotes.* Journal Of Biological Chemistry, 2003. **278**(42): p. 41148-41159.

344. Moore, S.J., et al., *Elucidation of the anaerobic pathway for the corrin component of cobalamin (vitamin B12).* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(37): p. 14906-14911.

345. Spencer, J.B., et al., *The Escherichia coli cysG gene encodes the multifunctional protein, siroheme synthase.* FEBS Letters, 1993. **335**(1): p. 57-60.

346. Lobo, S.A.L., et al., *Functional characterization of the early steps of tetrapyrrole biosynthesis and modification in Desulfovibrio vulgaris Hildenborough.* Biochemical Journal, 2009. **420**(2): p. 317-325.

347. Lübbe, Y.J., et al., *Siro(haem)amide in Allochromatium vinosum and relevance of DsrL and DsrN, a homolog of cobyrinic acid a,c-diamide synthase, for sulphur oxidation.* FEMS Microbiology Letters, 2006. **261**(2): p. 194-202.

348. Suh, S.J. and J.C. Escalante-Semerena, *Cloning, sequencing and overexpression of cobA which encodes ATP:corrinoid adenosyltranferase in Salmonella typhimurium.* Gene, 1993. **129**(1): p. 93-97.

349. Escalante-Semerena, J.C., S.J. Suh, and J.R. Roth, *cobA Function is required for both de novo cobalamin biosynthesis and assimilation of exogenous corrinoids in Salmonella typhimurium.* Journal of Bacteriology, 1990. **172**(1): p. 273-280.

350. Blatch, G.L. and M. Lässle, *The tetratricopeptide repeat: A structural motif mediating protein-protein interactions.* BioEssays, 1999. **21**(11): p. 932-939.

351. Mancia, F., et al., *How coenzyme B12 radicals are generated: The crystal structure of methylmalonyl-coenzyme A mutase at 2 Å resolution.* Structure, 1996. **4**(3): p. 339-350.

352. Sharma, P., et al., *On the function of the various quinone species in Escherichia coli.* FEBS Journal, 2012. **279**(18): p. 3364-3373.

353. Wallace, B.J. and I.G. Young, *Role of quinones in electron transport to oxygen and nitrate in Escherichia coli. Studies with a ubiA- menA- double quinone mutant.* BBA - Bioenergetics, 1977. **461**(1): p. 84-100.

354. Meganathan, R., *Biosynthesis of menaquinone (vitamin K2) and ubiquinone (coenzyme Q): A perspective on enzymatic mechanisms*, in *Vitamins and Hormones*. 2001. p. 173-218.

355. Zinoni, F., et al., *Nucleotide-sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from Escherichia-coli.* Proceedings of the National Academy of Sciences of the United States of America, 1986. **83**(13): p. 4650-4654.

356. Chambers, I., et al., *The structure of the mouse glutathione-peroxidase gene - the selenocysteine in the active-sites is encoded by the termination codon, TGA.* Embo Journal, 1986. **5**(6): p. 1221-1227.

357. Yoshizawa, S. and A. Böck, *The many levels of control on bacterial selenoprotein synthesis.* 2009. **1790**(11): p. 1404-1414.

358. Stock, T. and M. Rother, *Selenoproteins in Archaea and Gram-positive bacteria.* 2009. **1790**(11): p. 1520-1532.

359. Prat, L., et al., *Carbon source-dependent expansion of the genetic code in bacteria.* Proceedings of the National Academy of Sciences, 2012. **109**(51): p. 21070-21075.

360. Srinivasan, G., C.M. James, and J.A. Krzycki, *Pyrrolysine encoded by UAG in archaea: Charging of a UAG-decoding specialized tRNA.* 2002. **296**(5572): p. 1459-1462.

361. Hao, B., et al., *A new UAG-encoded residue in the structure of a methanogen methyltransferase.* Science, 2002. **296**(5572): p. 1462-1466.

362. Herring, S., et al., *Recognition of pyrrolysine tRNA by the Desulfitobacterium hafniense pyrrolysyl-tRNA synthetase.* 2007. **35**(4): p. 1270-1278.

363. Gaston, M.A., et al., *The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine.* Nature, 2011. **471**(7340): p. 647-650.

364. Lycklama a Nijeholt, J.A. and A.J.M. Driessen, *The bacterial Sec-translocase: Structure and mechanism.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2012. **367**(1592): p. 1016-1028.

365. Van Wely, K.H.M., et al., *Translocation of proteins across the cell envelope of Gram-positive bacteria.* FEMS Microbiology Reviews, 2001. **25**(4): p. 437-454.

366. Arthur, M., et al., *The vanZ gene of Tn1546 from Enterococcus faecium BM4147 confers resistance to teicoplanin.* Gene, 1995. **154**(1): p. 87-92.

367. Wadhams, G.H. and J.P. Armitage, *Making sense of it all: Bacterial chemotaxis.* Nature Reviews Molecular Cell Biology, 2004. **5**(12): p. 1024-1037.

368. Maphosa, F., et al., *Ecogenomics of microbial communities in bioremediation of chlorinated contaminated sites.* Frontiers in Microbiology, 2012. **3**(OCT).

369. Fleischmann, R., *Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.* Science, 1995. **269**: p. 496 - 512.

370. Regeard, C., et al., *Indications for acquisition of reductive dehalogenase genes through horizontal gene transfer by Dehalococcoides ethenogenes strain 195.* Applied And Environmental Microbiology, 2005. **71**(6): p. 2955-2961.

371. Boyer, A., et al., *Purification, cloning and sequencing of an enzyme mediating the reductive dechlorination of 2,4,6-trichlorophenol from Desulfitobacterium frappieri PCP-1.* Biochemical Journal, 2003. **373**(1): p. 297-303.

372. Pöritz, M., et al., *Genome sequences of two dehalogenation specialists – Dehalococcoides mccartyi strains BTF08 and DCMB5 enriched from the highly polluted Bitterfeld region.* FEMS Microbiology Letters, 2013. **343**(2): p. 101-104.

373. Wang, S., et al., *Draft Genome Sequence of Polychlorinated Biphenyl-Dechlorinating Dehalococcoides mccartyi Strain SG1, Which Carries a Circular Putative Plasmid.* Genome Announcements, 2014. **2**(5).

374. Wang, S., et al., *Genomic characterization of three unique Dehalococcoides that respire on persistent polychlorinated biphenyls.* 2014. **111**(33): p. 12103-12108.

375. Mansfeldt, C.B., et al., *Meta-Analyses of Dehalococcoides mccartyi Strain 195 Transcriptomic Profiles Identify a Respiration Rate-Related Gene Expression Transition Point and Interoperon Recruitment of a Key Oxidoreductase Subunit.* Applied and Environmental Microbiology, 2014. **80**(19): p. 6062-6072.

376. Lee, P.K.H., et al., *Comparative genomics of two newly isolated Dehalococcoides strains and an enrichment using a genus microarray.* Isme Journal, 2011.

377. Hug, L.A., et al., *Design and verification of a pangenome microarray oligonucleotide probe set for Dehalococcoides spp.* 2011. **77**(15): p. 5361-5369.

378. Waller, A.S., et al., *Transcriptional analysis of a Dehalococcoides-containing microbial consortium reveals prophage activation.* Applied and Environmental Microbiology, 2012. **78**(4): p. 1178-1186.

379. Men, Y., et al., *Characterization of four TCE-dechlorinating microbial enrichments grown with different cobalamin stress and methanogenic conditions.* Applied Microbiology And Biotechnology, 2013. **97**(14): p. 6439-6450.

380. McMurdie, P.J., et al., *Site-Specific Mobilization of Vinyl Chloride Respiration Islands by a Mechanism Common in Dehalococcoides.* Bmc Genomics, 2011. **12**.

381. Lohner, S.T. and A.M. Spormann, *Identification of a reductive tetrachloroethene dehalogenase in Shewanella sediminis.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1616).

382. Rupakula, A., et al., *Functional genomics of corrinoid starvation in the organohalide-respiring bacterium Dehalobacter restrictus strain PER-K23.* Frontiers in Microbiology, 2015. **5**.

383. Borths, E.L., et al., *The structure of Escherichia coli BtuF and binding to its cognate ATP binding cassette transporter.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(26): p. 16642-16647.

384. Saunders, N.F.W., et al., *Transcription regulation of the nir gene cluster encoding nitrite reductase of Paracoccus denitrificans involves NNR and NirI, a novel type of membrane protein.* Molecular Microbiology, 1999. **34**(1): p. 24-36.

385. Cuypers, H., A. Viebrock-Sambale, and W.G. Zumft, *NosR, a membrane-bound regulatory component necessary for expression of nitrous oxide reductase in denitrifying Pseudomonas stutzeri.* Journal of Bacteriology, 1992. **174**(16): p. 5332-5339.

386. Wunsch, P. and W.G. Zumft, *Functional domains of NosR, a novel transmembrane iron-sulfur flavoprotein necessary for nitrous oxide respiration.* Journal of Bacteriology, 2005. **187**(6): p. 1992-2001.

387. Nojiri, H., et al., *Genetic System of Organohalide-Respiring Bacteria*, in *Biodegradative Bacteria*. 2014, Springer Japan. p. 59-81.

388. Prat, L., *Identification and Characterization of Proteins Supporting Dehalorespiration in Desulfitobacterium hafniense strain TCE1*, in *Laboratoire de biotechnologie environnementale*. 2009, école polytechnique fédérale de Lausanne. p. 151.

389. Horwich, A.L., et al., *Two families of chaperonin: Physiology and mechanism*, in *Annual Review of Cell and Developmental Biology*. 2007. p. 115-145.

390. Thirumalai, D. and G.H. Lorimer, *Chaperonin-mediated protein folding*, in *Annual Review of Biophysics and Biomolecular Structure*. 2001. p. 245-269.

391. Lund, P.A., *Multiple chaperonins in bacteria - Why so many?: Review article.* FEMS Microbiology Reviews, 2009. **33**(4): p. 785-800.

392. Coates, A.R.M., T.M. Shinnick, and R.J. Ellis, *Chaperonin nomenclature [3].* Molecular Microbiology, 1993. **8**(4): p. 787.

393. Gray, M.J., N.K. Tavares, and J.C. Escalante-Semerena, *The genome of Rhodobacter sphaeroides strain 2.4.1 encodes functional cobinamide salvaging systems of archaeal and bacterial origins.* Molecular Microbiology, 2008. **70**(4): p. 824-836.

394. de Crécy-Lagard, V., et al., *Comparative genomic analysis of the DUF71/COG2102 family predicts roles in diphthamide biosynthesis and B12 salvage.* Biol Direct, 2012. **7**.

395. Gray, M.J. and J.C. Escalante-Semerena, *The cobinamide amidohydrolase (cobyric acid-forming) CbiZ enzyme: A critical activity of the cobamide remodelling system of Rhodobacter sphaeroides.* Molecular Microbiology, 2009. **74**(5): p. 1198-1210.

396. Finn, R.D., et al., *Pfam: the protein families database.* Nucleic acids research, 2014. **42**(D1): p. D222-D230.

397. Villemur, R., et al., *Occurrence of several genes encoding putative reductive dehalogenases in Desulfitobacterium hafniense/frappieri and Dehalococcoides ethenogenes.* Canadian Journal of Microbiology, 2002. **48**(8): p. 697-706.

398. Christiansen, N., et al., *Purification and characterization of the 3-chloro-4-hydroxy-phenylacetate reductive dehalogenase of Desulfitobacterium hafniense.* 1998. **436**(2): p. 159-162.

399. Suzuki, D., et al., *Phototactic and chemotactic signal transduction by transmembrane receptors and transducers in microorganisms.* Sensors, 2010. **10**(4): p. 4010-4039.

400. Mascher, T., J.D. Helmann, and G. Unden, *Stimulus perception in bacterial signal-transducing histidine kinases.* Microbiology and Molecular Biology Reviews, 2006. **70**(4): p. 910-938.

401. Gao, R. and A.M. Stock, *Biological insights from structures of two-component proteins.* Annu Rev Microbiol, 2009. **63**: p. 133-54.

402. Claas, K.R., et al., *tFunctional analysis of the nicotinate mononucleotide: 5,6- Dimethylbenzimidazole phosphoribosyltransferase (CobT) enzyme, involved in the late steps of coenzyme B12 biosynthesis in Salmonella enterica.* 2010. **192**(1): p. 145-154.

403. Trzebiatowski, J.R., G.A. O'Toole, and J.C. Escalante-Semerena, *The cobT gene of Salmonella typhimurium encodes the NaMN: 5,6- dimethylbenzimidazole phosphoribosyltransferase responsible for the synthesis of N1-(5-phospho-α-D-ribosyl)-5,6-dimethylbenzimidazole, an intermediate in the synthesis of the nucleotide loop of cobalamin.* 1994. **176**(12): p. 3568-3575.

404. Krajmalnik-Brown, R., et al., *Genetic identification of a putative vinyl chloride reductase in Dehalococcoides sp strain BAV1.* Applied and Environmental Microbiology, 2004. **70**(10): p. 6347-6351.

405. Padilla-Crespo, E., et al., *Identification and environmental distribution of dcpA, which encodes the reductive dehalogenase catalyzing the dichloroelimination of 1,2-dichloropropane to propene in organohalide-respiring Chloroflexi.* Applied and Environmental Microbiology, 2014. **80**(3): p. 808-818.

406. Chow, W.L., et al., *Identification and transcriptional analysis of trans-DCE-producing reductive dehalogenases in Dehalococcoides species.* ISME Journal, 2010. **4**(8): p. 1020-1030.

407. Waller, A.S., et al., *Multiple reductive-dehalogenase-homologous genes are simultaneously transcribed during dechlorination by dehalococcoides-containing cultures.* Applied and Environmental Microbiology, 2005. **71**(12): p. 8257-8264.

408. Mukherjee, K., et al., *Dehalogenimonas lykanthroporepellens BL-DC-9T simultaneously transcribes many rdhA genes during organohalide respiration with 1,2-DCA, 1,2-DCP, and 1,2,3-TCP as electron acceptors.* FEMS Microbiology Letters, 2014. **354**(2): p. 111-118.

409. Islam, M.A., et al., *New insights into Dehalococcoides mccartyi metabolism from a reconstructed metabolic network-based systems-level analysis of D. mccartyi transcriptomes.* 2014. **9**(4).

410. Rowe, A., et al., *Relating mRNA and protein biomarker levels in a Dehalococcoides and Methanospirillum-containing community.* Applied Microbiology and Biotechnology, 2015. **99**(5): p. 2313-2327.

411. Johnson, D.R., et al., *Transcriptional expression of the tceA gene in a Dehalococcoides-containing microbial enrichment.* Applied And Environmental Microbiology, 2005. **71**(11): p. 7145-7151.

412. Lee, P.K.H., et al., *Reductive dehalogenase gene expression as a biomarker for physiological activity of Dehalococcoides spp.* Applied and Environmental Microbiology, 2006. **72**(9): p. 6161-6168.

413. Rahm, B.G. and R.E. Richardson, *Correlation of respiratory gene expression levels and pseudo-steady-state PCE respiration rates in Dehalococcoides ethenogenes.* Environmental Science and Technology, 2008. **42**(2): p. 416-421.

414. Rowe, A.R., et al., *Relating chloroethene respiration rates in Dehalococcoides to protein and mRNA biomarkers.* Environmental Science and Technology, 2012. **46**(17): p. 9388-9397.

415. West, K.A., et al., *Global gene expression of Dehalococcoides within a robust dynamic TCE-dechlorinating community under conditions of periodic substrate supply.* 2013. **110**(5): p. 1333-1341.

416. Lee, P.K.H., et al., *Quantifying genes and transcripts to assess the in situ physiology of "Dehalococcoides" spp. in a trichloroethene-contaminated groundwater site.* Applied and Environmental Microbiology, 2008. **74**(9): p. 2728-2739.

417. Bælum, J., et al., *A conceptual model linking functional gene expression and reductive dechlorination rates of chlorinated ethenes in clay rich groundwater sediment.* Water Research, 2013. **47**(7): p. 2467-2478.

418. Behrens, S., et al., *Monitoring abundance and expression of "Dehalococcoides" species chloroethene-reductive dehalogenases in a tetrachloroethene-dechlorinating flow column.* Applied and Environmental Microbiology, 2008. **74**(18): p. 5695-5703.

419. Kranzioch, I., S. Ganz, and A. Tiehm, *Chloroethene degradation and expression of Dehalococcoides dehalogenase genes in cultures originating from Yangtze sediments.* Environmental Science and Pollution Research, 2015. **22**(4): p. 3138-3148.

420. Futamata, H., et al., *Phylogenetic and transcriptional analyses of a tetrachloroethene-dechlorinating "Dehalococcoides" enrichment culture TUT2264 and its reductive-dehalogenase genes.* Microbes and Environments, 2009. **24**(4): p. 330-337.

421. Cheng, D. and J. He, *Isolation and characterization of "Dehalococcoides" sp. strain MB, which dechlorinates tetrachloroethene to trans-1,2-dichloroethene.* Applied and Environmental Microbiology, 2009. **75**(18): p. 5910-5918.

422. Amos, B.K., et al., *Oxygen effect on Dehalococcoides viability and biomarker quantification.* Environmental Science and Technology, 2008. **42**(15): p. 5718-5726.

423. Fletcher, K.E., et al., *Effects of elevated temperature on Dehalococcoides dechlorination performance and DNA and RNA biomarker abundance.* Environmental Science and Technology, 2011. **45**(2): p. 712-718.

424. John, M., et al., *Retentive Memory of Bacteria: Long-Term Regulation of Dehalorespiration in Sulfurospirillum multivorans.* Journal of Bacteriology, 2009. **191**(5): p. 1650-1655.

425. Wilkinson, S.P. and A. Grove, *Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins.* Current Issues in Molecular Biology, 2006. **8**(1): p. 51-62.

426. Körner, H., H.J. Sofia, and W.G. Zumft, *Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: Exploiting the metabolic spectrum by controlling alternative gene programs.* FEMS Microbiology Reviews, 2003. **27**(5): p. 559-592.

427. Kemp, L.R., et al., *The transcriptional regulator CprK detects chlorination by combining direct and indirect readout mechanisms.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2013. **368**(1616).

428. Grostern, A. and E.A. Edwards, *A 1,1,1-Trichloroethane-Degrading Anaerobic Mixed Microbial Culture Enhances Biotransformation of Mixtures of Chlorinated Ethenes and Ethanes.* Appl. Environ. Microbiol., 2006. **72**(12): p. 7849-7856.

429. Grein, F., et al., *Unifying concepts in anaerobic respiration: Insights from dissimilatory sulfur metabolism.* Biochimica et Biophysica Acta - Bioenergetics, 2013. **1827**(2): p. 145-160.

430. Simon, J. and M. Kern, *Quinone-reactive proteins devoid of haem b form widespread membrane-bound electron transport modules in bacterial respiration.* Biochemical Society Transactions, 2008. **36**(5): p. 1011-1016.

431. Louie, T.M., et al., *Purification, characterization and gene sequence analysis of a novel cytochrome c co-induced with reductive dechlorination activity in Desulfomonile tiedjei DCB-1.* Archives Of Microbiology, 1997. **168**(6): p. 520-527.

432. Matias, V.R.F. and T.J. Beveridge, *Native cell wall organization shown by cryo-electron microscopy confirms the existence of a periplasmic space in Staphylococcus aureus.* Journal of Bacteriology, 2006. **188**(3): p. 1011-1021.

433. Zuber, B., et al., *Granular Layer in the Periplasmic Space of Gram-Positive Bacteria and Fine Structures of Enterococcus gallinarum and Streptococcus gordonii Septa Revealed by Cryo-Electron Microscopy of Vitreous Sections.* Journal of Bacteriology, 2006. **188**(18): p. 6652-6660.

434. Reichmann, N.T. and A. Gründling, *Location, synthesis and function of glycolipids and polyglycerolphosphate lipoteichoic acid in Gram-positive bacteria of the phylum Firmicutes.* Fems Microbiology Letters, 2011. **319**(2): p. 97-105.

435. Hazra, A.B., et al., *Anaerobic biosynthesis of the lower ligand of vitamin B12.* Proceedings of the National Academy of Sciences, 2015.

436. Schilhabel, A., et al., *The Ether-Cleaving Methyltransferase System of the Strict Anaerobe Acetobacterium dehalogenans: Analysis and Expression of the Encoding Genes.* Journal of Bacteriology, 2009. **191**(2): p. 588-599.

437. Sanford, R.A., et al., *Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils.* Proceedings of the National Academy of Sciences, 2012. **109**(48): p. 19709-19714.

438. Carlson, H.K., et al., *Surface multiheme c-type cytochromes from Thermincola potens and implications for respiratory metal reduction by Gram-positive bacteria.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(5): p. 1702-1707.

Summary

Organohalides have been and are still used for a wide range of applications such as wood preservation, degreasing of metallic surfaces, as pesticides or lubricants, and many of these compounds have been proven as, or are suspected to be, toxic for humans and other animals. Many highly chlorinated compounds, such as tetrachloroethene (PCE), have proven to be recalcitrant to aerobic degradation, but are often degraded under anaerobic conditions. Bacteria capable of coupling the dehalogenation of organohalides to growth, a process known as organohalide respiration (OHR), have therefore received considerable interest due to their potential for bioremediation. Members of the genera *Dehalobacter* and *Desulfitobacterium* are known to comprise obligate or facultative organohalide respiring bacteria (OHRB), respectively. We used a suite of tools, including genome sequencing, comparative genomics, and proteomic and enzyme assays, to study members of these two genera. We obtained the full genome sequence of *Dehalobacter restrictus* PER-K23^T, the first genome from a pure culture of a member of the *Dehalobacter* genus. The genome of *Dehalobacter restrictus* PER-K23^T consists of 2.9 Mbp, encoding 25 predicted reductive dehalogenase (RdhA)-encoding genes, of which the majority are full length and were found to be located in two genomic regions. In addition, this genome was predicted to encode three membrane-bound NiFe uptake hydrogenases and a complete cobalamin synthesis pathway, which was, however, impaired by a 101-bp deletion in c*biH*. We identified one of the three putative uptake hydrogenases as the most likely candidate as the hydrogenase active under OHR with H_2 as electron donor and PCE as electron acceptor. The PCE reductase-encoding gene was highly expressed throughout all growth phases, as were genes encoding several other RdhA enzymes, albeit at a lower level. We confirmed that the *de novo* corrinoid synthesis pathway is non functional, identified the genes encoding corrinoid transporters, and were able to show that their expression was regulated by the availability of corrinoids in the medium. These findings are described in **Chapters 2-4.** A novel vancomycin-resistance gene cluster, *vanI*, was found to be encoded on the chromosome of all *Desulfitobacterium hafniense* strains and *D. chlororespirans* DSM11544^T, and its functional expression was used as selection marker to separate *Desulfitobacterium hafniense* TCE-1 from a *Sedimentibacter* contaminant. We used a combination of genome analysis, enzyme activity assays and proteomics to identify elements of the electron transport chain from formate to 3-chloro-4-hydroxyphenyl acetate (Cl-OHPA) in *Desulfitobacterium dehalogenans* JW/IU-DC1^T. Unexpectedly, we identified a flavoprotein to be highly increased in

abundance when Cl-OHPA was used as terminal electron acceptor, suggesting its involvement in OHR. We described the genomes of nine newly sequenced *Desulfitobacterium* spp., among which the first genome of *Desulfitobacterium dichloroeliminans* and *Desulfitobacterium metallireducens*. Whole genome comparisons was used to assign strains to species that had previously been described as *Desulfitobacterium* sp. The average genome sizes are 5.5, 4.3 and 3.4 Mbp for *D. hafniense, D. dehalogenans and D. dichloroeliminans/metallireducens*, respectively. We compared and discussed the desulfitobacterial genomes obtained as part of this thesis work with the three genomes published previously. These genomes were found to encode none to seven RdhA genes, whereas the genomes of both *D. hafniense* DP7 and *D. metallireducens* 853-15A^T did not encode any reductive dehalogenases. The latter result was a surprise as *D. metallireducens* 853-15A^T previously had been reported to utilize several chlorinated aromatic and aliphatic compounds as terminal electron acceptors. All currently available desulfitobacterial genomes encoded a full cobalamin synthesis pathway. We presented and discussed the function of a selected set of relevant genes in detail. These findings are described in **Chapter 5-7. Chapter 8** provides an overview of the current state of the art regarding our understanding of the metabolism of OHRB obtained either directly or indirectly from genome sequencing**,** including the (predicted) function of genes associated with RdhA enzymes and genes, a unifying nomenclature for these, and the mechanism of transcriptional regulation of OHR. Finally, the results described in this thesis were discussed with specific attention to the lessons learned from genome sequencing regarding the metabolic repertoire and our current knowledge of the architecture of respiratory electron transport chains in *Desulfitobacterium* spp. (**Chapter 9**), and an outlook was provided on the road ahead towards fully understanding and exploiting these dedicated degraders of organohalides.

Acknowledgements

It has been a long time since I left the Shire to start the daily walk up and down the Wageningse berg, now it is finally time to throw the ring into the fire pit. It has been a long interesting journey, where I have met people from practically every corner of the world.

First I should thank my promotors, Hauke and Willem. Hauke, you started as daily supervisor and was later promoted to promotor. Thanks for giving me the opportunity to do my Ph.D. in your group, it has been a very interesting experience, where I have learned a lot. You have a remarkable ability to always see the glass as half full instead of half empty. Also thanks to Willem for your many comments, especially during the last weeks of the thesis writing.

Also thanks to all external collaborators, without you, this thesis would have looked very different. Especially, Julien, Aamani and Ute deserves thanks, we have exchanged countless e-mails with suggestions and updated manuscripts over the years.

Moving directly from the modern secular Scandinavia to Wageningen, naturally meant that there was a lot of "do" and "do not" to learn. Over the years I learned from early mistakes, such as the time when I caused a scandal by asking people not to bring children to my housewarming party. But Wageningen has changed a lot over the years, today many things that would have been highly controversial when I arrived here are now considered as less controversial. It has been very interesting, to witness this transformation from the first row.

So my dear MOLECOs, MICPHYS', BACGENs and SSBs Alas, 9 years is far too short a time to live among such excellent and admirable scientists. I don't know half of you half as well as I should like, and I like less than half of you half as well as you deserve. I have met far too many people here, in the south, to write a lengthy epistle for all but naturally everyone have contributed to the Wageningen experience and deserve to be thanked for this.

First I should mention the first people I met here, the so-called old MOLECO group. I will for sure never forget how you received me and introduced me to the southern way of life, when I first arrived to both the laboratory and the Netherlands. Especially the legendary MOLECO Christmas dinners are a thing I will always remember. But of course also the next generations of MOLECOs have all added their distinct flavour to the Wageningen experience as office mates, collaborators or just being around.

Then I should mention the people working on organo halide respiration: *Hauke*, *Farai*, *Tian*, *Yue*, *Siavash*, *Ying*, *Peng Peng* and *Tom*. It was nice to have dehalo-meetings now and then to exchange ideas, although we all worked on independent projects. *Yue* you took over the proteomics work when I was stuck in front of the computer, and did quite well which resulted in chapter 4 of this thesis.

The people from the MICPHYS group should be thanked for letting me do all the anaerobic work in the MICPHYS lab. Special thanks to *Ton*, for excellent technical assistance with all types of equipment, and interesting discussions about beer brewing, and *Caroline*, for teaching me the secret of operating the anoxic glove box. *Michael* "fat boy" thanks for many interesting and at times absurd discussions about cats, football and the definition of buitenlander. *Ana Paolo*, I always enjoyed our many discussions on science, surreal jokes, and the cod fish.

Also thanks to *Wim* for always being very enthusiastic and willing to help, keeping the PC in shape, finding some extra RAM when I needed it for all the genome analysis and taking care of the one million details that no one sees before they are not taken care of.

This thanks does naturally also apply to the rest of the backroom staff at Microbiology.

Over the years many people from all over the world have passed through office 20 also known as the "lunatic asylum", Starting with the Dutch contingent; Thanks to: *Eline*; *Hermien,* thanks for inviting the entire office for dinner followed by an invitation to attend the church service of your local church, that is for sure an aspect of Dutch life I else never would have experienced. *Coline*; *Lennart*, although you later moved to another office we still had the chance to have some interesting conversations; *Dennis*, I hope you manage to finalise your introduction mail before handing in your Ph.D. thesis; *Ludmilla*, It was very interesting sharing office with someone that has experienced so much in life; *Farai*, Since you now have a Dutch passport, I added you to the Dutch contingent. Thanks for the numerous discussions on everything from science, to life in general, for your help when I moved to my current flat. Especially for helping with project "sofa up the stairs", and being my main source of information about the Netherlands when I first arrived here. Then it is time to thank the Asian contingent; *Hao*, it was a pleasure sharing office with you, we have had many interesting talks; *Yangling* I enjoyed our discussions on the definition of "civilised enough" and which is the most exotic dish Chicken feet or Boerenkool?; *Tian*, always busy organising this and that, I hope you soon will be busy organising the defence of your

thesis; *Indra*, the only office mate that actually knew when something important happened in the badminton world; also thanks to *Ratnesh* and *Peng Peng*. And last but not least thanks to the many representatives from the islamic world *Sahar*, It was interesting to learn something about life in Egypt, which I knew practically nothing about before. Also thanks to *Kadrii*, *Friend*, *husband of Friend*, *Rosalyn* and *Gülhan*.

Also thanks to my current office mates in office K9, *Mark*; fun to meet one from the old gang again, for sure nice to share office with you and I certainly enjoyed our conversations about both science and non science, and the talks about the good old days; *Mauricio;* another member of the old gang that I found hiding in the basement, hope to see you soon at your own Ph.D. defence, *Florian;* the German master of multitasking. And lastly *Peer* who managed to sneak into the office at the very end, you should seriously find another hobby.

Also thanks to the organising committee of the Ph.D. trip 2009, *Matthijs*, *John*, *Marcel* and *Faab*. It was a lot of work, but in the end we managed to organise a quite good trip, that went without any major disasters.

Thanks to *Teresita*, for all the BBQs, and dinners, especially for organising the yearly buitenlander Christmas dinner, *Teunke*, the only person in the lab who knew that Denmark are not the capital of Legoland, *Peer* and *Elleke*, it was a pleasure to be paranymph with both of you.

Happily there has been more to life in the south than just work.

One thing I really enjoyed over the years was organising and playing in our legendary and immensely successful Lab. Football team "FC Kaas", once we even managed to win the league at "level very weak". We have had many famous players during the years so thanks to: *Marcel, Mark L., and Mark M., Bart, Bram, Mauricio, Edze, Magnus, Matthijs, Jose, Abishek, Femke, Farai, Philippe, Naim, Katrin, Monica, Farrakh, Corina, Stefan* and all those that I forgot to add to this list, I hope you had as much fun as I did.

Thanks to all who joined the microbiology team for the veluwe loop, especially the organisers *Anne-Meint*, *Matthijs* and *Tom*. It was always, regardless of the weather, a great day ending with everyone falling asleep during the dinner. Also thanks to the enthusiastic runners who joined the 7 hill run, *Bram*, *Edze*, *Mauricio*, *Serve* and *Tom*, it was fun, both the run the many discussions, absurd bets, and psychological warfare before and after the run...

Another highlight has been the board game sessions especially the "RISK" games and the "Axis and allies'" battles, usually the diplomacy, negotiations subtle and less subtle treats started several weeks before the actual games. The games especially the "axis and allies" usually turned into marathon sessions lasting up to 17 hours. Still we never managed to play a single game to the end. Although everyone could and did convincingly argue that he was in a clear winning position at the time when we decided to stop the game..... So thanks *Edze*, *Detmer, Mauricio, Bram, Martin, Tom, Alex,* and *Mark* for some exhausting but very entertaining battles, good food and company.

Edze, *Stineke*, it has been a pleasure meeting you, and I always enjoyed the gezelligheid when visiting you, the nice BBQs and all the other nice activities, such as the expedition to the Dutch country side, we have shared over the years. Hopefully I will manage to visit you on the island one day.

Mauricio, *Detmer, Martin*, *Bram* and *Mark* thanks for many interesting discussions, nice dinners and enjoyable moments with or without a cold beer. *Martin* thanks for helping me out when my basement suddenly had been transformed to a swimming pool.

Also thanks to the neighbours in the microbiology colony in the centre of the metropolis, *Nikolas*, *Pierepaolo* and *Marjet*. I really enjoyed our dinners, BBQs and the always good company.

Thanks to *Juanan*, *Irene*, *Maria, Jose, Corina, Svend* and the "*Spanish flock*" for always creating a good atmosphere and all the nice parties, watching football and handball championships together was a particular highlight. And also for being the native guides on the expedition to the extreme south AKA Spain.

Having flatmates from different countries have given me an interesting insight into lifestyles and food cultures of various countries, in chronological order: *Mari*; The only Nordic flatmate, I was truly delighted to finally meet a person familiar with the concept of having a beer and a chat Friday after work; *Anke*; my only OHR flatmate; *Jose*, although you every week announced your plans to move back to Spain soon, you still take the prize as the flat mate that stayed for the longest time, 3.5 years if I remember right. Besides the lectures on Spanish cuisine, I also enjoyed our many discussions on everything from Football, to politics, philosophy and science, and we almost managed to stick to the no mentioning of the lab. after 22 rule. I hope that I one day will get the chance to visit you again in the extreme south; *Andrea*, the Spanish psychrophile and the

only person I ever met that are even cheaper than myself; *Angela*, The "finnish" Italian, you have a great sense of humour; and finally my current tenant *Monika*, the silent pole.

If I was to give all the people the credit that they deserve this thesis would be twice as thick. So if after reading this, you feel left out and believe that you deserve to be mentioned, then you are undoubtedly right. Therefore feel free to give this book a personal twist by writing you name on the line below.

Thanks to _____for your valuable contribution to the Wageningen experience.

Irene and *Bram*, thanks for accepting the task as paranymphs.

Irene, you came to Wageningen for a short stay, but soon decided that this is the place where you want to live for ever. I have enjoyed all the dinners we have had over the years. I am happy that the Danish cuisine obviously falls in the taste of people from the extreme south. I am also intrigued by the many tales about the existence of a fourth Spanish dish. Although, I as a scientist, must consider this with some scepticism until I see more convincing data. We have had many discussions on the all kinds of topics such as, what bacteria is the coolest? *D. boringensis* or *Desulfitobacterium* spp., but in truth we both know the answer. I am already looking forward to visit you in Bennekom, when you have reached the goal of your dreams...

Bram, a few years ago I had the honour of being your paranymph and I am very happy that you agreed to return the favour. Over the years we have had many nice dinners, interesting discussions on science, culture, history, literature and much more. We also went on joint "hunting expeditions" for unguarded kroketten together... We both played in FC Kaas, joined the veluwe loop, 7 hills run, although not exactly running at the same speed, and even went for a few concerts together. You also taught me the secret of protein purification, which I am sure, will be useful also in the future. I, and many more people, have greatly enjoyed the Wednesday movie sessions that you organised, while you still were at microbiology. Recently you went for a new quest going from academia to industry, I am sure you will succeed, and that there are more dinners and unguarded kroketten waiting for us to share in the future.

Tilsidst, skal der siges tak til de vigtigste, familien.

Selvom vi ikke har set hinanden så meget de sidste år, så har i bestemt ikke været glemt. Den årlige jule tur til til både København og Sønderjylland, var altid et højdepunkt som jeg så frem til. Der var altid masser af god mad og hvad der er endnu vigtigere godt selskab af dem der betyder mest. Selvom man bor i udlandet så vil rødderne jo altid være i Danmark. Jeres støtte og opmuntring gennem årene har betydet meget. Jeg er osse glad for at ihvertfald nogle af jer har haft mulighed for at komme på bessøg her i syden, og håber at se så mange som muligt når jeg skal forsvare min afhandling.

Thomas Kruse

The central image on the front page shows a modified depiction of the Trundholm sun chariot or "Solvognen" in Danish, a bronze age artefact discovered in a peat bog in the Trundholm moor, Denmark, dated to 1800 to 1600 BC. The original sculpture showed a horse pulling a chariot carrying the sun, representing the best scientific explanation for the movement of the sun, available at the time. The research presented in the present thesis represents the state of our current understanding of organohalide respiring *Firmicutes*. Like everything in the solar system revolves around the sun, the work presented here revolves around genome analysis, thus a depiction of the genome of an organohalide respiring *firmicutes* has been inserted on the sun disc, surrounded by figures from various chapters of this thesis.

About the author

Thomas Kruse was born 5th of January 1977 in Aabenraa, Sønderjylland, Denmark. He obtained his gymnasium diploma from Aabenraa statsskole in 1997. Followed by military service at the Kings artillery regiment From 1997 to 1998. Where after he started studying biology at the University of Copenhagen in 1998. August 2001 to August 2002 he studied polar biology at the university centre on Svalbard (UNIS), Norway. He received his B.Sc. in Biology in 2003. Inspired by the stay at UNIS he continued with master studies at the department of geochemistry at the geological survey of Denmark and Greenland (GEUS). In 2005 he obtained his M.Sc. thesis entitled "Microbial responses to jet fuel pollution and fertilizer stimulation of pristine and pre-exposed arctic soils"

He received two scholar ships during his study time 10.000 DKK from Nordplus and 8.000 DKK from J.L Schmidts foundation.

In 2006 he moved again this time to the south to do his Ph.D. on the topic "organohalide respiring bacteria" at the laboratory of microbiology at Wageningen University, under supervision by Hauke Smidt

Throughout his study time, Thomas Kruse has been actively involved in both academic and social life. In 2001 he was elected member of the staff student committee at the Biology department at the University of Copenhagen, and in 2002 as a suppliant. From 2008 to 2010 he was member of the Ph.D. council for the research school SENSE. He was member of the organising committee for the Ph.D. trip to the north east coast of the U.S.A. in 2009.

Besides that he was editor of the student yearbook at UNIS 2001-2002, one of the founders of the football club for GEUS staff "HFF" in 2005 and initiated the formation of the laboratory of microbiology football team "FC Kaas" in Wageningen in 2008.

List of publications

Published

Kruse T, Maillard J, Holliger C, Smidt H, et al. (2013) Complete genome sequence of *Dehalobacter restrictus* PER-K23. Standards in Genomic Sciences

Rupakula A*, **Kruse T***, Boeren S, Holliger C, Smidt H, Maillard J (2013) The restricted metabolism of the obligate organohalide respiring bacterium *Dehalobacter restrictus*: lessons from tiered functional genomics. Philosophical Transactions of the Royal Society B: Biological Sciences 368 (1616). doi:10.1098/rstb.2012.0325

Kruse T, Levisson M, de Vos WM, Smidt H (2014) *vanI*: a novel d-Ala-d-Lac vancomycin resistance gene cluster found in *Desulfitobacterium hafniense*. Microbiol Biotechnol:n/a-n/a. doi:10.1111/1751-7915.12139

Goris T, Hornung B, **Kruse T**, Reinhold A, Westermann M, Schaap P, Smidt H, Diekert G (2015) Draft genome sequence and characterization of *Desulfitobacterium hafniense* PCE-S. Standards in Genomic Sciences 10 (1):15

Kruse T, Van de Pas BA, Atteia A, Krab K, Hagen WR, Goodwin L, Chain P, Boeren S, Maphosa F, Schraa G, de Vos WM, Van der Oost J, Smidt H, Stams AJ (2015) Genomic, proteomic and biochemical analysis of the organohalide respiratory pathway in *Desulfitobacterium dehalogenans* J Bacteriol 195 (5):12. doi:10.1128/JB.02370-14

Rupakula A, Lu Y, **Kruse T**, Boeren S, Holliger C, Smidt H, Maillard J (2015) Functional genomics of corrinoid starvation in the organohalide-respiring bacterium *Dehalobacter restrictus* strain PER-K23. Front Microbiol 5. doi:10.3389/fmicb.2014.00761

* Contributed equally

Accepted for publication

Kruse T, Smidt H, Lechner U (2015) Comparative genomics and transcriptomics of organohalide respiring bacteria and regulation of *rdh* gene transcription. In: Adrian L, Löffler FE (eds) Organohalide Respiring Bacteria. Springer

In preparation

Kruse T et al., The *Desulfitobacterium* genome

Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment

D I P L O M A

For specialised PhD training

The Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment
(SENSE) declares that

Thomas Kruse

born on 5 January 1977 in Aabenraa, Denmark

has successfully fulfilled all requirements of the
Educational Programme of SENSE.

Wageningen, 30 November 2015

the Chairman of the SENSE board	the SENSE Director of Education
Prof. dr. Huub Rijnaarts	Dr. Ad van Dommelen

The SENSE Research School has been accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW)

KONINKLIJKE NEDERLANDSE
AKADEMIE VAN WETENSCHAPPEN

The SENSE Research School declares that **Mr Thomas Kruse** has successfully fulfilled all requirements of the Educational PhD Programme of SENSE with a work load of 42.2 EC, including the following activities:

SENSE PhD Courses

- Environmental Research in Context (2007)
- Research in Context Activity: Organising PhD trip, United States (2009)
- Principles of Ecological genomics (2009)

Other PhD and Advanced MSc Courses

- Techniques for Writing and Presenting Scientific Papers, Wageningen University (2008)

Management and Didactic Skills Training

- Supervision of MSc student with thesis entitled 'PCR recovery of integrons from environments contaminated with chlorinated compounds' (2008)
- Teaching in the MSc course 'Microbial ecology' (2008-2010, 2014)
- Teaching in the BSc course 'Microbial physiology' (2007)
- Teaching in the MSc course 'Advanced methods in Microbiology' (2007)
- Member of the PhD SENSE Council (2008-2009)

Selection of Oral Presentations

- *Integrons: not only antibiotic resistance.* Ecogenomics Annual Meeting, 16-17 April 2008, Amsterdam, The Netherlands
- *Horizontal Gene Transfer & Adaption to Halogenated Organics in the Environment.* Center of Marine Biotechnology, PhD trip, 22 *April* 2009, Baltimore, United States
- *Horizontal Gene Transfer & Adaption to Halogenated Organics in the Environment.* National Center for Biotechnology Information (NCBI), PhD trip, 24 April 2009, Bethesda, United States
- *Early lessons from sequencing the Desulfitobacterium genus.* Bacsin Annual Meeting, 24-25 May 2012, Amsterdam, The Netherlands

SENSE Coordinator PhD Education

Dr. ing. Monique Gulickx

Work described in this thesis was performed at the Laboratory of Microbiology, Wageningen University, The Netherlands and funded financially supported by the Ecogenomics and EcoLinc projects of the Netherlands Genomics Initiative, the BE-BASIC foundation through projects F07.001.05 and F08.004.01, as well as the European Community program FP7 (grants KBBE-211684; BACSIN, and KBBE-222625; METAEXPLORE).

Cover design: Thomas Kruse

Thesis layout: Thomas Kruse

Printed by: GVO drukkers & vormgevers, Ede, The Netherlands

Financial support from the Laboratory of Microbiology, Wageningen University, for printing of this thesis if gratefully acknowledged