**Halorhabdus tiamatea**: proteogenomics and glycosidase activity measurements identify the first cultivated euryarchaeon from a deep-sea anoxic brine lake as potential polysaccharide degrader

Johannes Werner,1,2† Manuel Ferrer,3† Gurvan Michel,4 Alexander J. Mann,1,2 Sixing Huang,1 Silvia Juarez,5 Sergio Ciordia,5 Juan P. Albar,3 María Alcaide,3 Violetta La Cono,6 Michail M. Yakimov,6 André Antunes,7 Marco Taborda,8 Milton S. da Costa,9 Tran Hai,10 Frank Oliver Glöckner,1,2 Olga V. Golyshina,10 Peter N. Golyshin10† and Hanno Teeling1‡: The MAMBA Consortium

1Max Planck Institute for Marine Microbiology, Bremen, Germany. 2Jacobs University Bremen gGmbH, Bremen, Germany. 3Institute of Catalysis, CSIC, Madrid, Spain. 4UPMC University Paris 6 and CNRS, UMR 7139 Marine Plants and Biomolecules, Station Biologique, Roscoff, Bretagne, France. 5Proteomic Facility, CNB-National Centre for Biotechnology, CSIC, Madrid, Spain. 6Laboratory of Marine Molecular Microbiology, Institute for Coastal Marine Environment (IAML), CNR, Messina, Italy. 7Institute for Biotechnology and Bioengineering (IBB), Centre of Biological Engineering, Micoteca da Universidade do Minho, University of Minho, Braga, Portugal. 8Microbiology Unit, BIOCANT Biotechnological Park, Cantanhede, Portugal. 9Department of Life Sciences, University of Coimbra, Coimbra, Portugal. 10School of Biological Sciences, Bangor University, Bangor, Gwynedd, UK.

Summary

**Euryarchaeota from the genus Halorhabdus** have been found in hypersaline habitats worldwide, yet are represented by only two isolates: *Halorhabdus utahensis* AX-2T from the shallow Great Salt Lake of Utah, and *Halorhabdus tiamatea* SARL4BT from the Shaban deep-sea hypersaline anoxic lake (DHAL) in the Red Sea. We sequenced the *H. tiamatea* genome to elucidate its niche adaptations. Among sequenced archaea, *H. tiamatea* features the highest number of glycoside hydrolases, the majority of which were expressed in proteome experiments. Annotations and glycosidase activity measurements suggested an adaptation towards recalcitrant algal and plant-derived hemicelluloses. Glycosidase activities were higher at 2% than at 0% or 5% oxygen, supporting a preference for low-oxygen conditions. Likewise, proteomics indicated quinone-mediated electron transport at 2% oxygen, but a notable stress response at 5% oxygen. *Halorhabdus tiamatea* furthermore encodes proteins characteristic for thermophiles and light-dependent enzymes (e.g. bacteriorhodopsin), suggesting that *H. tiamatea* evolution was mostly not governed by a cold, dark, anoxic deep-sea habitat. Using enrichment and metagenomics, we could demonstrate presence of similar glycoside hydrolase-rich *Halorhabdus* members in the Mediterranean DHAL Medee, which supports that *Halorhabdus* species can occupy a distinct niche as polysaccharide degraders in hypersaline environments.

Introduction

Hypersaline habitats are found worldwide, for example in the form of terrestrial and deep-sea brine lakes or man-made solar salterns. The salinities of these habitats range from just above seawater to salt saturation, and their salt compositions range from concentrated seawater with sodium chloride as major salt (thalassohaline habitats) to compositions where other salts such as magnesium chloride dominate (athalassohaline habitats). Despite harsh conditions, microorganisms inhabit hypersaline habitats with a spectrum from species that merely tolerate hypersalinity to true halophiles that require 0.5–2.5 M of salt for growth (Andrei et al., 2012).
Two major strategies have evolved to cope with high salinities and prevent enzymes from denaturing and salt-out precipitation (Galinski, 1995). The first, the organic osmolyte strategy, consists of countering high osmolarities by intracytoplasmic accumulation of compatible solutes like quaternary amines or sugars such as trehalose. The second, the salt-in strategy, relies on accumulation of high levels of internal potassium (and to lesser extents sodium) chloride.

Among the most peculiar hypersaline habitats are deep-sea brine lakes, like the Orca Basin in the Northern Gulf of Mexico (Pilcher and Blumstein, 2007), the ice-sealed Antarctic Vida lake (Murray et al., 2012), the numerous deep-sea hypersaline anoxic lakes (DHALs) in the Eastern Mediterranean Sea (Bortoluzzi et al., 2011) and the Red Sea (Antunes et al., 2011a). The thalassohaline DHAL Shaban Deep in the Red Sea was discovered in 1984 (Pautot et al., 1984), and since several novel species were isolated from this location (Antunes et al., 2003; 2007; 2008a,b). Halorhabdus tiamatea SARL4BT stems from the brine–sediment interface of the Shaban Deep’s Eastern basin (26°13.9′ N, 35°21.3′ E, −1447 m depth, pH 6.0, salinity: 244) and features pleomorphic, non-pigmented cells that grow chemooorganoheterotrophically under anoxic to micro-oxic conditions [optimum: 45°C; pH 5.6–7.0; 27% NaCl (w/v)], but poorly under oxic conditions (Antunes et al., 2008a). The only other Halorhabdus (Hrd.) (Oren et al., 2007) species with a validly published name so far is Halorhabdus utahensis AX-2T (DSM 12940T), a sediment isolate from the southern arm of the shallow thalassohaline Great Salt Lake in Utah, USA (Wainø et al., 2000). Halorhabdus utahensis also features pleomorphic but pigmented cells that can grow under anoxic and oxic conditions (Table 1). Both Hrd. species exhibit a 16S ribosomal RNA (rRNA) sequence identity of 99.3% (Fig. 1). The genome of

| Table 1. General characteristics of the H. tiamatea and H. utahensis genomes. |
|---------------------------------|-----------------|
|                                 | H. tiamatea      | H. utahensis   |
| Contigs                         | 1 chromosome, 1 plasmid | 1 chromosome |
| Chromosome size (G + C content) | 2 815 791 bp (63.4%) | 3 116 795 bp (62.9%) |
| Plasmid size (G + C content)    | 330 369 bp (57.4%) | – |
| Total genes (coding density)    | 3023 (83.2%)     | 2998 (86.2%)  |
| Genes with annotated functions rRNAs | 1974 (65.3%) | 2243 (74.8%) |
| tRNAs                           | 3 (one rRNA operon) | 3 (one rRNA operon) |
| rRNAs                           | 46 (all 20 amino acids) | 45 (all 20 amino acids) |

Fig. 1. Maximum likelihood tree of the family Halobacteriaceae. The tree was calculated with RAxML v. 7.0.3 (Stamatakis et al., 2005) with *Methanospirillum hungatei* JF-1 as outgroup. The scale bar represents 10% estimated sequence divergence.

© 2014 The Authors. Environmental Microbiology published by Society for Applied Microbiology and John Wiley & Sons Ltd, Environmental Microbiology, 16, 2525–2537
**H. utahensis** has been completely sequenced (Anderson et al., 2009), whereas until now only a draft sequence was available for **H. tiamatea** (Antunes et al., 2011b). Both genomes share a large proportion of genes, but also exhibit notable niche differentiations, such as an increased number of genes for membrane transport and utilization of maltose, maltodextrin, phosphonate, and di- and oligopeptides in **H. tiamatea** (Antunes et al., 2011b). **Halorhabdus tiamatea** and **H. utahensis** belong to those Halobacteriaceae species that can degrade plant polysaccharides (Anderson et al., 2011). For instance, **H. utahensis** has proven β-xylanase, β-xylosidase (Waïnø and Ingvorsen, 2003) and cellulase activities (Zhang et al., 2011).

We sequenced and closed the **H. tiamatea** genome *de novo*, annotated it manually and analysed its niche adaptations with an emphasis on polysaccharide degradation and response to oxygen. This included in-depth phylogeny-based annotations of its carbohydrate-active enzymes (CAZymes), corresponding glycosidase activity measurements and proteome analyses of **H. tiamatea** cultures grown in liquid media under 0%, 2% and 5% of oxygen-containing headspace.

**Results and discussion**

**Genome features**

The genome of **H. tiamatea** (Fig. 2A and B) consists of a 2.8 Mbp chromosome and a putative 330 kbp plasmid with 2743 and 280 predicted genes respectively (Table 1). The chromosome exhibits high overall collinearity with the **H. utahensis** chromosome, but differs by multiple larger inversions and minor genome rearrangements (Supporting Information Fig. S1). Conversely, the putative **H. tiamatea** plasmid exhibits no such collinearity, and mostly harbours hypothetical and conserved hypothetical genes as well as genes for transposases, DNA-associated proteins and restriction enzymes. The transposase density of the putative plasmid is 12.5% whereas it is only ~ 2% for the **H. tiamatea** chromosome. The chromosome furthermore features three regions with putative phage genes that are characterized by dissimilar tetranucleotide usage patterns (Fig. 2A). **Halorhabdus tiamatea**, like **H. utahensis**, contains one clustered regularly interspaced short palindromic repeats (CRISPR) element (**H. tiamatea**: 4703 bp with 71 spacers; **H. utahensis**: 3381 bp with 51 spacers).

**Monosaccharide utilization**

**Halorhabdus tiamatea** can degrade hexoses via the semi-phosphorylated Entner–Doudoroff (ED) pathway. Resulting D-glyceraldehyde-3-phosphate can be further oxidized to pyruvate via the lower part of the Embden–Meyerhof–Parnas (EMP) pathway. The upper part lacks 6-phosphofructokinase. Instead the genome encodes a 1-phosphofructokinase that is co-located with a fructose-1,6-bisphosphate aldolase gene. Such incompleteness or variations of the EMP and gluconeogenesis pathways are common in Archaea. Gluconeogenesis has been deemed non-operational in **H. utahensis** because of a lack of pyruvate phosphate dikinase (Anderson et al., 2011), which in **H. tiamatea** is also not present. **Halorhabdus tiamatea** has the potential to use fructose by conversion to fructose-1,6-bisphosphate (1-phosphofructokinase), and galactose via the Leloir pathway.

The pentose-5-phosphate (PP) pathway in **H. tiamatea** is missing the oxidative branch, but the non-oxidative branch is present. The latter is likely used to convert pentoses to fructose-6-phosphate. Without the PP pathway’s oxidative branch, glucose cannot be converted to ribulose-5-phosphate. However, both **Hrd.** genomes harbour genes to convert xylose and arabinose to ribulose-5-phosphate. **Halorhabdus tiamatea** has a ribokinase that **H. utahensis** lacks, which implies that **H. tiamatea** in contrast to **H. utahensis** can also utilize ribose.

**Polysaccharide utilization**

Both **Hrd.** species contain high numbers of CAZyme genes, i.e. genes for enzymes that synthesize, modify or breakdown glycosides (Henrissat and Coutinho, 2001). **Halorhabdus tiamatea** SARLB1 has in total 50 glycoside hydrolase genes (15.9 GHs Mbp−1), 42 on its chromosome and eight on its putative plasmid (Supporting Information Table S1) – the highest numbers so far observed in Archaea (Supporting Information Fig. S2). **Halorhabdus utahensis** has 44 GH genes (14.1 GHs Mbp−1) according to the CAZy database as of 10 December 2013 (Cantarel et al., 2009; Lombard et al., 2013).

**Halorhabdus tiamatea** features genes for the degradation of xylan, arabinan, arabinoxylan and galactan-containing hemicelluloses, pectin and likely cellulose. These polysaccharides occur in land plant cell walls, algae (Popper et al., 2011) and seagrass [e.g. pectin in Zostera marina (Zaporozhets, 2003; Khotimchenko et al., 2012)]. **Halorhabdus tiamatea** furthermore has the genomic potential to degrade exogenous storage carbohydrates such as sucrose or α-1,4-glucans [e.g. starch (Antunes et al., 2008a) or glyogen].

Xylans can be hydrolysed to xylose by concerted action of seven GH10 endo-β-1,4-xylanases and three GH43 β-xylosidases. A dedicated transporter can subsequently import the xylose monomers. Arabinans can be cleaved to L-arabinose by concerted action of a GH43 endo-α-1,5-L-arabinosidase and six GH51 exo-acting α-L-arabinofuranosidases. The latter can also remove
Fig. 2. Circular representation of the (A) chromosome and (B) plasmid of *H. tiamatea*. From inside to outside: GC content, GC skew, DNA curvature, DNA bending, deviation from the average tetranucleotide composition, CAZymes (blue: glycoside hydrolase, red: glycosyl transferase, green: carbohydrate esterase, orange: polysaccharide lyase, cyan: carbohydrate binding module), RNAs (red: rRNA, green: tRNA, orange: other RNA), genes in reverse direction and genes in forward direction. GC content and GC skew were calculated with a self-written PERL script (sliding windows: 5 kbp for chromosome; 0.5 kbp for plasmid). DNA curvature and bending were calculated with the program banana from the EMBoss package (Rice et al., 2000). TETRA (Teeling et al., 2004) was used for the calculation of the deviation from the average tetranucleotide composition (sliding windows: 5 kbp for chromosome; 1 kbp for plasmid).
Table 2. Glycoside hydrolases in the Hrd. genomes.

<table>
<thead>
<tr>
<th></th>
<th>H. tiamatea</th>
<th>H. utahensis</th>
</tr>
</thead>
<tbody>
<tr>
<td>GH2</td>
<td>4 (1.27), 3</td>
<td>4 (1.28)</td>
</tr>
<tr>
<td>GH3</td>
<td>6 (1.91), 4</td>
<td>7 (2.25)</td>
</tr>
<tr>
<td>GH4</td>
<td>1 (0.32), 1</td>
<td>2 (0.64)</td>
</tr>
<tr>
<td>GH5</td>
<td>2 (0.64)</td>
<td>7 (2.25)</td>
</tr>
<tr>
<td>GH9</td>
<td>1 (0.32)</td>
<td>1 (0.32)</td>
</tr>
<tr>
<td>GH10</td>
<td>7 (2.22), 4</td>
<td>4 (1.28)</td>
</tr>
<tr>
<td>GH11</td>
<td>0 (0.00)</td>
<td>2 (0.64)</td>
</tr>
<tr>
<td>GH13</td>
<td>7 (2.22), 6</td>
<td>1 (0.32)</td>
</tr>
<tr>
<td>GH31</td>
<td>1 (0.32), 1</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>GH32</td>
<td>2 (0.64)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>GH42</td>
<td>1 (0.32)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>GH43</td>
<td>3 (0.95), 2</td>
<td>4 (1.28)</td>
</tr>
<tr>
<td>GH51</td>
<td>6 (1.91), 1</td>
<td>1 (0.32)</td>
</tr>
<tr>
<td>GH67</td>
<td>1 (0.32), 1</td>
<td>1 (0.32)</td>
</tr>
<tr>
<td>GH77</td>
<td>1 (0.32), 1</td>
<td>1 (0.32)</td>
</tr>
<tr>
<td>GH88</td>
<td>1 (0.32)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>GH93</td>
<td>1 (0.32)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>GH94</td>
<td>3 (0.95)</td>
<td>2 (0.64)</td>
</tr>
<tr>
<td>GH95</td>
<td>1 (0.32)</td>
<td>1 (0.32)</td>
</tr>
<tr>
<td>GH97</td>
<td>1 (0.32), 1</td>
<td>1 (0.32)</td>
</tr>
</tbody>
</table>

GH abundances in the genomes of H. tiamatea and H. utahensis (according to the CAZy database as of 10 December 2013). The first number represents absolute counts, the number in parentheses counts per Mbp and numbers in boldface GHs detected in proteome data.

decorating L-arabinose side chains from arabinoxylans and arabinogalactans. Arabinoxylans are likely degraded by concerted action of GH10 xylanases and GH51 arabinosidas. The resulting L-arabinose is then likely taken up, isomerized to L-ribulose and subsequently funnelled into the PP pathway. Halorhabdus tiamatea lacks any obvious galactanase and thus probably cannot use the backbones of galactans and arabinogalactans. However, its genome codes for a GH4 α-galactosidase and a GH42 β-galactosidase, which likely enable H. tiamatea to cleave galactose side chains from hemicelluloses (Popper et al., 2011).

The H. tiamatea genome encodes a single PL1 family polysaccharide lyase with pectate lyase function, and a GH88 enzyme. The latter resembles the d-4,5-unsaturated β-glucuronyl hydrolase from Bacillus sp. GL1, which participates in the hydrolysis of unsaturated glycosaminoglycan oligosaccharides released by glycosaminoglycan lyases (Itoh et al., 2004). Similarly, the H. tiamatea GH88 likely cleaves unsaturated oligopeptins released by its PL.

Like H. utahensis, H. tiamatea has a modular GH9 with a C-terminal family 3 carbohydrate-binding module. This architecture resembles the endo-processive cellulase E4 from Thermomonospora fusca (Sakon et al., 1997). Two GH5 enzymes with possible glucanase functions could provide complementary cellulytic activity. Released oligoglucans may be further degraded by GH3 β-glucosidas. Alternatively, resulting cellobiose dimers might be processed by two GH94 cellobiose phos-
cytoplasmic heterotetrameric [Ni-Fe] hydrogenase, which agrees with the finding that *H. tiamatea* produces gas from sugars (Antunes et al., 2008a).

Sulphur stimulates growth of *H. utahensis* and is reduced to hydrogen sulphide. It has been suggested that this is facilitated fermentation rather than respiration that serves as hydrogen sink without producing energy (Wainé et al., 2000). Conversely, sulphur reduction has not been reported for *H. tiamatea* (Antunes et al., 2008a), and its genome does not seem to contain any respective genes. The bidirectional tetrameric hydrogenase (I) might be able to reduce sulphur as shown in *Pyrococcus furiosus* (Ma et al., 1993; 2000), but this has not been observed in *H. tiamatea*.

*Halorhabdus tiamatea* produces acids when grown on maltose (Antunes et al., 2008a) and is known to possess an L-lactate dehydrogenase (Antunes et al., 2011b) and an L-lactate permease, likely for lactate export. Besides, *H. tiamatea* also features a D-lactate dehydrogenase. Lactate fermentation seems to be the sole mechanism for recycling of reduced pyridine and flavin adenine dinucleotides under anoxic conditions. Acetate is a second likely fermentation product as the *H. tiamatea* genome encodes an AMP-forming acetyl-CoA synthetase, whose reverse reaction releases acetate from acetyl-CoA while conserving energy as ATP.

*Halorhabdus tiamatea* has all genes for anaerobic glycerol degradation: a glycerol kinase and an anaerobic glycerol-3-phosphate dehydrogenase (*glpABC*) (Rawls et al., 2011). In *Escherichia coli*, the anaerobic oxidation of glycerol-3-phosphate to dihydroxyacetone phosphate by GpABC is coupled to the reduction of fumarate to succinate (Schryvers and Weiner, 1981), but other halophilic archaea such as representatives of the genera *Haloferax* and *Haloarcula* have been shown to metabolize glycerol under anoxic conditions to D-lactate, acetate and pyruvate (Oren and Gurevich, 1994).

**Krebs cycle**

*Halorhabdus tiamatea* has a complete Krebs cycle (without glyoxylate shunt). The only anaplerotic reaction seems to be the carboxylation of phosphoenolpyruvate (PEP) by PEP carboxylase. *Halorhabdus tiamatea* features a malate:quinone-oxidoreductase that funnels electrons directly in the quinone pool and a ferredoxin-dependent 2-oxoglutarate oxidoreductase. Both enzymes might facilitate to run the Krebs cycle in reverse from oxaloacetate to the precursor 2-oxoglutarate, as in some methanogenic archaea (Sakai et al., 2011). The malate:quinone-oxidoreductase could operate reversely in terms of thermodynamics, as has been discussed for *Helicobacter pylori* (Kather et al., 2000). *Halorhabdus tiamatea* lacks a distinct fumarate reductase such as the membrane-bound type found in *E. coli* or the coenzyme M-reducing cytoplasmic type found in many methanogenic archaea; hence, a reverse Krebs cycle would involve the regular succinate dehydrogenase.

**Respiration**

*Halorhabdus tiamatea* grows under hypoxic, but poorly under oxic conditions (Antunes et al., 2008a). Its genome encodes all genes for the archaeal membrane-bound NADH:ubiquinone oxidoreductase, as well as cytochrome bd and bc ubiquinol oxidase subunits, a complete 3-subunit copper-containing cytochrome oxidase, together with cytochrome c biogenesis and copper transport genes, and a V-type ATPase.

It is known that *H. tiamatea* reduces nitrate and nitrite (Antunes et al., 2008a). However, its genome does not encode any membrane-associated (Nar) or periplasmic (Nap) respiratory nitrate reductase. It also lacks a membrane-bound Nrf-type cytochrome c nitrite reductase as it is typically found in anaerobes employing dissimilatory nitrate reduction to ammonium. *Halorhabdus tiamatea* does possess genes for a Nir-type cytoplasmic nitrite reductase, which, however, is non-respiratory as it typically acts only as an electron sink to cope with excess reductants under anoxic conditions. Nir activity is strictly anaerobic and if constitutively required might contribute to the oxygen sensitivity of *H. tiamatea*.

It is not clear whether respiration of endogenous fumarate produced by carboxylation of PEP to oxalacetate via a reversely operating Krebs cycle is functional. However, exogenous fumarate can likely be reduced to succinate as in other halophilic archaea (Oren, 1991).

**Response to oxygen**

Proteomics on extracts of *H. tiamatea* cultures grown under three oxygen conditions (0%, 2% and 5% in the culture headspace) revealed an increasing stress response with increasing oxygen exposure. Of 699 identified proteins (~24% of the cytosolic proteome), 435 were quantified using a label-dependent method, with at least two peptides in triplicate experiments (Supporting Information Table S2). Compared with the anoxic condition, expression ratios (O2/anoxic) varied from 0.36 to 5.22-fold at 2% oxygen, and from below 0.01 to 10.19-fold at 5% oxygen. The 2% oxygen condition caused a mild response with induction of six (including a catalase) and repression of two proteins with at least ± twofold changes.

The 5% oxygen condition caused a more pronounced shift in the protein expression pattern, with higher abundances of 32 (7.0%) and lower abundances of 38 (8.4%) proteins. Notably expressed were a catalase (7.5-fold), a superoxide dismutase (6.8-fold), a thiosulfate-
sulfurtransferase-like rhodanese (4.9-fold) likely for scavenging oxidative thyl-radicals (Remelli et al., 2012), a thioredoxin reductase (2.7-fold) probably with antioxidant functions and a chlorite dismutase (2.4-fold) countering oxidative hypochlorite. In contrast, expression of the chaperonine Hsp20 was more than 100-fold lower at 5% oxygen than at the anoxic condition.

The energy metabolism was down-regulated at the 5% oxygen condition, most notably subunits of the pyruvate : ferredoxin oxidoreductase (> 100-fold), which seemed to act as major regulation unit for controlling the intracellular carbon flux, but also enzymes from the semi-phosphorylated ED (2.0-fold) and lower EMP pathways (up to 2.3-fold), the Krebs cycle (up to 4.0-fold) and a putative hydrogenase component (3.6-fold). Likewise, most subunits of the NADH-quinone dehydrogenase were down regulated (up to 2.5-fold).

Almost half (24/50) of the glycoside hydrolases (3 × GH2, 4 × GH3, 1 × GH4, 4 × GH10, 6 × GH13, 1 × GH31, 2 × GH43, 1 × GH51, 1 × GH67 and 1 × GH77) as well as three polysaccharide deacetylases were identified, which stresses their relevance. The 5% oxygen condition led to a down-regulation of six GHs (2 × GH10, 3 × GH13 and 1 × GH31 ranging from 2.2 to 3.9-fold), including the trehalose synthase. Additional GHs seemed to be down-regulated, albeit with a lower than twofold change in expression ratio. Likewise, a subunit of the maltose transporter was down regulated (> 100-fold).

**Glucosidase activities**

We tested hydrolysis of 18 p-nitrophenol (pNP) glycosides with *H. tiamatea* protein extracts from the above-mentioned three oxygen conditions with and without 3 M KCl. Glucosidase activities were not detected for three pNP glycosides (α-maltose, α-xylose and β-arabinose) and positive for the remaining 15 (Fig. 3). The latter support that *H. tiamatea* can utilize α-glucans (starch, amylose and amylopectin), β-glucans (cellulose), β-xylans, arabinans and galactose-containing hemicelluloses, as suggested by the genome analysis. Activity was also positive for pNP glycosides of the typical plant saccharides α-fucose, α-rhamnose and α-mannose, as well as for β-lactose.

Total activity was highest for the 2% and lowest for the 5% oxygen condition, which agrees with the proteome experiments and supports an adaptation towards micro-oxic conditions (Fig. 3A). Individual activities varied with salt and oxygen concentrations, indicating that *H. tiamatea* actively regulates its sugar metabolism.

Alpha-arabinose activity was positive, and a GH51 α-L-arabinofuranosidase (located on the putative plasmid) was expressed in the proteome experiments. This conflicts with previous findings that *H. tiamatea* does not grow...
on arabinose (Antunes et al., 2008a). Such discrepancies among genomic potential, functional assays and growth experiments are common. For instance, it is known that sugar oligomers rather than monomers induce genes for polysaccharide degradation in Flavobacteria (Martens et al., 2011). Hence, growth experiments with monomers might constitute an artificial situation for polymer degraders under which they do not exhibit their normal in situ physiology.

Light-dependent enzymes

Some halobacteria use light-driven ion pumps: bacteriorhodopsin as a proton pump for chemiosmotic ATP generation and halorhodopsin for importing chloride against the concentration gradient. Halorhabdus tiamatea has two genes with rhodopsin domains, one bacteriorhodopsin-like gene and 10 genes containing the bacterio-opsin activator domain (including one located on the putative plasmid). Both rhodopsin domain proteins have the R-X-X-D proton acceptor and D-X-X-X-K retinal Schiff-base binding sites (Jiao et al., 2006). One of these genes likely has a sensory function, as it is co-located with a methyl-accepting chemotaxis signal transducer gene (see Supporting Information Text for motility). The second rhodopsin gene is co-located with genes encoding the bacteriorhodopsin-like protein, one of the bacterio-opsin activator domain proteins, two isoprenoid biosynthesis enzymes (likely for the retinal cofactor) and the exinuclease subunit UvrA. It is noteworthy that the genome encodes not only the complete UvABC DNA repair system, but also the blue light-dependent deoxyribodipyrimidine photolyase (Park et al., 1995), both of which repair ultraviolet-induced DNA damages.

Traits associated with a thermophilic lifestyle

As many other halophilic archaea, both Hrd. species share traits that are typically associated with a thermophilic lifestyle. Both feature a four-subunit pyruvate:ferredoxin oxidoreductase, which is usually found in hyperthermophilic anaerobes (Mai and Adams, 1996) and in methanogenic archaea, many of which are also moderately thermophilic. In addition, the known β-xylanases, β-xylosidase and cellulase of H. utahensis feature remarkably high temperature optima of 55/70°C, 65°C and 70°C (3 M NaCl) respectively (Waite and Ingvorsen, 2003; Zhang et al., 2011). Both species also feature surprisingly high optimum growth temperatures [H. utahensis: 50°C (Waite et al., 2000); H. tiamatea: 45°C (Antunes et al., 2008a)]. Likewise, both species harbour a gene for LysW, an enzyme that bypasses thermolabile diaminopimelate in lysine biosynthesis in Thermus thermophilus and hyperthermophilic Archaea (Horie et al., 2009).

Storage compounds and osmoregulation

Halorhabdus tiamatea is known to produce poly-β-hydroxy-alkanoates (PHA) (Antunes et al., 2008a). A class III PHA synthase is encoded by clustered phaE and phaC genes, and likely forms short-chained polymers with up to five carbon moieties in the hydroxyacyl backbone. Some bacteria are known to build a membrane complex from poly-β-hydroxy-butyrate, polyphosphate and calcium ions. This complex is believed to function as a calcium/phosphate symporter (Reusch and Sadoff, 1988) and if present in H. tiamatea might take part in osmoregulation.

Halorhabdus tiamatea also encodes a family 35 glycosyltransferase (GT) glycogen phosphorylase and a GH77 α-1,4-glucanotransferase. Both are essential for endogenous glycogen usage. However, H. tiamatea lacks known enzymes for de novo glycogen biosynthesis, i.e. a glycogen synthase (GT3 or GT5) or glycogen branching enzyme (GH13). This suggests that H. tiamatea either uses unknown iso-functional enzymes or employs a novel pathway for internal α-1,4-glucan biosynthesis. A possible candidate is a GH13 gene (HTIA_0925) with high similarity to the amylosucrase from Neisseria polysaccharea (Supporting Information Fig. S3). The latter can build an internal storage α-1,4 glucan by simultaneously adding glucose units of maltose and sucrose [maltose + sucrose + α−1,4 glucan\textsubscript{n} = fructose + α−1,4 glucan\textsubscript{n+3}] (Okada and Hehre, 1974; De Montalk et al., 1999). Such a storage polysaccharide would probably remain linear because H. tiamatea lacks any obvious glycogen branching and GH13 debranching enzyme (isomylase). Recycling of the α-1,4-glucan could be mediated by concerted action of a GT35 glycogen phosphorylase and a GH77 α-1,4-glucanotransferase via glucose-1-phosphate to glucose-6-phosphate. Almost all respective enzymes are co-located in a single cluster in the H. tiamatea genome, which supports that they act together.

Halorhabdus tiamatea also possesses a trehalose synthase-like GH13 (HTIA_0926). Trehalose synthase (TreS) converts maltose to the non-reducing disaccharide trehalose, which can serve as additional storage compound, but is mainly known for its function as compatible solute. Presence of trehalose in H. tiamatea has already been hypothesized before (Antunes et al., 2011b) and corroborates recent findings that Halobacteriales not only use the salt-in, but also the organic osmolyte strategy, e.g. by internal accumulation of trehalose and glycine-betaine (Youssef et al., 2013).
It is noteworthy that in *H. tiamatea* storage compounds and osmoregulation are possibly interconnected, similarly as in species that use the alternate TreY/TreZ trehalose biosynthesis pathway such as *Sulfolobus* spp. (Avonce *et al.*, 2006). In addition to trehalose, also maltose might play a dual role as precursor for both trehalose and a possible storage glucan. Thus, *H. tiamatea* could use storage glucan biosynthesis to regulate its pool of osmoregulating trehalose.

**Habitat-specific adaptations**

*Halorhabdus tiamatea* and *H. utahensis* exhibit some notable differences that are likely linked to their isolation sites. For example, the lack of pigmentmentation of *H. tiamatea* has been interpreted as adaptation to a light-deprived environment (Antunes *et al.*, 2008a). *Halorhabdus tiamatea* features mercuric ion and arsenate reductases, which mirrors the specific ion compositions in Red Sea anoxic brine pools (Antunes *et al.*, 2011b). *Halorhabdus tiamatea* also contains genes for the degradation of methylphosphonate that *H. utahensis* lacks (see Supporting Information Text). The Shaban Deep is a phosphate-limited environment (Antunes *et al.*, 2011b), which favours species with alternative phosphorus acquisition strategies.

*Halorhabdus tiamatea* has sugar transporters with annotated specificities (ribose, xylulose, maltose and maltodextrin), whereas *H. utahensis* has been reported not to contain any known sugar transporter (Anderson *et al.*, 2011). Likewise, *H. tiamatea* has the capacity to target a broader spectrum of polysaccharides than *H. utahensis*. Deep-sea habitats typically receive only little land plant, algal and seagrass carbohydrate substrates. Detritus sinking in from the photic zone will be largely consumed when it reaches the deep sea, and mostly the more recalcitrant components such as xylan and arabinan hemicelluloses, cellulose and connecting pectin will prevail. In case of DHALs, such compounds accumulate at the boundary layer of the seawater and the denser brine, and only little ultimately reaches the sediment. Hence, *H. tiamatea* is likely carbon limited in the Shaban Deep. In this context, the ability to store PHA and possibly α-1,4-glucans might be crucial for its survival. *Halorhabdus utahensis* in contrast stems from the sediment of the shallow Great Salt Lake of Utah. This lake has influxes from the Bear, Webber and Jordan Rivers, and thus features higher availabilities of suitable substrates.

One distinction of *H. tiamatea* from *H. utahensis* is its higher GH13 gene number (Table 2). *Halorhabdus tiamatea* has seven such genes (six of which were expressed) with functions that indicate (i) the degradation of exogenous α-glucans, (ii) the possible synthesis/turnover of a storage α-glucan and (iii) synthesis/turnover of trehalose. We found similarly high GH13 gene frequencies in *Hrd.* species that we enriched from the brine of another DHAL, the Mediterranean thalassohaline DHAL Medee (SW off the Western coast of Crete). Shotgun metagenomics of the enrichment (termed ANR26) yielded a 184 kbp contig carrying a 16S rRNA gene with 99.7% sequence identity to *H. tiamatea* (Fig. 1 and Supporting Information Text). Metagenome analyses demonstrated that the enrichment contained 80–95% *Hrd.* species and that its CAZyme profile correlated much better to *H. tiamatea* than to *H. utahensis* (Supporting Information Table S4 and Supporting Information Text).

**Concluding remarks**

Hypersaline habitats in general and DHALs in particular seem geographically isolated. Nonetheless, strains of *Haloquadratum walsbyi* with highly similar genomes have been isolated from salterns in Spain and Australia, indicating mechanisms for exchange (Dyall-Smith *et al.*, 2011). Such exchange would also explain the highly similar genomes of *H. tiamatea* and *H. utahensis*. It is therefore possible that *H. tiamatea* has been transported to the Shaban Deep, but does not belong to its autochthonous microbial community. While *H. tiamatea* has a gene repertoire that is distinct from that of *H. utahensis* and allows survival in DHALs, its preference for micro-oxic conditions, its light-dependent enzymes as well as its relatively high optimum growth temperature are much more in line with life in the upper sediment of terrestrial shallow warm hypersaline lakes. Likewise, specialization of *H. tiamatea* on plant-derived polysaccharides contradicts preference for DHALs because these habitats receive only little such material. This would also explain why we found only barely detectable abundances of *Hrd.* species in Medee brine samples by catalysed reporter deposition fluorescence *in situ* hybridization with a novel *Hrd.*-specific probe (Supporting Information Fig. S4 and Supporting Information Text).

Genomic and functional data from the as yet investigated *Hrd.* members point towards an ecological niche that involves the capability to degrade complex polysaccharides. Recent identifications of *Hrd.* members in various other hypersaline habitats (see Supporting Information Text) might even suggest that *Hrd.* species occupy such a niche in many hypersaline habitats worldwide.

**Experimental procedures**

**Sampling and sequencing of H. tiamatea**

DNA was extracted from a *H. tiamatea* pure culture using the modified phenol-chloroform protocol (Urakawa *et al.*, 2010) and subsequently sequenced on a 454 FLX Ti sequencer.
Blosum62 matrix. Phylogenetic trees were computed from MAFFT (Katoh, 1998) and CONSED (Gordon, 2003) were used for full assembly. Repeat mis-assemblies were corrected with DupFinder (Han and Chain, 2006), and a single scaffold was generated and verified using 384 Sanger end-sequenced fosmids. Direct adjacent sequencing reactions were necessary for full closure (Malysky et al., 2004). Illumina reads (GA II; PE library, ~160 mio reads) were used to correct 454 base calling base errors. This provided 197× coverage of the chromosome and 311× coverage of the plasmid, with an error rate of less than 1:100 000. Sequences have been deposited at the European Nucleotide Archive with the accession numbers HF571520-HF571521.

Gene prediction and annotation

Gene prediction and annotation of protein-coding genes was done as described elsewhere (Mann et al., 2013). Ribosomal RNA genes were identified via BLAST searches (Altschul et al., 1997) against public nucleotide databases and transfer RNA genes using tRNAscan-SE (1.21) (Lowe and Eddy, 1997). CRISPRs were identified with CRISPRFinder (Grisa et al., 2007). Selected CAZymes from multi-functional CAZyme families were subjected to in-depth phylogenetic analyses to uncover their substrate specificities. For each family, a set of experimentally characterized proteins was selected and aligned with their homologues using MAFFT (Katoh et al., 2002) with iterative refinement and the Blosum62 matrix. Phylogenetic trees were computed from these alignments using PhyML (Guindon and Gascuel, 2003) with 100 replicates for bootstrapping and annotations derived based on proximities to experimentally characterized proteins.

Glucosidase activity measurements and proteomics

Growth conditions. Halorhabdus tiamatea cultures were grown as described previously (Antunes et al., 2008a) in HBM liquid medium (Halobacteria medium; DSMZ (German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany) medium 372) in 200 ml glass vials until an optical density (600 nm) of 0.1. For anoxic conditions, vials were filled with medium and subsequently incubated in sealed cylinders with an anoxic gas phase (80% N2/20% CO2) and anaerobic container system sachets (MBraun UNILab, Garching, Germany). Foroxic conditions, the vials were filled to one quarter with medium and incubated in sealed cylinders with a gas phase with 2% and 5% oxygen respectively. Peptide solutions were obtained from all three conditions as described elsewhere (Yakimov et al., 2011). All subsequent measurements were carried out in triplicates.

Glucosidase activity measurements. Protein extracts (~0.23 mg ml⁻¹) were stored at −80°C until use. Activity was measured in 96-well plates using 18 pNP sugars (Sigma Chemical, St. Louis, MO, USA) and a BioTek Synergy HT spectrophotometer (BioTek, Highland Park, VT, USA) as described previously (Alcaide et al., 2012; Hernández et al., 2013). Reactions contained 4 µg total protein and 1 mg ml⁻¹ substrate in a 20-mM HEPES buffer (pH 7, T = 30°C, final volume: 150 µl) with or without 3 M KCl. The release of pNP was measured at 410 nm in 1 min intervals for 130 min with derivatives of α-D-glucose, β-D-glucose, α-D-maltose, α-D-maltotriose, α-D-maltohexose, β-D-celllobiose, α-L-galactose, β-D-galactose, α-D-xylose, β-D-xylose, α-L-arabinopyranose, β-L-arabinopyranose, α-L-arabinofuranose, α-L-rhamnose, α-D-mannose, β-D-lactose, α-L-fucose and β-D-acetylgaluronide.

Protein digestion and tagging with iTRAQ-4-plex®. Peptide solutions (50 µg) were labelled for 2 h at room temperature with half a unit of iTRAQ Reagent Multi-plex kit (AB SCIEX, Foster City, CA, USA), previously reconstituted with 70 µl of ethanol. In the iTRAQ labelling, tags 114, 115 and 116 were used for 0%, 2% and 5% oxygen conditions respectively. Afterwards, labelled samples containing the same protein content were combined, and the labelling reaction was stopped by evaporation in a Speed Vac (Eppendorf, Madrid, Spain).

Peptide fractionation at basic pH and mass spectrometry analysis. The digested, labelled and pooled samples were studied by RP-LC-MArDi TOF/TOF MS as described elsewhere (Yakimov et al., 2011), using 150 µg of digested and labelled peptides and a Fortis C18 column, 100 mm × 2.1 mm, 5 µm (Fortis Technologies, Marl, Germany). In brief, a MALDI TOF/TOF 4800 mass spectrometer (AB SCIEX) was used for acquisition and processing of the peptides. The resulting raw peak lists of precursors and fragment ions were filtered and exported with ABI-Extractor (PeakS-informatics Solutions, Waterloo, ON, Canada). Protein identification and quantitation were done with MASCOT v. 2.3.01 (Matrix Science, London, UK) and PHENIX v. 2.6 (GeneBio, Geneva, Switzerland). The search was performed against the predicted protein sequences of H. tiamatea. The concatenated target-decoy database search strategy was used to estimate the false positive rate (below 1%) to improve reliability of the data (P-value <0.01). A minimum of two unique peptides and three sets of spectra obtained in triplicated RP-LC-MArDi TOF/TOF MS data was required for protein identification and further quantification.

Acknowledgments

We thank the crew of the RV Urania for sampling of the Medee DHAL, Jörg Wulf for catalysed reporter deposition fluorescence in situ hybridization analyses, Rafael Bargiela for helping with figures and Rudolf Amann for critical reading of the manuscript. This study was supported by the EU FP7 project MAMBA (‘Marine Metagenomics for New Biotechnological Applications’, FP7-KBBE-2008–226977; http://mamba.bangor.ac.uk/), the Spanish Ministry of Economy and Competitiveness (grant BIO2011–25012) and the Max Planck Society. H.T., O.V.G. and P.N.G. acknowledge the support of EU FP7 for the project MicroB3 (OCEAN-2011-287589).

References


Kitaoka, M., Matsuoka, Y., Mori, K., Nishimoto, M., and Hayashi, K. (2012) Characterization of a bacterial...


Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

Supporting Information Text. Transporters; motility and chemotaxis; phosphonate utilization; fermentations; biotechnological aspects; Hrd. species in Medee; biogeographical aspects.

Fig. S1. Whole genome alignment between Halorhabdus utahensis and H. tiamatea. The genome of H. tiamatea was split into 50 bp fragments and subsequently mapped on the H. utahensis with SSAHA2 v. 2.5 (Ning et al., 2001).

Fig. S2. Numbers of GHs in archaeal genomes according to the CAZyme database (Cantarel et al., 2009). From top to bottom: Halorhabdus tiamatea SARL4β, Haloterrigena turkmenica 4Kα, Halorhabdus utahensis AX-2, Halopiger xanaduensis SH-6, Caldivinga maquiltingensis IC-167, Igniophaga aggregans AQ1.S1, Sulfolobus islandicus REY15A, Sulfolobus solfataricus P2, Sulfolobus islandicus Y.N.15.51, Sulfolobus islandicus M.16.27.

Fig. S3. Unrooted phylogenetic tree of characterized GH13 family proteins and GH13 family proteins of H. tiamatea (see Supporting Information Table S3 for details). Sequences were aligned with MAFFT (Katoh et al., 2002), and the tree was computed with the program PhyML (Guindon and Gascuel, 2003). Numbers indicate the bootstrap values (100 replicates). Proteins from H. tiamatea are marked by black diamonds.

Fig. S4. Identification of Hrd. species in a sample from the Eastern Mediterranean DHAL Medee, A: DAPI staining; B: hybridization with the probe Halo178 and helper probes Halo178-h1 and Halo178-h2, showing typical pleomorphic cells of Hrd. species. Images were post-processed with AUTOQUANT X (A and B; MediaCybernetics, Inc., Rockville, MD, USA) and IMARIS 7.4.0 (only B; Bitplane AG, Zurich, Switzerland).

Table S1. Annotations of all 50 glycoside hydrolases in H. tiamatea, including gene identifiers (locus tags), closest characterized homologues and EC numbers.

Table S2. Differentially expressed proteins of H. tiamatea grown on HBM liquid medium with 2% and 5% oxygen in the headspace compared with 0% oxygen. Protein annotations and their associated biological processes are shown. A shows 435 proteins unambiguously identified with at least two peptides. B shows 183 proteins identified with on one peptide. Protein identification and quantitation was conducted with MASCOT v. 2.3.01 (A contains protein and B peptide scores). Log2 ratios of the relative protein/peptide abundances at 2% versus 0% oxygen and 5% versus 0% oxygen are shown.

Table S3. NCBI GenPept labels and accession numbers of the GH13 proteins that were used in the phylogenetic analysis.

Table S4. Glycoside hydrolases in the genomes of H. utahensis (according to the CAZy database as of 10 December 2013), H. tiamatea and the ANR26 enrichment from the Eastern Mediterranean DHAL Medee. The first number represents absolute counts, the number in parentheses the relative number per Mbp and the number in boldface reflects GHs detected in the proteome.