environmental microbiology

Environmental Microbiology (2015) **17**(2), 332–345 doi:10.1111/1462-2920.12660

Pressure adaptation is linked to thermal adaptation in salt-saturated marine habitats

María Alcaide,[1‡] Peter J. Stogios,[2‡] Álvaro Lafraya,[1†] Anatoli Tchigvintsev,[2] Robert Flick,[2] Rafael Bargiela,[1] Tatyana N. Chernikova,[3] Oleg N. Reva,[4] Tran Hai,[3] Christian C. Leggewie,[5] Nadine Katzke,[6] Violetta La Cono,[7] Ruth Matesanz,[8] Mohamed Jebbar,[9] Karl-Erich Jaeger,[6] Michail M. Yakimov,[7] Alexander F. Yakunin,[2] Peter N. Golyshin,[3] Olga V. Golyshina,[3*] Alexei Savchenko,[2**] Manuel Ferrer[1*] and The MAMBA Consortium**

[1]*Institute of Catalysis, Consejo Superior de Investigaciones Científicas (CSIC), Madrid 28049, Spain.*

[2]*Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON M5S 3E5, Canada.*

[3]*School of Biological Sciences, University of Bangor, Gwynedd LL57 2UW, UK.*

[4]*Department of Biochemistry, University of Pretoria, Pretoria, South Africa.*

[5]*Evocatal GmbH, Monheim am Rhein 40789, Germany.*

[6]*Institute of Molecular Enzyme Technology Heinrich-Heine-University Düsseldorf and Institute of Bio- and Geosciences IBG-1: Biotechnology Forschungszentrum Jülich GmbH, Jülich D-52426, Germany.*

[7]*Institute for Coastal Marine Environment, CNR, Messina 98122, Italy.*

[8]*Centro Investigaciones Biológicas, CSIC, Madrid 28040, Spain.*

[9]*Université de Bretagne Occidentale, Laboratoire de Microbiologie des Environnements Extrêmes-UMR 6197 (CNRS-Ifremer-UBO), Institut Universitaire Européen de la Mer, Plouzané, France.*

Received 7 August, 2014; revised 1 October, 2014; accepted 2 October, 2014. For correspondence. *E-mail mferrer@icp.csic.es; Tel. +34915854872; Fax +34915854760. **E-mail alexei.savchenko@utoronto.ca; Tel. 416 978 3925; Fax 416 978 8605. ***E-mail: o.golyshina@bangor.ac.uk; Tel. 4401248383629; Fax 4401248382569. †Present address: Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (LISBP), Institut National des Sciences Appliquées, 135 Avenue de Rangueil, F-31077 Toulouse Cedex 04, France. ‡The authors have contributed equally to this work.

Summary

The present study provides a deeper view of protein functionality as a function of temperature, salt and pressure in deep-sea habitats. A set of eight different enzymes from five distinct deep-sea (3040–4908 m depth), moderately warm (14.0–16.5°C) biotopes, characterized by a wide range of salinities (39–348 practical salinity units), were investigated for this purpose. An enzyme from a 'superficial' marine hydrothermal habitat (65°C) was isolated and characterized for comparative purposes. We report here the first experimental evidence suggesting that in salt-saturated deep-sea habitats, the adaptation to high pressure is linked to high thermal resistance (*P* value = 0.0036). Salinity might therefore increase the temperature window for enzyme activity, and possibly microbial growth, in deep-sea habitats. As an example, Lake *Medee*, the largest hypersaline deep-sea anoxic lake of the Eastern Mediterranean Sea, where the water temperature is never higher than 16°C, was shown to contain halopiezophilic-like enzymes that are most active at 70°C and with denaturing temperatures of 71.4°C. The determination of the crystal structures of five proteins revealed unknown molecular mechanisms involved in protein adaptation to poly-extremes as well as distinct active site architectures and substrate preferences relative to other structurally characterized enzymes.

Introduction

The deep oceanic/sea regions (below 200 m depth) form the largest marine subsystem by volume and comprise 1.3×10^{18} m^{-3} or approximately 80% of the oceanic/sea volume (De Corte *et al.*, 2012). However, the bathy (1000–4000 m depth), abyssopelagic (4000–6000 m depth) and hadopelagic (below 6000 m depth) regions are by far the least explored systems on Earth, although they are the largest reservoirs of organic carbon in the biosphere and also home to largely enigmatic food webs (Nagata *et al.*, 2010). The habitability window in deep realms is shaped mostly as a function of salinity, pressure and temperature, and the individual and collective effects of these characteristics on life have been extensively

investigated (Harrison *et al*., 2013). Microorganisms are able to grow in a wide range of salt concentrations ranging from sea water (De Corte *et al*., 2012) to salt-saturated lakes (Daffonchio *et al*., 2006; Smedile *et al*., 2013; Yakimov *et al*., 2013), pressures up to 120 MPa (e.g. Zeng *et al*., 2009) and temperatures from 4°C (De Corte *et al*., 2012) to ~ 60–108°C in active chimney walls (Zeng *et al*., 2009; Eloe *et al*., 2011; Wang *et al*., 2011). However, attempts to define the collective influence of these environmental conditions on protein function are scarce, and in particular, the extent of poly-extremes remains mostly undefined. Poly-extremes are of particular importance as life adaptation is a focus of intense research interest, much of which is centred on dissecting the changes in the composition and genomic content of the communities under different environmental constraints (Daffonchio *et al*., 2006; Schlitzer, 2010; Eloe *et al*., 2011; La Cono *et al*., 2011; Smedile *et al*., 2013; Yakimov *et al*., 2013). One limitation of taxonomic and genomic data based on the analysis of total extracted DNA is that these studies cannot address whether an organism is alive or has succumbed to such multiple extreme conditions. Another major limitation of using small-subunit ribosomal gene surveys and shotgun data is the large number of organisms and genes that are anonymous; the majority have not yet been cultured or sequenced (Puspita *et al*., 2012; Akondi and Lakshmi, 2013). However, these data can be complemented with naïve screens, which directly analyse the enzymes of metagenomes (Martínez-Martínez *et al*., 2013 and references therein).

The importance of defining the combined effects of salt, pressure and temperature extremes on protein function contrasts with the limited information available in the literature. Therefore, as shown in Table S1, the majority of enzymes characterized thus far from deep-sea realms were mostly isolated either from slightly saline, cold (4°C) habitats or from slightly saline, high-temperature hydrothermal vents, and only one study described enzymes from salt-saturated basins (Ferrer *et al*., 2005). In addition, only six deep-sea proteins, all from single cultivated organisms (De Vos *et al*., 2007; Shirai *et al*., 2008; Xu *et al*., 2008; Shin *et al*., 2009; Sineva and Davydov, 2010; Pietra, 2012), have been structurally characterized; none of which originated from sites experiencing the three stressors (Table S1). This dearth of three-dimensional structures of proteins from deep-sea inhabitants, particularly from salt-saturated biotopes, precludes a thorough understanding of the structural adaptations necessary for life in poly-extremophilic environments and stifles the discovery and optimization of useful enzymes for structural-functional, engineering and industrial purposes. Nevertheless, it has been recently shown that DNA-based methodologies appear to be inaccurate approaches by which to study the 'adaptation signatures' in the brines of deep-sea salt-saturated lakes (Hallsworth *et al*., 2007; Yakimov *et al*., 2013). Rather, biochemical-based methodologies might be more accurate approaches by which to study such 'signatures', as enzyme activities represent the highest level of the functional hierarchy regardless of the heterogeneities that commonly appear at the DNA and amino acid levels.

This investigation takes a step beyond descriptive studies of microbial cultivation and gene repositories and the utilization of naïve screens, biochemical tests and structure determinations provided deeper insights into the combined effects of salinity, high pressure and temperature on marine enzymes. We used the bio-resources from two collaborative projects, BEEM (http://www.beem.utoronto.ca; funded by Genome Canada) and MAMBA (http://mamba.bangor.ac.uk; funded by the EU FP7 program), which are focused on investigating moderate-temperature deep-sea sites in the Mediterranean Sea. We particularly focused here on the salt-saturated deep-sea lakes *Medee*, *Bannock* and *Kryos* (Daffonchio *et al*., 2006; Yakimov *et al*., 2013) and the hadopelagic seawater column at Station Matapan-Vavilov Deep (Smedile *et al*., 2013) in the Eastern Mediterranean Sea (Fig. 1). A 'superficial' marine hydrothermal habitat was used for comparison. The extensive characterization of a set of nine different enzymes isolated from those sites by naïve screens and the determination of five crystal structures provided, to the best of our knowledge, the first experimental evidence linking pressure adaptation to thermal adaptation in salt-saturated habitats, by as yet unknown molecular mechanisms. The results are discussed in the context of expanding the thermal window for growth in deep-sea realms, and novel electrostatic charges and active site architectures are also examined.

Results and discussion

Deep-sea metagenome libraries, screening and general features of selected sequence-encoded esterases (ESTs)

Samples were collected from two distinct deep-sea marine environments. First, brine/interface fluid was collected from deep hypersaline anoxic basins (DHAB): Lake *Medee* (3040 m depth; 15.5°C), which is the largest known DHAB, Lake *Bannock* (3342 m depth; 14.5°C) and Lake *Kryos* (3340 m depth; 16.5°C) (Daffonchio *et al*., 2006; Yakimov *et al*., 2013). Note that brine fluids were collected from *Medee* and *Bannock* basins whereas alive interface was collected from Lake *Kryos*. Second, seawater was collected from the hadopelagic Station Matapan-Vavilov Deep (4908 m depth; 14.5°C), which is the deepest site of the Mediterranean Sea (Smedile *et al*.,

A

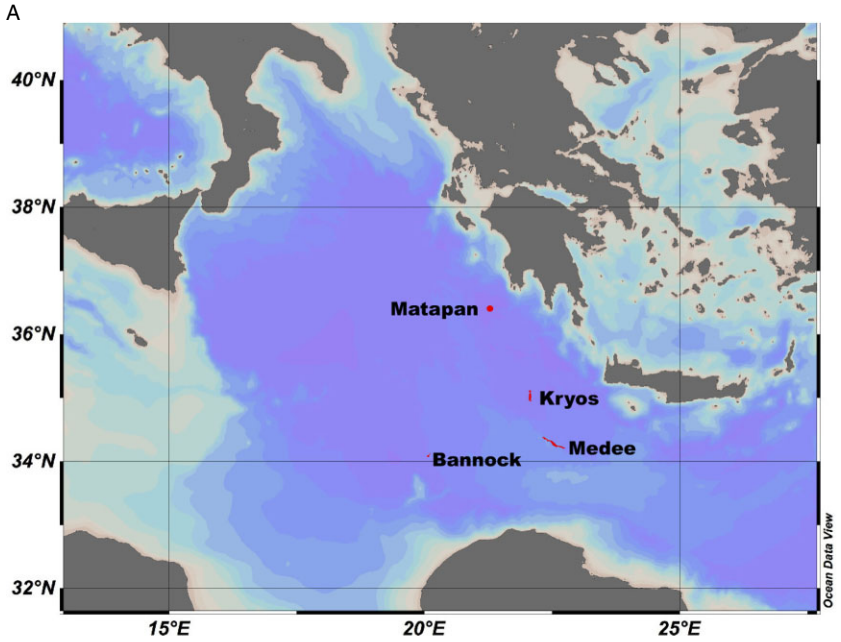

Fig. 1. Global map displaying the locations of the main deep-sea sites in the Mediterranean Sea investigated in this study. Hydrochemistry of selected deep-sea regions (A) is shown in (B) (for extensive details, see Table S2).

B

	Medee	Kryos	Bannock	Matapan
Depth (m):	3010	3340	3342	4908
Salinity (g kg^{-1}) :	347.80	313.92	263.20	38.6
Temperature (ºC) :	14.75–15.46	16.5	14.5	14.29

2013). At these sites, the water temperature is never below 13.0°C or higher than 16.5°C. Total DNA was extracted, and subsets of 40 024 clones from the four libraries generated in this study harbouring nearly 1.3 Gbp of community genomes were scored for the ability to hydrolyse α-naphthyl acetate (αNA) and tributyrin, as previously reported (Reyes-Duarte *et al*., 2012), which is indicative of EST/lipase activity. A total of five unique clones were selected as active. Hydrochemistry of selected deep-sea regions and EST screening statistics are shown in Fig. 1 and Table S2. The inserts were sequenced, analysed and compared with the sequences available in the National Center for Biotechnology Information (NCBI) non-redundant public database (Hall, 1999). Five (one per active clone) predicted metagenome sequence (MGS)-encoding ESTs with the α/β hydrolase fold were identified and successfully produced as soluble proteins when expressed in *Escherichia coli*, and their properties were investigated. They were named based on the source ID followed by a serial number: MGS-M1 and MGS-M2 (from the *Medee* basin), MGS-B1 (from the *Bannock* basin), MGS-K1 (from the *Kryos* basin), and MGS-MT1 (from the Matapan basin).

According to BLAST searches of the NCBI non-redundant database, the five studied protein sequences were 44–62% similar to homologous proteins in the database (Table S3A–B). The deduced molecular masses and estimated *p*I values of these proteins ranged from 31.6 to 56.2 kDa and from 4.4 to 7.3 respectively. The pairwise amino acid sequence identity ranged from 8.3% to 21.8%; MGS-B1 and MGS-MT1 were the most similar enzymes (21.8% sequence identity), whereas MGS-K1 and MGS-M1 were the most divergent at the sequence level (8.3% sequence identity). The selected α/β hydrolases contain a classical Ser-Asp-His catalytic triad, but the catalytic elbow and oxyanion hole (i.e. the GXSXGG and H/N-GGG(A)/P-X motifs) often diverged from the consensus regions as identified by an extensive sequence analysis. Nonetheless, there was adequate sequence conservation among these catalytic motifs and the overall enzyme sequence to categorize the enzymes into the following accepted lipase/EST subfamilies (Kourist *et al*., 2010): family IV (MGS-MT1, MGS-B1 and MGS-M1), family V (MGS-M2) and family VII (MGS-K1). The sequence-based features and amino acids participating in the predicted catalytic sites are listed in Table S3A–B.

Salt-saturated deep-sea brines contain biochemical signatures indicating adaptation to salinity and thermal extremes

According to the standard assay conditions described in Experimental procedures, and summarized in Table 1, the five proteins were fully characterized. We first confirmed that the purified proteins, which were most active at pH values ranging from 7.0 to 8.5 (Fig. S1), exhibited the expected EST activity, tested over a set of 101 structurally different esters. Extensive differences in activity level as well as substrate profiles and preferences were noticeable according to specific activity (units mg^{-1}) determinations (Appendix S1; Fig. S2). Using *p*-nitrophenyl propionate (*p*NP-propionate) as a model substrate, Lake *Medee* enzymes were found to be the least active enzymes, i.e. the enzyme from Matapan-Vavilov Deep, which was the most active enzyme, exhibited specific activities 900-fold greater than those of Lake *Medee* enzymes (Table 1).

The enzymatic activities of purified proteins were stimulated by the addition of NaCl and KCl to the reaction mixture (Fig. 2). The optimal concentration of Na^+/K^+ for activity was the lowest for MGS-B1 and MGS-K1 (optimal at 0.8–1.2 M), whereas the other enzymes were most active at concentrations greater than 3.0 M. At the optimal concentration, MGS-B1 from Lake *Bannock* exhibited the greatest increase in activity (14-fold) compared with reactions not containing salts. $MgCl_2$ triggered stronger effects on enzyme activities. However, while MGS-K1, MGS-B1 and MGS-M1 (in order of greatest inhibition to least inhibition by $MgCl_2$) were strongly inhibited by this salt at concentrations greater than 0–0.8 M, MGS-M2 and MGS-MT1 were stimulated with maximal activity at 2.4 and 1.6 M salt respectively. The activation of all enzymes by $Na^+/K^+/Mg^{2+}$ indicates that the properties of the ESTs herein reported reflect specific habitat characteristics and that activation by sodium, potassium, and, to a lesser extent, magnesium, may be common in enzymes from the deep-sea sites examined, independently of the habitat environmental constraints.

Three distinct profiles of temperature optima were further observed (Fig. 3). The first profile was a thermophilic-like profile, as exemplified by MGS-M2, in which the enzyme was most active at 70°C and retained ≤ 5% activity at 4–16°C (in the presence or absence of NaCl). Calculation of denaturing temperature (71.4°C) by circular dichroism (CD; Table 1) confirmed the high protein stability. The high optimal temperature of MGS-M2 was unexpected because this enzyme was isolated from Lake *Medee*, a permanently moderate-temperature site (15.5°C) (Yakimov *et al*., 2013). To prove that the MGS-M2 enzyme showed a temperature profile typical for thermophilic proteins, an EST with the α/β hydrolase fold (referred to as MGS-HA1), which originated from a clone library created from a superficial seawater sample at a hydrothermal vent at Saint Paul Island (100 m depth) and maintained at 60–65°C, was purified and characterized for comparative purposes. The enzyme was derived from a clone (out of 20 000 total clones) active towards αNA. The MGS-HA1 enzyme, which can be categorized into the lipase/EST family VI (Kourist *et al*., 2010) and most likely originated from *Geobacillus* as determined by BLAST homology search and GOHTAM (Tables S3A and S4), showed maximal activity at pH 8.0 (Fig. S1) and concentrations of $Na^+/K^+/Mg^{2+}$ of up to 1.6–3.2 M (Fig. 1). Its optimal temperature for activity (70–75°C) and residual activity at low temperatures, e.g. < 0.4% at 4°C (Fig. 3),

Table 1. Specific activity and protein denaturation temperature (T_d) of proteins as determined by circular dichroism.

Enzyme	Activity	Protein fold	Specific activity (units g^{-1})[a]	Standard assay conditions for activity determination [pH/T(°C)/NaCl (M)][a]	T_d (°C)[b]
MGS-M1	EST	α/β Hydrolase	105.73 ± 2.34	8.0/25/3.6	65.2
MGS-M2	EST	α/β Hydrolase	188.32 ± 8.76	8.0/70/4	71.4
MGS-B1	EST	α/β Hydrolase	24 077 ± 85	8.0/25/0.8	52.4
MGS-K1	EST	α/β Hydrolase	18 094 ± 270	7.0/30/0.8	40.3
MGS-MT1	EST	α/β Hydrolase	94 994 ± 460	8.0/40/2.8	55.7
MGS-M3	GLY	α/β Hydrolase	12 471 ± 895	8.0/45/ 0.4	59.3
MGS-M4	AKR	TIM-barrel	5371 ± 28	8.0/16/0.8	63.3
MGS-M5	LDH	Rossman	1072 ± 12	8.0/30/0	n.d
MGS-HA1	EST	α/β Hydrolase	408.2 ± 18.0	8.0/70/3.2	79.5

a. *p*NP-propionate (1 mM; for EST activity), *p*NP-β-D-glucose (30 mM; for GLY activity), sodium pyruvate (5 mM; for LDH activity) and methyl glyoxal (1 mM; for AKR activity) were used as standard assay substrates. The following buffers were used for activity determinations: 50 mM Tris-HCl for pH 8.0 and 50 mM 4-(2-hydroxyethylpiperazine-1-ethanesulphonic acid (HEPES) for pH 7.0, which were used as standard assay buffers. Note that activity determinations were performed at the optimal parameters and conditions specifically cited in the Experimental Procedures, and summarized in this Table.

b. Standard deviation < 0.1.

Activity and protein fold associated to each of the enzymes are also summarized. AKR, aldo-keto reductase; EST, esterase; GLY, glycosidase; LDH, lactate dehydrogenase; n.d., not determined.

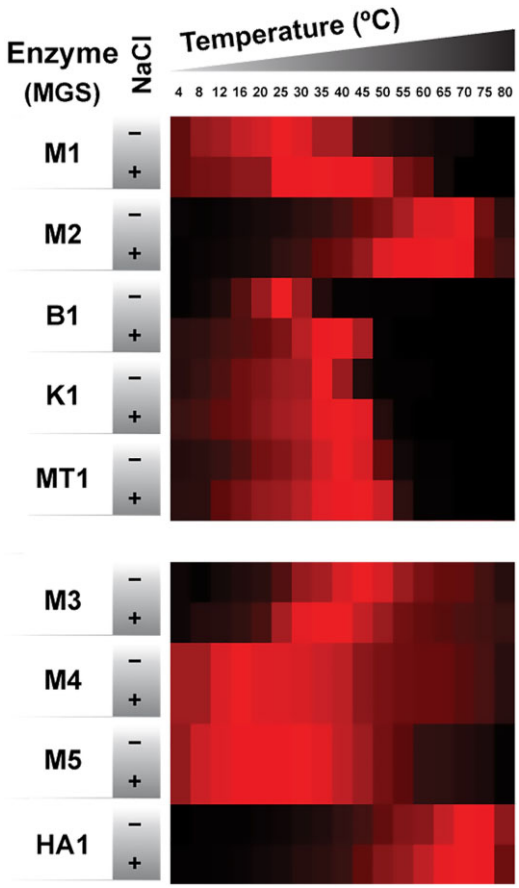

Fig. 2. Temperature profiles of the enzymes in the absence (–) or presence (+) of salt (NaCl). Profiles for deep-sea esterases are shown in the top panel, whereas those for other enzymes are shown in the bottom panel. The heat map colours represent the relative percentages of specific activity (units g^{-1}) compared with the maximum (100%) given in Table 1. The specific activities were calculated in triplicate (SD < 0.5%) using the standard assay substrates (see Table 1) and conditions described in the Experimental procedures. Note that due to the significant differences in specific activities in the presence and absence of salt, the assays were performed using the same units of enzymes on the basis of *p*NP-propionate transformation. This effect is particularly noticeable for the MGS-B1 esterase, which retains only 3.4% of its activity in the absence of salt compared with the optimum level (0.8 M NaCl). The colour code ranges from black (no activity) to intense red (100% activity). Heat maps were constructed in R (http://www.r-project.org) using the 'heatmap.2' function within the 'gplots' package.

allowed it to be categorized as a typical thermophilic enzyme, which was further evidenced by the determination of its denaturing temperature (79.5°C; Table 1). The comparative analysis of MGS-M2 and MGS-HA1 showed that MGS-M2 reassembled a thermophilic protein. The second profile was a mesophilic-like profile, as exemplified by MGS-MT1, in which the enzyme was most active at 40–45°C and retained less than 40% of its activity at ≥ 50°C (Fig. 3) in the presence or absence of NaCl; this result is in agreement with its denaturing temperature (55.7°C; Table 1). The third profile consisted of a psychrophilic-like profile, as exemplified by MGS-M1, MGS-B1 and MGS-K1, with enzymes that were most active at 25–40°C (Fig. 3). Notably, compared with mesophilic (MGS-MT1) and thermophilic (MGS-M2) ESTs that showed a salt-independent thermal profile (Fig. 3), a positive impact of salinity on thermal activation/stabilization in these three enzymes was observed in the presence of optimal concentrations of NaCl. This result was particularly noticeable for MGS-M1 and MGS-B1, as they displayed a shift in the optimal temperature from 25 to 35°C and from 25 to 40°C respectively. It was also true to a lesser extent for MGS-K1 (from 35 to 40°C). The higher stabilization levels for MGS-M1 and MGS-B1 agreed with their higher denaturing temperatures (65.2°C and 52.4°C respectively) compared with MGS-K1 (40.3°C). Together, the data demonstrated that deep-sea salt-saturated biotopes might contain enzymes adapted to work under multiple temperature extremes (80% activity retained in the range from 16 to 70°C), despite these sites being moderately warm (14.0–16.5°C).

Adaptation to high pressure is linked to high thermal resistance in deep-sea brines

As pressure is one of the most representative characteristics of deep-sea habitats, the influence of hydrostatic pressure (350 bar) on enzyme performance was further evaluated in high-pressure 2 ml reactors as described in the Experimental procedures, and presumptive links between the site and EST characteristics were evaluated. The data presented in Fig. 4 revealed that the relative percentage of activity at 350 bar compared with that at atmospheric pressure positively correlated with the optimal temperature for activity in the two ESTs from the salt-saturated Lake *Medee*; thus, the highest adaptation to pressure was obtained for the MGS-M2 EST, which had the highest temperature optima (70°C) compared with MGS-M1 (25°C). No such correlation could be evaluated for the other deep-sea sites, which showed different activation levels, as only one enzyme candidate was characterized per site; no additional enzymes could be tested because no additional active clones could be obtained in our library screen tests, or we were unable to produce additional soluble proteins from sequenced positive fosmids.

To prove that such a thermal-pressure correlation exists in Lake *Medee*, we polymerase chain reaction-amplified candidate genes from the two EST-positive clones (the ones containing MGS-M1 and MGS-M2 ESTs), and we were able to successfully express and produce in soluble form three proteins: a glycosidase (GLY; herein named MGS-M3), an aldo-keto reductase (AKR; MGS-M4) and an (*L*)-lactate dehydrogenase

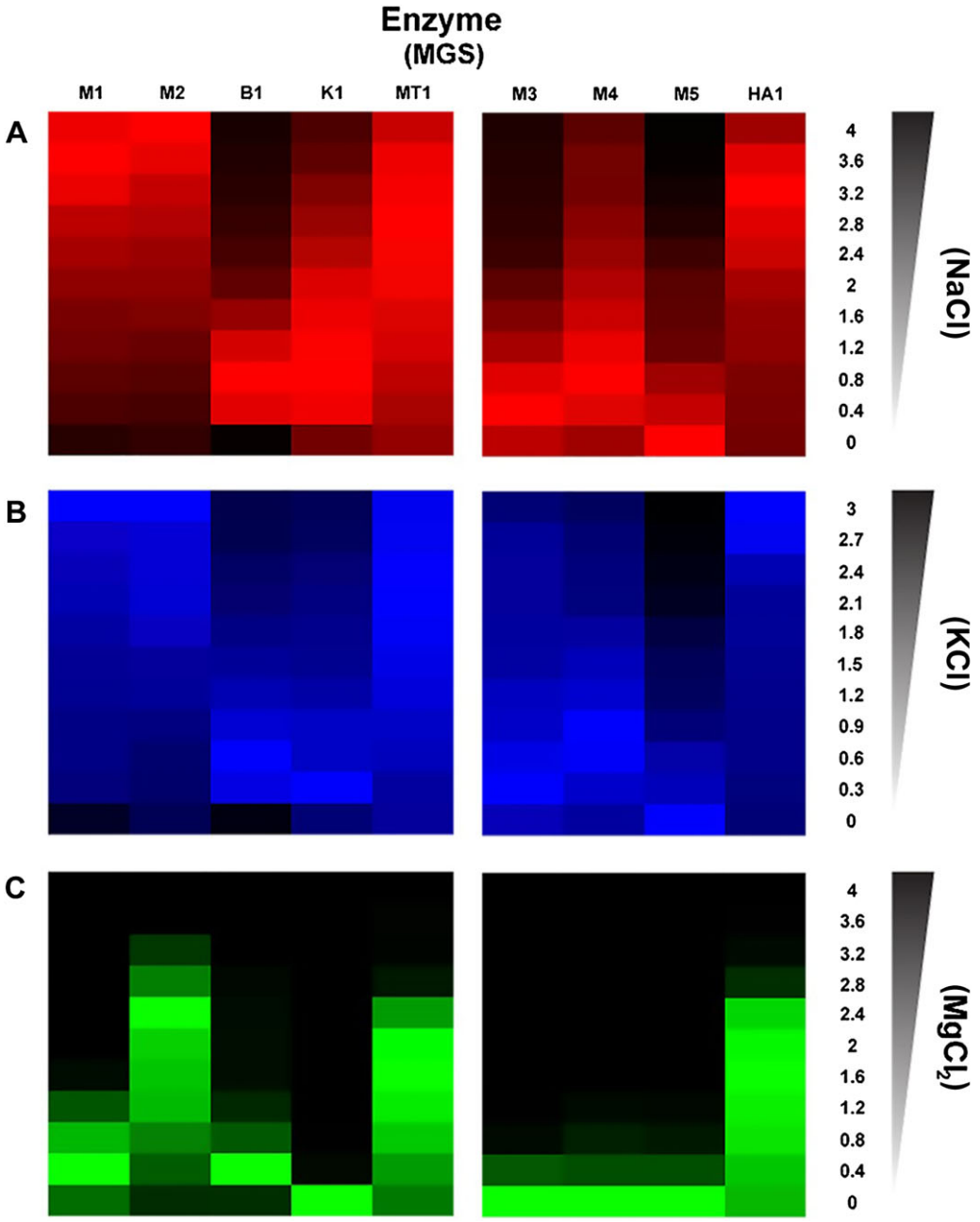

Fig. 3. Heat maps displaying the activities of enzymes at different concentrations of NaCl (A, red), KCl (B, blue) and $MgCl_2$ (C, green). Profiles for deep-sea esterases are shown in the left panel, whereas those for other enzymes are shown in the right panel. The heat map colours represent the relative percentages of specific activity (units g^{-1}) compared with the maximum activity (100%). The specific activities were calculated in triplicate [standard deviation (SD) < 0.5%] using the standard assay substrates (see Table 1) and conditions described in the Experimental procedures. The colour code ranges from black (no activity) to intense red, blue and green (100% activity). The 100% levels for NaCl/KCl/$MgCl_2$ are as follows: MGS-M1 (105.7/105.6/38.1 units/g), MGS-M2 (188.3/111.6/212.5 units g^{-1}), MGS-M3 (12 471/ 8142/6481 units g^{-1}), MGS-M4 (5371/4715/362.0 units g^{-1}), MGS-M5 (1072/1000/153.7 units g^{-1}), MGS-B1 (24 077/11 685/4496 units g^{-1}), MGS-K1 (18 094/17 829/8239 units g^{-1}), MGS-MT1 (94 994/88 259/117 186 units g^{-1}) and MGS-HA1 (408.2/388.4/246.1 units g^{-1}). Heat maps were constructed in R (http://www.r-project.org) using the 'heatmap.2' function within the 'GPLOTS' package.

(LDH; MGS-M5). According to the CAZY database (http://www.cazy.org; Cantarel *et al.*, 2009) and considering structural similarities, MGS-M3 is related to family 3 of GLYs, whereas MGS-M4 belongs to the gluconic reductase subfamily 5 of AKRs, and MGS-M5 matches lactate/malate dehydrogenases rather than other Rossman fold-containing enzymes. The sequence-based features and amino acids participating in the predicted

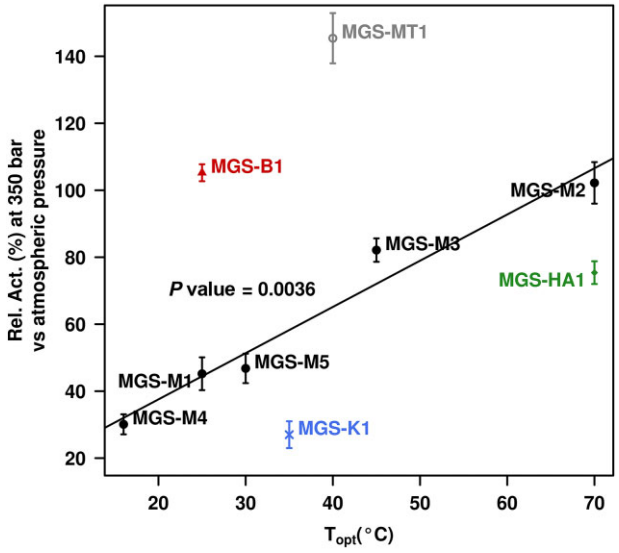

Fig. 4. Activity levels of enzymes as a function of pressure and optimal temperature for activity. The specific activities using the standard assay substrates and pH (see Table 1) and 25°C were calculated in triplicate [standard deviation (SD) is shown] at atmospheric pressure or 350 bar as described in the Experimental procedures. One hundred percent activity refers to the activity value at atmospheric pressure for each of the enzymes. The optimal temperature for activity of each of the enzymes studied is reported in Fig. 2. Only enzymes from *Medee* Lake were considered for correlation analysis (see Results and Discussion section for details), and thus the enzymes from other site are shown in different colours.

catalytic sites are listed in Table S3C. According to the standard assay conditions described in Experimental procedures (see also Table 1), the corresponding activities, found to be maximal at pH 8.0 (Fig. S1), were first confirmed against a total of 47 model substrates (Table 1; Appendix S1; Fig. S1). Activity measurement further showed that while MGS-M5 was inhibited by Na^+ and K^+ (Fig. 1), MGS-M3 and MGS-M4 were activated by low salt concentrations (0.3–0.4 M and 0.8–1.2 M respectively). $MgCl_2$ strongly inhibited all enzymes (Fig. 1). The influence of temperature and hydrostatic pressure (350 bar) on enzyme performance was further evaluated for these three enzymes, which confirmed that the highest resistance to pressure was obtained for enzymes with the highest temperature optima (Fig. 4).

Overall, we found that pressure adaptation, at least in Lake *Medee*, is linked to thermal adaptation and that the correlation (*P* value = 0.0036; *t*-test; Fig. 4) was accurate for five proteins with different types of fold, such as MGS-M1, MGS-M2 and MGS-M3 with an α/β-hydrolase fold, MGS-M4 with a triosephosphate isomerase TIM-barrel fold, and MGS-M4 with a Rossman fold (MGS-M4).

Analysis of temperature, salt and pressure adaptations in enzyme primary and tertiary structures

Given the diverse thermo- and halo-tolerance of the MGS proteins (for summary see Figs 2 and 3), we were interested in identifying primary sequence and tertiary structure elements that might reflect adaptations conferring these properties. It has been hypothesized that intracellular proteins of thermophilic, psychrophilic and halophilic organisms undergo multiple, distinct adaptations to retain activity in such environments. For example, enzymes of thermophilic origin may increase their stability to resist thermal-induced unfolding by increasing the size of their hydrophobic cores, increasing the number of disulphide bonds, forming additional salt-bridge interactions or increasing the number of charged residues on their exposed surfaces (Reed *et al.*, 2013). Enzymes of halophilic origin must increase their hydration to compensate for increased extracellular salt concentrations (Madern *et al.*, 2000; Delgado-García *et al.*, 2012; Reed *et al.*, 2013). Observed adaptations included an increased negative surface charge and/or lower isoelectric point (*p*I) caused by an increase in acidic residues and a decrease in lysine residues (Ferrer *et al.*, 2012), a decrease in aliphatic amino acids (Leu/Ile) and an increase in small hydrophobic amino acids (Gly/Ala/Val), and a decrease in the extent of buried non-polar amino acids. We searched for these adaptations by obtaining the crystal structures and/or primary sequences of the thermophilic enzyme MGS-M2, the psychrophilic/mesophilic enzymes MGS-M1, MGS-M4, MGS-M5 and MGS-MT1, the halophilic enzymes MGS-M1 to MGS-M4 and MGS-MT1 and comparing these enzymes with their closest homologues lacking these adaptations (Fig. 5; Appendix S1). We also analysed MGS-M5, which does not exhibit salt tolerance, as a baseline comparison. To remove any signal from substrate-binding amino acids correlated with functional diversification, we excluded residues close to the catalytic sites.

An analysis of 16 known temperature-dependent adaptations showed that the MGS-M2 enzyme contains six adaptive features typical of thermophilic proteins (Table S3D). However, the analysis also showed that the psychrophilic enzymes MGS-M1 and MGS-M4 contain 10 and 8 adaptations, respectively, that are typically observed in thermophiles. In addition, the mesophilic enzyme MGS-M4 contains nine of these so-called thermophilic adaptations. The presence of these thermophilic adaptations may explain the preservation of the activity of these enzymes at temperatures as high as 60–75°C (Fig. 2). These results suggest that other characteristics beyond those studied here are involved in conferring the temperature dependence of these MGS enzymes.

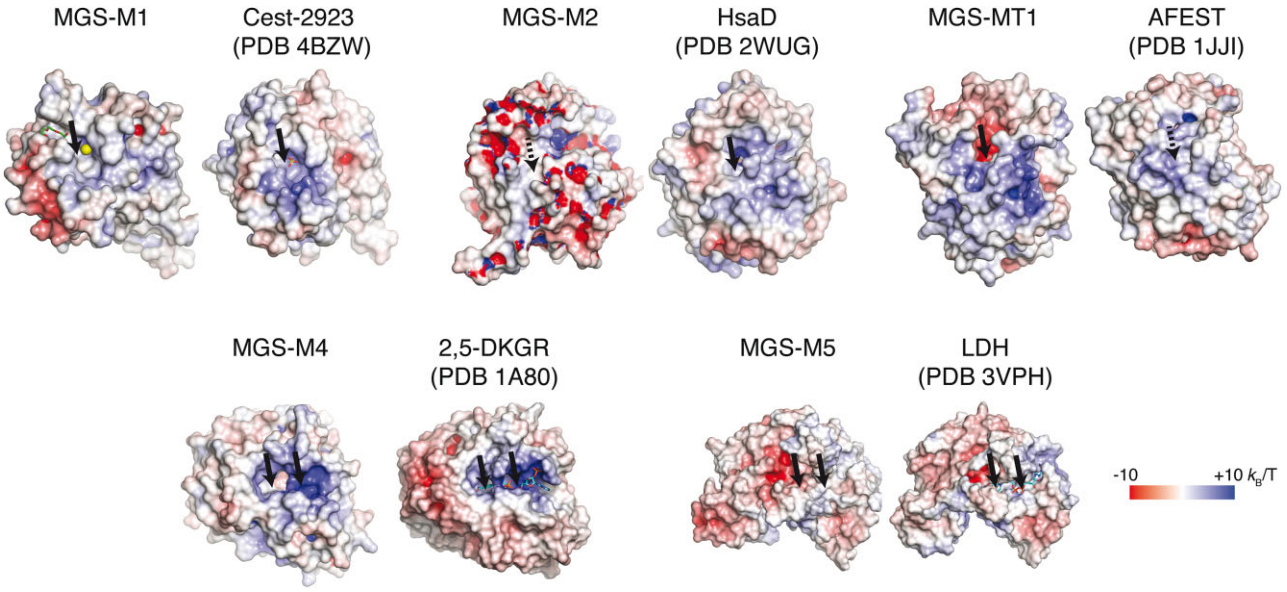

Fig. 5. Comparison of surface features of the MGS enzymes crystallized in this study and their structural homologues revealed by structure similarity searches. The solvent-exposed surfaces of the enzymes are shown, coloured by electrostatic potential as indicated by the scale at the bottom right (in units k_B/T = Boltzmann constant over temperature). The arrows refer to the locations of the catalytic serines (for esterases in the top row), with dashes indicating that this residue is hidden within the protein, or the NADPH/NADH and substrate binding sites (for enzymes in the bottom row).

As the activity of MGS-M4 was preserved even at 1 M KCl, we repeated the analysis of structural adaptations on this enzyme by comparison with a non-halophilic homologue, excluding a region surrounding the substrate-binding site. For this enzyme, the analysis identified only two of the seven adaptations (Table S3E, a decrease in overall *p*I and a decrease in lysine residues). This result suggests that this enzyme contains other unknown characteristics that preserve its activity in high-salt environments. Of the seven halophilic-dependent adaptations examined in the MGS enzymes, MGS-M1, MGS-MT1 and MGS-M2 contained five, three and two of these adaptations respectively (Table S3E). Thus, although the salt adaptation of MGS-M1 is well explained by previously identified adaptations, the salt adaptation of the MGS-M2 and MGS-MT1 enzymes is not. Notably, an increase in surface acidic residues is the most commonly observed adaptation in halophilic enzymes, but MGS-MT1 exhibited the opposite feature: it has a high *p*I value (7.29), reflecting a decrease in surface acidic residues. This evidence suggests that other undiscovered adaptations are present in this enzyme.

Piezophilic proteins appear to exhibit a smaller hydrophobic core, fewer large and more small amino acids in the hydrophobic core, increased multimerization and fewer intra-molecular ionic interactions (Michels and Clark, 1997). An analysis of these four known pressure-dependent adaptations (Table S3E) did not reveal that MGS-MT1 and MGS-M2 have more of these adaptations than the proteins that are not piezophilic (Fig. 4).

An analysis of the crystal structures further revealed that the conformation and/or overall electrostatic charge (Fig. 5; Fig. S3 and S4) were distinct relative to other structurally characterized homologues in four of the five proteins. Indeed, the overall electrostatic charge on the surface of MGS-M1 was similar to the charges of Cest-2923 and bacillibactin esterase (BES), with no obvious patches of concentrated charge; however, MGS-M1 contained a deletion of 14 amino acids relative to Cest-2923 (residues 141–150 in MGS-M1 and 158–181 in Cest-2923), which resulted in alterations in the conformation of the active site. MGS-M2 has different electrostatic characteristics and residue composition: the active sites of the closest structural homologue (HsaD) are noticeably positively charged, whereas those of MGS-M2 contain a mix of positive and negative features. The MGS-MT1 active site is long, deep, negatively charged, and open to solvent; these electrostatic characteristics are in sharp contrast to the properties of Archaeoglobus fulgidus esterase (AFEST). Compared with its closest structural homologues (2,5-DKGR), MGS-M4 contains a small (153 $Å^3$), slightly positively charged nicotinamide adenine dinucleotide phosphate (NADPH)-binding site and a smaller active site, which is missing some structural loops involved in substrate contacts. Finally, a comparison of the electrostatic surface and active site composition of the

MGS-M5 structure revealed that it closely resembled the structures of its closest homologues (LDH), including the four residues involved in interacting with the pyruvate/L-lactate ligands. Such differences are consistent with variations at the level of substrate profiles and preferences (Fig. S2), but they may also be implicated in yet unknown differential adaptations to environmental constraints. Full details about substrate fingerprint and crystal structures and active site architecture are given in Appendix S1.

Source organisms of investigated sequence-encoded ESTs

We finally investigated the presumptive microbial origin of the genes encoding enzymes under investigation (Table S4). A search against the GOHTAM database (Menigaud *et al*., 2012) revealed compositional similarities between the DNA fragments containing the genes for MGS-M1 to MGS-M5 and several bacterial strains that are most likely from uncharacterized species belonging to *Firmicutes*. Genes in the DNA fragment encoding MGS-M2 to MGS-M5 share similarity with the corresponding genes of the *Tenericutes*-like bacterium *Haloplasma contractile* (Antunes *et al*., 2008). This bacterium was recently isolated from deep-sea sediment samples and is a unique 'transiting form' of bacteria that has been placed on the phylogenetic tree between *Firmicutes* (*Bacilli* and *Clostridia*) and *Mollicutes* (*Mycoplasmas* and other strictly symbiotic bacteria). Interestingly, the closest neighbour to *H. contractile* is '*Candidatus lumbricincola*', an uncultured bacterium associated with earthworms (Nechitaylo *et al*., 2009). However, unlike *Mollicutes*, *H. contractile* grows axenically on 'simple' media, suggesting a free-living lifestyle (Antunes *et al*., 2011). MGS-B1 was not identified by GOHTAM, but the DNA fragment encoding this enzyme shares genes with *Betaproteobacteria* of the genus *Variovorax*. MGS-K1, which is part of an insertion sequence element or prophage, shares DNA similarity with *Aspergillus*, although TBLASTX analysis indicated some weak similarity with rare *Betaproteobacteria* species. Hence, similarly to other fragments, MGS-K1 could be derived from an unknown species. The DNA fragment encoding the MGS-MT1 enzyme exhibits compositional similarity to the *Alteromonas* genomes, most likely *Alteromonas macleodii* 'deep-ecotype', which dominates the Matapan-Vavilov Deep (Smedile *et al*., 2013).

Experimental procedures

Source of enzymes

Total DNA was extracted from selected sites using the G'NOME DNA extraction kit (BIO 101/Qbiogene, Morgan Irvine, CA, USA) according to the manufacturer's instructions for each of the microbial communities, and large-insert pCCFOS1 fosmid (for all but *Bannock* Lake) or bacteriophage lambda-based ZAP phagemid (*Bannock* Lake) libraries were generated (Ferrer *et al*., 2005; Alcaide *et al*., 2013) and scored for the ability to hydrolyse αNA and tributyrin (Reyes-Duarte *et al*., 2012). Positive clones were selected, and their DNA inserts were sequenced using a Roche 454 GS FLX Ti sequencer (454 Life Sciences, Branford, CT, USA) at Life Sequencing SL (Valencia, Spain) or completely sequenced using universal primers and subsequent primer walking. In all cases, upon completion of sequencing, the reads were assembled to generate non-redundant metasequences using Newbler GS De Novo Assembler v.2.3 (Roche). The GENEMARK software (Lukashin and Borodovsky, 1998) was employed to predict potential protein-coding regions (open reading frames with ≥ 20 amino acids) from the sequences of each assembled contig, and deduced proteins were screened via BLASTP and PSI-BLAST (Altschul *et al*., 1997). MGS-encoding enzymes were deposited in GenBank under the accession numbers KF831414 – KF831421.

General methods

The cloning, expression and purification of selected proteins using the p15TV-Lic vector and *E. coli* BL21(DE3) Codon Plus-RIL (for MGS-M4 and MGS-M5) and Ek/LIC 46 and *E. coli* BL22 (for MGS-M1, MGS-M2, MGS-B1, MGS-K1 and MGS-HA1) or *Rosseta* (for MGS-MT1) were performed as described elsewhere (Alcaide *et al*., 2013) using the primer pairs described in Table S5. Purity was assessed as > 98% by SDS-PAGE. Note that after purification, MGS-MT1, MGS-K1 and MGS-B1 were required to be maintained in 50 mM HEPES [4-(2-hydroxyethylpiperazine-1-ethanesulphonic acid] pH 7.0 containing 0.8 M NaCl to ensure protein stability. All chemicals used for enzymatic tests were of the purest grade available and were purchased from Fluka-Aldrich-Sigma Chemical Co. (St Louis, MO, USA) or Apin Chemicals (Oxon, UK) (Alcaide *et al*., 2013). EST activity was assayed using 1 mM *p*-nitrophenyl (*p*NP) esters (at 410 nm) and structurally diverse esters other than *p*NP esters (at 540 nm) using substrates and conditions as previously described (Alcaide *et al*., 2013). GLY activity was determined using 1 mM *p*NP sugars (at 410 nm) as described elsewhere (Del Pozo *et al*., 2012). LDH activity was assayed by a colorimetric assay in which the conversion of pyruvate to lactate was determined using a LDH kit (Sigma Chemical Co., St. Louis, MO, USA) to follow the release of NAD^+ at 340 nm in a reaction mixture containing 0.4 mM nicotinamide adenine dinucleotide (NADH) and 5 mM sodium pyruvate. AKR activity was determined routinely by monitoring the decrease in absorbance at 340 nm in an assay mixture containing the corresponding substrate (1 mM) and 0.4 mM NADPH. Unless otherwise stated, *p*NP-propionate (1 mM; for EST activity), *p*NP-β-D-glucose (30 mM; for GLY activity), sodium pyruvate (5 mM; for LDH activity) and methyl glyoxal (1 mM; for AKR activity) were used as standard assay substrates. The optimal pH, temperature and salt (NaCl) concentrations (see Table 1) were used for standard assay reactions and specific activity determinations (see Fig. S2). pH values between 4.0 and 11.0 (at the optimal temperature),

temperatures between 4 and 80°C (at the optimal pH), and NaCl, KCl and $MgCl_2$ concentrations of up to 4 M (at the optimal pH and temperature) were tested for optimal parameter determinations using the standard assay substrates (for details Table 1). For pH and temperature optima determinations, the assay buffers were supplemented with NaCl, which was used as model salt, at the optimum concentration or not supplemented (Table 1).

Unless stated otherwise, reactions were conducted using 0.005–2 μg of pure proteins; the absorbance was determined every 1 min for a total time of 15 min. All reactions were performed in triplicate, and one unit (U) of enzyme activity was defined as the amount of enzyme transforming 1 μmol of substrate in 1 min under the assay conditions. All values were corrected for non-enzymatic hydrolysis (background rate).

Pressure perturbation studies

The effect of hydrostatic pressure was further analysed by placing the reactions into a high-pressure incubating system that consists of 2 ml of high-pressure cell connected to a pressure generator (High Pressure Equipment, Erie, PA, USA) that is capable of generating a pressure of up to 10 000 psi. Briefly, 20 ml of standard assay buffer (see Table 1) containing 30 mM *p*NP-β-D-xylose (for MGS-M3), 1.0 mM *p*NP-propionate (for ESTs), 5 mM sodium pyruvate and 0.4 mM NADH (for MGS-M4) or 1.0 mM methyl glyoxal and 0.4 M NADH (for MGS-M5) was freshly prepared. Next, 0.32–540 μg of pure protein (depending on the enzyme) was added. Two millilitres of the reaction mixture was immediately transferred to high-pressure-maintaining reactors, and another 2 ml was held at atmospheric pressure. The experiments were performed at 25°C and at a salt (NaCl/KCl) concentration similar to that found at the sampling site [Fig. 1; 39–348 practical salinity units (PSUs)]. In all cases, reactions without protein were used as negative controls. After incubation for a total of 5–30 min (depending on the enzyme), the samples were depressurized, and the extent of the reaction was monitored by spectrophotometer measurements at 410 nm. All experiments were performed in triplicate.

Protein purification, crystallization and structure determination

The ESTs MGS-M1, MGS-M2 and MGS-MT1, the LDH MGS-M4, and the AKR MGS-M5 were expressed and purified according to the procedures described by Lai and colleagues (2011). The purified His_6-tagged MGS-M1, MGS-M2, MGS-MT1, MGS-M4 and MGS-M5 enzymes were crystallized using the sitting drop method, Intelliplate 96-well plates and a Mosquito liquid handling robot (TTP LabTech), which mixed 0.5 μl of protein between 21 to 27 mg ml^{-1} with 0.5 μl of the reservoir solution. The reservoir solutions were as follows: MGS-M1 – 20% potassium fluoride, 20% (w/v) polyethylene glycol (PEG) 3350; MGS-M2 – 0.1 M sodium HEPES pH 7.5, 1.4 M sodium citrate, thermolysin protease; MGS-MT1 – 0.1 M MES pH 6.0, 20% (w/v) PEG 10K; MGS-M4 – 0.1 M Tris pH 8.5, 0.2 M ammonium sulphate, 25% (w/v) PEG 3350, tobacco etch virus protease; MGS-M5 – 0.1 M sodium cacodylate pH 6.5, 0.2 M calcium acetate, 9% PEG8 K, trypsin protease. The crystals were cryo-protected with reservoir solution supplemented with either 12% glycerol, 15% ethylene glycol or Paratone-N oil prior to flash freezing in an Oxford Cryosystems Cryostream. Diffraction data were collected at 100K and the Cu Kα emission wavelength using a Rigaku HF-0007 rotating anode with a Rigaku R-AXIS IV++ detector. Diffraction data were reduced with either HKL3000 (Minor *et al*., 2006) or XDS (Kabsch, 2010) and SCALA (Evans, 2006). The structures were determined by molecular replacement using the following Protein Databank (PDB) codes: MGS-M1 – 3HXK, MGS-M2 – 2XUA, MGS-MT1 – 3V9A, MGS-M4 – 4FZI, and MGS-M5 – 1LDN. Electrostatic surface representations were calculated using the PDB2PQR server (Dolinsky *et al*., 2004) with the Assisted Model Building and Energy Refinement (AMBER) force field and otherwise default settings. Substrate-binding cavity volumes were calculated by the CASTp server (Dundas *et al*., 2006). Structural homologues in the PDB were identified using the PDBeFold server (Krissinel and Henrick, 2004). For halophilic and thermophilic adaptation analysis, we excluded from all analyses residues within 10 Å of the catalytic serine in the α/β hydrolase enzymes or within 5 Å of the bound substrate in the AKR prostaglandin F synthase (PDB 1RY0) to remove sequence substitutions due to substrate-binding differences. Surface or core classification was assigned after calculation of the solvent-accessible surface area (SASA) by NACCESS (http://www.bioinf.manchester.ac.uk/naccess); residues with > 10 $Å^2$ SASA were assigned as surface residues. The difference in hydrophobic core burial was calculated by comparing the apolar buried surface area in the hydrophobic core using NACCESS with the unfolded state SASA calculated by the unfolded server (http://folding.chemistry.msstate.edu/utils/unfolded.html) (Creamer *et al*., 1997). Multiple sequence alignments were constructed using the CLUSTALW2 tool (http://www.ebi.ac.uk/clustalw/index.html) integrated into the BIOEDIT 7.0.9.1 software (Hall, 1999) and GENO3D (Combet *et al*., 2002). X-ray diffraction statistics can be found in Table S6. The structural coordinates reported will appear in the PDB under accession codes 4Q3K, 4Q3L, 4Q3M, 4Q3N and 4Q3O.

CD

CD spectra were acquired between 190 and 255 nm with a Jasco J-720 spectropolarimeter equipped with a Peltier temperature (from 4 to 95°C) controller, employing 0.1 cm of path cell, at 25°C. The protein concentration was determined spectrophotometrically (at 280 nm) according to the corresponding amino acid sequence (http://www.expasy.org/tools/protparam.html). Protein solutions were prepared in 50 mM HEPES buffer pH 7.0; for MGS-B1 and MGS-MT1, 0.8 M NaCl was added to the buffer to ensure protein stability during the assay. Spectra were analysed and denaturation temperatures were determined at 220 nm, as reported (Pace and Scholtz, 1997; Schmid, 1997).

Construction of a neighbour-joining tree and oligonucleotide usage pattern analysis

Multiple protein alignments were performed using the CLUSTALW program built into the BIOEDIT software version

7.0.9.0 (Hall, 1999). Phylogenetic analysis of protein sequences was conducted with the MEGA 4.0 software (Tamura *et al*., 2007) using the neighbour-joining tree method and sampling of 1000 trees for bootstrapping and Poisson correction. DNA sequences of contigs were searched against all sequenced bacterial chromosomes, plasmids and phages for oligonucleotide compositional similarity using the GOHTAM web tool (Menigaud *et al*., 2012).

Conclusions

Since their discovery in 1983 (De Lange and Ten Haven, 1983; MEDRIFF Consortium, 1995; Wallmann *et al*., 1997), the number of Mediterranean deep-sea hypersaline lakes that have been described has grown constantly (Chamot-Rooke *et al*., 2005; Yakimov *et al*., 2007; La Cono *et al*., 2011; Yakimov *et al*., 2013). The surface of these brine lakes lies 3.0–3.5 km below sea level, and the salinity of the brines is 5–13 times higher than that of seawater. These lakes are characterized by a moderate-temperature gradient that consistently ranges from 13 to 16.5°C. Microbial populations inhabiting such ancient ecosystems are adapted to operate under harsh physical and chemical conditions, particularly high salinities and high pressures (Daffonchio *et al*., 2006; Smedile *et al*., 2013; Yakimov *et al*., 2013). These conditions are incompatible with life for common marine microorganisms (Harrison *et al*., 2013); however, although increasing evidence suggests that these environmental constraints may impact organism and protein evolution and properties, how and why this process occurs remains to be fully elucidated. The main reason for this lack of knowledge is that the overwhelming majority of autochthonous microbiota resist cultivation, and only a few isolates and enrichments have been obtained from these basins thus far (Antunes *et al*., 2003; 2007; 2008; Albuquerque *et al*., 2012; Yakimov *et al*., 2013; Werner *et al*., 2014).

The biochemical knowledge generated in this study demonstrated that pressure had a marked and consistent effect on the temperature profiles of enzymes from microorganisms inhabiting deep-sea salt-saturated habitats. Thus, by examining the two variables of pressure resistance and optimal temperature, we noticed that the effect of salt, e.g. 348 practical salinity units (PSU), is demonstrated by the link between high pressure and high thermal adaptations. Crystal structure analysis of five enzymes further demonstrated that the salinity level and the protein sequence/structure may play additional significant roles in defining the temperature profile by unknown structural adaptation mechanisms. Manipulating these factors may allow expansion of the lower and upper thermal tolerance limits of microbes inhabiting deep-sea salt-saturated lakes. Notably, marine enzymes that are most active at temperatures as high as 85–130°C have only been identified in deep-sea hydrothermal vent chimneys, whereas the optimal temperatures range from 4 to 60°C for the other deep-sea enzymes reported thus far (Table S1). Nevertheless, although pressure-enhanced activity has been reported in some hyperthermophilic proteins (Michels and Clark, 1997), no examples of proteins that are most active at ≥ 70°C (resembling thermophiles), such as the MGS-M2 enzyme herein reported, have been identified in deep-sea regions other than deep-sea hydrothermal vent chimneys with salinities < 40 PSU.

We hypothesize that the data generated herein may help with the design of new cultivation strategies for the isolation of new thermophiles from moderately warm (14.0–16.5°C) salt-saturated deep-sea lakes. Furthermore, the present study reported the largest set of structures of deep-sea proteins from uncultivable bacteria inhabiting hypersaline lakes (348 PSU) and the hadopelagic water column of the Eastern Mediterranean Sea; these results may provide future implications for our understanding of deep-sea protein adaptation, reaction mechanisms and substrate preferences.

Acknowledgements

The authors gratefully acknowledge the financial support provided by the European Community project MAMBA (FP7-KBBE-2008-226977). This work was further funded by grant BIO2011-25012 from the Spanish Ministry of Economy and Competitiveness (formerly MICINN). HT, OVG and PNG acknowledge European Commission for 'MicroB3' grant (FP7-OCEAN.2011-2 (contract Nr 287589)). This work received support from the Government of Canada through Genome Canada and the Ontario Genomics Institute (grant 2009-OGI-ABC-1405 to A.F.Y. and A.S.) and from the U.S. National Institutes of Health (grants GM074942 and GM094585 to A.S. through the Midwest Center for Structural Genomics). We thank Dr. C. Méndez-García for her excellent support for the preparation of heat map figures and Dr. T.Y. Nechitaylo for the sequence submission to databases. We also thank José M. Andreu and María A. Olivia (CIB-CSIC) for their excellent support in relation to the circular dichroism experiments and analyses. The authors declare that they have no competing interests.

References

Akondi, K.B., and Lakshmi, V.V. (2013) Emerging trends in genomic approaches for microbial bioprospecting. *OMICS* **17:** 61–70.

Albuquerque, D.M., Lopes, J.B., Segundo, L.F., Brandão, T.M., Ribeiro, M.N., Ramos, L., *et al*. (2012) Dehydrated brewery residue for pigs in the growth phase under high temperature conditions. *Rev Bras Zool* **41:** 1784–1788.

Alcaide, M., Tornes, J., Stogios, P.J., Xu, X., Gertler, C., Di Leo, R., *et al*. (2013) Single residues dictate the co-evolution of dual esterases – MCP hydrolases from the α/β hydrolase family. *Biochem J* **454:** 157–166.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402.

Antunes, A., Eder, W., Fareleira, P., Santos, H., and Huber, R. (2003) *Salinisphaera shabanensis* gen. nov., sp. nov., a novel, moderately halophilic bacterium from the brine-seawater interface of the Shaban Deep, Red Sea. *Extremophiles* **7:** 29–34.

Antunes, A., França, L., Rainey, F.A., Huber, R., Nobre, M.F., Edwards, K.J., and da Costa, M.S. (2007) *Marinobacter salsuginis* sp. nov., isolated from the brine-seawater interface of the Shaban Deep, Red Sea. *Int J Syst Evol Microbiol* **57:** 1035–1040.

Antunes, A., Taborda, M., Huber, R., Moissl, C., Nobre, M.F., and da Costa, M.S. (2008) *Halorhabdus tiamatea* sp. nov., a non-pigmented, extremely halophilic archaeon from a deep-sea, hypersaline anoxic basin of the Red Sea, and emended description of the genus *Halorhabdus*. *Int J Syst Evol Microbiol* **58:** 215–220.

Antunes, A., Rainey, F.A., Wanner, G., Taborda, M., Pätzold, J., Nobre, M.F., *et al*. (2008) A new lineage of halophilic, wall-less, contractile bacteria from a brine-filled deep of the Red Sea. *J Bacteriol* **190:** 3580–3587.

Antunes, A., Alam, I., El Dorry, H., Siam, R., Robertson, A., Bajic, V.B., and Stingl, U. (2011) Genome sequence of *Haloplasma contractile*, an unusual contractile bacterium from a deep-sea anoxic brine lake. *J Bacteriol* **193:** 4551–4552.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res* **37:** D233–D238.

Chamot-Rooke, N., Rabaute, A., and Kreemer, C. (2005) Western Mediterranean Ridge mud belt correlates with active shear strain at the prism-backstop geological contact. *Geology* **33:** 861–864.

Combet, C., Jambon, M., Deléage, G., and Geourjon, C. (2002) Geno3D: automatic comparative molecular modelling of protein. *Bioinformatics* **18:** 213–214.

Creamer, T.P., Srinivasan, R., and Rose, G.D. (1997) Modeling unfolded states of proteins and peptides. II. Backbone solvent accessibility. *Biochemistry* **36:** 2832–2835.

Daffonchio, D., Borin, S., Brusa, T., Brusetti, L., van der Wielen, P.W., Bolhuis, H., *et al*. (2006) Biodeep scientific party. Stratified prokaryote network in the oxic-anoxic transition of a deep-sea halocline. *Nature* **440:** 203–207.

De Corte, D., Sintes, E., Yokokawa, T., Reinthaler, T., and Herndl, G.J. (2012) Links between viruses and prokaryotes throughout the water column along a North Atlantic latitudinal transect. *ISME J* **6:** 1566–1577.

De Lange, G.J., and Ten Haven, H.L. (1983) Recent sapropel formations in the eastern Mediterranean. *Nature* **305:** 797–798.

De Vos, D., Xu, Y., Hulpiau, P., Vergauwen, B., and Van Beeumen, J.J. (2007) Structural investigation of cold activity and regulation of aspartate carbamoyltransferase from the extreme psychrophilic bacterium *Moritella profunda*. *J Mol Biol* **365:** 379–395.

Del Pozo, M.V., Fernández-Arrojo, L., Gil-Martínez, J., Montesinos, A., Chernikova, T.N., Nechitaylo, T.Y., *et al*. (2012) Microbial β-glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail. *Biotechnol Biofuels* **5:** 73.

Delgado-García, M., Valdivia-Urdiales, B., Aguilar-González, C.N., Contreras-Esquivel, J.C., and Rodríguez-Herrera, R. (2012) Halophilic hydrolases as a new tool for the biotechnological industries. *J Sci Food Agric* **92:** 2575–2780.

Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **32:** W665–W667.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res* **34:** W116–W118.

Eloe, E.A., Shulse, C.N., Fadrosh, D.W., Williamson, S.J., Allen, E.E., and Bartlett, D.H. (2011) Compositional differences in particle-associated and free-living microbial assemblages from an extreme deep-ocean environment. *Environ Microbiol Rep* **3:** 449–458.

Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62:** 72–82.

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

Ferrer, M., Werner, J., Chernikova, T.N., Bargiela, R., Fernández, L., La Cono, V., *et al*. (2012) Unveiling microbial life in the new deep-sea hypersaline Lake Thetis. Part II: a metagenomic study. *Environ Microbiol* **14:** 268–281.

Hall, T.A. (1999) Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41:** 95–98.

Hallsworth, J.E., Yakimov, M.M., Golyshin, P.N., Gillion, J.L., Auria, D., de Lima, G., *et al*. (2007) Limits of life in $MgCl_2$-containing environments: chaotropicity defines the window. *Environ Microbiol* **9:** 801–813.

Harrison, J.P., Gheeraert, N., Tsigelnitskiy, D., and Cockell, C.S. (2013) The limits for life under multiple extremes. *Trends Microbiol* **21:** 204–212.

Kabsch, W. (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66:** 125–132.

Kourist, R., Jochens, H., Bartsch, S., Kuipers, R., Padhi, S.K., Gall, M., *et al*. (2010) The alpha/beta-hydrolase fold 3DM database (ABHDB) as a tool for protein engineering. *Chembiochem* **11:** 1635–1643.

Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60:** 2256–2268.

La Cono, V., Smedile, F., Bortoluzzi, G., Arcadi, E., Maimone, G., Messina, E., *et al*. (2011) Unveiling microbial life in new deep-sea hypersaline Lake Thetis. Part I: prokaryotes and environmental settings. *Environ Microbiol* **13:** 2250–2268.

Lai, K.K., Stogios, P.J., Vu, C., Xu, X., Cui, H., Molloy, S., *et al*. (2011) An inserted α/β subdomain shapes the catalytic pocket of *Lactobacillus johnsonii* cinnamoyl esterase. *PLoS ONE* **6:** e23269.

Lukashin, A.V., and Borodovsky, M. (1998) GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res* **26:** 1107–1115.

Madern, D., Ebel, C., and Zaccai, G. (2000) Halophilic adaptation of enzymes. *Extremophiles* **4:** 91–98.

Martínez-Martínez, M., Alcaide, M., Tchigvintsev, A., Reva, O., Polaina, J., Bargiela, R., *et al*. (2013) Biochemical diversity of carboxyl esterases and lipases from Lake Arreo (Spain): a metagenomic approach. *Appl Environ Microbiol* **79:** 3553–3562.

MEDRIFF Consortium (1995) Three brine lakes discovered in the seafloor of the eastern Mediterranean. *EOS Trans AGU* **76:** 313–318.

Menigaud, S., Mallet, L., Picord, G., Churlaud, C., Borrel, A., and Deschavanne, P. (2012) GOHTAM: a website for genomic origin of horizontal transfers, alignment and metagenomics. *Bioinformatics* **28:** 1270–1271.

Michels, P.C., and Clark, D.S. (1997) Pressure-enhanced activity and stability of a hyperthermophilic protease from a deep-sea methanogen. *Appl Environ Microbiol* **63:** 3985–3991.

Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution – from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* **62:** 859–866.

Nagata, T., Tamburini, C., Aristegui, J., Baltar, F., Bochdansky, A.B., Fonda-Umani, S., *et al*. (2010) Emerging concepts on microbial processes in the bathypelagic ocean – ecology, biogeochemistry and genomics. *Deep Sea Res Part II Top Stud Oceanogr* **57:** 1519–1536.

Nechitaylo, T.Y., Timmis, K.N., and Golyshin, P.N. (2009) *'Candidatus Lumbricincola*', a novel lineage of uncultured *Mollicutes* from earthworms of family Lumbricidae. *Environ Microbiol* **11:** 1016–1026.

Pace, N., and Scholtz, J.M. (1997) Measuring the conformational stability of a protein, Chapter 12. In *Protein Structure: A practical Approach*. Creighton, T.E. (ed.). Heidelberg, Germany: IRL Press at Oxford University Press, pp. 299–322.

Pietra, F. (2012) On 3LEZ, a deep-sea halophilic protein with in vitro class-a β-lactamase activity: molecular-dynamics, docking, and reactivity simulations. *Chem Biodivers* **9:** 2659–2684.

Puspita, I.D., Kamagata, Y., Tanaka, M., Asano, K., and Nakatsu, C.H. (2012) Are uncultivated bacteria really uncultivable? *Microbes Environ* **27:** 356–366.

Reed, C.J., Lewis, H., Trejo, E., Winston, V., and Evilia, C. (2013) Protein adaptations in archaeal extremophiles. *Archaea* **2013:** 373275.

Reyes-Duarte, D., Ferrer, M., and García-Arellano, H. (2012) Functional-based screening methods for lipases esterases and phospholipases in metagenomic libraries. In *Lipases and Phospholipases: Methods and Protocols Methods in Molecular Biology*, Vol. 861. Sandoval, G. (ed.). New York, NY, USA: © Springer Science+Business Media, pp. 101–113.
doi:10.1007/978-1-61779-600-5_6.

Schlitzer, R. (2010) *Ocean data view*. SeaDataNet, Pan-European Infrastructure for Oceand & Marine Data Management, IFREMER/SISMER, Plouzane, France. URL http://odv.awi.de.

Schmid, F.Z. (1997) Optical spectroscopy to characterize protein conformation, Chapter 11. In *Protein Structure: A Practical Approach*. Creighton, T.E. (ed.). Oxford, UK: IRL Press, pp. 261–298.

Shin, D.S., Didonato, M., Barondeau, D.P., Hura, G.L., Hitomi, C., Berglund, J.A., *et al*. (2009) Superoxide dismutase from the eukaryotic thermophile *Alvinella pompejana*: structures, stability, mechanism, and insights into amyotrophic lateral sclerosis. *J Mol Biol* **385:** 1534–1555.

Shirai, T., Hung, V.S., Morinaka, K., Kobayashi, T., and Ito, S. (2008) Crystal structure of GH13 alpha-glucosidase GSJ from one of the deepest sea bacteria. *Proteins* **73:** 126–133.

Sineva, E.V., and Davydov, D.R. (2010) Cytochrome P450 from *Photobacterium profundum* SS9, a piezophilic bacterium, exhibits a tightened control of water access to the active site. *Biochemistry* **49:** 10636–10646.

Smedile, F., Messina, E., La Cono, V., Tsoy, O., Monticelli, L.S., Borghini, M., *et al*. (2013) Metagenomic analysis of hadopelagic microbial assemblages thriving at the deepest part of Mediterranean Sea, Matapan-Vavilov Deep. *Environ Microbiol* **15:** 167–182.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24:** 1596–1599.

Wallmann, K., Suess, E., Westbrook, G.H., Winckler, G., and Cita, M.B. (1997) Salty brines on the Mediterranean Sea floor. *Nature* **387:** 31–32.

Wang, H., Gong, Y., Xie, W., Xiao, W., Wang, J., Zheng, Y., *et al*. (2011) Identification and characterization of a novel thermostable gh-57 gene from metagenomic fosmid library of the Juan de Fuca Ridge hydrothemal vent. *Appl Biochem Biotechnol* **164:** 1323–1338.

Werner, J., Ferrer, M., Michel, G., Mann, A.J., Huang, S., Juarez, S., *et al*. (2014) *Halorhabdus tiamatea*: proteogenomics and glycosidase activity measurements identify the first cultivated euryarchaeon from a deep-sea anoxic brine lake as potential polysaccharide degrader. *Environ Microbiol* **16:** 2525–2537.

Xu, M., Xiao, X., and Wang, F. (2008) Isolation and characterization of alkane hydroxylases from a metagenomic library of Pacific deep-sea sediment. *Extremophiles* **12:** 255–262.

Yakimov, M.M., Giuliano, L., Cappello, S., Denaro, R., and Golyshin, P.N. (2007) Microbial community of a hydrothermal mud vent underneath the deep-sea anoxic brine lake Urania (eastern Mediterranean). *Orig Life Evol Biosph* **37:** 177–188.

Yakimov, M.M., La Cono, V., Slepak, V.Z., La Spada, G., Arcadi, E., Messina, E., *et al*. (2013) Microbial life in the Lake Medee, the largest deep-sea salt-saturated formation. *Sci Rep* **3:** 3554.

Zeng, X., Birrien, J.L., Fouquet, Y., Cherkashov, G., Jebbar, M., Querellou, J., *et al*. (2009) *Pyrococcus* CH1, an obligate piezophilic hyperthermophile: extending the upper pressure-temperature limits for life. *ISME J* **3:** 873–876.

Supporting information

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

Fig. S1. pH profiles of wild-type enzymes. The specific activities were calculated in triplicate as described in the Experimental procedures. The standard deviation (SD) is shown. The 100% activity is as shown in Table 1. Note: due to protein instability at low pH, the pH profile for MGS-M5 could not be obtained; preliminary test reactions indicated pH 8.0 (50 mM Tris-HCl) as being the most suitable buffer for activity

determinations, and this value was used as the standard buffer for this enzyme.

Fig. S2. Substrate profiles of the enzymes with a set of structurally diverse substrates. The specific activities were calculated in triplicate as described in the Experimental procedures, using the standard assay conditions (see also summary conditions in Table 1). Mean values (in log scale) are given. The standard deviation (SD) is not shown due to the logarithmic scale, but it is ≤ 0.23%. Note: using standard conditions for MGS-M3, no activity was detected using *p*NP-α-glucose, *p*NP-α-maltooligosaccharides (C2 to C6), *p*NP-α-galactose, *p*NP-β-galactose, *p*NP-α-xylose, *p*NP-β-arabinopyranose, *p*NP-α-rhamnose, *p*NP-α-mannose, *p*NP-β-mannose, *p*NP-α-fucose, *p*NP-β-glucuronide, carboxymethyl cellulose, laminarin, lichenan and crystalline cellulose.

Fig. S3. Comparison of the structures of MGS enzymes crystallized in this study and their structural homologues revealed by structure similarity searches. Enzymes are shown as cartoon representations. Arrows refer to the locations of the catalytic serine (for esterases in top row) or the NADPH/NADH and substrate-binding sites (for bottom row enzymes).

Fig. S4. Comparison of putative active sites of MGS enzymes crystallized in this study and their structural homologues revealed by structure similarity searches. Sticks are shown for bound substrates and residues predicted to participate in catalytic reactions and/or interact with substrates. The catalytic triads for the esterases in the top row are labelled. The substrate-binding canals for MGS-M2 and HsaD are also shown in solvent-accessible surface representations, coloured by electrostatic potential, highlighting the disparate charge features, which are shown under their respective cartoon images. The non-NADH/NADPH substrate-binding residues for 2,5-DKGR and LDH, plus the equivalent residues in MGS-M4 and MGS-5, are labelled.

Table S1. General features of reported enzymes isolated from deep-sea regions. The data are based on bibliographic records that are specifically cited.

Table S2. Hydrochemistry of selected deep-sea regions, deep-sea libraries and esterase screening statistics.

Table S3. General features and residues potentially involved in catalysis, substrate recognition and thermal and halophilic adaptations in the proteins investigated.

A. General features of esterase-like proteins.

B. Percentage of identity between esterases with the α/β hydrolase fold as determined by the Matcher (EMBOSS package). Matches/alignment lengths (% identity) are specifically indicated.

C. General features of proteins characterized from the *Medee* basin other than esterases.

D. Thermophilic adaptations for proteins with determined crystal structures.

E. Halophilic adaptations for proteins with determined crystal structures.

F. High-pressure adaptations for proteins with determined crystal structures.

Table S4. Compositional similarities between the DNA fragments containing the genes of interest and bacterial genomes as shown by GOHTAM and TBLASTX analyses.

Table S5. List of primers used in the study.

Table S6. X-ray diffraction statistics.

Appendix S1. Supplementary Results and Discussion.