1 2	Diversity and evolution of nitric oxide reduction
3	Ranjani Murali ^{1,a*} , Laura A. Pace ² , Robert A. Sanford ³ , Lewis M. Ward ⁴ , Mackenzie
4	Lynes ⁵ , Roland Hatzenpichler ^{5,6} , Usha F. Lingappa ⁷ , Woodward W. Fischer ⁸ , Robert B.
5	Gennis ⁹ , James Hemp ^{2,8,a*}
6	
7	¹ Division of Biology and Biological Engineering, California Institute of Technology,
8	Pasedena, CA, USA
9	² Metrodora Institute, Salt Lake City, UT, USA
10	³ Department of Geology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
11	⁴ Smith College, Northampton, MA, USA
12	⁵ Department of Chemistry and Biochemistry, Thermal Biology Institute, and Center for
13	Biofilm Enginering, Montana State University, MT, USA
14	⁶ Department of Microbiology and Cell Biology, Montana State University, MT, USA
15	⁷ Department of Plant and Microbial Biology, University of California, Berkeley, CA,
16	USA
17	⁸ Division of Geological and Planetary Sciences, California Institute of Technology,
18	Pasadena, USA
19	⁹ Department of Biochemistry, University of Illinois, Urbana-Champaign, Urbana, USA
20	^a James Hemp and Ranjani Murali contributed equally to the work in this manuscript.
21	* Corresponding authors: E-mail: james.hemp@metrodora.org, <u>m.ranjani@gmail.com</u>
22	
23	
24	

25 Nitrogen is an essential element for life, with the availability of fixed nitrogen 26 limiting productivity in many ecosystems. The return of oxidized nitrogen species to 27 the atmospheric N_2 pool is predominately catalyzed by microbial denitrification $(NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2)^1$. Incomplete denitrification can produce N₂O as 28 29 a terminal product, leading to an increase in atmospheric N₂O, a potent greenhouse and ozone-depleting gas^2 . The production of N₂O is catalyzed by nitric oxide 30 reductase (NOR) members of the heme-copper oxidoreductase (HCO) superfamily³. 31 32 Here we use phylogenomics to identify a number of previously uncharacterized 33 HCO families and propose that many of them (eNOR, sNOR, gNOR, and nNOR) 34 perform nitric oxide reduction. These families have novel active-site structures and 35 several have conserved proton channels, suggesting that they might be able to 36 couple nitric oxide reduction to energy conservation. We isolated and biochemically 37 characterized a member of the eNOR family from Rhodothermus marinus, verifying 38 that it performs nitric oxide reduction both in vitro and in vivo. These newly 39 identified NORs exhibit broad phylogenetic and environmental distributions, 40 expanding the diversity of microbes that can perform denitrification. Phylogenetic 41 analyses of the HCO superfamily demonstrate that nitric oxide reductases evolved 42 multiple times independently from oxygen reductases, suggesting that complete 43 denitrification evolved after aerobic respiration.

The HCO superfamily is extremely diverse, with members playing crucial roles in both aerobic (oxygen reductases) and anaerobic (nitric oxide reductases) respiration. The superfamily currently consists of at least three oxygen reductase families (A, B and C) and three NOR families (cNOR, qNOR, and qCu_ANOR)⁴. The oxygen reductases

catalyze the reduction of O_2 to water $(O_2 + 4e_{out} + 4H_{in}^+ + nH_{in}^+ \rightarrow 2H_2O + nH_{out}^+)$ and 48 share a conserved reaction mechanism^{5,6}, where three of the electrons required to reduce 49 50 O_2 are provided by the active-site metals, heme-Fe and Cu_B , while the fourth electron is derived from a redox-active cross-linked tyrosine cofactor⁷ (Figure 1). The free energy 51 of the reaction is converted to a transmembrane proton electrochemical gradient by two 52 different mechanisms, charge separation across the membrane and proton $pumping^{8-10}$. 53 54 Both the chemical and pumped protons are taken up from the electrochemically negative 55 side of the membrane (bacterial cytoplasm) via conserved proton-conducting channels 56 that are comprised of conserved polar residues and internal water molecules. The 57 different oxygen reductase families exhibit differential proton pumping stochiometries $(n=4 \text{ for the A-family, and } n=2 \text{ for the B and C-families})^{8-10}$, depending on the details of 58 59 their conserved proton channels. The oxygen reductases also vary in their secondary 60 subunits that function as redox relays from electron donors to the active-site, with the A and B-families utilizing a Cu_A-containing subunit^{11–13} and the C-family containing one or 61 more cytochrome c subunits¹⁴ (Figure 1). 62

63

Nitric oxide reductases (NORs) catalyze the reduction of nitric oxide to nitrous oxide (2NO + $2H_{out}^{+}$ + $2e_{out}^{-}$ + $nH_{in}^{+} \rightarrow N_2O$ + H_2O + nH_{in}^{+}). Nitric oxide reduction is only a 2electron reaction, so it does not require the cross-linked tyrosine cofactor for catalysis. There are currently three known HCO NORs, the cNOR, qNOR, and qCu_ANOR. The cNOR and qNOR families have a four amino acid coordinated Fe_B ion in their activesites, in contrast to the three amino acid coordinated Cu_B found in oxygen reductases^{15,16}. The cNOR and qNOR families are closely related to the C-family oxygen reductases and 71 likely evolved from this oxygen reductase family. In accordance with this relationship, 72 the cNORs have a secondary cytochrome *c*-containing subunit, while in the qNORs the 73 two subunits corresponding to those in the cNORs have been fused to a single subunit that lacks the heme c binding. Although, the qNOR from N. meningitidis¹⁶ is proposed to 74 75 take up protons from the cytoplasm for NO reduction, neither the cNOR nor qNOR have 76 conserved residues that could form a proton channel from the cytoplasm, therefore it is 77 unlikely that they conserve energy via either charge separation or proton pumping (n=0 78 for the cNOR and qNOR families). Phylogenomic analysis shows that the qCu_ANOR from *Bacillus azotoformans*^{17,18} is unrelated to the qNOR family, and has been 79 80 reclassified here as the bNOR family. The bNOR active-site structure is fundamentally 81 different than those from the cNOR and qNOR families (Figure 1). It also contains a 82 conserved proton channel that is very similar to that found in the B-family oxygen reductases (Figure 1), and has been shown to be $electrogenic^{18}$. This has important 83 consequences for the efficiency of energy conservation associated with denitrification¹⁹. 84

85

86 Novel heme-copper oxygen reductase homologs

Phylogenomic analyses of genomic and metagenomic data have identified at least six new families belonging to the HCO superfamily (**Figure 2**). All of these families are missing the active-site tyrosine, suggesting that they do not catalyze O₂ reduction. Furthermore, their active-sites exhibit structural features never seen before within the superfamily (**Figure 1**). One of these families is closely related to qNOR and has been proposed to be a nitric oxide dismutase contributing to O₂ production in *'Candidatus* Methylomirabilis oxyfera'²⁰. Another family is closely related to cNOR and might be a sulfide and acetylene insensitive nitrous oxide reductase^{21,22}. The remaining four families (eNOR, sNOR, nNOR and gNOR) are closely related to the B-family of oxygen reductases (**Figure 2**) and encode for homologs of the Cu_A-containing secondary subunits, consistent with this evolutionary relationship (**Figure 1, Table S1**). Based on modeled active-site structures and genomic context we propose that these four families perform nitric oxide reduction (**Figure 1**).

100

101 Biochemical characterization of eNOR

102 To validate these predictions we isolated and biochemically characterized a member of 103 the eNOR family from *Rhodothermus marinus* DSM 4252, a thermophilic member of the 104 Bacteroidetes phylum. Rhodothermus marinus was originally classified as a strict 105 aerobe²³, but its genome encodes a periplasmic nitrate reductase (NapA), two nitrite 106 reductases (NirS and NirK), and a nitrous oxide reductase (NosZ), suggesting that it may 107 also be capable of denitrification (Extended Data Figure 1). Denitrification was not 108 observed under strictly anaerobic conditions, however under microoxic conditions isotopically labeled ${}^{15}NO_3$ was converted to ${}^{30}N_2$ (Extended Data Figure 2) 109 110 demonstrating that *R. marinus* DSM 4252 is capable of complete aerobic denitrification $(NO_3 \rightarrow N_2)$. Blockage of the nitrous oxide reductase with acetylene led to the 111 112 accumulation of N₂O (Figure 3), suggesting that a nitric oxide reductase was also present 113 in *R. marinus* DSM 4252. No known NORs (cNOR, qNOR, qCuANOR) were found in 114 the genome. However, R. marinus DSM 4252 does encode for a member of the eNOR 115 family (Extended Data Figure 1).

116

117 Isolation and biochemical characterization of the R. marinus DSM 4252 eNOR protein verified that it catalyzes NO reduction (at 25° C, $k_{cat} = 0.68 \pm 0.21$ NO.s⁻¹ (n = 4)) 118 119 (Figure 3). eNOR was unable to catalyze O_2 reduction using various electron donors 120 (Extended Data Figure 3), clearly showing that it functions solely as a nitric oxide 121 reductase. UV-Vis spectroscopy and heme characterization via mass spectrometry 122 demonstrated that the R. marinus DSM 4252 eNOR contains a novel modified heme a 123 that is used in both heme sites (Figure 3 and Extended Data Figures 3 and 4). A 124 member of the eNOR family was previously isolated from the aerobic denitrifier Magnetospirillum magnetotacticum MS-1^{16,17}, however its function was never 125 determined. The UV-Vis spectra of the *M. magnetotacticum* $eNOR^{17}$ is identical to the *R.* 126 127 *marinus* eNOR, suggesting that the modified heme *a* is a general feature of the family. 128 Mass spectrum analysis of the hemes extracted from eNOR suggest that this heme is A_s, a 129 previously isolated heme a with a hydroxyethylgeranylgeranyl side chain first identified in the B-family oxygen reductase from *Sulfolobus acidocaldarius*²⁴. Many eNOR operons 130 contain a CtaA homolog, an O₂-dependent enzyme that converts heme o to heme a^{25} . 131 This is consistent with the observation that eNOR requires microoxic conditions for 132 133 expression and suggests that aerobic denitrification might be common in nature.

134

135 Active-site features of novel NORs

In addition to the experimental evidence that both eNOR and bNOR enzymes are NO reductases, there is good evidence that the other newly identified families also perform nitric oxide reduction. The sNOR family has the same active-site structure as the bNOR family, strongly suggesting that it also performs nitric oxide reduction. However, the

140 sNOR and bNOR families are not closely related, providing an example of convergent 141 evolution of active-site structures within the HCO superfamily (Figures 1 and 2). 142 Another example of convergent evolution is the nNOR family. It has the same conserved 143 active-site residues as the cNOR and qNOR families (Figure 1), but is very distantly 144 related to them (nNOR is related to the B-family, while cNOR and qNOR are related to 145 the C-family). Interestingly, the low-spin heme in nNOR is ligated by a histidine and methionine, which likely raises its redox potential by $\sim 150 \text{ mV}^{26}$. This is similar to a 146 147 modification found in some eNORs, where the low spin heme is ligated by histidine and 148 lysine. The gNOR is the first example of a HCO family that has replaced one of the 149 active-site histidines, residues completely conserved in all other families. A bioinorganic mimic of the gNOR active-site exhibited nitric oxide reduction capability²⁷, suggesting 150 151 that it is likely a functional NOR in vivo. The gNOR has a secondary subunit with a 152 cupredoxin fold that is missing the residues required to bind Cu_A, similar to the quinol 153 oxygen reductase from E. coli. Conserved residues that could bind quinol have been 154 identified in gNOR, suggesting that it is a quinol nitric oxide reductase similar to qNOR.

155

The biochemically characterized (eNOR and bNOR) and proposed (sNOR and gNOR) HCO NORs have active-sites that differ significantly from those utilized by cNOR and qNOR (**Figure 1**). This demonstrates that while oxygen reduction chemistry is constrained to require a redox active tyrosine cofactor, multiple HCO active-site structures are compatible with nitric oxide reduction chemistry. Interestingly, in the currently characterized HCOs Cu_B is utilized for O_2 reduction chemistry and Fe_B for NO reduction chemistry. If this pattern is verified for the other predicted NOR families it would suggest that the chemistry performed by HCOs is partially determined by theelectronic properties of the active-site metal.

165

166 Bioenergetics of novel denitrification pathways

167 Although both denitrification and aerobic respiration are highly exergonic processes, 168 most of the enzymes in the denitrification pathway are not coupled to energy conservation, making denitrification significantly less efficient than aerobic respiration²⁸. 169 170 In the HCO oxygen reductases conserved proton channels deliver protons from the 171 cytoplasm to the active-site for chemistry. These same channels are used to pump protons 172 to the periplasmic side. The cNORs and qNORs do not have conserved proton channels from the cytoplasm^{15,29}, making them incapable of conserving energy. The eNOR has 173 174 conserved residues that closely resemble those found in the proton-conducting K-channel within the B-family of oxygen reductases³⁰ (Supplementary Table 1, Extended Data 175 176 Figure 5). The sNOR family also has conserved residues in the K-channel region, 177 however they are different than those found in the eNOR and bNOR families. 178 Interestingly, the nNOR family, which has the same active site as cNOR and qNOR, also 179 has a conserved proton channel (Supplementary Table 1, Extended Data Figure 5), 180 suggesting that the lack of a proton channel in the cNOR and qNOR may not be due to energy constraints³¹. The conserved proton channels in the eNOR, bNOR, sNOR, and 181 182 nNOR families would allow them to conserve energy via charge separation, and 183 potentially by proton pumping. Characterization of these new NOR families will be 184 critical for a full understanding of the mechanism of proton pumping in the HCO 185 superfamily, one of the most important unanswered question in bioenergetics.

186 Distribution and environmental relevance of NORs

187 The new HCO NOR families have broad phylogenetic and environmental distributions 188 (Table 1, Supplementary Tables 2, 3). The eNOR, sNOR, gNOR, and nNOR families 189 are found in both Bacteria and Archaea. To date, the bNOR family has been found only 190 in the Bacillales order of Firmicutes (Supplementary Table 2). Phylogenetic analysis of 191 metagenomic data shows that the majority of eNOR, sNOR, gNOR, and nNOR enzymes 192 are from uncharacterized organisms, suggesting that many more organisms are capable of 193 nitric oxide reduction than previously suspected. Furthermore, the new HCO NOR 194 families are found in a wide variety of environments (**Table 1**, **Supplementary Table 2**). 195 The sNOR are found in the majority of ammonia oxidizing bacteria sequenced to date 196 and likely play a role in this process. The gNORs are predominantly found in sulfidic 197 environments and may be an adaptation that allows for denitrification in the presence of 198 sulfide, which inhibits other NOR families. The eNOR family is very common in nature, 199 having a broad distribution similar to the cNOR and qNOR families. The eNOR family 200 appears to play key roles in a number of important microbiological processes. They are 201 found in many strains of *Candidatus Accumulibacter phosphatis*, a microbe utilized for 202 phosphate accumulation in wastewater treatment plants during enhanced biological 203 phosphorus removal. The eNOR is highly expressed in transcriptomic datasets from 204 these facilities, demonstrating that Accumulibacter phosphatis is capable of complete denitrification in situ³². eNOR has also been found in microbes capable of performing 205 206 autotrophic nitrate reduction coupled to Fe(II) oxidation (NRFO). Gallionellaceae KS 207 and related strains express an eNOR under denitrifying conditions, suggesting that an individual organism is capable of complete NRFO³³. eNOR is also commonly found in 208

209 hypersaline environments (Supplementary Table 2) where is might play a role in the210 adaptation of denitrification to high salt conditions.

211

212 Many organisms encode NORs from multiple families (e.g., Ca. Methylomirabilis 213 oxyfera has a qNOR, sNOR and gNOR, and Bacillus azotoformans has a qNOR, sNOR, 214 and bNOR). This suggests that selection for different enzymatic properties (NO affinity, 215 enzyme kinetics, energy conservation, or sensitivity to inhibitors) or the concentration of 216 O₂ may be important factors in determining their distribution, similar to what is observed for the HCO oxygen reductase families⁸. Analysis of the presence of denitrification genes 217 218 (nitrite reductases, nitric oxide reductases, and nitrous oxide reductases) within 219 sequenced genomes indicates that many more organisms are capable of complete 220 denitrification than previously realized. Our current understanding of the diversity of 221 organisms capable of performing denitrification is far from complete.

222

223 Our evolutionary analysis shows that nitric oxide reductases have evolved many times 224 independently from oxygen reductases (Figure 2). The current data show that NORs 225 have originated from both the B and C-families of oxygen reductases, enzymes that are 226 adapted to low O₂ environments. These oxygen reductases can reduce NO at high concentrations in vitro³⁴ so it is not surprising that small evolutionary modifications 227 228 would lead to enzymes capable of NO reduction at the lower NO concentrations 229 produced during denitrification. The fact that NO reductases are derived from oxygen 230 reductases strongly suggests that complete denitrification evolved after aerobic

- respiration. This places important constraints on the nitrogen cycle before the rise ofoxygen.
- 233

240

241

234 Competing Interests

- 235 The authors declare no competing interests.
- 236 Acknowledgment

237 We would like to thank NIH for funding this research (Grant # , Principle Investigator:

238 Dr. Robert Gennis). We thank Sylvia Choi for providing pure ba_3 oxygen reductase from 239 *Thermus thermophilus* to use as a control for oxygen reductase assays and for heme

extraction, Lici Schurig-Briccio for guidance in performing nitric oxide reductase assays

with the Clark Electrode and Peter Yau at the University of Illinois' Mass spectrometric

242 facility for protein identification. We thank Alon Philosof and Connor Skennerton for

valuable discussions on bioinformatics analysis. A portion of this research was performed

244 under the Facilities Integrating Collaborations for User Science (FICUS) initiative (award

503546 to R.H.) and used resources at the DOE Joint Genome Institute and theEnvironmental Molecular Sciences Laboratory, which are DOE Office of Science User

- 247 Facilities. Both facilities are sponsored by the Office of Biological and Environmental
- 248 Research and operated under Contract Nos. DE-AC02-05CH11231 (JGI) and DE-AC05-
- 249 76RL01830 (EMSL).

250

251 References

- 252
- 1. Thamdrup, B. New Pathways and Processes in the Global Nitrogen Cycle. *Annu. Rev.*
- 254 Ecol. Evol. Syst. 43, 407–428 (2012).

255	2.	Klotz, N	Л. (G. &	Stein.	L.	Y.	Nitrifier	genomics and	d ev	olution	of th	ne nitr	ogen	cvc	ele.

- 256 *FEMS Microbiology Letters* **278**, 146–156 (2008).
- 257 3. Zumft, W. G. Nitric oxide reductases of prokaryotes with emphasis on the
- respiratory, heme-copper oxidase type. J. Inorg. Biochem. 99, 194–215 (2005).
- 259 4. Hemp, J. & Gennis, R. B. Diversity of the Heme–Copper Superfamily in Archaea:
- 260 Insights from Genomics and Structural Modeling. *Bioenergetics* 1–31 (2008).
- 261 5. Pereira, M. M., Sousa, F. L., Veríssimo, A. F. & Teixeira, M. Looking for the
- 262 minimum common denominator in haem–copper oxygen reductases: Towards a
- 263 unified catalytic mechanism. *Biochim. Biophys. Acta BBA Bioenerg.* 1777, 929–934
- 264 (2008).
- 265 6. Hemp, J., Christian, C., Barquera, B., Gennis, R. B. & Martínez, T. J. Helix
- 266 Switching of a Key Active-Site Residue in the Cytochrome *cbb*₃ Oxidases.
- 267 Biochemistry 44, 10766–10775 (2005).
- 268 7. Hemp, J. et al. Evolutionary migration of a post-translationally modified active-Site
- residue in the proton-pumping heme-copper oxygen reductases. *Biochemistry* **45**,
- 270 15405–15410 (2006).
- 8. Han, H. *et al.* Adaptation of aerobic respiration to low O₂ environments. *Proc. Natl. Acad. Sci.* 108, 14109–14114 (2011).
- 9. Hemp, J. et al. Comparative Genomics and Site-Directed Mutagenesis Support the
- 274 Existence of Only One Input Channel for Protons in the C-Family (cbb3 Oxidase) of
- Heme–Copper Oxygen Reductases. *Biochemistry* **46**, 9963–9972 (2007).
- 10. Chang, H.-Y., Hemp, J., Chen, Y., Fee, J. A. & Gennis, R. B. The cytochrome ba₃
- 277 oxygen reductase from *Thermus thermophilus* uses a single input channel for proton

- delivery to the active site and for proton pumping. *Proc. Natl. Acad. Sci.* **106**, 16169–
- 279 16173 (2009).
- 280 11. Qin, L., Hiser, C., Mulichak, A., Garavito, R. M. & Ferguson-Miller, S. Identification
- of conserved lipid/detergent-binding sites in a high-resolution structure of the
- 282 membrane protein cytochrome *c* oxidase. *Proc. Natl. Acad. Sci.* **103**, 16117–16122
- 283 (2006).
- 12. Soulimane, T. *et al.* Structure and mechanism of the aberrant *ba*₃-cytochrome *c*
- 285 oxidase from *Thermus thermophilus*. *EMBO J.* **19**, 1766–1776 (2000).
- 13. Tiefenbrunn, T. *et al.* High resolution structure of the ba_3 cytochrome c oxidase from
- 287 *Thermus thermophilus* in a lipidic environment. *PLOS ONE* **6**, 1-12 (2011).
- 288 14. Buschmann, S. *et al.* The Structure of *cbb*₃ Cytochrome Oxidase Provides Insights
- into Proton Pumping. *Science* **329**, 327–330 (2010).
- 15. Hino, T. *et al.* Structural basis of biological N₂O generation by bacterial nitric oxide
 reductase. *Science* 330, 1666–1670 (2010).
- 292 16. Gonska, N. et al. Characterization of the quinol-dependent nitric oxide reductase
- from the pathogen *Neisseria meningitidis*, an electrogenic enzyme. *Sci. Rep.* 8, 3637
 (2018).
- 295 17. Suharti, Strampraad, M. J., Schröder, I. & de Vries, S. A novel copper A containing
- 296 menaquinol NO reductase from *Bacillus azotoformans*. *Biochemistry* 40, 2632–2639
 297 (2001).
- 298 18. Al-Attar, S. & de Vries, S. An electrogenic nitric oxide reductase. *FEBS Lett.* 589,
 2050–2057 (2015).

300	19. Strohm, T. O., Griffin, B., Zumft, W. G. & Schink, B. Growth Yields in Bacterial
301	Denitrification and Nitrate Ammonification. Appl. Environ. Microbiol. 73, 1420-
302	1424 (2007).
303	20. Hanson, B. T. & Madsen, E. L. In situ expression of nitrite-dependent anaerobic
304	methane oxidation proteins by Candidatus Methylomirabilis oxyfera co-occurring
305	with expressed anammox proteins in a contaminated aquifer. Environ. Microbiol.
306	<i>Rep.</i> 7 , 252–264 (2015).
307	21. Jones, A. M., Hollocher, T. C. & Knowles, R. Nitrous oxide reductase of Flexibacter
308	canadensis: a unique membrane-bound enzyme. FEMS Microbiol. Lett. 92, 205-209
309	(1992).
310	22. Jones, A. M., Adkins, A. M., Knowles, R. & Rayat, G. R. Identification of a
311	denitrifying gliding bacterium, isolated from soil and able to reduce nitrous oxide in
312	the presence of sulfide and acetylene, as Flexibacter canadensis. Can. J. Microbiol.
313	36 , 765–770 (1990).
314	23. Alfredsson, G. A., Kristjansson, J. K., Hjrleifsdottir, S. & Stetter, K. O.
315	Rhodothermus marinus, gen. nov., sp. nov., a Thermophilic, Halophilic Bacterium
316	from Submarine Hot Springs in Iceland. J. Gen. Microbiol. 299-306 (1988).
317	24. Lübben, M. & Morand, K. Novel prenylated hemes as cofactors of cytochrome
318	oxidases. Archaea have modified hemes A and O. J. Biol. Chem. 269, 21473-21479
319	(1994).
320	25. Brown, K. R., Allan, B. M., Do, P. & Hegg, E. L. Identification of Novel Hemes
321	Generated by Heme A Synthase: Evidence for Two Successive Monooxygenase
322	Reactions. Biochemistry 41, 10906–10913 (2002).

323	26. Fufezan, C., Zhang, J. & Gunner, M. R. Ligand preference and orientation in b- and
324	c-type heme-binding proteins. Proteins Struct. Funct. Bioinforma. 73, 690-704
325	(2008).
326	27. Lin, YW. et al. Introducing a 2-His-1-Glu nonheme iron center into myoglobin
327	confers nitric oxide reductase activity. J. Am. Chem. Soc. 132, 9970-9972 (2010).
328	28. Chen, J. & Strous, M. Denitrification and aerobic respiration, hybrid electron
329	transport chains and co-evolution. Biochim. Biophys. Acta BBA - Bioenerg. 1827,
330	136–144 (2013).
331	29. Matsumoto, Y. et al. Crystal structure of quinol-dependent nitric oxide reductase
332	from Geobacillus stearothermophilus. Nat. Struct. Mol. Biol. 19, 238–245 (2012).
333	30. Chang, HY. et al. Exploring the proton pump and exit pathway for pumped protons
334	in cytochrome ba3 from Thermus thermophilus. Proc. Natl. Acad. Sci. 109, 5259-
335	5264 (2012).
336	31. Blomberg, M. R. A. & Siegbahn, P. E. M. Why is the reduction of NO in cytochrome
337	c dependent nitric oxide reductase (cNOR) not electrogenic? Biochim. Biophys. Acta
338	BBA - Bioenerg. 1827, 826–833 (2013).
339	32. Camejo, P. Y., Oyserman, B. O., McMahon, K. D. & Noguera, D. R. Integrated Omic
340	Analyses Provide Evidence that a "Candidatus Accumulibacter phosphatis" Strain
341	Performs Denitrification under Microaerobic Conditions. mSystems 4, e00193-18.
342	33. Huang, YM., Straub, D., Blackwell, N., Kappler, A. & Kleindienst, S. Meta-omics
343	Reveal Gallionellaceae and Rhodanobacter Species as Interdependent Key Players
344	for Fe(II) Oxidation and Nitrate Reduction in the Autotrophic Enrichment Culture
345	KS. Appl. Environ. Microbiol. 87, e00496-21.

- 346 34. Loullis, A. & Pinakoulaki, E. Probing the nitrite and nitric oxide reductase activity of
- 347 *cbb*₃ oxidase: resonance Raman detection of a six-coordinate ferrous heme–nitrosyl
- 348 species in the binuclear b_3 /Cu_B center. Chem. Commun. 51, 17398–17401 (2015).
- 349

350 List of Figures

351 Figure 1. Comparison of HCO active sites. a) Active-site and proton channel properties 352 of the five characterized HCO families (A-family, B-family, C-family, cNOR, and 353 qNOR). The oxygen reductases all have an active-site composed of high-spin heme, a 354 redox-active cross-linked tyrosine cofactor, and a copper (Cu_B) ligated by three 355 histidines. The A-family has two conserved proton channels, whereas the B and C-356 families only have one. The active-sites of the nitric oxide reductases are composed of a 357 high-spin heme and an iron (Fe_B) that is ligated by three histidines and a glutamate. 358 Notably they are missing the tyrosine cofactor. The cNOR and qNOR are also missing 359 conserved proton channels, making them non-electrogenic. b) Sequence alignment of the 360 active-sites of the newly discovered HCO families that are related to the B-family. c) 361 Predicted active-sites and proton channels for the new HCO families. The eNOR, bNOR, 362 sNOR, and nNOR families contain completely conserved proton channels shown here as 363 arrows. The putative proton channel in the bNOR and eNOR families are highly similar 364 to the K-channel from the B-family oxygen reductase and are colored in red. The K-365 channel in the B-family is also similar to the K-channel in the A-family oxygen reductase 366 which is colored in dark red. The proton channel in the C-family is different from these 367 channels and is marked in yellow. The putative proton channels in sNOR and nNOR are 368 marked in cyan and differentiated from the other channels with a dashed black outline.

369

Figure 2. Evolution of nitric oxide reductases. Phylogenetic tree of HCO families. All
of the new NOR families are derived from oxygen reductase ancestors. Oxygen
reductases are in shades of blue, whereas nitric oxide reductases are in shades of yellow,
green and red.

374 Figure 3. Biochemical Characterization of the eNOR from *Rhodothermus marinus*.

a) UV-Vis spectrum of isolated eNOR indicates the presence of an unusual heme *a*signature at 589 nm. b) NO reductase activity was measured with the use of a Clark
electrode in the presence of TMPD and ascorbate as electron donor. c) N₂O accumulation
observed in a culture of *Rhodothermus marinus*, in the presence of acetylene d) N₂O
production by eNOR from *R. marinus* and qNOR from *Persephonella marina*.

380 Extended Data Figure 1: Genome of *R. marinus* encodes for the complete

381 **denitrification pathway.** a. The genes for NapAB, the periplasmic nitrate reductase

382 (Rmar_0413), the nitrite reductases nirK (Rmar_1208) and nirS (Rmar_0652) and nitrous

383 oxide reductase, nosZ (Rmar_2012) are encoded in the *R. marinus* genome. b. The gene

neighborhood of eNOR in *R. marinus*. c. The gene neighborhood of eNOR in *Halovivax*

385 *ruber* includes ctaB, an enzyme involved in the biosynthesis of heme *a*.

386

387 Extended Data Figure 2: *Rhodothermus marinus* does perform complete

denitrification. a. *R. marinus* converts ${}^{15}NO_3$ to ${}^{30}N_2$. Ratio of ${}^{30}N_2$ to ${}^{28}N_2$ for each sample. Air is ambient atmosphere as a standard. C0 is a nitrate-free control. 14 0-17 are

390 cultures grown with unlabeled nitrate, transferred to sealed vials after 0-17 hours,

respectively. 15 0-17 are the equivalent samples grown with ¹⁵N-labeled nitrate. Error

bars represent two standard deviations from three replicate GC/MS measurements. ${}^{30}N_2$ enrichments from the ${}^{15}N$ -labeled samples are over 30-60x higher than background

394 atmospheric ratios, while unlabeled samples have no significant enrichment over

- background. b. Growth of *R. marinus*, measured using OD_{600nm} over 39 hours. NO₃
- 396 utilization was established by measuring the concentrations of nitrate in the media using a
- 397 calorimetric assay. c. *R. marinus* growth in rich media was compared under denitrifying
- and non-denitrifying conditions using OD_{600nm} . d. Phenotypic differences of *R. marinus* cultures, under denitrifying and non-denitrifying conditions.
- 400

401 Extended Data Figure 3: Characteristics of eNOR from *Rhodothermus marinus* a.

402 SDS-PAGE gel electrophoresis of eNOR shows two bright bands which are estimated to 403 subunits of I and II of the complex. Both subunits appear to run faster than their

subunits of I and II of the complex. Both subunits appear to run faster than their
estimated molecular weight. This is typical for membrane proteins. For comparison, an

404 estimated molecular weight. This is typical for membrane proteins. For comparison, an 405 SDS-PAGE gel of cytochrome bo_3 oxidase from *E. coli* is included. b. Mass

405 SDS-PAGE get of cytochrome *b0*₃ oxidase from *E. coll* is included. b. Mass
 406 spectrometric identification of eNOR is confirmed by LC/MS/MS analysis. c,d. Absence

407 of O_2 reduction by *R. marinus* eNOR, in comparison to robust O_2 reduction by *T.*

407 of O_2 reduction by *K*. *martnus* eNOK, in comparison to robust O_2 reduction by *T*. 408 *thermophilus ba*₃-type oxygen reductase. e. UV-visible spectrum of eNOR f. Pyridine

409 hemochrome-spectra of extracted hemes from eNOR showing a peak which is atypical of 410 hemes a, b or c.

411

412 **Extended Data Figure 4: Identification of hemes extracted from eNOR.** Comparing 413 the elution profile of extracted hemes from partially purified *R. marinus* eNOR to bovine 414 cytochrome *c* oxidase (A-type, t=16 min), *T. thermophilus ba*₃-type oxygen reductase (*b*-415 and A_s-type hemes, t=12 min and t=19 min) reveals that the heme is most likely an A_s-416 type heme. Mass spectra of the peak at ~19 min from the eNOR hemes elution profile 417 confirms that the heme is an A_s-type heme with a molecular weight of 920 Da.

418

419 Extended Data Figure 5: Proton channel in eNOR of *Rhodothermus marinus* and in

420 **the NOR families bNOR, sNOR and nNOR. a.** eNOR contains conserved residues in 421 Helix VII, similar to the location of K-proton channel residues in *T. thermophilus ba*₃-

type oxygen reductase. b. A multiple sequence alignment of the NOR families eNOR,

423 bNOR, sNOR and nNOR show conserved amino acids in analogous location to the K-424 channel in the B-type oxygen reductase. Some conserved residues are also identified in

425 gNOR and may indicate the presence of a conserved proton channel but they do not map

- 426 to corresponding residues in the B-type oxygen reductase.
- 427

428 Extended Data Figure 6: Conserved amino acids in the eNOR family of enzymes.

429 Multiple sequence alignment of 23 eNOR sequences from various taxonomically

430 divergent organisms reveals conserved residues that correspond to the active site ligands,

431 proton channel residues and other sequence features that are unique to eNOR. The active

site residues are highlighted in maroon while the proton channel residues are highlightedin blue.

- 434
- 435
- 436
- 437
- 438 List of Tables
- 439

Table 1. Environmental distribution of the HCO NOR families. Distribution of NOR
families in sequenced genomes versus environmental datasets. The newly discovered
NOR families account for approximately 2/3 of currently known diversity and 1/2 of the
abundance of NORs in nature.

444

Supplementary Table 1. Putative proton channels in the new NOR families – eNOR,
bNOR, sNOR, nNOR. A list of conserved residues is noted in the table for each family
with a reference sequence according to which the residues are numbered. These
conserved residues are compared with amino acids found in analogous positions in the Btype oxygen reductase.

451 Supplementary Table 2. Distribution of NOR families by phylum in GTDB. A
452 distribution of all the NOR families within various bacterial and archaeal phyla within the
453 genomes in release 202 of GTDB was analyzed using HMMs that are specific to each
454 NOR family.

455

456 Supplementary Table 3. Distribution of NOR families in various ecosystems as per
457 the IMG database. A distribution of various NOR families in over 2000 metagenomes
458 on the IMG database was evaluated, and then tabulated according to the environment
459 from which each metagenome is sourced.

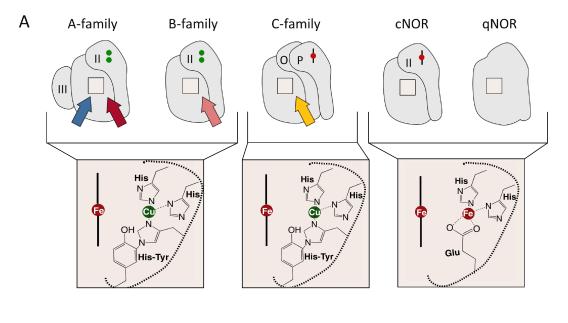
460

461 Supplementary Table 4. Denitrification pathways in bacteria and archaea. An
462 analysis of denitrification pathways in bacterial genomes and archaeal genomes in release
463 202 of GTDB was performed by searching for the presence and absence of NarGHI,
464 NapAB, NirK, NirS, NosZ, NosD and the NORs in each genome using curated HMMs
465 for each of the proteins.

Table 1. Environmental distribution of the HCO NOR families. Distribution of NOR families in sequenced genomes versus environmental datasets. The newly discovered NOR families account for approximately 2/3 of currently known diversity and 1/2 of the abundance of NORs in nature.

	NCBI-Genomes	IMG-metagenomes	GTDB-genomes
A-family	20290	102368	45135
B-family	1238	4683	2021
C-family	13976	23015	14981
qNOR	4388	7680	3458
cNOR	2801	4824	2594
eNOR	68	2709	547
sNOR	95	872	344
bNOR	51	12	200
nNOR	6	289	32
gNOR	10	913	156
NOD	8	539	108
N2O	25	597	n.a.

Figure 1. Comparison of HCO active sites. a) Active-site and proton channel properties of the five characterized HCO families (A-family, B-family, C-family, cNOR, and qNOR). The oxygen reductases all have an active-site composed of high-spin heme, a redox-active cross-linked tyrosine cofactor, and a copper (Cu_B) ligated by three histidines. The A-family has two conserved proton channels, whereas the B and C-families only have one. The active-sites of the nitric oxide reductases are composed of a high-spin heme and an iron (Fe_B) that is ligated by three histidines and a glutamate. Notably they are missing the tyrosine cofactor. The cNOR and qNOR are also missing conserved proton channels, making them non-electrogenic. b) Sequence alignment of the active-sites of the newly discovered HCO families that are related to the B-family. c) Predicted active-sites and proton channels for the new HCO families. The eNOR, bNOR, sNOR, and nNOR families contain completely conserved proton channels shown here as arrows. The putative proton channel in the bNOR and eNOR families are highly similar to the K-channel from the B-family oxygen reductase and are colored in red. The K-channel in the B-family is also similar to the Kchannel in the A-family oxygen reductase which is colored in dark red. The proton channel in the C-family is different from these channels and is marked in yellow. The putative proton channels in sNOR and nNOR are marked in cyan and differentiated from the other channels with a dashed black outline.



В	A-family_1M56	PVLYQHILWFFGHPEVYIIVLPAFGIVSHVIATFAKGYLPMVYAMVAIGVLGFVVWAHHMYT
_	B-family_3S8F	PLVARTLFWWTGHPIVYFWLLPAYAIIYTILPKQAGSDPMARLAFLLFLLLSTPVGFHHQFA
	e N O R	AAWYRQMYWIIGHGSQQ <mark>Q</mark> INLAAMITVWYFLTHVVGGSEKLSRTAFILYLFFINMGAAHHLLA
	b N O R	VMVARTLFWAFGHTAVNIWYLTAVSAWYVIVPKIIGSDMLTRVVIIALVIMNITGGFHHQII
	s N O R	PLLSKNLIYAFGHIFANSIIYMGVIAVYEIFPKYTGVYGNFLIAWNASTLFTMIIYPHHLLM
	gNOR	ALLYKNVYWWGLDLIADGLVLIYVAGSWYLLAMLLT MQHIARAALFVELVVSWFVWSHHLLS
	n N O R	PWPFNVAFWLFAHNLM <mark>E</mark> AMGIMALAAVYALVPLYTR SPGMGVLAVGLYTLSAIPAFGHHLYT
	C-family_3MK7	GATDAMVQWWYGHNAVGFFLTAGFLGIMYYFVPKQASYRLSIVHFWALITVYIWAGPHHLHY
	c N O R _ 3 O Ø R	LTRDKFYWWWVVHLWVEGVWELIMGAILAFVLVKITIEKWLYVIIAMALISGIIGTGHHYFW
	q N O R _ 3 A Y F	FTMADFWRWWIIHLWVEGIFEVFAVVVIGFLLVQLRTVRALYFQFTILLGSGVIGIGHHYYY
		Helix VIHelix VII

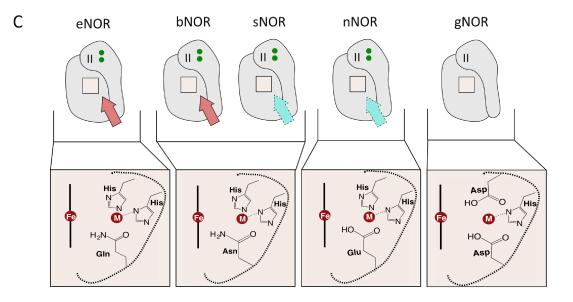


Figure 2. Evolution of nitric oxide reductases. Phylogenetic tree of HCO families. All of the new NOR families are derived from oxygen reductase ancestors. Oxygen reductases are in shades of blue, whereas nitric oxide reductases are in shades of yellow, green and red. The biochemically characterized eNOR from *Rhodothermus marinus* is marked with a black star.

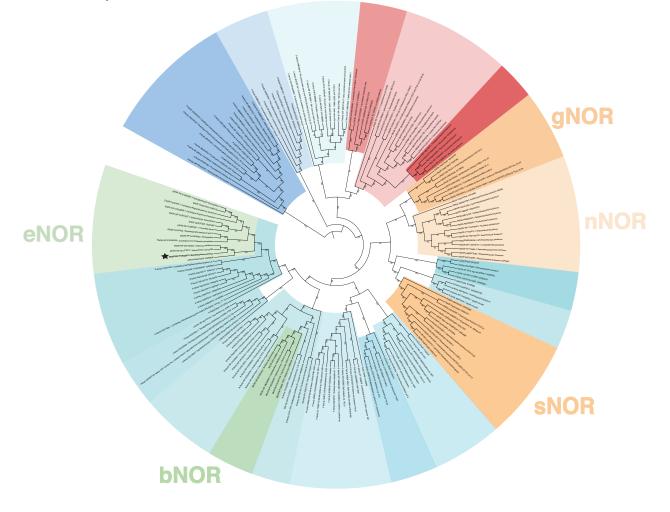
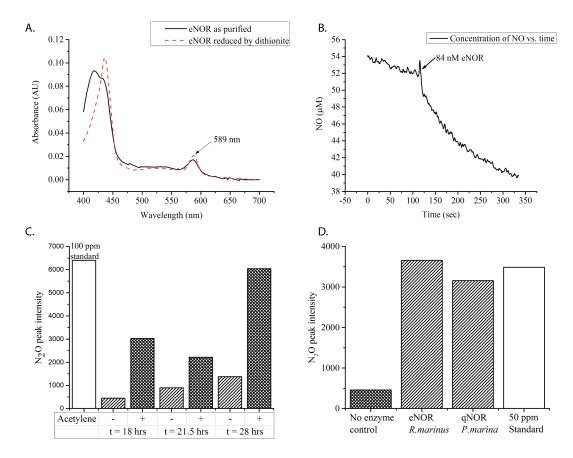
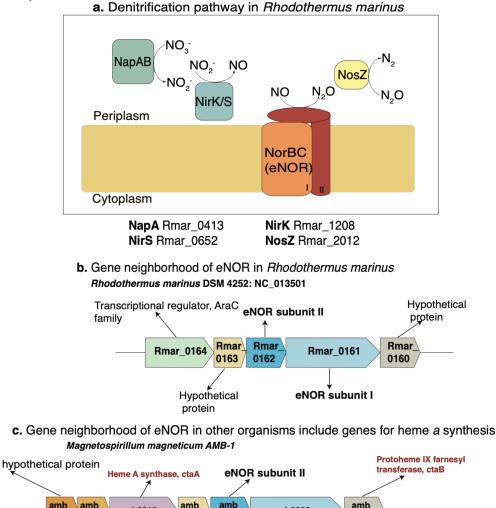
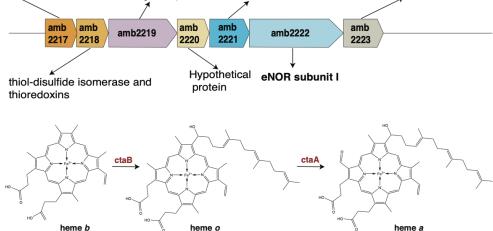


Figure 3. Biochemical Characterization of the eNOR from *Rhodothermus marinus.* a) UV-Vis spectrum of isolated eNOR indicates the presence of an unusual heme *a* signature at 589 nm. b) NO reductase activity was measured with the use of a Clark electrode in the presence of TMPD and ascorbate as electron donor. c) N₂O accumulation observed in a culture of *Rhodothermus marinus*, in the presence of acetylene d) N₂O production by eNOR from *R. marinus* and qNOR from *Persephonella marina*.

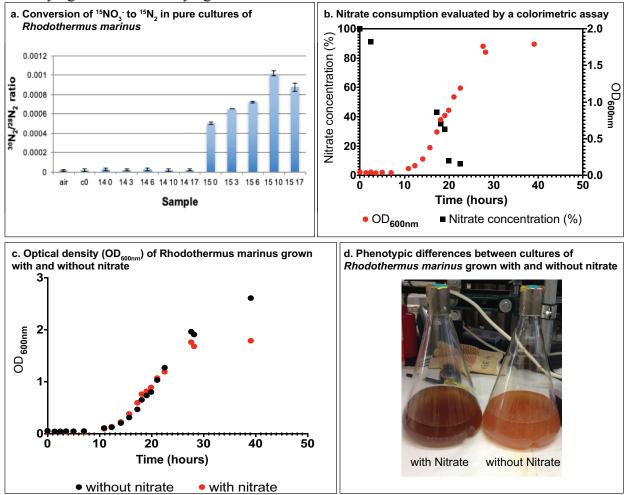


Extended Data Figure 1: Genome of *R. marinus* encodes for the complete denitrification pathway. a. The genes for NapAB, the periplasmic nitrate reductase (Rmar_0413), nitrite reductases nirK (Rmar_1208) and nirS (Rmar_0652), nitric oxide reductase eNOR(Rmar_0161) and nitrous oxide reductase, nosZ (Rmar_2012) are encoded in the *R. marinus* genome. b. The gene neighborhood of eNOR in *R. marinus*. c. The gene neighborhood of eNOR in *Magnetospirillum magneticum AMB-1* includes ctaA and ctaB, enzymes involved in the biosynthesis of heme *a*.

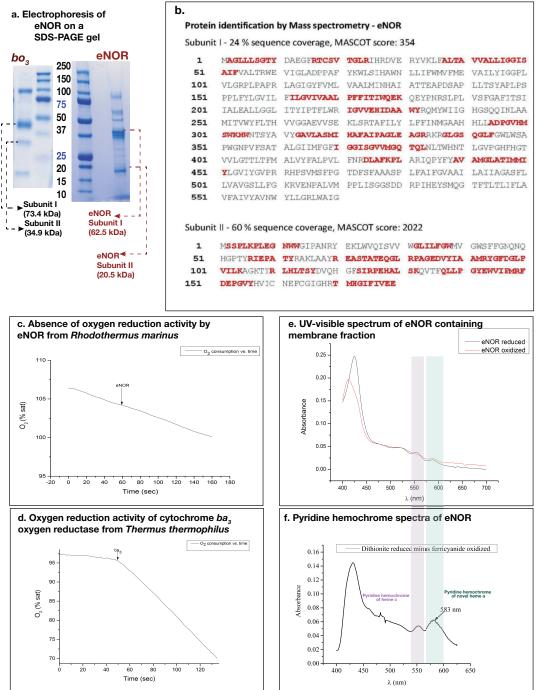




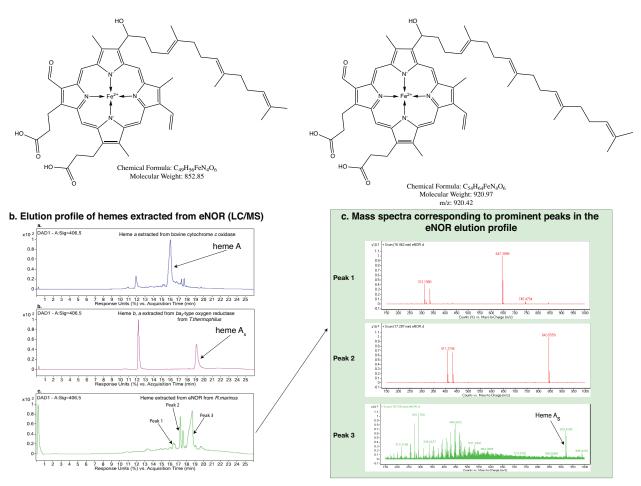
Extended Data Figure 2: *Rhodothermus marinus* does perform complete denitrification. a. *R. marinus* converts ¹⁵NO₃⁻ to ³⁰N₂. Ratio of ³⁰N₂ to ²⁸N₂ for each sample. Air is ambient atmosphere as a standard. C0 is a nitrate-free control. 14 0-17 are cultures grown with unlabeled nitrate, transferred to sealed vials after 0-17 hours, respectively. 15 0-17 are the equivalent samples grown with ¹⁵N-labeled nitrate. Error bars represent two standard deviations from three replicate GC/MS measurements. ³⁰N₂ enrichments from the ¹⁵N-labeled samples are over 30-60x higher than background atmospheric ratios, while unlabeled samples have no significant enrichment over background. b. Growth of *R. marinus*, measured using OD_{600nm} over 39 hours. NO₃⁻ utilization was established by measuring the concentrations of nitrate in the media using a calorimetric assay. c. *R. marinus* growth in rich media was compared under denitrifying and non-denitrifying conditions.



Extended Data Figure 3: Characteristics of eNOR from *Rhodothermus marinus* a. SDS-PAGE gel electrophoresis of eNOR shows two bright bands which are estimated to subunits of I and II of the complex. Both subunits appear to run faster than their estimated molecular weight. This is typical for membrane proteins. For comparison, an SDS-PAGE gel of cytochrome bo_3 oxidase from *E. coli* is included. b. Mass spectrometric identification of eNOR is confirmed by LC/MS/MS analysis. c,d. Absence of O₂ reduction by *R. marinus* eNOR, in comparison to robust O₂ reduction by *T. thermophilus ba*₃-type oxygen reductase. e. UV-visible spectrum of eNOR f. Pyridine hemochrome-spectra of extracted hemes from eNOR showing a peak which is atypical of hemes *a, b* or *c*.

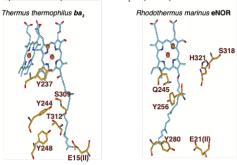


Extended Data Figure 4: Identification of hemes extracted from eNOR. Comparing the elution profile of extracted hemes from partially purified *R. marinus* eNOR to bovine cytochrome *c* oxidase (A-type, t=16 min), *T. thermophilus ba*₃-type oxygen reductase (*b*- and A_s-type hemes, t=12 min and t=19 min) reveals that the heme is most likely an A_s-type heme. Mass spectra of the peak at ~19 min from the eNOR hemes elution profile confirms that the heme is an A_s-type heme with a molecular weight of 920 Da. a. Chemical structures of hemes A and heme As.



Extended Data Figure 5: Proton channel in eNOR of *Rhodothermus marinus* and in the **NOR families bNOR, sNOR and nNOR. a.** eNOR contains conserved residues in Helix VII, similar to the location of K-proton channel residues in *T. thermophilus ba*₃-type oxygen reductase. b. A multiple sequence alignment of the NOR families eNOR, bNOR, sNOR and nNOR show conserved amino acids in analogous location to the K-channel in the B-type oxygen reductase. Some conserved residues are also identified in gNOR and may indicate the presence of a conserved proton channel but they do not map to corresponding residues in the B-type oxygen reductase.





b. Conserved residues forming putative proton channels in eNOR, sNOR, bNOR and nNOR. Conserved proton channel residues are marked in black, while the active site residues are marked in **red**. Each group is numbered according to the protein whose accession number is in **blue**.

according to the	protein whose accession number	is in plue.	
nNOB	H255 E259 Y270 Y276 Y284	H307	T338 T341 T345
nNOR Thermomicrobium sp002898255 GBD20489.1	FWLFAHNLMEAMGIMTLGAIYAIVPRYTR-SGQLY-SPR	AAVVAMILYTMAAIPAFG <mark>HH</mark> LYTWVTGNPEVLQNVSH	R-STSWATGFIAATLTAFNVGLTVWRNGL
Sedimenticola selenatireducens WP 084609916.1	FYIFAHNLMEAMAIMVISAVYATLPLYLADGTRKLYSDKI	LANLALWILLVTSVTSFF <mark>HH</mark> FYTTNPGLPSALAY-HO	SNFMSWATGVGAAL-STFTILATIWKHGI
Sedimenticola selenatireducens PLX61751.1	FYIFAHNLMEAMAIMVISAVYATLPLYLADGTRKLYSDKI	LANLALWILLVTSVTSFF <mark>HH</mark> FYTTNPGLPSALAY-HO	GNFMSWATGVGAAL-STFTILATIWKHGI
Rhodocyclaceae bacterium UTPRO2 OQY64980.1	FFFFAHNLMEAMAIMVASAIYATLPLYLADGSRKLFSDKM	MANLALWILLVTSITSGL <mark>HH</mark> FITFYPNQPAALSYWG:	S-IMSWGTGIGAAL-SIFTVFATIWQHGL
aNOB	D285 D289 Y301	H335	<u>H373</u>
Sulfurimonas autotrophica WP_013326534.1	WWGLDLVADGLVLIYVAGSWYLLATLITGQKLF-MENV	VARAALMLELLVSWMVWSHHLLA-DQGQPEMMKLIS	SEMVTAFELLTQGL-ALFITLVTLWKARP
Ignavibacteria bacterium GWA2_36_19 OGU38604.1			GEMVTAFELITQGL-AFFITLATLWSARP
Ignavibacteria bacterium CG2_30_36_16 0IP63421.1			GEMVTAFELITQGL-AFFITLATLWSARP
candidate division WOR-1 bacterium OGW14262.1			GEMVTAFELVTMGI-AIFITLKTLWEARP
candidate division WOR-1 bacterium ogc04573.1	WWGLDLIADGLVLIYVAGTWYLLAMIITGRQIF-MQNH	FARAALFVELVVSWFVWSHHLLS-DQTQPVMMRIFS(GEMITAFELVTSGI-AVFLTLATLWQARP
sNOB	H232 N236 Y240 Y247E248	H288	T319
SNOR Bacteroidetes bacterium 37-13 0JV27025.1			Q-IASYFATIPSVVVTIISIVTLLYNNKI
Geobacillus sp. WP_023634191.1			GQVFSYLNGLPVLVVTAFGALMIVYRSGI
Rhodanobacter denitrificans WP_015448124.1			GQIISYAAGFPVFLVTAYGVLTNIHRSGL
Nitrosococcus halophilus WP_013031504.1			GQVLSYGSGVPVMVVTGYGALMIVYRSGI
Thioalkalivibrio sp. ALRh wp_019592254.1	TYFFGHVFINATIYMAVIGVYEILPRYTGRP-WKVSRV	VFLAAWAASTVMVLLVYPHHLLM-DFSQPTSLHVLG	Q-VISYTSGLPVLLVTAWGALTNVYRSGI
bNOR	H233 N237 Y240 Y248	H281	S309 T312
Salinicoccus qingdaonensis WP_092985759.1	FWAFGHTLVNVWYLVAVSAWYLVVPKVIGGKLF-SDSI	LARAVVILIVVLNVPGGFHHQIV-DPGFTEGLKFMH	L-FMSLAIAVPSLL-TAFALFATLERTGRRRG-G-KGLLGWFWKL
Salinicoccus qingdaonensis WP_092985759.1 Salinicoccus sp. YB14-2WP_092985759.1	FWAFGHTLVNVWYLVAVSAWYLVVPKVIGGKLF-SDSI FWAFGHTLVNVWYLVAVSAWYLVVPKVIGGKLF-SDSI	LARAVVILIVVLNVPGGFHHQIV-DPGFTEGLKFMH LARAVVILIVVLNVPGGFHHQIV-DPGFTEGLKFMH	L-FMSLAIAVPSLL-TAFALFATLERTGRRRG-G-KGLLGWFWKL L-FMSLAIAVPSLL-TAFALFATLERTGRRRG-G-KGLLGWFWKL
Salinicoccus qingdaonensis WP_092985759.1 Salinicoccus sp. YB14-2WP_092985759.1 Jeotgalicoccus psychrophilus WP_026860023.1	FWAFGHTLVNVWYLVAVSAWYLVVPKVIGGKLF-SDSI FWAFGHTLVNVWYLVAVSAWYLVVPKVIGGKLF-SDSI FWSFGHTLVNVWYLVAVSAWYIVLPKVIGGKIF-SDSI	LARAVVILIVVLNVPGGFHHQIV-DPGFTEGLKFMHJ LARAVVILIVVLNVPGGFHHQIV-DPGFTEGLKFMHJ LARLVVILIVILNVPGGFHHQIV-DPGFTEGLKFMHJ	L-FMSLAIAVPELL-TAFALFATLERTGRRRG-G-KGLLGWFWKL L-FMSLAIAVPELL-TAFALFATLERTGRRRG-G-KGLLGWFWKL L-FMSLAIGFPELM-TAFALFATLERAGRNKG-G-KGLFGWFFKL
Salinicoccus qingdaonensis WP_092985759.1 Salinicoccus sp. YB14-2WP_092985759.1 Jeotgalicoccus psychrophilus WP_026860023.1 Virgibacillus dakarensis WP_08049698.1	FWAFGHTLVNVWYLVAVSAWYLVVPKVIGGKLF-SDSI FWAFGHTLVNVWYLVAVSAWYLVVPKVIGGKLF-SDSI FWSFGHTLVNVWYLVAVSAWYVVPKIIGGKIF-SDSI FWSFGHTLVNUWYLVAVSAWYVVVPKIIGGKVF-SDKI	LARAVVILIVVLNVPGGFHHQIV-DPGFTEGLKFMHI LARAVVILIVVLNVPGGFHHQIV-DPGFTEGLKFMHI LARLVVILIVILNVPGGFHHQIV-DPGFTEGLKFMHI LARLVVVLLVILNIPGGFHHQII-DPGISESVKFLHV	L-FMSLAIAVPSLL-TAFALFATLERTGRRRG-G-KGLLGWFWKL L-FMSLAIAVPSLL-TAFALFATLERTGRRRG-G-KGLLGWFWKL L-FMSLAIGFPSIM-TAFALFATLERAGRNKG-G-KGLFGWFFKL /-FMSISIAFPSIM-TAFAMFAVFERAGRKLG-G-KGLLGWFKKL
Salinicoccus gingdanonensis WP 092985759.1 Salinicoccus psychrophilus WP 092985759.1 Jeotgalicoccus psychrophilus WP 026860023.1 Virgibacillus dakarensis WP 088049598.1 Sporosarcia sp. HYC09 & KUB97068.1	FWAFGITLYIWILVAVIAWILVVFKVIGCKLF-SDSI FWAFGITLYIWILVAVIAWILVVFKVI-CGKLF-SDSI FWSFGITLYIWILVAVIAWILVVFKVIGGKIF-SDSI FWSFGITLYIWILVAVIAWITVFKIIGGKRF-SDKI FWSFGITLYIWILVIATIAWITVFKIIGGRRF-SDKI	LARAVVILIVVLAVPGGFHHQIV-DPGFTEGLKFMHI LARAVVILIVVLAVPGGFHHQIV-DPGFTEGLKFMHI LARLVVILIVILAVPGGFHHQIV-DPGFTEGLKFMHI LARLVVVLLVILANIPGGFHHQII-DPGISESVKFLHI LTRVVVLLVITNIPGGFHHQIV-DPGMGEALKYMW	L-PHSLATAVPSILL-TAPALFATLERTGRRG-G-KGLIGNFWKL L-FMSLATAVPSILL-TAPALFATLERTGRRG-G-KGLIGNFWKL J-MSLATGPPSIM-TAPALFATLERTGRRG-G-KGLIGNFFKL J-MSLSTAFPSIM-TAPAMFAVFERAGRKIG-G-KGLIGNFFKL
Salinicoccus gingdaonensis WP_022985759.1 Salinicoccus sp. YB14-2 WP_092985759.1 Jeotgalicoccus sp. yB14-2 WP_026860023.1 Virgibacillus dakarensis WP_088049698.1 Sporosarcina sp. HY008 KXH87068.1 Sporosarcina ureihidra BW 075529285.1	FWAFGUTLVIVWILVAWILVVEKVIGGKLF-5DSI FWAFGUTLVIVWILVAVIAWILVVEKVIGGKLF-5DSI FWSFGUTLVIVWILVAVIAWILVLEKVIGGKLF-SDSI FWSFGUTLVIIWILVAVIAWILVLEKVIGGRF-SDTI FWSFGUTLVIIWILTAVIAWILVPKIIGGRF-SDTI FWSFGUTLVIIWILTAVIAWILVPKIIGGRF-SDTI	Laravvilivvlnvpggfhhqiv-dpgfteglkfmhi Laravvilivvlnvpggfhqiv-dpgfteglkfmhi Laruvilivilivggfhqiv-dpgfteglkfmhi Laruvvlivilinipggfhqiv-dpgfteglkfmhi Liruvvlivilinipggfhqiv-dpgmgbalkymhi Liruvvliviinipggfhqiv-dpgmgbalkymhi	L-FMSLATAVPSLL-TAPALFATLERTGRRRG-G-KGLLGWFWKL -TMSLATAVPLLL-TAPALFATLERTGRRRG-G-KGLLGWFWKL -TMSLATGPSLM-TAPALFATLERAGRRKG-G-KGLLGWFKKL /-FMSLSTAPPSLM-TAPAMFAVFERAGRKLG-G-KGLLGWFKKL /-FMSLSTAPPSLM-TAPAMFVFERTGRRKG-G-KGLLGWFKKL
Salinicoccus gingdaonensis WP 092985759.1 Salinicoccus p. VB14-2WP 092985759.1 Jeotgalicoccus psychrophilus WP 026860023.1 Virgibacillus dakarensis WP 088049598.1 Sporosarcina ureilytica WP 075529285.1 Bacillus ps. JATA-27445 wP 059173536.1	FWAFGITLVIWWILVAX AWILVVFKI-GGKLF-5DSI FWAFGITLVIWILVAX AWILVFKI-GGKLF-5DSI FWAFGITLVIWILVAX AWILVFKI-GGKLF-5DSI FW3FGITLVIWILVAX AWILVFKI-GGRVF-5DT FW3FGITLVIUTLTXA AWILVFKII-GGRRF-5DT FW3FGITLVIWILTXAX AWILVFKII-GGRRF-5DT	LARAVUILUVUNVEGGFHIQIV-DEGTEGLEFMHI LARAVUILUVINVEGGFHIQIV-DEGTEGLEFMHI LARLVVILUVILIVIENTEGFHIQIT-DEGTEGLEFMHI LARLVVULLVIINIEGFHIQIT-DEGSESSIVEFHIU LARLVVULLVIINIEGFHIQIT-DEGKESALKYMHI LRVVVMLVIINIEGFHIQIT-DEGKESALKYMHI LRVVVMLVIINIEGFHIQIT-DEGSESVEFHIN	PMSLAIAVPEILPAPLAFATLEFTGRRG-C-KGLIGNFWKL PMSLAIVPEILPAPLAFATLEFTGRRG-C-KGLIGNFWKL PMSLAIGPPIMFAFALFATLEFTGRRG-C-KGLIGNFKKL PMSLAIGPFIMFAFALFATLEFTGRRG-C-KGLIGNFKKL PMSLAIGPFIMFAFALFYVFENTGRKG-C-KGLIGNFKKL PMSLAIGPFIMFAFALFYVFENTGRKG-C-KGLIGNFKKL
Salinicoccus gingdaonensis WF 092985759.1 Salinicoccus sp. VB14-2WF 092985759.1 Jeotgalicoccus sp. VB14-2WF 092985759.1 Virgbacilus dakarensis WF 088049598.1 Sporosarcina sp. HVO08 XXIIB7068.1 Sporosarcina ureightea WF 075529285.1 Bacillus sp. FJAT-27445 WF 0759173536.1 Bacillus sp. FJAT-27445 WF 0459173536.1	FWAFGUTLVI VWILVAXI AWILVVFKVIGGKLF-SDSI FWAFGUTLVI VWILVAVI AWILVVFKVIGGKLF-SDSI FWSFGUTLVI VWILVAVI AWILVLFKVIGGKLF-SDSI FWSFGUTLVI IWILVAXI AWIVVVFKIIGGKRF-SDTI FWSFGUTLVI IWILTAXI AWIVIVFKIIGGKRF-SDTI FWAFGUTLVI IWILTAXI AWIVIVFKIIGGKRF-SDTI FWAFGUTLVI IWILTAXI AWIVITKIMGGRW-SDTI FWAFGUTLVI IWILTAXI AWIVITKIMGGRW-SDTI	LARAWULILVULNVEGGFHEQIV-DEGFEGLEFHH LARAWULIVULNVPGGFHEQIV-DEGFEGLEFHH LARLWULIVULNVPGGFHEQIV-DEGFEGLEFHH LARLWULLVILNIPGGFHEQIV-DEGFEGLEFHH LARLWVLLVILNIPGGFHEQIV-DEGGEALKYMH LTRWULAUNIFIGGFHEQIV-DEGGEALKYMH LTRWUILALVNNITGGFHEQIV-DEGIESAVKYMH	L-PHSLATAVPSIL-TAPALFATLERTGRRG-G-KGLLGNFWKL L-PMSLATAVPSIL-TAPALFATLERTGRRG-G-KGLLGNFWKL -PMSLATGPP IM-TAPALFATLERAGRNKG-G-KGLLGNFWKL -PMSLSTAFP IM-TAPAMFAVFERAGRKIG-G-KGLLGNFKKL -PMSLSTAFPSIM-TAPAMFYVFERTGRAKG-G-KGLLGNKKKL -PMSLSTAFPSIM-TAPAMFYVFERTGRAKG-G-KGLLGNKKKL -PMSLSTGPP IM-TAPAMFYVFERTGRAKG-G-KGLLGNKKKL
Salinicoccus gingdaonensis WP 092985759.1 Salinicoccus gi, V614-2WP 092985759.1 Jeotgalicoccus givento-philus WP 024660023.1 Virgibacillus dakarensis WP 08049698.1 Sporosarcina ureilytica WP 075529285.1 Bacillus sp. FA7C2445WP 059173536.1 Bacillus sp. EB01 WP 05923172.1 Thalascobacillus cyrup 039362123.1	FWAFGITLVIWWILVAX AWILVVFKVI-GCKLF-5DSI FWAFGITLVIWWILVAX AWILVLFKVI-GCKLF-5DSI FWAFGITLVIWWILVAX AWILVLFKVI-GCKLF-5DSI FWAFGITLVIWILVAX AWILVFKII-GCRNF-SDTI FWAFGITLVIIWILVAX AWILVFKII-GGRNF-SDTI FWAFGITLVIIWILVAX AWILVFKII-GGRNF-SDTI FWAFGITLVIIWILVAX AWILVFKII-GGRNF-SDTI FWAFGITLVIIWILVAX AWILVFKII-GGRNF-SDTI FWAFGITLVIIWILVAX AWILVFKII-GGRNF-SDTI	LARAVILIVVINVEGEHIQIV-DEGTEGLEFHH LARAVILIVINVEGEHIQIV-DEGTEGLEFHH LARAVILVILVINVEGEHIQIV-DEGTEGLEFHH LARAVINVILVILNINSEEHIQIV-DEGTEGLEFHH LARAVVVILVILNINIFGEHIQIV-DEGTESLEFHH LARAVINVILVINIFGEHIQIV-DEGTESVEHH LRAVVALVINIFGEHIQIV-DEGTESVEHH LRAVVILSIVVNITGEFHIQIV-DEGTESSVEHH LRAVVILSIVVNITGEFHIQI-DEGTESSVEHH	PMSLAITAVPEILPAPLIPATLEPTGRRAG-G-KGLLGNFWKL PMSLAITAVPEILPAPLIPATLEPTGRRAG-G-KGLLGNFWKL PMSLAIGPPIMPAPAMPAVTERAGRNKG-G-KGLLGNFFKL -PMSLAIGPPIMPAPAMPAVTERAGRKLG-G-KGLLGNFFKL -PMSLAIGPPIMPAPAMPSVFENTGRLKG-G-KGLLGNFFKL -PMSLAIGPPIMPAPAMPSVFENTGRLKG-G-KGLLGNFKKL -PMSLAIGPPIMPAPAMPSVFENTGRLKG-G-KGLLGNFKKL -PMSLAIGPPIMPAPAMPSVFENTGRLKG-G-KGULGNFKKM -PMSLAIGPPIMPAPAMPSVFENTGRLKG-G-KGULGNFKKM
Salinicoccus gingdaonensis WP 092985759.1 Salinicoccus sp. V514-24 PP 092985759.1 Jeotgalicoccus sp. V514-24 PP 092985759.1 Virgibacillus dakarensis WP 088049598.1 Sporosarcina ureilytica WP 075529285.1 Bacillus sp. IA/T22445 PP 059173536.1 Bacillus sp. IED01 WP 049923172.1 Thalascobacillus cytwp 099923172.1 Neobacillus bataviensis WP 007083098.1	FWAFGITLVI WWILVAX AWILVVFKII-GGKLF-BDSI FWAFGITLVI WWILVAX AWIVLVFKII-GGKLF-BDSI FW3FGITLVI WWILVAX AWIVLVFKII-GGKLF-BDSI FW3FGITLVI WILVAX AWIVVVFKII-GGKRF-SDT FW3FGITLVI WILTAX AWIVLVFKII-GGKRF-SDT FWAFGITLVI WILTAX AWIVLFKIM-GGKRW-SDT FWAFGITLVI WILTAX AWIVVIFKIM-GGKRW-SDT FWAFGITLVI WILTAX AWIVVIFKII-SGKRW-SDT	LARAVULIUVUNVEGGFINGUV-DEGFEGLEFMH LARAVULIVUNVPGGFINGUV-DFGFEGLEFMH LARLVVLLVUNVEGFINGUV-DFGFEGLEFMH LARLVVULVUNVEGFINGUV-DFGFEGLEFMH LARLVVULVUNVIFGFINGUV-DFGKEGLEFMH LARLVVULVUNVIFGFINGUV-DFGKESLKYMM LTRVVVMIVUTNIFGFINGUV-DFGKESVKFMM LTRVVIALVINNIFGFINGUV-DFGSSVKFMM LTRVVIALVINNIFGFINGUV-DFGSSVKFMM	1-PHSLATAVPSIL-TAPALFATLERTGRRG-G-KGLLGNFWKL 1-PHSLATAVPSIL-TAPALFATLERTGRRG-G-KGLLGNFWKL 1-PHSLATGPP IM-APALFATLERTGRRG-G-KGLLGNFWKL 1-PHSLSTAFP IM-APAMFAVFERAGRKIG-G-KGLLGNFKKL 1-PHSLATGPF IM-APAMFAVFERTGRKG-G-KGLGNFKKKL 1-PHSLATGPF IM-APAMFAVFERTARRNG-G-KGLGNFKKK 1-PHSLATGPP IM-APALFATFERTARRNG-G-KGLGNFKKM 1-PHSLATGPP IM-APAMFAVFERTARRNG-G-KGLGNFKKM 1-PHSLATGPP IM-APAMFAVFERTARRNG-G-KGLGNFKKM
Salinicoccus gingdaonensis WP 092985759.1 Salinicoccus gr. V614-2WP 092985759.1 Jeotgalicoccus gr. V614-2WP 092985759.1 Virgibacilus dakarensis WP 088049598.1 Sporosarcina unijitida wP 0015529285.1 Bacillus gr. FJAT-27445 WP_059173536.1 Bacillus gr. EB01 WP 043992172.1 Thalassobacillus gr. WP 043946123.1 Neobacillus bataviensis WP_007983088.1 eNOR	FWAFGITLVIWUTLVAN ANTLVVFKII-GGKLF-5DSI FWAFGITLVIWUTLVAN ANTIVLFKII-GGKLF-5DSI FWAFGITLVIWITLVAN ANTIVLFKII-GGKLF-5DSI FWAFGITLVIIWITLVAN ANTIVLFKII-GGKRF-SDTI FWAFGITLVIIWITLAN ANTIVFKII-GGKRF-SDTI FWAFGITLVIIWITLAN ANTVIFKIM-GGKRM-SDTI FWAFGITLVIIWITLAN ANTVIFKIM-GGKRM-SDTI FWAFGITLVIIWITLAN ANTVIFKII-SGKRM-SDTI FWAFGITLVIIWITLAN ANTVIFKII-SGKRM-SDTI FWAFGITLVIIWITLAN ANTVIFKII-SGKRM-SDTI FWAFGITLVIIWITLAN ANTVIFKII-SGKRM-SDTI FWAFGITLVIIWITLAN ANTVIFKII-SGKRM-SDTI FWAFGITLVIIWITLAN ANTVIFKII-SGKRM-SDTI	LARAVUILUVU.NVFGGFHIQIV-DGGTFEGLKFMH LARAVUILUVI.NVFGGFHIQIV-DGGTFEGLKFMH LARAVUILUVI.NVFGGFHIQIV-DGGTFEGLKFMH LARLVVULLVI.NVFGGFHIQIV-DFGGTFEGLKFMH LARLVVULLVI.NIFIGGFHIQIV-DFGGTFESLKFMH LARVVILLVI.NIFIGGFHIQIV-DFGGTFESLKFMH LARVVILLVI.NIFIGGFHIQIV-DFGGTFESLKFMH LARVVILLVI.NIFIGGFHIQIV-DFGGTFSVKFMH LARVVILLVI.NIFIGGFHIQIV-DFGGTSSVKFMH V2800 H2300	PMSLAITAVPEILPAPLIPATLEPTGRRRG-G-KGLLGNFWKL PMSLAITAVPEILPAPLIPATLEPTGRRRG-G-KGLLGNFWKL PMSLAIGPFIMTAPAMFAVTERAGRKIG-G-KGLLGNFFKL PMSLAIGPFIMTAPAMFSVFERTGRLKG-G-KGLLGNFFKL PMSLAIGPFIMTAPAMFSVFERTGRLKG-G-KGLLGNFFKL PMSLAIGPFIMTAPAMFSVFERTGRLKG-G-KGLLGNFKKL PMSLAIGPFIMTAPAMFSVFERTGRLKG-G-KGLLGNFKKM PMSLAIGPFIMTAPAMFYFERTARKGG-G-KGLLGNFKKM S338 H221
Salinicoccus gingdaonensis WP 092985759.1 Salinicoccus gp. V814-2WP 092985759.1 Jeotgalicoccus gp. V814-2WP 092985759.1 Virgibacillus dakarensis WP 088049598.1 Sporosarcina ureilytica WP 075529285.1 Bacillus sp. FA/722445W 059173536.1 Bacillus sp. FA/72445W 059173536.1 Thalascobacillus cyripp 093906123.1 Neobacillus bataviensis wP 007083088.1 eNOR Rhoddhermus marinus WP_012842681.1	FWAFGITLVI WWILVAX AWILVVEKII-GGKLF-BDSI FWAFGITLVI WWILVAX AWIVLVEKII-GGKLF-BDSI FWSFGITLVI WWILVAX AWIVLPKVI-GGKLF-BDSI FWSFGITLVI WILVAX AWIVVVEKII-GGKRF-SDTI FWSFGITLVI WILTAX AWIVIVEKII-GGKRF-SDTI FWAFGITLVI WILTAX AWIVIVEKII-GGKRM-SDTI FWAFGITLVI WILTAX AWIVVIEKIM-GGKRM-SDTI FWAFGITLVI WILTAX AWIVVIEKIM-GGKRM-SDTI FWAFGITLVI WILTAX AWIVVIEKII-GGKRM-SDTI BWAFGITLVI WILTAX AWIVVIEKII-GGKRM-SDTI BWAFGITLVI WILTAX AWIVVIEKII-GGKRM-SDTI BWAFGITLVI WILTAX AWIVVIEKII-GGKRM-SDTI BWAFGITLVI WILTAX AWIVVIEKII-GGKRM-SDTI	LARAVULIUVU.NVPGGFHIQIV-DEGTFGLKFMHI LARAVULIVU.NVPGGFHIQIV-DFGTFGLKFMHI LARIVVLIVI.NVPGGFHIQIV-DFGTFGLKFMHI LARIVVULIVI.NIPGGFHIQIV-DFGTFGLKFMHI LARIVVULIVI.NIPGGFHIQIV-DFGTSSLKYFMHI LTRVVVLIVI.NIPGFHIQIV-DFGTSSVKFMHI LTRVVI.NIVINITGGFHIQIV-DFGTSSVKFMHI LTRVVI.NIVINITGGFHIQIV-DFGTSSVKFMHI LTRVVI.NIVINITGGFHIQIV-DFGTSSVKFMHI Z280 1290 STAFFILJFFINGAAHHLA-DFGVHNSWKFMMI	1- PMSLAITAVPSILL - TAPALFATLERTGRREG-G-KGLLGNFWKL 1- PMSLAITAVPSILL - TAPALFATLERTGRREG-G-KGLLGNFWKL 1- PMSLAIGPP IM - APALFATLERTGRRKG-G-KGLLGNFWKL 1- PMSLSIAFP IM - TAPAMFAVFERAGRKIG-G-KGLLGNFKKL 1- PMSLSIAFP IM - TAPAMFAVFERTARRKG-G-KGLGNFKKL 1- PMSLAIGPF IM - TAVALFATFERTARRNG-G-KGLGNVKKM 1- PMSLAIGPF IM - TAVALFATFERTARRNG-G-KGLGNVKKM 5318 H321 574VYGAVLAM H-HAPITPALEAGRKRKGG-G-KGLGNVKKM
Salinicoccus gingdaonensis WP 092985759.1 Salinicoccus gp. V614-2WP 092985759.1 Jeotgalicoccus gp. V614-2WP 092985759.1 Virgibacilus dakarensis WP 086049598.1 Sporosarcina unijkirda ump 015529285.1 Bacillus sp. FJAT-27445 WP_059173536.1 Bacillus sp. FJAT-27445 WP_059173536.1 Thalassobacillus sylw 0 039046123.1 Neobacillus bataviensis WP_0193046123.1 Neobacillus bataviensis WP_0193046123.1 Chodothermus marinus WP_013264066.1	FWAFGITLVIWWILVAX ANTLVVFKI-GCKLF-SDSI FWAFGITLVIWWILVAX ANTLVVFKI-GCKLF-SDSI FWAFGITLVIWILVAX ANTIVLFKI-GCKLF-SDSI FWAFGITLVIWILVAX ANTIVLFKI-GCRKF-SDTI FWAFGITLVIWILTAX ANTIVVFKI-GCRKF-SDTI FWAFGITLVIWILTAX ANTVVIFKLM-GCRKF-SDTI FWAFGITLVIWILTAX ANTVVIFKLM-GCRKF-SDTI FWAFGITLVIWITLTAX ANTVVIFKLM-GCRKF-SDTI FWAFGITLVIWITLTAX ANTVVIFKI-GCRKF-SDTI FWAFGITLVIWITLTAX ANTVVIFKI-GCRKF-SDTI FWAFGITLVIWITLTAX ANTVVIFKI-SCRKF-SDTI HE2102625021	LARAVILIVVIANVEGEHIGIV-DEGTEGLEFHH LARAVILIVVIANVEGEHIGIV-DEGTEGLEFHH LARAVILIVIANVEGEHIGIV-DEGTEGLEFHH LARAVILIVIANVEGEHIGIV-DEGTEGLEFHH LARAVVILIVIINIFGEHIGIV-DEGTESSVEFH LTRVVVALVITNIFGEHIGIV-DEGTESSVEFH LTRVVVALVITNIFGEHIGIV-DEGTESSVEFH LTRVVIALVINITGEHIGIV-DEGTESSVEFH LTRVVIALVINITGEHIGIV-DEGTESSVEFH LTRVVIALVINITGEHIGIV-DEGTESSVEFH LTRVVIALVINITGEHIGIV-DEGTESSVEFH SCADE LTRVVIALVINITGEHIGIV-DEGTESSVEFHI V280 L290	PMSLAITAVPEILPAPLIPATLEPTGRRRG-G-KGLLGNFWKL PMSLAITAVPEILPAPLIPATLEPTGRRRG-G-KGLLGNFWKL PMSLAIGPPIMTAYMAYDTEENGRKKG-G-KGLLGNFKKL PMSLAIGPPIMTAYMAYDTEENGRKG-G-KGLLGNFKKL PMSLAIGPPIMTAYMAYDTEENGRKG-G-KGLGNFKKL PMSLAIGPPIMTAYMAYDTEENGRKG-G-KGLGNFKKL PMSLAIGPPIMTAYMAYDTEENGRKG-G-KGLGNFKKM PMSLAIGPPIMTAYMAYTEENTARKG-G-KGLGNFKKM SIB H321 SYAYYGAVLAMIINFAIDLEAGRKRGLGSQGLFONLMAA
Salinicoccus gingdaonensis WP 092985759.1 Salinicoccus gi, VB14-2WP 092985759.1 Jeotgalicoccus gi, VB14-2WP 092985759.1 Virgibacillus dakarensis WP 08049598.1 Sporosarcina ureilytica WP 075529285.1 Bacillus sp. FA7-27445WP 05917336.1 Bacillus sp. FA7-27445WP 05917336.1 Thalasobacillus cyripte 093942172.1 Neobacillus bataviensis WP 001983088.1 eNOR Rhodothermus marinus WP 012842691.1 Alicycliphilus denitrificans WP 013822406.1 Candidatus Kryptonum thompsoni WP 075226648.1	$\label{eq:response} \begin{split} & FWAFGITUVIWITUXANTUVPKIT-GGKLF-BDSI \\ & \mathsf{FWAFGITUVIWITUXANTUVPKIT-GGKLF-BDSI \\ & \mathsf{FWAFGITUVIWITUXANTUVPKIT-GGKLF-BDSI \\ & \mathsf{FWSFGITUVIMITUTANANTUVPKIT-GGKRF-SDT \\ & \mathsf{FWSFGITUVIMITUTXANANTUVPKIT-GGKRF-SDT \\ & \mathsf{FWAFGITUVIWITUTXANTUVPKIT-GGKRM-SDT \\ & \mathsf{FWAFGITUVIWITUTXANTUVIPKIT-SGKRM-SDT \\ & \mathsf{FWAFGITUVIIWITATANTVVIPKIT-SGKRM-SDT \\ & \mathsf{FWAFGITUVIIWITATANTVVIPKIT-SGKRM-SDT \\ & \mathsf{FWAFGITUVIIWITATANTVVIPKIT-SGKRM-SDT \\ & \mathsf{FWAFGITUVIIWITATANTVVIPKIT-SGKRM-SDT \\ & \mathsf{FWAFGITUVIIWITATANTVIVPKIT-SGKRM-SDT \\ & \mathsf{FWAFGITUVIIWITATANTVIVPKIT-SGKRM-SDT \\ & \mathsf{FWAFGITUVIIWITATANTVIVPKIT-SGKRM-SDT \\ & \mathsf{FWAFGITUVIIWITATANTVIVPKIT-SGKRM-SDT \\ & \mathsf{FWAFGITUVIIWITATANTVITTITVVG-GAEV-SEKI \\ & \mathsf{WAFGISSQUIVAANISIWAVANAAF-GARM-SGKTM-SGKTM \\ & \mathsf{FWGFGIPAQUIVIANISIWAVANAAF-GARM-SGKTM-SGKTM \\ & \mathsf{FWGFGIPAQUIVIANTISIWAVANAAF-GARM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTISIWIAVANAAF-GARM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTISIWIAVANAAF-GARM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTISIMANTMANAAF-GARM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTISIMANTANTAAF-GARM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTANTAAF-GARM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTSANTANTAAF-GARM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTSANTANTAAF-GARM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTANTAAFGITUSSUTTM-SGKTM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTANTAAFGITM-SGKTM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTANTAAFGITM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTANTAAFGITM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTANTAAFGITM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTANTAAFGITM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTANTAAFGARTM-SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTANTAATT SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTIANTS SSKT \\ & \mathsf{FWGFGIPAQUIVIANTIANTIANT \\ & \mathsf{FWGFGIPAQUIVIANTIANTIANT \\ & \mathsf{FWGFGIPAQUIVIANTIANT \\ & \mathsf{FWGFGIPAQUIVIIANTIANTIANTAATT SSKT \\ & \mathsf{FWGFGIPAQUIVIIANTIANTIANT \\ & \mathsf{FWGFGIPAQUIVIANTIANTIANT \\ & \mathsf{FWGFQUIPAQUIVIANTIANT \\ & \mathsf{FWGFQUIPAQUIVIANTIANT \\ & \mathsf{FWGFUPAQUIVIANTIANTIANT \\ & \mathsf{FWGFUPAQUIVIANTIANT \\ & \mathsf{FWGFUPAQUIVIANTIANTIANT \\ & \mathsf{FWGFUPAQUITANT \\ & \mathsf{FWGFUPAQUITANT \\ & \mathsf{FWGPUPAQUITANT \\ & \mathsf{FWGFUPAQUIVIANTIANT \\ & \mathsf{FWGPUPAQUIPAQUIPAQUIPAU \\ & \mathsf{FWGUPAQUIPAQUIVI \\ & \mathsf{FWG$	LARAVU LLVVLNVPGGFHIQIV-DEGTEGLEFMH LARAVU LLVVLNVPGGFHIQIV-DGFTEGLEFMH LARAVVLLVILNIPGGFHIQI-DGFTEGLEFMH LARAVVLLVILNIPGGFHIQI-DGFTEGLEFMH LARAVVLLVILNIPGGFHIQIV-DFGTESLEFMH LARAVVLLVILNIPGGFHIQIV-DFGTESLEFMH LTRVVILNINTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH LTRVVILININTIGGFHIQIV-DFGTESLEFMH SERTELLVILFINGAAHHLLA-DFGYSTEMHINN SERTELLVILFINGAAHHLLA-DFGTEMHKINN	1-PMSLAITAVPSILI-TAPALIPATLEETGRREG-C-KGLLGNEWIKL 1-PMSLAIGPP IM-TAPALFATLEETGRREG-C-KGLLGNEWIKL 1-PMSLAIGPP IM-TAPALFATLEETGRREG-G-KGLLGNEWIKL 1-PMSLAIGPP IM-TAPAMFAVFERAGRKIG-G-KGLLGNEWIKL 1-PMSLAIGPP IM-TAPAMFAVFERTAGRKG-G-KGLGNEWIKL 1-PMSLAIGPP IM-TAYAFAVFERTAGRKG-G-KGLGNEWIKL 1-PMSLAIGPP IM-TAYAFAVFERTAGRKG-G-KGLGNEWIKL 1-PMSLAIGPP IM-TAYAFAVFERTAGRKG-G-KGLGNEWIKL 1-PMSLAIGPP IM-TAWAFAVFERTARKGG-G-KGLGNEWIKL 5318 #21 STAYGAULABMT-HAPAIFAGLEAGRRKRGLGSQGLFGNEMSA SYTMITAVLGMT-HAPAIFAGLEAGRRKRGLGSQGLFGNEMSA SYTMITAVLGMT-HAPAIFAGLEAGRRKRGLGSQGLFGNEMSA
Salinicoccus gingdaonensis WP 092985759.1 Salinicoccus gp. V614-2WP 092985759.1 Jeotgalicoccus gp. V614-2WP 092985759.1 Virgibacilus dakarensis WP 086049598.1 Sporosarcina unijkirda ump 015529285.1 Bacillus sp. FJAT-27445 WP_059173536.1 Bacillus sp. FJAT-27445 WP_059173536.1 Thalassobacillus sylw 0 039046123.1 Neobacillus bataviensis WP_0193046123.1 Neobacillus bataviensis WP_0193046123.1 Chodothermus marinus WP_013264066.1	$\label{eq:response} \begin{split} & PARFCITUV (WUTUVAS AWTLUVPEXI'-GCKLF-SDSI\\ & PARFCITUV (WUTUVAS AWTUVPEXI'-GCKLF-SDSI\\ & PARFCITUV (WUTUVAS AWTUVPEXI'-GCKUF-SDKI\\ & PARFCITUV (WUTUVAS AWTUVPEXI'-GCRF-SDT\\ & PARFCITUV (WUTUVTX AWTVUPEXI'-GCRF-SDT\\ & PARGCITUV (WUTUPEXI'AS AWTVUPEXI'-GCRW-SDT\\ & PARGCITUV (WUTEXA AWTVUPEXI'-GCRW-SDT\\ & PARGCITUV (WUTEXA AWTVUPEXI'-GCRW-SDT\\ & PARGCITV (WATA AWTVUPEXI'-GCRW-SDT\\ & PARGCITV (WARIAWTVWFITWVG-GCRV-SEK\\ & WAFGUSSQU'IVAAIIATVWFITWVAAIAI-CGCTP-SEK\\ & \mathsf{WGLGISQU'IVAAIISTVWSFLAVCAGTS' - GCTP-SEK\\ & \mathsf{WGLGISQU'IVAANISTVWSFLAVCSTV-GCTS' - NEXX \\ & \mathsf{WGLGISSQU'IVAANISTVWSFLAVCGS' - SCR'S' - SCR'S' \\ & \mathsf{WGLGISSQU'IVAANISTVWSFLAVCSSU'SI - SCR'S' - SCR'S' \\ & \mathsf{WGLGISSQU'IVASSW'IVASSW'ISSTAV-GCTS' - SCR'S' - SCR'S' \\ & \mathsf{WGLGISSQU'IVASSW'IVASSW'ISSTAV-GCTS' - SCR'S' \\ & \mathsf{WGLGISSQU'IVASSW'IVASSW'ISSTAV-SCST' - SCR'S' - SCR'S' \\ & \mathsf{WGLGISSQU'IVASSW'IVASSW'ISSTAV-SCST' - SCR'S' - SCR'S' \\ & \mathsf{WGLGISSQU'IVASSW'IVASSW'ISSTAV-SCST' - SCR'S' - SCR'S' \\ & \mathsf{WGLGISSQU'IVASSW'IVASSW'ISSU'' - SCST' - SCR'S' \\ & \mathsf{WGLGISSQU'IVASSW'IVASSW'' SSTAV-SCST' - SCR'S' - SCR'S' \\ & \mathsf{WGLGISSQU'IVASSW'' SSTAVSSW'' SSTAVSSW'' SSTAVSS'' - SCST' - SCR''' \\ & \mathsf{WGLGU'SSQU'IVASSW''' SSTAVSSW''' SSTAVSS''' - SCST' - SCR'''' \\ & \mathsf{WGLGU'''''' SSU'''''' SSTAVSSW'''''''''''''''''''''''''''''''''''$	LARAVU LIVVINVEGEHIGIV-DEGTEGLEFHH LARAVU LIVVINVEGEHIGIV-DEGTEGLEFHH LARAVU LIVVINVEGEHIGIV-DEGTEGLEFHH LARIVVILIVINVEGEHIGIV-DEGTEGLEFHH LARIVVILIVINIFGEHIGIV-DEGTEGLEFHH LIRVVILIVINIFGEHIGIV-DEGTESVEFHI LIRVVILIVINIFGEHIGIV-DEGTESVEFHI LIRVVILIVINIFGEHIGIV-DEGTESSVEFHI LIRVVILIVINIFGEHIGIV-DEGTESSVEFHI LIRVVILIVINIFGEHIGIV-DEGTESSVEFHI LIRVVILIVINIFGEHIGIV-DEGTESSVEFHI SCRAFLI LIFLGLASAHILLA-DEGVSTEMETIMI VSRAFLI LIFLGLASAHILLA-DEGVSTEMETIMI SSRAFLI LIFLGLASAHILLA-DEGVSTEMETIMI SSRAFLI LIFLGLASAHILLA-DEGVSTEMETIMI	PMSLAITAVPEILPAPLIPATLEPTGRRRG-G-KGLLGNFWKL PMSLAITAVPEILPAPLIPATLEPTGRRRG-G-KGLLGNFWKL PMSLAIGPPIMTAYMAYDTEENGRKKG-G-KGLLGNFKKL PMSLAIGPPIMTAYMAYDTEENGRKG-G-KGLLGNFKKL PMSLAIGPPIMTAYMAYDTEENGRKG-G-KGLGNFKKL PMSLAIGPPIMTAYMAYDTEENGRKG-G-KGLGNFKKL PMSLAIGPPIMTAYMAYDTEENGRKG-G-KGLGNFKKM PMSLAIGPPIMTAYMAYTEENTARKG-G-KGLGNFKKM SIB H321 SYAYYGAVLAMIINFAIDLEAGRKRGLGSQGLFONLMAA

Extended Data Figure 6: Conserved amino acids in the eNOR family of enzymes. Multiple

sequence alignment of 23 eNOR sequences from various taxonomically divergent organisms reveals conserved residues that correspond to the active site ligands, proton channel residues and other sequence features that are unique to eNOR. The active site residues are highlighted in maroon while the proton channel residues are highlighted in blue.

a. Sequence alignment of eNOR - Subunit I showing conserved amino acid residues according to R.marinus eNOR numbering R16 H77 F88 E90 A92 Image: Control of the second of the VTGLRIHRDVERYVKLFALTAVVAL VTGFDVHRSVEHNVKLFGVTAVVFL VTGFDVHRTVEHNVKLFGVTAVVFL VTGLRIHRTVEHNVKLFGLTAVVAL VTGLEVHRSAEDLVKLFGLTAVVAL VTGLTIHKSAEDLVKLFGLTAIVSIA ATGLRFGVAETLVKWNAVAVVFL FWMV FME VA FWMV FME IA FWM V FME IA FWM V FFE GAC FWM V FFE GAC Photochermus marinus, WP_0128426811. Haloxiper golemasilinasis, WP_00987204.1 Haloxibas asilancis, WP_00987204.1 Haloxibas asilancis, WP_0080687204.1 Haloxibas atismus, WP_0080687204.1 Harban themsoletaras, WP_00780471. Thermas horokanus, WP_01787471.1 Anaecolineae bacterium UTCPX1_0078834.1 Anaecolineae bacterium UTCPX1_0078834.1 Anaecolineae bacterium UTCPX1_0078834.1 Anaecolineae bacterium UTCPX1_0078834.1 bacteria bacterium UTCPX1_0078834.1 bacteria bacterium Imcoline_WP_00896084.1 bacteria bacterium Imcoline_WP_00896084.1 bacteria bacterium Imcoline_WP_00896084.1 AWN L AWFM V S P N L L P T D L L P T D L AWN AWI AWI FWI L NA L FW II FF U A GLYF I NM I LW LF FY VA LYF I A MVW II FF I A VLYF I A GLYF. LNM VFW II FF VA GLYF. LNM VFW II FF VA GLYF. LNM VFW II FF VA GLYF. LNM IFW IFF VA LYF. LNM IFF II FF MA LYF. MM IFF II FF MA LYF. MM IFF II FF MA LYF. MM VFF ILFF MA LYF. Bacterioldetes bacterium OLB10, XXK457611, Lacunisphates Imporfue, WP, 059960641, Lacunisphates Imporfue, WP, 15, 0671049841, SERUMO2, 01, FULL, 39, 10, 000W64851, SSPL0WO2, 01, FULL, 39, 10, 000W64851, SSPL0WO2, 01, FULL, 39, 10, 000W64851, SSPL0WO2, 01, FULL, 39, 17, 000W64851, SSPL0WO2, 01, FULL, 39, 17, 000W64851, SSPL0WO2, 02, 616, 64, 62, 000W64851, SSPL0WO2, 02, FULL, 39, 17, 000W54851, SSPL0WO2, 02, FULL, 71, 17, 000W5481, SSPL0WO2, 01, FULL, 71, 000W5481, SSPL0WO2, 01, 500W548, SSPL0WO2, 01, 500W548, SSPL0WO2, 01, 500W5481, SSPL0W04, 01, 500W54 19 V.A.L.LIN HAAVY AVAN TYUL 19 V.A.L.LIN HAAVY AVAN TYUL 10 V.A.L.LIN HAAVY AVAN TYUL 11 A.G.LIN HAAVY AVAN TYUL 12 V.A.L.LIN HAAVY AVAN TYUL 13 V.A.L.LIN HAAVY AVAN TYUL 14 A.G.LIN HAAVY AVAN TYUL 15 V.A.L.LIN HAAVY AVAN TYUL 16 V.A.L.LIN HY AND TY AS CAP HAAVAN TYUL 17 V.A.L.LIN HAAVY AVAN TYN AND TYN AG ANN YN YN LLG CAUNTYN TYN AND TY WQ A V H L L P A DM T K F L T L F WQ A V H L M D A EW F Y R I L T V H LI DC IAALFYLLT WOQAVHLOD Y145 D147 LITTA VIEN FN FYAGAIVYL LITTA VIEN FN FYAGAIVYL LITTA VIEN FN FYAGAIVYL LITTA VY MP Y15Y LFYAGAIVFL LITTA VY MP Y15Y LFYAGAIVFL YTTY VIEN FAPYFAGAVFL YTTY VIEN FAPYFAGAIVFL HTY VIEN FAPYFL HTY VIEN FAPYFAGAIVFL HTY VIEN FAPYFL HTY VIEN FAPYFL F R I L T V Merce P172 VAALP F I T I WQ EKQ EYP VVAALP F F LA LWF EK RENP VVAALP F F LA LWF EK RENP VVAALP F F LA LWF EK RENP VVAALP F F ATMWF EK RENP VVAALP F F ATMWF EK RENP VVAALP F F ATMWF EK RENP nium RIFCSPLOWO2_02_FULL_71_11_0CT058941 Rhodothermus marinus_WP_018242681.1 Halobjorg optimassilinesi, WP_049927041 Halobiorma halotenrestris_SFC35919.1 Haloberigana thermotolerans_WP_005965421.1 Haloberigana thermotolerans_WP_059565421.1 Candidates Kryptonium thompsort, WP_05956543.1 Candidates Kryptonium thompsort, WP_05956543.1 Candidates Kryptonium thompsort, USP58541.1 Asaeclineae bacterium UTCFX1_02195651.1 proteobacteria bacterium UTCFX1_02195651.1 groteobacteria bacterium UTCFX1_0219551.1 Bacteriotebacteria bacterium UTCFX1_0219551.1 Bacteriotebacteria bacterium UTCFX1_0219551.1 Bacteriotebacteria bacterium UTCFX1_0219551.1 groteobacteria bacterium GM2_5_5_8_050P0150.1 Salikonellake bacterium GM2_5_5_8_050P0150.1 min RIFCSPLOW02_0_1FULL_39_10_0CW0563.1 Nitrospine bacterium RI62_16_4_2_2_0201464331 Nitrospine bacterium RI62_4_17_2_0201464331 Nitrospine bacterium RI62_4_5_12_0201464331 Nitrospine bacterium RI62_4_5_12_0201464331 Nitrospine bacterium RI62_4_5_12_0201464331 Nitrospine bacterium RI64_4_5_12_0201464331 Nitrospine bacterium RI64_4_5_12_0201464331 M119 G164 P189 L102 .P A P R L A G I G Y .P Q Q WMA K V G Y .P L P S L A K V G Y .P W T K V A T A G W .P F T K I A K A G W .S V P K L G WA G Y 'T LWR EKR EHP – . T LHRAWKTKVY . NVY NAKR EGTY . NVY NAKR EGTY LIGA ALTACGI ALTACGI ALTACGI ALTACGI CALLEVCV FATLVVA ALTOVAL CALLEVCV FATLVVA ALTOVAL Q244 G240H241 Q245 Rhodothermus mainus. WP 012842581.1 Halopper goleimasaliensis. WP 049927301. Halopper goleimasaliensis. WP 049927301. Halobraka asiaticus. WP 04770404.1 Halobraka asiaticus. WP 04770404.1 Halobrakarium Nubeinse, WP 059055523.1 Therms borksinus. WP 077167471.1 Therms borksinus. WP 077167471.1 Therms borksinus thrompson (VP 075456484.1 Therms borksinus thrompson (VP 075456484.1 Halobrake bacterium UTC7X1, 04798551.1 Bactonoldetise bacterium 01072X1, 04798554.1 Bactonoldetise bacterium 01072X1, 04798541. Bactonoldetise bacterium 010710, 047045854.1 Bactonoldetise bacterium 010710, 047045854.1 Bactonoldetise bacterium 010710, 047045854.1 Bactonoldetise bacterium 010710, 047045854.2 Bactonoldetise bacterium 010710, 047045854 Bactonoldetise bacterium 010710, 047045854 Bactonoldetise bacterium 010710, 04704574 Bactonoldetise bacterium 010710, 04704574 Bactonoldetise bacterium 010710, 04704574 Bactonoldetise bacterium 010710, 0470457 Bactonoldetise bacterium 010710, 0470457 Bactonoldetise bacterium 010710, 04704 Bactonoldetise bacterium 010710, 0470 Bactonoldetise bacterium 010710 A202 W255 Y256 H290 H291 3 G240H241 AAAWYRQMYWIICHGTS GGIYRQWYWIICHGTG AGVYRQWFWIICHGTG AAWYRQMYWIICHGTG AAWYRQMYWIICHGTG AAWYRQMYWIICHGTG AAWYRQMFWIICHGS PUWWRHTYWGFGHGA ASLYRLLFWGFGHPAC GIYRNFFWGFGHPAC AMIA. FDPCIVRNFWCL IDPSAWRLYWCL IDPAWRLIWWCL IDPAWRLIWWCL IDPAWRLIWWCL IDPAWRLIWWCL IDPAWRLIWWCL IDPAWRLIWCL IDPAWRLIWCL IDPAWRLIWCL IDPAWRLIWCL IDPAWRLIWCL IDPAWRLIWCL IDPAWRLIWCL IDAMRUWCL IDACHRYRWC VDPQIYRWWAL VDPQIYRWWWC F FWG V FWG I WWG I WWG I FWG I FWG I FWG VWWG VWWG V FWG VCAMV VSAMV VAAMV VCAMV VCAMV ales bacharlum RBC, 16, 57, 15, 0GT04934.1 Schrie bacharlum WH2, 42, 85, 0GP09150.1 CSPLOWO2, 01, FULL, 39, 10, 0GW069151.1 GRPLOWO2, 01, FULL, 39, 17, 0GW069151.1 frae bacharlum RBC, 16, 64, 22, 0GW069151.1 frait bacharlum RBC, 18, 71, 12, 0GQ10601.1 fraits bacharlum WH2, 45, 12, 0GQ019071.1 FCSPLOWO2, 02, FULL, 71, 11, 0GT96594.1 I AVFILACGATTY IP I AVVALAHGATTY IP I AVVALAHGATTY IP I AVTILLHGAATFVP I AVTILLHGAATFVP I AVTILGHGAATY IP SBOD IVAAMVS VIVI LUGGIT VGAVULIN KI SGGAFFALLET VASE VSSE UFGUUSASVIVI LUGGIT VGAVULIN KI SGGAFFALLET VSSE UFGUUSASVIVI LUGGIT VGAVULIN KI SGGAFFALLET UFGUUSASVIVI LUGGIT VGAVULIN KI SGGAFFALLET UFGUUSASVIVI VSST UFGUUSASVIVI UFGUUSASV M - V DP 0 I I R V V WG C L M S : 3 230 R 334 F A I P A C L A A R K R R C C S 0 F I P A C L A A R K R K R C C -T I P A C L A R K R K R K C C -T I P A C L A R K R K R K C C -T I P A C L A R K R K R K C C -T I P A C L A R K R K R K C C -T I P A C L A R K R K R K C -F V A A L A C R K R K R K C -F V A A L A C R K R K R K C -F V A A L A C R K R K R K C -F V A A L A C R K R K R K C -F V A A L A C R K R K R K C -F S I A A V A L A C R K R K C -F S I A A V A L A C R K R K C -F I A A V A L A C R K R K C -F I A A V A L A C R K R K C -F I A A V A L K C R K -F I A A V A L K C R K -F I A A V A L K C R F R -F I A S I V A L K C R F R num Hir CSP4 UW02, 02, PUL, 71, 1, 01169394.1 Rhodotheimus marinus, WP, 012842861.1 Halobiorma haloterrestris, SFC35919.1 Halobiorma haloterrestris, SFC35919.1 Halobiorma haloterrestris, SFC35919.1 Halobiorgan thermotolerans, WP, 01954644.1 Candidatus Kryptonium thompson, UWP, 015174511.1 Candidatus Kryptonium thompson, UWP, 015165611.1 Salikonellales bacterium GIW2, 25, 25, 02, 014934.1 proteobacteria bacterium GIW2, 25, 28, 03, 00490563.1 nium RiFCSP, UWO, 21, FL, LL, 39, 10, 020W663.1 Nitrospinea bacterium RiBd, 16, 64, 22, 0014633.1 Nitrospinea bacterium RiBd, 21, 20, 00145351.1 proteobacteria bacterium RiBd, 21, 20, 00145351.1 Nitrospinea bacterium RiBd, 16, 64, 22, 0014633.1 Nitrospinea bacterium RiBd, 16, 64, 22, 0014633.1 Nitrospinea bacterium RiBd, 21, 20, 00145351.1 proteobacteria bacterium RiBd, 24, 20, 20, 014413.3 Nitrospinea bacterium RiBd, 24, 20, 20, 01411.3 Nitrospinea bacterium RiBd, 24, 12, 001471.3 proteobacteria bacterium RiBd, 24, 12, 001471.3 nium RiFCSPLOWO2, 02, FULL, 27, 11, 001596394.1 proteobacteria bacterium RiBd, 24, 12, 001471.3 nium RiFCSPLOWO2, 02, FULL, 27, 11, 001596394.1 N306 T307 Y309 HM SWK HWNT SY A VY ST GWR I WNT SY A FY SV GWR VWNT SY A FY ST GWR I WNT SY A FY S318 H321 H388 N389 309 S318 H321 AVY GAVLASMI FAF AFY GATFASLI FAT AAY GAVVASMI FAF LI HVAVLGSLI FA VI YMAVLASLI FC VI YMAVLASLI FC H38.QTQLNLTW, . dVLMQQLQLNMSW TGVMMGQLQLNMSW TGVMMGQLQLNMTW TGVMMGQLQLNMTW TGVMMGQLQLNMTW TGVMMGQLQLNMTW TGVMMGQLQLNMTW - AVSTGWRI - AVSTGWRI - ALSSAWRM GALSSAWTW - GFGAPWKI - GFGAPWKI - GLSFAWKA - VLSSWHKV V L S SWHK V L S TWHK G V S S AWK G V G A TWK G L S S AWK T L S P Y HK nium RIFCSPLOWO2_02_FULL_71_11_0CT96394.1 Rhodothermus marinus_WP_018242631.1 Habotiprogramsalinesia, WP_019270201.1 Habotimaniana, Buoternestris_SFC39391.9 Habotimaniana, Buoternestris_SFC39391.9 Habotimaniana, Buoternestris_SFC39391.1 Candidatus Koptonium thompsoni, UWP_5534.1 Candidatus Koptonium thompsoni, UWP_5534.1 Candetatus Koptonium thompsoni, UWP_5534.1 Satestraides bacterium GWR2_55.26, 20 GP09150.1 Bacteriotedes bacterium GWR2_55.28, 20 GP09150.1 Imm RIFCSPLOW02_01, FULL_39.10, OW08603.1 Nitrosgines bacterium RIR, 16, 64.22, 00L049333.1 Nitrosgines bacterium RIR, 16, 64.22, 00L049334.1 Nitrosgines bacterium RIR, 16, 71, 72, 00021061.1 perioheadertaide bacterium RIR, 16, 74.22, 00L04933.1 Nitrosgines bacterium RIR, 16, 74.12, 00T96579.1 Nitrosgines bacterium RIR, 16, 74.12, 00T96739.1 Nitrosgines bacterium RIR, 16, 74.12, 00T96739.1 Nitrosgines bacterium RIR, 16, 74.12, 00T96739.1 Nitrosgines bacterium RIR, 16, H396 H/K398 T405 R460 R461 461 HSVMSFPCT --- DFSFAAASPLFA- I GOVA TAQV -- VENIPTDFSLGAAPLFN- VFGVV TAV -- VENIPTDFSLGAAPLFN- VFGV TVEV -- VSNIGCTEFDLAASPLFA- VFGI TAV -- VSNIGCTEFSLGAAPLFA- VFGI FGDTFOGGSFF-GFAVPFASTLLAAMGIC WWET-SALVVPY ----- FELLOATLGIC WWET-SALVVPY ----- FELLOATLGIC WWET-FSGAPFSFTFDFAIDUFFMSMMCVC HYD I- TFSGAPFSFTFDFAIDUFFMSMMCVC 5 TFMAL TFMGV TFMGV (AFMGI (AFMG VYFALPVLFNRDLAFK AVA AG<mark>G</mark>MAL /CSAMAV YAGAMG VYFVIRTMFMRKFVSGI VYFVIRTMFMREFVSSK LFFLIRTMFRRQWLSTR VLRLVFGRDW FRRKVVLPRLBAW IFRREWLSKFWARI IFRREWLSKFWARI IFRRELKLKKWAVW IFRRELIWKGFAAI SFSGGPF-THNFNPAVDF SW01-DAQC [------YCAAHWLC SADW-FFACQPL-AYEYPE LAKCMM HWD1-FFACQPL-AYEYPE LAKCMM HWD1-FFACNJ-PFSFDPTVLFTT HWDV-FFADSII-PFSFDPTVLFTT HWDV-FFADSII-PFSFDPTVLFTT HWD1-AFTCAPF-QLAIDPTALLENS HWD1-FFSCAPF-OVFFWD LAFMG MAFMA LAFMG LAFMG LAFMG LAFMG LAFMG LAFMG FRREIIGK FRREIIGK FRREIIGK - AFTGAP F - QLA I DP TA LLML SVMG I G - QFTNAL F - QP P V EAAAY V F LG I MG LG - TF SGAP F - T LE FP AAT NL FF AM FG I G - TF SQAP F - QV E FNP V V N LLQA SMG LG - SFAQAP F - D FQ F SP T V D LM I G I MA I G LOWCE, U.F. Pol. Loodchermus marinus, WP, 012842811. jerg zpolansasilensis, WP, 049927204.1 Halolitoma haldernestris, SFC358101. Spans harmoholena, WP, 006644757.1 bacterium hubelense, WP, 059056826.1 Spans harmoholena, WP, 00547574.1 Spans harmoholena, WP, 01574574.1 Spans harmoholena, WP, 0157474.1 Spans harmoholena, WP, 0 N LA H I AGA SFLLVAVGS LLF GKRV - ENPAL--- YMPPL-VF ALT GGA FFV LAVGS LLF GERI - EFGONDALVADGG-LLA I AGA VF LLAVGS LLF GERI - EFGCNDAVADGG-LLA I AGV FFI LLAVGS LLF GKR I DDESNVS DLVADGG-ALSAV AGV FFV LAVGS LLF GKR - DDESNVS DLVADGG-ALSAV GALFVV AVGS LLF GGF - TPF I VE DRA I S PG-ALSA VGGL FVV AVGS LLF GGF - TPF I VE DRA I S PG-ALSAV GALFVV I VV SV LG GF - TPF I VE DRA I S PG-NVAFI GALVF I LHTVLT I I FGKRV - EGK----- FK FFE-NLAVI GGI MFVTVVFVS VLG FRT - EANKLL - - VSGD-VAAFI GALVF I LHTVLT I I FGKRV - EGK----- FK FFE-NLAVI GGI MFVTVFVS VLG FRT - EANKLL - - VSGD-TI AF GL TFV LV VSV LG GR - TPG V- NGPDMQLK I S PL TI AF GL THVL I VLG VSV FGGV - NGPDMQLK I S PL VI AFI GL I HTVL I VI S VF GGV - NGPDMQLK I S PL - I HEY SMQGT FT LT LI F LAV FA I VY AV NWY LLGR LWA I G-- VHAY EMRGT FV LCI VF LGV FV VA YA LWW FL LT QLWS I GA - VHAY EMRGT FV LCI FF LGV FV VA YA LWW FL LT QLWS I GA - VHAY EMRGT FV LCI FF LGV FV VA YA LWW FL LT QLWS I GA - PHOS MRGT FV LCI VF LAAFV I TY LLWW FL TQLWS I GA - PHOS MRGT FV LCI VF LAAFV LI YY LLWW FL TQLWS I GA - - - - TST FGT I VI LI FF FVL AMF HAV HY FL SQAWFI GP - - - TST FGT VI LAI FF FL I MY FL MY HWY FL SQAWFI GP Rhodothermus marinus WP lopiger goleimassiliensis WP Halobiforma haloterrestris, Haloviax asiaticus WP lerrigena thermotolerans WP lalobacterium hubeiense WP Thermus brockianus WP s Kryptonium thompsoni WP L GR L WA T G - -L T Q L WS T G A -L S Q V WE T G P -L S Q A WP T G T R L S R V WE V R - -VVQVSNPDEK-----GNPVLESTKLTDHMME I NWYWL SR VWEVR T NWWLLGR TWFIR APQAKEHKG allionellales bacterium 7880, 16, 57, 15, OGT0984.1 interobacteria bacterium GW2, 42, 48, OGP09150.1 um RIFCSPLOWO2, 01, FULL, 39, 10, OGW04603.1 um RIFCSPLOWO2, 01, FULL, 39, 17, OGW05810.1 Nitrospitae bacterium RBG, 16, 64, 22, OGW58853.1 um RIFCSP-RIORC, 12, FULL, 37, 22, OGL04331.1 toobacteria bacterium GW2, 42, 12, OGP09171.1 ium RIFCSPHICHUM, 042, 02, FULL, 71, 11, OGT96394.1 GAMY GAMY GAMY GALY Deltapro Delta

GLIAVTGGAN

Materials and Methods

Growth and Expression Conditions

Rhodothermus marinus DSM 4252 was inoculated from frozen stock and grown in 5 ml of DSM Medium 630 with 10 g/L NaCl at 60 °C for 36 hrs. It was then inoculated into a larger secondary culture and grown overnight. 25 ml of the culture was inoculated into 1 L of medium with 30 mM nitrate added. The cells were shaken at 75 rpm and grown at 60 °C. The cells were pelleted by centrifugation at 8000 rpm. The cell pellet was either directly used for protein purification or frozen at -80 °C until the time of use.

Labeled ¹⁵NO experiments

We used labeled nitrate (¹⁵NO₃²⁻) to verify that *Rhodothermus marinus* DSM 4252 was capable of complete denitrification (NO₃⁻ to N₂) using eNOR as the sole nitric oxide reductase. Cultures were inoculated into flasks containing media with either ¹⁴NO₃²⁻, ¹⁵NO₃²⁻, or no nitrate. The cultures were then allowed to grow microaerobically for a period of time before being subsampled for transfer to sealed vials in order to allow accumulation of gaseous end products. Samples were taken from each media composition after 0, 3, 6, 10, and 17 hours. The headspace was sampled after 20 hours of growth in sealed vials via gastight GC syringe and immediately injected into a Hewlett Packard 5972 gas chromatograph/mass spectrometer. Chromatogram peaks corresponding to isotopologues of NO, N₂O, and N₂ were identified by their mass spectra and peak areas were quantified relative to ambient air. As ¹⁵N cultures were grown in isotopically pure ¹⁵NO₃²⁻, complete denitrification should result in accumulation of ³⁰N₂ at a 1:2 ratio relative to nitrate consumption. ³⁰N₂ should only accumulate if eNOR is functioning as part of a complete denitrification pathway. If eNOR does not function effectively as a nitric oxide reductase, then ¹⁵NO should be seen to accumulate. Instead, only the ³⁰N₂ peak was observed, indicating the eNOR

functioned effectively as a nitric oxide reductase for denitrification. Over the course of incubations, ${}^{30}N_2$ was seen to accumulate to more than 50x background. ${}^{14}N$ samples showed no significant accumulation of ${}^{30}N_2$ above background, confirming that the ${}^{30}N_2$ in ${}^{15}N$ samples was due to denitrification of labeled nitrate. NO was not seen to accumulate in any of the cultures. These results demonstrate that eNOR is a functional nitric oxide reductase and can be used as part of a complete denitrification pathway.

Purification of eNOR

The culture of *Rhodothermus marinus*, once harvested, was re-suspended in 100 mM Tris-HCl, pH 8 with 10 mM MgCl₂ and 50 μ g/ml DNase, using a Bamix homogenizer. The resulting solution was spun down at 42000 rpm in a Beckman Ultracentrifuge. The membrane pellet was collected and re-suspended in 20 mM Tris-HCl, pH 7.5, 1 % CHAPS (Affymetrix) to a final concentration of 40-50 mg/ml. The solution was stirred at 4 °C for 1 hr. In this step a lot of peripheral membrane proteins appear to be solubilized and the remaining protein is pelleted by spinning down at 42000 rpm for 1 hr. The remaining pellet is then solubilized in 20 mM Tris-HCl, pH 7.5, 1 % DDM (Affymetrix) at a final protein concentration of around 5-10 mg/ml. The DDM solubilized fraction was once again centrifuged at 42000 rpm to pellet down protein that was not solubilized.

The solubilized protein was then loaded on a DEAE CL-6B (Sigma) column, pre-equilibrated in 20 mM Tris-HCl, pH 7.5, 0.05 % DDM, and subjected to a linear gradient spanning from 0 to 500 mM NaCl. The fraction containing the eNOR, identified using a peak at 591nm, corresponding to the peak of cytochrome 'a1' in *Magnetospirillum magnetotacticum*¹, eluted at around 200 mM salt. This fraction was then loaded on a Q Sepharose High Performance (GE Healthcare) column, pre-equilibrated with 20 mM Tris-HCl, pH 7,5, 0.05 % DDM and then eluted in a gradient from 0 to 1 M NaCl. The eNOR containing fraction was eluted at around 250 mM salt and the eluted

fraction was then loaded on a Chelating Sepharose (GE Healthcare) column, loaded with Cu^{2+} and equilibriated with 20 mM Tris-HCl, 500 mM NaCl, as previously described for cytochrome *caa*₃ from *Rhodothermus marinus*². The eNOR fraction was once again loaded on a Q Sepharose High performance column, and a gradient was run between 0 and 300 mM NaCl at low flow rates (0.5 ml/min) and the first peak was found to be the eNOR.

Gel Electrophoresis

The purified eNOR was run on a Tris-Hepes 4-20 % acrylamide gel (NuSep) in the recommended Tris-Hepes-SDS running buffer at 120 V for ~1 hr. The protein was visualized and compared to the Precision Plus Protein[™] Dual Color Standards (BIO-RAD).

UV-Visible Spectroscopy

All spectra were recorded on a HP Agilent 8453 UV-Vis spectrophotometer using a quartz cuvette from Starna Cells (No. 16.4-Q-10/Z15). Potassium Ferricyanide was used to obtain the oxidized spectrum, and dithionite was used to obtain the reduced spectrum.

Pyridine Hemochrome Assay

The hemes in eNOR were analyzed using a pyridine hemochrome assay³. A stock solution of 200 mM NaOH with 40 % pyridine was prepared. The stock solution was mixed 1:1 with the protein and an oxidized spectrum was obtained by adding 3 μ l of 100 mM K₃Fe(CN)₆. A reduced spectrum was similarly prepared by adding a few crystals of sodium dithionite. The reduced minus oxidized spectrum was used to identify the heme co-factors

Heme extraction and HPLC Analysis

The hemes from eNOR were extracted and analyzed using an HPLC elution profile according to established protocols^{4,5}. 50 μ l of eNOR was mixed with 0.45 ml of acetone / HCl (19:1) and incubated for 20 minutes at room temperature after shaking. The mixture was centrifuged at 14,000 rpm for 2 minutes, followed by addition of 1 ml of ice cold water, and 0.3 ml of 100% ethyl acetate to the supernatant. The water/ethyl acetate mixture was vortexed and centrifuged again for 2 minutes. The ethyl acetate phase was recovered and concentrated using a speed vac.

The extracted hemes were analyzed using an Agilent 1290 Infinity LC attached to an Agilent 6230 TOF LC/MS equipment by separation using an Agilent Eclipse Plus C18 column (2.1x300 mm, 1.8 µm, 600 bar) and an acetonitrile (0.05 %TFA) / water (0.05 % TFA) gradient from 20 to 95 %.

NO reductase activity verification using GC

Anaerobic reaction conditions were set up in a 5 mL clear serum vial (Voigt Global Distribution, Inc) sealed with a 20 mm rubber stopper, by passing N₂ through 2 ml of 20 mM KPi, 0.05 % DDM, pH 7.5 with 1 mM TMPD, 5 mM Ascorbate. A control was performed by adding only 50 μ M NO. Sample reactions were begun by adding eNOR to a final concentration of 100 nM. The reaction was incubated at 42 °C for half an hour before the headspace was injected into an HP Agilent 5890 Series GC, fitted with a TCD and ECD (SRI Instruments) for verification of N₂O production.

Turnover measurement using a Clark electrode

A sealed chamber fitted with an ISO-NO (World Precision Instruments) electrode was used for NO reductase activity measurements. 1 mM TMPD or 100 μ M PMS and 4 mM Ascorbate were was added to 2 ml 50 mM Citrate, pH 6, 0.05 % DDM in the reaction chamber and all traces of oxygen were removed by passing water-saturated Argon for 20 minutes through the solution. This

is similar to the protocol described for cNOR from *Thermus thermophilus*⁶. The buffer system also contained an oxygen scavenging system constituting 100 nM catalase, 35 nM Glucose oxidase and 90 nM Glucose. The NO reduction traces were recorded using a Duo-18 (World Precision Instruments), and activities calculated from the slope of the traces.

LC/MS/MS analysis

Mass spectrometric analysis was conducted at the Protein Sciences Facility, Roy J Carver Biotechnology Center, University of Illinois, Urbana, IL 61801 using a Thermo LTQ Velos ETD pro mass spectrometer. For liquid samples, the samples were cleaned up using G-Biosciences Perfect Focus (St. Louis MO) prior to digestion with trypsin. Digestion was done using proteomics grade trypsin 1:20 (G-Biosciences, St. Louis, MO) and a CEM Discover Microwave Reactor (Mathews, NC) for 15 minutes at 55° C at 50 Watts. Digested peptides were extracted 3X using 50% acetonitrile containing 5% formic acid, pooled and dried using a Speedvac (Thermo Scientific). The dried peptides were suspended in 5% acetonitrile containing 0.1% formic acid and applied to LC/MS.

HPLC for the trypsin digested peptides was performed with a Thermo Fisher Dionex 3000 RSLCnano using Thermo Acclaim PepMap RSLC column (75 μ m x 15 cm C-18, 2 μ m, 100Å) and a Thermo Acclaim PepMap 100 Guard column (100 μ m x 2 cm, C-18, 5 μ m, 100Å), solvents were water containing 0.1% formic acid (A) and acetonitrile containing 0.1% formic acid (B) at a flow rate of 300 nanoliters per minute at 40° C. Gradient was from 100% A to 60% B in 60 minutes. The effluent from the UHPLC was infused directly into a Thermo LTQ Velos ETD Pro mass spectrometer.

Control and data acquisition of the mass spectrometer was done using Xcalibur 2.2 under data dependent acquisition mode, after an initial full scan, the top five most intense ions were subjected to MS/MS fragmentation by collision induced dissociation. The raw data were processed by Mascot Distiller (Matrix Sciences, London, UK) and then by Mascot version 2.4. The result was searched against NCBI NR Protein database.

Analysis of heme-copper oxygen reductase phylogeny and distribution in environmental datasets

We performed a large-scale analysis of heme-copper oxygen reductase (HCO) protein sequences in the NCBI and IMG databases with BLASTP using an e-value of 1e-³ to generate a database of HCO sequences that had at least some of the conserved amino acids previously identified in subunit I⁷. We then used the database of HCOs, filtered it with a sequence cut-off of 50% to generate the multiple sequence alignment, **MSA1**. A phylogenetic tree (**Figure 2**) was inferred using IQ-TREE 2⁸ with the Dayhoff substitution matrix, Gamma model of rate heterogeneity and 1000 ultrafast bootstraps. Using the curated HMMs for each of the HCO family oxygen reductases⁹, we probed release 202 of the Genome Taxonomy Database¹⁰ for distribution of the NOR families – eNOR, bNOR, sNOR, nNOR, gNOR, cNOR and qNOR – in bacteria and archaea. Curated HMMs for the nitrate reductases (NapAB, NarGH), nitrite reductases (NirK, NirS) and nitrous oxide reductases (NosD and NosZ) were sourced from the HMMs database of MagicLamp¹¹. Analysis of HCO distribution in various ecosystems were performed using the metagenomes in the IMG database. Approximately 2300 metagenomes were identified which were sourced from 44 environments identified by IMG. The number of different HCOs in each of these environments were extracted using BLASTP and query sequences that belong to each of the different HCO families.

References

- Tamegai, H., Yamanaka, T. & Fukumori, Y. Purification and properties of a 'cytochrome a1'-like hemoprotein from a magnetotactic bacterium, Aquaspirillum magnetotacticum. *Biochim. Biophys. Acta BBA - Gen. Subj.* 1158, 237–243 (1993).
- Pereira, M. M. *et al.* The caa3 terminal oxidase of the thermohalophilic bacterium Rhodothermus marinus: a HiPIP:oxygen oxidoreductase lacking the key glutamate of the Dchannel. *Biochim. Biophys. Acta BBA - Bioenerg.* 1413, 1–13 (1999).
- 3. Berry, E. A. & Trumpower, B. L. Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. *Anal. Biochem.* **161**, 1–15 (1987).
- Sone, N. & Fujiwara, Y. Haem O can replace haem A in the active site of cytochrome c oxidase from thermophilic bacterium PS3. *FEBS Lett.* 288, 154–158 (1991).
- Reinhold-Hurek, B. & Zhulin, I. B. Terminal oxidases of Azoarcus sp. BH72, a strictly respiratory diazotroph. *FEBS Lett.* 404, 143–147 (1997).
- Schurig-Briccio, L. A. *et al.* Characterization of the nitric oxide reductase from Thermus thermophilus. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 12613–12618 (2013).
- Hemp, J. & Gennis, R. B. Diversity of the Heme–Copper Superfamily in Archaea: Insights from Genomics and Structural Modeling. *Bioenergetics* 1–31 (2008).

- Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37, 1530–1534 (2020).
- Murali, R., Hemp, J., Orphan, V. & Bisk, Y. FIND: Identifying Functionally and Structurally Important Features in Protein Sequences with Deep Neural Networks. *bioRxiv* 592808 (2019) doi:10.1101/592808.
- 10. Parks, D. H. *et al.* A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).
- Garber, A. MagicLamp: toolkit for annotation of 'omics datasets using curated HMM sets.
 (2020).