Fe oxidation by a fused cytochrome-porin common to diverse Fe-oxidizing bacteria

Clara S. Chan[1,2]*, Sean M. McAllister[2]^, Arkadiy Garber[1]^, Beverly J. Hallahan[1], Sharon Rozovsky[3]

[1]Dept. of Geological Sciences, University of Delaware, Newark, DE

[2]School of Marine Science and Policy, University of Delaware, Newark, DE

[3]Dept. of Chemistry and Biochemistry, University of Delaware, Newark, DE

*Corresponding author (cschan@udel.edu, Dept. of Geological Sciences, University of Delaware, 103 Penny Hall, Newark, DE 19716)

^Equal contributions

Running title: Fe oxidation by a fused cytochrome-porin

Summary

Fe oxidation is one of Earth's major biogeochemical processes, key to weathering, soil formation, water quality, and corrosion. However, our ability to track the contributions of Fe-oxidizing microbes is limited by our relatively incomplete knowledge of microbial Fe oxidation mechanisms, particularly in neutrophilic Fe-oxidizers. The genomes of many Fe-oxidizers encode homologs to an outer-membrane cytochrome (Cyc2) that has been shown to oxidize Fe in two acidophiles. Here, we demonstrate the Fe oxidase function of a heterologously expressed Cyc2 homolog derived from a neutrophilic Fe oxidizer. Phylogenetic analyses show that Cyc2 from neutrophiles cluster together, suggesting a common function. Sequence analysis and modeling reveal the entire Cyc2 family is defined by a unique structure, a fused cytochrome-porin, consistent with Fe oxidation on the outer membrane, preventing internal Fe oxide encrustation. Metatranscriptomes from Fe-oxidizing environments show exceptionally high expression of *cyc2*, supporting its environmental role in Fe oxidation. Together, these results provide evidence that *cyc2* encodes Fe oxidases in diverse Fe-oxidizers and therefore can be used to recognize microbial Fe oxidation. The presence of *cyc2* in 897 genomes suggests that microbial Fe oxidation may be a widespread metabolism.

Introduction

Fe oxidation occurs in virtually all near-surface environments, producing highly reactive Fe oxyhydroxides that often control the fate of carbon, phosphorous, and other metals (Borch *et al.*, 2010). It is commonly assumed that abiotic mechanisms are sufficient to account for Fe oxidation, particularly at near-neutral pH. However, Fe-oxidizing microbes are increasingly observed in a wide range of environments (Emerson *et al.*, 2010; Kappler *et al.*, 2015), leading

us to ask to what extent microbes drive Fe oxidation. To address this, we need to confidently identify the Fe oxidase. But unlike other major microbial metabolisms, we have relatively incomplete knowledge of Fe oxidation pathways (Bird *et al.*, 2011; Hedrich *et al.*, 2011; Ilbert and Bonnefoy, 2013), and no candidates for a broadly distributed Fe oxidasehave emerged until now.

Rising interest in Fe-oxidizing microbes has resulted in a surge of Fe-oxidizer sequencing, including isolate genomes, single cell genomes, metagenomes, and metatranscriptomes (refs. in Table 1) (Moya-Beltrán *et al.*, 2014; Quaiser *et al.*, 2014; Fukushima *et al.*, 2015; Jewell *et al.*, 2016; Fullerton *et al.*, 2017), enabling us to search for the genes involved in microbial Fe oxidation. Analyses of chemolithotrophic and phototrophic Fe-oxidizer isolate genomes have revealed that most possess homologs to *cyc2* (Emerson *et al.*, 2013; Kato *et al.*, 2015; Chiu *et al.*, 2017; Mori *et al.*, 2017; Crowe *et al.*, 2017), which encodes an Fe-oxidizing outer membrane cytochrome first characterized in *Acidithiobacillus ferrooxidans* (Castelle *et al.*, 2008). Fe oxidation has also been shown for a distant Cyc2 homolog, Cyt_{572}, purified from an acid mine drainage *Leptospirillum sp.* (Jeans *et al.*, 2008). In all, this suggests that Cyc2 may be an Fe oxidase in a wide range of Fe-oxidizers.

To prove this, we need functional information on Cyc2 from neutrophilic chemolithotrophic Fe-oxidizing bacteria (FeOB). Dark neutral pH environments are prevalent, and to date, these FeOB have been found in a wide variety of marine, terrestrial, and engineered environments, including aquifers, soils, sediments, hydrothermal vents, and water treatment systems (Kappler *et al.*, 2015; Emerson and de Vet, 2015). In these environments, FeOB grow by coupling Fe oxidation to the

65 reduction of O_2 or nitrate, using this energy to fuel carbon fixation, thus serving as primary
66 producers (Emerson *et al.*, 2010). Known neutrophilic chemolithoautotrophic FeOB mostly fall
67 within the marine Zetaproteobacteria (*Mariprofundus spp*., *Ghiorsea spp*.) and freshwater
68 Betaproteobacteria (Gallionellales genera *Gallionella*, *Sideroxydans,* and *Ferriphaselus*)
69 (Emerson *et al.*, 2010; Kato *et al.*, 2015; Mori *et al.*, 2017). All sequenced genomes of
70 Zetaproteobacteria and Gallionellales FeOB have *cyc2* homologs, and these sequences are
71 among the closest homologs to one another despite being from separate classes of Proteobacteria,
72 forming a cluster separate from the acidophilic *cyc2* (Kato *et al.*, 2015; He *et al.*, 2017). A
73 second potential Fe oxidase gene, *mtoA*, was found in the Gallionellales *Sideroxydans*
74 *lithotrophicus* ES-1 (Emerson *et al.*, 2013), and functional and genetic information supports the
75 role of MtoA and its homolog PioA in Fe oxidation (Jiao and Newman, 2007; Liu *et al.*, 2012).
76 However, few other FeOB genomes contain *mtoA*, suggesting that Cyc2 is potentially a more
77 widespread Fe oxidase.
78
79 Thus, we set out to demonstrate the function of a Cyc2 from a neutrophilic Fe-oxidizer. We first
80 analyzed the Cyc2 family phylogeny and then made structure-function predictions, which
81 informed the design of the gene constructs that we expressed in *E. coli*. To support the Fe
82 oxidase function, we performed whole cell Fe oxidation assays on Cyc2-expressing *E. coli*. To
83 determine the environmental relevance, we analyzed *cyc2* expression in a new marine Fe mat
84 metatranscriptome and reanalyzed a published Fe-rich aquifer metatranscriptome (Jewell *et al.*,
85 2016). Finally, we compare the genomic distribution and expression of *cyc2* and *mtoA*, to better
86 understand the relative significance of these two putative Fe oxidases.
87

88 **Phylogeny of Cyc2**

89 We started by producing a comprehensive phylogenetic tree of Cyc2 sequences acquired from
90 databases (National Center for Biotechnology Information (NCBI) and
91 Integrated Microbial Genomes (IMG)). To ensure that we were analyzing true homologs, we
92 screened the sequences for appropriate length (352 to 587 aa, average/median 446 aa), the
93 cytochrome c binding motif CXXCH, and beta barrel porin portion (see Cyc2 structure and
94 conservation below). Our search yielded 897 unique near-full length sequences, which were
95 reduced to 530 sequences when closely-related sequences were removed. The resulting Cyc2 fell
96 into three distinct clusters, with sequences distributed amongst various bacterial taxa, largely
97 Proteobacteria (**Fig. 1, Supplemental File 1**). Cyc2 homologs are present in all well-established
98 neutrophilic microaerophilic chemolithotrophic FeOB (**Table 1**), many of which are obligate
99 FeOB that lack other apparent Fe oxidase candidates. These microaerophilic FeOB Cyc2
100 sequences form a well-supported cluster (Cluster 1 in **Fig. 1**), with the marine Zetaproteobacteria
101 within one subcluster, and the freshwater Gallionellales forming a separate subcluster that also
102 includes the neutrophilic photoferrotrophic *Chlorobi*. This clustering suggests a common
103 function for the Cyc2 in these neutrophilic FeOB.

104

105 Beyond neutrophilic FeOB, *cyc2* is found in the genomes of various acidophilic FeOB, which
106 suggests a common adaptation to Fe oxidation. There are homologs in various acidophilic FeOB
107 genomes: *Ferrovum spp*., *Thiomonas spp*., and Burkholderiales GJ-E10, in addition to the
108 functionally-verified Cyc2 from *A. ferrooxidans* and *L. ferriphilum*. Unlike the neutrophiles, the
109 Cyc2 sequences from acidophiles do not form a single cluster, and instead are scattered across
110 the Cyc2 tree. Notably, the two functionally verified Cyc2 from *A. ferrooxidans* and *L.*

111 *ferriphilum* fall in different regions of the tree, Clusters 2 and 3 respectively. This presents the
112 intriguing possibility that many or all of the Cyc2 homologs are Fe oxidases.
113
114 At first glance, a common Fe oxidation pathway for both neutrophiles and acidophiles might not
115 be expected, due to the drastically different redox potential of Fe(II)/Fe(III) at acidic versus
116 neutral pH (770 mV at pH 2 vs. 24 mV at pH 7 (Bird *et al.*, 2011; Majzlan, 2013)). However, the
117 Cyc2 tree shows horizontal transfer between various lineages, and furthermore, certain
118 neutrophilic and acidophilic FeOB show signs of horizontal transfer of other electron transport
119 genes. Specifically, *S. lithotrophicus* ES-1, *Gallionella acididurans*, *M. ferrooxydans* PV-1,
120 *Ferrovum spp*., and Burkholderiales GJ-E10 all share a gene cassette that includes cbb_3-type
121 cytochrome c oxidase genes and the periplasmic cytochrome *cyc1* gene (**Supplemental Fig. 1A**).
122 The conservation in synteny and sequence homology signifies horizontal transfer, strongly
123 suggesting a common Fe oxidation pathway amongst diverse FeOB, involving Cyc2, Cyc1, and
124 cytochrome c oxidase (**Supplemental Fig. 1B**).
125
126 **Cyc2 structure and conservation**
127 To better understand the potential role of Cyc2 homologs and to prepare for functional studies,
128 we performed sequence and structure predictions, focusing on Cyc2 from FeOB. Despite the
129 great sequence diversity, the FeOB Cyc2 are all predicted to have a unique structure, a fused
130 cytochrome-porin. All contain a signal sequence, a c-type cytochrome, and a porin
131 (**Supplemental Fig. 2**). The signal sequence was predicted by SignalP, indicating that the protein
132 is exported to the periplasm. The cytochrome portion is identifiable by a single CXXCH heme-
133 binding motif, and is by far the most conserved part of the sequence (**Fig. 2; Supplemental Fig.**

134 **2, 3**). Conserved residues include a AXPXFAR[Q/K][T/Y] motif located 5 amino acids upstream
135 of the CXXCH heme binding site (AXPXFARQT in Clusters 1 and 2 sequences; AXPXFARKY
136 in Cluster 3). There is also a PXL motif 4 amino acids downstream of the CXXCH. This PXL
137 motif can be found in many other cytochromes, such as the structurally characterized MtoD
138 (Beckwith *et al.*, 2015) and Cyc1 gene in *Acidithiobacillus ferrooxidans* (CYC41 in (Abergel *et*
139 *al.*, 2003)); the proline and lysine appear to help stabilize the heme (Abergel *et al.*, 2003). In
140 contrast, the AXPXFAR[Q/K][T/Y] motif is unique to Cyc2, and therefore could be used to
141 distinguish Cyc2-like outer membrane cytochromes.
142
143 The rest of the sequence corresponds to a porin, based on the presence of beta strands predicted
144 by PSIPRED (McGuffin *et al.*, 2000) (**Supplemental Fig. 4**) and homology matching by
145 HHpred (Söding *et al.*, 2005). This C-terminal section has low sequence homology, but high
146 structural homology to the outer membrane phosphate-selective porins OprO and OprP (PDB
147 structures 4RJW and 2O4V (Moraes *et al.*, 2006; Modi *et al.*, 2015); **Supplemental Table 1**).
148 Poor sequence conservation is typical of porins (Nikaido, 2003), and since the porin portion
149 constitutes most of the sequence, this explains why Cyc2 homologs tend to have low amino acid
150 identity (**Supplemental Fig. 3**). The porin structure is further supported by consistent 3D models
151 calculated by iTasser, MODELLER, and Phyre (Söding *et al.*, 2005; Zhang, 2008; Kelley *et al.*,
152 2015; Webb and Sali, 2016) (**Fig. 3**; **Supplemental Fig. 5**), which predict a barrel of 16 beta
153 strands, a common size porin (Schulz, 2004). Because porins are located in the outer membrane
154 (Hancock, 1987), Cyc2 is clearly an outer membrane cytochrome. This location is consistent
155 with our previous observations that Fe oxidation occurs at the cell surface, preventing internal Fe
156 oxide encrustation (Chan *et al.*, 2011; Comolli *et al.*, 2011).

157

158 **Cyc2 heterologous expression and functional assay**

159 Cyc2 of neutrophiles are in a distinct cluster from functionally-characterized Cyc2 homologs, so
160 its function still requires experimental evidence. Neutrophilic FeOB are challenging to grow in
161 quantities sufficient for protein assays, so we took a heterologous expression approach. The *cyc2*
162 gene sequence from *Mariprofundus ferrooxydans* PV-1 ($cyc2_{PV\text{-}1}$) was prepared for expression in
163 *E. coli* by (1) codon optimization, (2) replacing the signal sequence with that of the *E. coli* outer
164 membrane protein OmpA, and (3) adding a StrepII-tag at the C-terminus (**Supplemental Fig.**
165 **6A**). The resulting sequence was synthesized, cloned into pMAL-p4X, transformed into *E. coli*
166 C43(DE3), and co-expressed with the pEC86 plasmid containing the *ccm* cytochrome c
167 maturation genes under a constitutive promoter, to ensure proper cytochrome maturation under
168 aerobic conditions (Arslan *et al.*, 1998). The protein appeared to be somewhat toxic to *E. coli*, as
169 the yield of Cyc2-expressing cells (OD_{600} = 1.1) was much lower than that of cells with empty
170 vectors (2x-diluted OD_{600} = 1.9) even at low IPTG concentrations of 0.5 mM. Nevertheless,
171 expression was successful, as shown by western blot using Strep-Tactin antibody against the
172 protein N-terminal tag (**Fig. 4A**). This band runs close to the expected molecular weight of 43
173 kDa and contains heme, as established by heme-specific staining (**Fig. 4B**).

174

175 We tested the function by assaying Fe oxidation by Cyc2-expressing *E. coli* cells. Because of the
176 low expression levels, we tested relatively dense cell suspensions (OD=2), washed and
177 resuspended in fresh LB medium, buffered to pH 6 to help slow abiotic Fe oxidation. Fe(II) was
178 added from an anoxic stock solution of $FeCl_2$, to a concentration of 100 μM. The dense cell
179 suspension appeared to stabilize Fe(II), as cells with an empty vector (i.e. plasmid without *cyc2*)

180 showed considerably slower Fe oxidation relative to cell-free medium (**Supplemental Fig 7A**).
181 Cyc2-expressing cells did indeed oxidize Fe(II): cells oxidized 41% of the Fe(II) within 2
182 minutes, and 73% within 10 min (**Fig. 4C**). On further addition of Fe(II) at 45 min, the Cyc2-
183 expressing cells continued to oxidize Fe(II). In contrast, the empty vector control oxidized 14%
184 of the Fe(II) in 10 min (**Fig. 4C**). Consistent results were obtained in triplicate experiments from
185 cells taken from one specific expression stock (**Fig. 4**), as well as replicate assays using cells
186 from different expressions (**Supplemental Fig. 7B-D**). Azide (3 mM) reduced Fe oxidation by
187 50%. Azide inhibits cytochromes by binding Fe in heme (Yoshikawa *et al.*, 1998), suggesting
188 that Cyc2 and/or the terminal cytochrome c oxidase were partially inhibited, though 3 mM azide
189 may not have been sufficient to completely inhibit such a dense cell suspension. To confirm that
190 Fe oxidation is due to the Cyc2 cytochrome, we expressed and assayed the porin portion of
191 $Cyc2_{PV-1}$ ($Cyc2_{porin}$, 37 kDa as expected; **Fig. 4A**). These $Cyc2_{porin}$-expressing cells oxidized
192 Fe(II) very slowly (17% in 10 min; **Fig. 4C**), demonstrating that the cytochrome is required for
193 Fe oxidation. Taken together, the data show that Cyc2-expressing *E. coli* accelerate Fe oxidation,
194 so we conclude that Cyc2 confers the ability to oxidize Fe(II).
195
196 **Expression of *cyc2* in the environment**
197 If Cyc2 is an Fe oxidase, we would expect high expression of *cyc2* in Fe-oxidizing
198 environments. To investigate this, we analyzed metatranscriptomes from two ecosystems
199 dominated by Fe-oxidizing bacteria. We present a new dataset from the Loihi Seamount
200 (Hawaii) hydrothermal vent Fe microbial mat and reanalyze an existing dataset from an Fe-rich
201 alluvial aquifer in Rifle, Colorado (Jewell *et al.*, 2016).
202

203 At the Loihi seamount, Fe-oxidizing microbial mats thrive where hydrothermal vents emit
204 Fe(II)-rich fluids into oxygenated seawater (up to 700 μM Fe, <3 to 52 μM O_2 in the mats
205 (Glazer and Rouxel, 2009)). Here, Fe(II) is by far the most abundant electron donor, with
206 relatively low or localized sulfide (Glazer and Rouxel, 2009). Using a remotely operated vehicle
207 and syringe-based biomat sampler (Breier *et al.*, 2012) containing RNA Later, we obtained a 17
208 mL sample of a surface mat for metagenomic and metatranscriptomic analyses. The microbial
209 community was almost entirely Zetaproteobacteria (94.4% based on metagenome coverage), a
210 class in which all cultured representatives are microaerophilic FeOB (n=15 (Emerson and
211 Moyer, 2002; Emerson *et al.*, 2007; McAllister et al., 2011; McBeth et al., 2011; Field *et al.*,
212 2015; Makita *et al.*, 2016; Mumford *et al.*, 2016; Mori *et al.*, 2017; Chiu *et al.*, 2017; Laufer et
213 al., 2017; Barco et al., 2017; Beam *et al.,* in press)). One specific Zetaproteobacteria comprised
214 79.4% of the community. Overall, Zetaproteobacteria were also the most active, representing
215 90.5% of the mapped transcripts, with the dominant bin accounting for 78.9% of the transcripts.
216 For this dominant Zetaproteobacteria bin, *cyc2* is among the most highly expressed genes (91-
217 99% percentile; **Supplemental Fig. 8A**). We looked for genes that indicated other possible
218 respiratory metabolisms, but found none. This, combined with the high expression of *cyc2*,
219 strongly suggests that *cyc2* is a key gene in this marine Fe oxidation-based ecosystem.
220
221 In contrast to Zetaproteobacteria, freshwater Gallionellaceae genomes contain a wider range of
222 electron transport genes: *cyc2*, *mtoA*, and in some cases the sulfur oxidation *sox* and reverse *dsr*
223 genes. To see which were most highly expressed, we reanalyzed existing metatranscriptomes
224 from an oxidation experiment at the well-studied Rifle aquifer. Situated next to the Colorado
225 River, this aquifer was the subject of a long-term experimental study on uranium remediation by

226 Fe- and S-reducing microbes (Williams *et al.*, 2011). Acetate amendments resulted in a highly
227 reduced zone with a large reservoir of Fe(II) minerals (Williams *et al.*, 2009). Subsequently,
228 Jewell et al. (2016) re-oxidized some of this Fe(II) by injecting nitrate-amended oxic
229 groundwater, causing a bloom of Gallionellaceae. As the authors reported, *cyc2* was among the
230 most highly expressed genes, at 99.99-100th percentile in all three post-injection samples
231 (**Supplemental Fig. 8B**). However, the *cyc2* expression levels were not compared to *mtoA*, *sox*,
232 and *dsr* genes, so their relative importance was unclear. Our re-analysis shows that the *cyc2* gene
233 was expressed at much higher levels than *mtoA*, *sox* genes, and *dsr* genes (**Supplemental Table**
234 **2**). In particular, *cyc2* expression was approximately two orders of magnitude higher than *mtoA*
235 (**Table 2; Supplemental Fig. 9**). This was true in individual bins when both were expressed, and
236 also in total across all bins expressing these genes. Thus, while both putative Fe oxidases were
237 expressed when Fe oxidation was stimulated, *cyc2* was clearly preferentially expressed.
238
239 We can gain insight into the different niches of *cyc2* and *mtoA* by examining the temporal
240 expression patterns of different Gallionellaceae spp., (i.e. genomic bins; **Supplemental Fig. 9**).
241 Four species/bins expressed both *cyc2* and *mtoA* while >5 bins only expressed either *cyc2* or
242 *mtoA*. Overall, both putative Fe oxidase genes increased in expression level over time, but *mtoA*
243 slightly peaked at the 3rd time point, when the aquifer was largely anoxic (as indicated by a low
244 in terminal oxidase expression). The bin with the highest *mtoA* expression (22.6), showed peak
245 *mtoA* expression during this anoxic period. This suggests that *mtoA* may be more useful under
246 conditions of electron acceptor limitation.
247

248 The difference in *cyc2* and *mtoA* expression at the Rifle site led us to ask whether *cyc2* and
249 *mto/mtr* homologs commonly co-occur in genomes (*mtrABC* encodes for an Fe reductase system
250 homologous to the proposed MtoAB Fe oxidase system). We found that although *cyc2* and *mto*
251 genes co-occur in Gallionellaceae genomes, they are very rarely found in the same genome
252 (**Supplemental Fig. 10**), suggesting that *cyc2* and *mto/mtr* do indeed correspond to different
253 niches.
254
255 **Cyc2 homologs in other Fe-cycling and extracellular electron-transporting organisms**
256 Of the 897 Cyc2 homologs analyzed (represented in **Fig. 1**, fully labeled tree in **Supplemental**
257 **File 1**), very few are from genomes of well-established Fe-oxidizing taxa, in part because Fe
258 oxidation is not typically tested in isolates. This brings up the question of whether or not all
259 Cyc2 represent Fe oxidases. Although we do not know yet, we can gain insight from examining a
260 few homologs in other organisms known to cycle Fe and/or engage in extracellular electron
261 transport.
262
263 At least one of the organisms with *cyc2*, *Dechloromonas aromatica* RCB (Coates *et al.*, 2001), is
264 reported to be an anaerobic FeOB, coupling Fe oxidation with denitrification (Salinero *et al.*,
265 2009). Nitrate-dependent Fe oxidation is controversial because these organisms often require
266 organics, so it is not always clear if the Fe oxidation is enzymatic, or indirect via nitrite produced
267 by heterotrophic denitrification (reviewed by Kappler et al. (2015)). We can now hypothesize
268 that *cyc2* encodes an Fe oxidase in *D. aromatica*, as well as some other denitrifiers not yet tested
269 for Fe oxidation. If true, this would provide evidence in favor of enzymatic Fe oxidation in

270 heterotrophs, and give a mechanism for studying this metabolism in isolates and the
271 environment.
272
273 Another possibility is that Cyc2 functions more generally as a mechanism of extracellular
274 electron transport (EET), i.e. to transfer electrons to and from a cell via the cell surface. This is
275 not exclusive of Fe oxidation, as *M. ferrooxydans* PV-1 has been shown to oxidize a cathode
276 (Summers *et al.*, 2013) (though it is unknown whether $Cyc2_{PV-1}$ is involved). Homologs of *cyc2*
277 are also present in the genomes of organisms known to conduct EET, but not proven to oxidize
278 Fe. An example is the Gammaproteobacteria *Tenderia electrophaga*, which is the most active
279 organism in a stable cathode-oxidizing consortia (Eddie *et al.*, 2017). Curiously, *T. electrophaga*
280 also has the conserved gene cassette with *cyc1* and cytochrome c oxidases, found in FeOB
281 (**Supplemental Fig. 1**), but this organism has not yet been isolated or shown to oxidize Fe.
282 Homologs of *cyc2* are also found in *Geobacter spp.*, specifically *G. bemidjiensis*, *G.*
283 *uraniireducens*, and *Geobacter* M18, M21, and Rf64, members of a clade that predominates in
284 aquifers (Holmes *et al.*, 2007; Merkley *et al.*, 2015). These organisms reduce Fe(III), though it is
285 thought that this happens via the Omc system. A study on the *G. bemidjiensis* proteome showed
286 that the Cyc2 homolog Gbem_3353 was expressed at same levels during Fe(III) and fumarate
287 reduction (Merkley *et al.*, 2015), so function remained unclear, but could involve a more general
288 EET role.
289
290 **Conclusions**
291 This study has provided multiple lines of evidence that Cyc2 is an Fe oxidase in diverse FeOB,
292 including the first functional evidence that neutrophilic FeOB Cyc2 oxidizes Fe. Verifying the

293 function of neutrophilic FeOB Cyc2 was important because this cluster of Cyc2 sequences is
294 distinct and distant from the previously characterized Cyc2 and Cyt_{572} from acidophiles. It is still
295 not clear if *cyc2* is differentially expressed. If so, we would have a genetic marker of microbial
296 Fe oxidation activity, which is otherwise difficult to distinguish from abiotic Fe oxidation at
297 circumneutral pH. If not, the *cyc2* gene is still a valuable way of recognizing the Fe oxidation
298 potential in genomes and transcriptomes.
299
300 It is striking that there are so many Cyc2 homologs, including many from organisms not known
301 to oxidize Fe. Phylogenetic and genomic analyses show that *cyc2* has been horizontally
302 transferred between known FeOB and with other organisms. The addition of Cyc2 alone
303 conferred some Fe oxidation ability on *E. coli*, with Cyc2 likely interfacing with the existing
304 electron transport system. This suggests that acquisition of *cyc2* alone can allow an organism to
305 oxidize Fe. If this is generally true, the abundance of Cyc2 homologs suggests that microbial Fe
306 oxidation is more widespread than we currently recognize.
307
308 **Experimental Procedures**
309 **Cyc2 phylogeny and amino acid identity calculations**
310 Sequences with homology to Cyc2 were collected using blastp (Camacho *et al.*, 2009) against
311 the NCBI and IMG databases (maximum e-value $1*10^{-5}$). Query sequences were chosen to
312 represent all major groupings of the Cyc2 tree, including *Mariprofundus ferrooxydans* PV-1,
313 *Gallionella capsiferriformans* ES-2, *Acidithiobacillus ferrooxidans* ATCC 23270, *Tenderia*
314 *electrophaga*, and *Geobacter* sp. FRC-32. To remove identical sequences, the resulting 2,413
315 sequences were clustered at 100% identity using CD-HIT, reducing the database to 977 cluster

316 representatives (Li and Godzik, 2006). These sequences were imported into Geneious v.7.1.7,
317 where they were aligned using MUSCLE (Edgar, 2004). The resulting alignment was used to
318 filter out sequences of partial length, resulting in 897 sequences. Alignment columns with greater
319 than 30% gaps were removed, and a maximum likelihood phylogenetic tree was built using
320 RAxML (392 alignment columns, 100 bootstraps, CAT model of rate heterogeneity, JTT amino
321 acid substitution model (Stamatakis, 2014)). From the resulting tree, sequences that were closely
322 related and highly sampled (primarily from the *Burkholderia* and *Xanthomonas* genera) were
323 removed, resulting in 530 full-length sequences that were re-aligned, and a final phylogenetic
324 tree was built in RAxML (357 alignment columns, 300 bootstraps, CAT and JTT models). The
325 resulting phylogenetic tree was colored and names customized using the Iroki program (Moore *et*
326 *al.*, 2017).
327
328 To calculate amino acid identities of the Cyc2 homologs, pairwise alignments between the 530
329 full-length Cyc2 sequences were constructed using Muscle (Edgar, 2004). Amino acid identity
330 (AAI) values were calculated from these pairwise alignments based on the full-length sequence,
331 as well as the cytochrome and porin domains separately. The cytochrome domain was defined as
332 the conserved region starting 14 residues upstream from the heme binding site, and ending 21
333 residues downstream. The rest of the sequence downstream of this was defined as the β-barrel
334 porin domain. The AAI values were imported into R and a histogram was plotted using the
335 ggplot2 package (Wickham, 2009).
336
337 **Structural modeling**
338 Signal peptides were predicted using SignalP(Petersen *et al.*, 2011) and secondary elements were

339 predicted using PSIPRED (McGuffin *et al.*, 2000). For identification of structural homologs to
340 Cyc2, we uploaded sequences to the HHPRED tool, available as part of the Max Planck Institute
341 Bioinformatics Toolkit (Söding *et al.*, 2005). Information for the structural homologs are
342 compiled in Supplementary Table 1. Structural modeling was carried out by
343 HHpred/MODELLER (Söding *et al.*, 2005; Webb and Sali, 2016), iTasser (Zhang, 2008), and
344 Phyre (Kelley *et al.*, 2015). The structural models from all three platforms were found to be in
345 close agreement.
346
347 **Cloning and heterologous expression of Cyc2**
348 The sequence of *cyc2* was optimized for expression in *Escherichia coli*, and the signal sequence
349 replaced with the signal sequence of a native *E. coli* gene, *ompA* (sequences in Supplementalry
350 Fig. 6). The gene was synthesized by Genscript (Piscataway, NJ, USA). The *cyc2* gene was
351 cloned into the EcoRI/ HindIII sites of the pMal-p4X plasmid (with the *malE* gene removed).
352 This *cyc2*-containing plasmid was co-transformed into *E. coli* C43(DE3) with pEC86, a plasmid
353 containing the cytochrome c maturation (*ccm*) genes under a constitutive promoter, to ensure
354 heme insertion into Cyc2 (Arslan *et al.*, 1998). We also co-transformed pEC86 and the pMalp4X
355 plasmid without the *cyc2* gene (and without *malE*) as an empty-vector control; cells with the
356 empty vector control received the same treatment throughout heterologous expression and Fe-
357 oxidation assays. For expression of *cyc2*, *E. coli* was grown aerobically at 37°C (with shaking at
358 200 RPM) in Lysogeny Broth (LB), buffered with 10 mM 2-(N-morpholino)ethanesulfonic acid
359 (MES), pH 6; in addition, the medium also contained ampicillin (100 µg/mL) for propagation of
360 the pMal-p4X plasmid, and chloramphenicol (30 µg/mL) for the pEC86 plasmid. After reaching
361 mid-log phase (2.5 h), cultures were amended with 0.5 mM isopropyl β-D-1-

362 thiogalactopyranoside (IPTG) for de-repression of the lac operon to induce *cyc2* expression;
363 induction proceeded for 20 hours at 18°C with shaking at ~200 RPM. Expression of the $Cyc2_{porin}$
364 was carried out in an identical manner, except for the absence of the pEC86 plasmid from cells
365 expressing this construct, and therefore the absence of chloramphenicol from the media.
366
367 *SDS-PAGE, western and heme staining.* Cells were harvested and from cultures before IPTG
368 induction and following 20 hours after induction. Biomass was adjusted to an optical density of
369 0.6 for the porin-only control and 1.2 for the empty-vector and Cyc2-expressing cells. Cells were
370 lysed by re-suspension in 100 µL of 5X sodium dodecyl sulfate (SDS) running buffer (125 mM
371 Tris, 1.25 M glycine, 0.5% SDS, pH 8.3) and passing the resuspension through a 27.5-gauge
372 needle 10 times. Cells were then combined with gel loading buffer (50 mM Tris-HCl, 12.5 mM
373 EDTA, 2% SDS, 10% glycerol, 0.02% bromophenol blue, pH 6.8), and centrifuged for 10 min at
374 15 000 x g. Fifteen µL of the sample was then loaded onto a 16% Tris-glycine SDS-PAGE gel,
375 and ran at 100 V for 30 min, then 160 V for 40 min. The gel was then either Coomasie-stained (1
376 g Coomasie Brilliant Blue Stain in 10% acetic acid, 40% ethanol), or transferred to a PVDF
377 membrane for heme stain and western blot. For heme and StrepII-tag detection, the SDS-PAGE
378 gel was transferred to a PVDF membrane at 30 V for 16 h (4°C) in transfer buffer (25 mM Tris,
379 192 mM glycine, 20% methanol, pH 7.2). Heme peroxidase activity was assessed by washing the
380 membrane with TBST buffer (20 mM Tris, 137 mM NaCl, 0.1% Tween-20, pH 7.6), and
381 incubating for 30 minutes with Pierce ECL luminol and substrate (Carlson *et al.*, 2013), before
382 imaging on the Typhoon FLA 9500 (GE Healthcare Life Sciences). For StrepII-tag detection, the
383 PVDF membrane was blocked for one hour with 5% bovine serum albumin (BSA) and 50
384 µg/mL avidin (from egg white) in TBST buffer, rinsed with TBST, and incubated for one hour

385 with Precision Protein Streptactin-HRP conjugate (1:60,000 dilution, Biorad). The membrane
386 was then washed with four 15-minute washes in TBST, then incubated with Pierce ECL luminol
387 and substrate for 5 minutes before imaging on the Typhoon FLA 9500.
388
389 **Fe oxidation assay**
390 After induction of Cyc2 in *E. coli*, triplicate cultures were centrifuged at 3200 x g, and washed
391 with sterile LB (supplemented with 10 mM MES, pH 6) before resuspension and incubation in
392 50 mL glass beakers in the same medium at an OD of 2 (final volume = 25 mL); these cell
393 suspensions were stirred to ensure homogeneity and aeration. In the case of azide experiments,
394 75 µL of a fresh 1M sodium azide stock solution was added to 25 mL cell suspension (to a final
395 concentration of 3 mM), and incubated for 5 min. In all experiments, $FeCl_2$ was added from an
396 anoxic, filter-sterilized 100 mM stock solution to a target concentration of 100 µM. Ferrozine
397 measurements were taken over a period of one hour (ferrozine method adapted from (Stookey,
398 1970)). At each time point, a sample was taken from each of the triplicate cultures, and
399 centrifuged at 15,000 x g for 20-30 s to remove cells. A 150 µL portion of the supernatant was
400 then combined with 40 µL of 1.225 mM ferrozine, 50 µL of 6.87 M acetate buffer, and 10 µL of
401 H_2O inside 200 µL 96-well plates, and incubated for 15 min in the dark before absorption was
402 measured at 562 nm using a plate reader (Perkin Elmer 1420 Multilabel Counter Victor^{3}V). pH
403 of the cell suspensions was monitored before and after each assay, and was found to be stable at
404 pH 6 (within one-tenth of a pH unit).
405
406 **Loihi Fe mat sampling and metatranscriptome analysis**
407 *Sampling*. Sample J2-674-BM1-C123456 was collected from the Pohaku vents (Mkr 57) at Loihi

408 Seamount, Hawaii using the Jason II remotely operated vehicle in 2013. Using the mat sampler
409 designed by Breier et al. (2012), 60 mL of the top 1 cm of Fe mat was collected into each of six
410 syringes, each pre-loaded with 50 mL of RNALater (Ambion, United States) for immediate
411 DNA/RNA preservation. Upon recovery, samples were allowed to settle (~17 mL total mat
412 material), overlying fluids were decanted, and samples were stored at -80˚C until extraction.
413
414 *Metagenome and metatranscriptome sequencing*. Genomic DNA was extracted using the
415 FastDNA SPIN kit for soil (MP Bio, Solon, OH, USA) following the manufacturer's protocol,
416 with the addition of 250 µL 0.5 M sodium citrate pH 5.8 prior to cell lysis. RNA was extracted
417 using the NucleoSpin RNA kit (Macherey-Nagel, Bethlehem, PA, USA) following the
418 manufacturer's protocol. Metagenomic sequencing was performed on sample J2-674-BM1-C3.
419 An Illumina library with paired reads of 251 bp length with 450 bp insert size was prepared
420 following standard protocols and sequenced on a HiSeq 2500 (Illumina Inc., San Diego, CA).
421 Metatranscriptomic sequencing was performed on sample J2-674-BM1-C6. Ribosomal RNA-
422 depleted total RNA (Ribo-Zero, Epicenter, Madison, WI) was prepared for sequencing using
423 Illumina's NEXTflex Rapid RNA-Seq Library Prep Kit (Bioo Scientific, Austin, TX), and was
424 sequenced on a HiSeq 2500 producing single-end 51 bp reads.
425
426 *Metagenome assembly and binning*. Metagenomic reads were quality controlled (QC'ed) using
427 trimmomatic and merged using FLASH (see pipeline: https://github.com/mooreryan/qc). QC'ed
428 reads were then assembled using metaSPAdes (Nurk *et al.*, 2017), with a k-mer sweep from 21
429 to 127. In total 82,765 contigs were produced, from 128 bp to 79,748 bp (average 1,203 bp).
430 Contigs were then QC'ed so that only contigs with at least 1X coverage over 90% of their length

431 were used. Contigs over 2,000 bp in length (10,352 total) were binned using Binsanity (Graham
432 *et al.*, 2017), CONCOCT (Alneberg *et al.*, 2014), MaxBin (Wu *et al.*, 2016), and MetaBAT
433 (Kang *et al.*, 2015). The best resulting bins were chosen using DAS Tool (Sieber *et al.*, 2017).
434 Bin completeness and redundancy were calculated using CheckM (Parks *et al.*, 2015). The
435 relative abundance of each bin was calculated using a length-normalized average of contig read
436 coverage.
437
438 *Manual bin curation.* Loihi metagenome sample 674-BM1-C3 was dominated by a single
439 Zetaproteobacteria OTU2 bin (S1_binsanity019), accounting for 79.4% of the total binned
440 average read coverage (50.3% including unbinned contigs). This bin was estimated by CheckM
441 to be 92.7% complete (3.03% redundancy, 0% strain heterogeneity), yet lacked any protein
442 BLAST hits to *cyc2*. Because previous research has identified *cyc2* within at least 9 genomic
443 representatives from ZOTU2 (3 SAGs, 6 MAGs (Field *et al.*, 2015; Fullerton *et al.*, 2017)), we
444 used two methods to find the cyc2 belonging to this bin: 1) subsampled reads for a simplified
445 assembly and 2) used information from the assembly graph to locate cyc2 connected with this
446 dominant bin. Quadruplicate, randomly-sampled read subsets at 10%, 2%, and 1% of the total
447 number of quality-controlled reads were independently assembled and binned to simplify the
448 assembly of the dominant ZOTU2 bin (starting at 1,352X coverage). Subsampling in this way
449 allowed us to increase the quality of the dominant ZOTU2 bin, with a maximum completeness of
450 98.1% (2.27% redundancy, 0% strain heterogeneity), though unfortunately cyc2 was again not
451 recovered from any of the binned contigs. However, the subsampled assembly was helpful in
452 recovering *cyc2* through use of the simplified assembly graph. Using blastp, 12 contigs were
453 identified that had homology to Cyc2. All twelve of these contigs were found within one section

454 of the assembly graph, connected through a minimal k-mer overlap of 21 bp to the assembly
455 network containing the binned contigs of interest. No other bins were contained within this
456 assembly graph network. Taking the section of the assembly graph containing these *cyc2* contigs,
457 overlap consensus assembly produced five unique contigs. Four of these contigs had sufficient
458 length to confirm their clustering within the dominant ZOTU2 bin using VizBin (Laczny *et al.*,
459 2015). These four contigs were combined with the dominant subsampled bin and used in further
460 analysis.
461
462 *Metatranscriptome recruitment and analysis.* Metatranscriptome reads were recruited to
463 metagenomic contigs using bowtie2 (Langmead and Salzberg, 2012). Expression estimates were
464 calculated by dividing the total read count by the gene length and total sequencing effort (reads
465 per thousand bp per million reads; RPKM). The relative abundance of the expression in each bin
466 was calculated as the percent of total reads mapping to that bin, normalized to bin length. RPKM
467 expression estimates from all expressed genes were then imported into R and plotted using the
468 ggplot2 package (Wickham, 2009).
469
470 *Nucleotide submission*. Raw sequence data were submitted to the sequence read archive (SRA) at
471 the National Center for Biotechnology Information (NCBI), with all appropriate metadata under
472 project accession number PRJNA412510.
473
474 **Reanalysis of Rifle aquifer metatranscriptome**
475 To quantify *cyc2* and other gene expression in an FeOB-dominated terrestrial ecosystem, we
476 extracted information from the supplementary dataset of Jewell et al.(2016), a

477 metatranscriptomic/metagenomic study of a nitrate/O_2-amended ferruginous Rifle aquifer. This
478 dataset consists of over 200,000 translated open reading frames (ORFs), grouped by bin and
479 associated with gene expression data in units of RPKM (reads per thousand base pairs per
480 million reads). From this dataset, amino acid sequences corresponding to ORFs in
481 Gallionellaceae bins 22.1-22.9 were used as a protein sequence database, which we used for
482 blastp to search for Cyc2, and Mto, Sox, Dsr protein sequences (**Supplementary table 2**).
483 RPKM expression values were then imported into R, log transformed, and plotted using the
484 ggplot2 package.
485

486 ***cyc2* and *mto/mtr* gene distribution in genomes**

487 To look for *cyc2* and *mtoAB/mtrABC* gene distribution in genomes, we used blastp against the
488 non-redundant NCBI database to search for homologs of these genes from a representative set of
489 organisms: *Sideroxydans lithotrophicus* ES-1, *Shewanella oneidensis* MR-1, *Rhodoferax*
490 *ferrireducens* T118, *Magnetospirillum magneticum* AMB-1, and *Rhodopseudomonas palustris*
491 TIE-1. Blast results were analyzed using a custom Python script to identify which *cyc2* and
492 *mto*/*mtr* homologs co-occurred within the same genomes. Results were imported in R and plotted
493 using the VennDiagram package.
494

495 **Conserved gene cassette identification**

496 We used a custom Python script to identify organisms that encode the conserved gene cassette
497 described by Field et al. (2015). This cassette includes 2 subunits of the cbb3-type cytochrome c
498 oxidase, *cyc1*, spermidine synthase, ferredoxin, and several other c-type cytochromes; we blasted
499 these cassette genes against the non-redundant NCBI database (Release 84), and clustered close

homologs according to gene ID, which allowed us to detect genomes that included genes of interest in close proximity, defined as within 20 genes of each other.

Acknowledgments

We thank Shawn Polson for assistance in bioinformatic analyses and the crews of R/V Thompson and ROV Jason for support during sampling at the Loihi Seamount. This research was funded by the National Science Foundation (EAR-1151682, OCE-1155290) and the Office of Naval Research (N00014-17-1-2640). The authors declare no conflict of interest.

References

Abergel, C., Nitschke, W., Malarte, G., Bruschi, M., Claverie, J.-M., and Giudici-Orticoni, M.-T. (2003) The Structure of *Acidithiobacillus ferrooxidans c4*-cytochrome. *Structure* **11**: 547–555.

Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., et al. (2014) Binning metagenomic contigs by coverage and composition. *Nat Methods* **11**: 1144–1146.

Arslan, E., Schulz, H., Zufferey, R., Künzler, P., and Thöny-Meyer, L. (1998) Overproduction of the *Bradyrhizobium japonicum c*-type cytochrome subunits of the cbb3 oxidase in *Escherichia coli*. *Biochemical and Biophysical Research Communications* **251**: 744–747.

Barco, R.A., Emerson, D., Sylvan, J.B., Orcutt, B.N., Meyers, M.E.J., Ramírez, G.A., et al. (2015) The proteomic profile of an obligate iron-oxidizing chemolithoautotroph reveals new insight into microbial iron oxidation. **81**: 5927–5937.

Barco, R.A., Hoffman, C.L., Ramírez, G.A., Toner, B.M., Edwards, K.J., and Sylvan, J.B. (2017) *In-situ* incubation of iron-sulfur mineral reveals a diverse chemolithoautotrophic community and a new biogeochemical role for *Thiomicrospira*. *Environ. Microbiol.* **19**: 1322–1337.

Beam, J., Scott, J.J., McAllister, S.M., Chan, C.S., McManus, J., Meysman, F., and Emerson, D. Potential for biological rejuvenation of iron oxides in bioturbated marine sediments. *in press*.

Beckwith, C.R., Edwards, M.J., Lawes, M., Shi, L., Butt, J.N., Richardson, D.J., and Clarke, T.A. (2015) Characterization of MtoD from *Sideroxydans lithotrophicus*: a cytochrome c

529 electron shuttle used in lithoautotrophic growth. *Front Microbiol* **6**:.

530 Bird, L.J., Bonnefoy, V., and Newman, D.K. (2011) Bioenergetic challenges of microbial iron
531 metabolisms. *Trends in Microbiology* **19**: 330–340.

532 Borch, T., Kretzschmar, R., Kappler, A., Cappellen, P.V., Ginder-Vogel, M., Voegelin, A., and
533 Campbell, K. (2010) Biogeochemical redox processes and their impact on contaminant
534 dynamics. *Environ Sci Technol* **44**: 15–23.

535 Breier, J.A., Gomez-Ibanez, D., Reddington, E., Huber, J.A., and Emerson, D. (2012) A
536 precision multi-sampler for deep-sea hydrothermal microbial mat studies. *Deep Sea Res Pt I*
537 **70**: 83–90.

538 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden,
539 T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics 2016 17:1* **10**: 421.

540 Cardenas, J.P., Lazcano, M., Ossandon, F.J., Corbett, M., Holmes, D.S., and Watkin, E. (2014)
541 Draft genome sequence of the iron-oxidizing acidophile *Leptospirillum ferriphilum* type
542 strain DSM 14647. *Genome Announc.* **2**: e01153–14–e01153–14.

543 Carlson, H.K., Clark, I.C., Blazewicz, S.J., Iavarone, A.T., and Coates, J.D. (2013) Fe(II)
544 oxidation is an innate capability of nitrate-reducing bacteria that involves abiotic and biotic
545 reactions. *Journal of Bacteriology* **195**: 3260–3268.

546 Castelle, C., Guiral, M., Malarte, G., Ledgham, F., Leroy, G., Brugna, M., and Giudici-Orticoni,
547 M.T. (2008) A new iron-oxidizing/O_2-reducing supercomplex spanning both inner and outer
548 membranes, isolated from the extreme acidophile *Acidithiobacillus ferrooxidans*. *Journal of*
549 *Biological Chemistry* **283**: 25803–25811.

550 Chan, C.S., Fakra, S.C., Emerson, D., Fleming, E.J., and Edwards, K.J. (2011) Lithotrophic iron-
551 oxidizing bacteria produce organic stalks to control mineral growth: implications for
552 biosignature formation. *ISME J* **5**: 717–727.

553 Chiu, B.K., Kato, S., McAllister, S.M., Field, E.K., and Chan, C.S. (2017) Novel pelagic iron-
554 oxidizing Zetaproteobacteria from the Chesapeake Bay oxic–anoxic transition zone. *Front*
555 *Microbiol* **8**: 4206.

556 Coates, J.D., Chakraborty, R., Lack, J.G., O'Connor, S.M., Cole, K.A., Bender, K.S., and
557 Achenbach, L.A. (2001) Anaerobic benzene oxidation coupled to nitrate reduction in pure
558 culture by two strains of *Dechloromonas*. *Nature* **411**: 1039–1043.

559 Comolli, L.R., Luef, B., and Chan, C.S. (2011) High-resolution 2D and 3D cryo-TEM reveals
560 structural adaptations of two stalk-forming bacteria to an Fe-oxidizing lifestyle. *Environ*
561 *Microbiol* **13**: 2915–2929.

562 Crowe, S.A., Hahn, A.S., Morgan-Lang, C., Thompson, K.J., Simister, R.L., Llirós, M., et al.
563 (2017) Draft genome sequence of the pelagic photoferrotroph *Chlorobium*
564 *phaeoferrooxidans*. *Genome Announc.* **5**: e01584–16.

565 Eddie, B.J., Wang, Z., Hervey, W.J., IV, Leary, D.H., Malanoski, A.P., Tender, L.M., et al.
566 (2017) Metatranscriptomics supports the mechanism for biocathode electroautotrophy by
567 "Candidatus *Tenderia electrophaga*." *mSystems* **2**: e00002–17.

568 Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high
569 throughput. *Nucleic Acids Research* **32**: 1792–1797.

570 Emerson, D. and de Vet, W. (2015) The role of FeOB in engineered water ecosystems: a review.
571 *J Am Water Works Ass* **107**: E47–E57.

572 Emerson, D., Field, E.K., Chertkov, O., Davenport, K.W., Goodwin, L., Munk, C., et al. (2013)
573 Comparative genomics of freshwater Fe-oxidizing bacteria: implications for physiology,
574 ecology, and systematics. *Front Microbiol* **4**: 254.

575 Emerson, D., Fleming, E.J., and McBeth, J.M. (2010) Iron-oxidizing bacteria: an environmental
576 and genomic perspective. *Annu Rev Microbiol* **64**: 561–583.

577 Emerson, D. and Moyer, C.L. (2002) Neutrophilic Fe-oxidizing bacteria are abundant at the
578 Loihi Seamount hydrothermal vents and play a major role in Fe oxide deposition. *Appl.*
579 *Environ. Microbiol.* **68**: 3085–3093.

580 Emerson, D., Rentz, J.A., Lilburn, T.G., Davis, R.E., Aldrich, H., Chan, C., and Moyer, C.L.
581 (2007) A novel lineage of Proteobacteria involved in formation of marine Fe-oxidizing
582 microbial mat communities. *PLoS ONE* **2**: e667.

583 Fabisch, M., Beulig, F., Akob, D.M., and Kusel, K. (2011) New *Thiomonas* and *Bordetella*
584 strains involved in iron oxidation at a slightly acidic, heavy metal contaminated creek.
585 *Mineralogical Magazine* **75**: 825.

586 Field, E.K., Sczyrba, A., Lyman, A.E., Harris, C.C., Woyke, T., Stepanauskas, R., and Emerson,
587 D. (2015) Genomic insights into the uncultivated marine Zetaproteobacteria at Loihi
588 Seamount. *The ISME Journal* **9**: 857–870.

589 Fukushima, J., Tojo, F., Asano, R., Kobayashi, Y., Shimura, Y., Okano, K., and Miyata, N.
590 (2015) Complete genome sequence of the unclassified iron-oxidizing, chemolithoautotrophic
591 Burkholderiales bacterium GJ-E10, isolated from an acidic river. *Genome Announc.* **3**:
592 e01455–14.

593 Fullerton, H., Hager, K.W., McAllister, S.M., and Moyer, C.L. (2017) Hidden diversity revealed
594 by genome-resolved metagenomics of iron-oxidizing microbial mats from Loihi Seamount,
595 Hawaii. *The ISME Journal* **11**: 1900–1914.

596 Glazer, B.T. and Rouxel, O.J. (2009) Redox speciation and distribution within diverse iron-
597 dominated microbial habitats at Loihi Seamount. *Geomicrobiol J* **26**: 606–622.

598 Graham, E.D., Heidelberg, J.F., and Tully, B.J. (2017) BinSanity: unsupervised clustering of
599 environmental microbial assemblies using coverage and affinity propagation. *PeerJ* **5**:
600 e3035.

601 Hallberg, K.B., González-Toril, E., and Johnson, D.B. (2009) Acidithiobacillus ferrivorans, sp.
602 nov.; facultatively anaerobic, psychrotolerant iron-, and sulfur-oxidizing acidophiles isolated
603 from metal mine-impacted environments. *Extremophiles* **14**: 9–19.

604 Hancock, R.E. (1987) Role of porins in outer membrane permeability. *Journal of Bacteriology*
605 **169**: 929–933.

606 He, S., Barco, R.A., Emerson, D., and Roden, E.E. (2017) Comparative genomic analysis of
607 neutrophilic iron(II) oxidizer genomes for candidate genes in extracellular electron transfer.
608 *Front Microbiol* **8**: W400.

609 He, S., Tominski, C., Kappler, A., Behrens, S., and Roden, E.E. (2016) metagenomic analyses of
610 the autotrophic Fe(II)-oxidizing, nitrate-reducing enrichment culture KS. **82**: 2656–2668.

611 Hedrich, S., Schlömann, M., and Johnson, D.B. (2011) The iron-oxidizing proteobacteria.
612 *Microbiology* **157**: 1551–1564.

613 Heising, S., Richter, L., LUDWIG, W., and SCHINK, B. (1999) *Chlorobium ferrooxidans* sp.
614 nov., a phototrophic green sulfur bacterium that oxidizes ferrous iron in coculture with a
615 "*Geospirillum*" sp. strain. *Arch Microbiol* **172**: 116–124.

616 Holmes, D.E., O'Neil, R.A., Vrionis, H.A., N'Guessan, L.A., Ortiz-Bernad, I., Larrahondo, M.J.,
617 et al. (2007) Subsurface clade of Geobacteraceae that predominates in a diversity of Fe(III)-
618 reducing subsurface environments. *The ISME Journal* **1**: 663–677.

619 Ilbert, M. and Bonnefoy, V. (2013) Insight into the evolution of the iron oxidation pathways.
620 *Biochim. Biophys. Acta* **1827**: 161–175.

621 Jeans, C., Singer, S.W., Chan, C.S., VerBerkmoes, N.C., Shah, M., Hettich, R.L., et al. (2008)
622 Cytochrome 572 is a conspicuous membrane protein with iron oxidation activity purified
623 directly from a natural acidophilic microbial community. *The ISME Journal* **2**: 542–550.

624 Jewell, T.N.M., Karaoz, U., Brodie, E.L., Williams, K.H., and Beller, H.R. (2016)
625 Metatranscriptomic evidence of pervasive and diverse chemolithoautotrophy relevant to C,
626 S, N and Fe cycling in a shallow alluvial aquifer. *ISME J* **10**: 2106–2117.

627 Jiao, Y. and Newman, D.K. (2007) The pio operon is essential for phototrophic Fe(II) oxidation
628 in *Rhodopseudomonas palustris* TIE-1. *Journal of Bacteriology* **189**: 1765–1773.

629 Johnson, D.B., Hallberg, K.B., and Hedrich, S. (2014) Uncovering a microbial enigma: isolation
630 and characterization of the streamer-generating, iron-oxidizing, acidophilic bacterium
631 "Ferrovum myxofaciens". **80**: 672–680.

632 Kang, D.D., Froula, J., Egan, R., and Wang, Z. (2015) MetaBAT, an efficient tool for accurately
633 reconstructing single genomes from complex microbial communities. *PeerJ* **3**: e1165.

634 Kappler, A., Emerson, D., Gralnick, J.A., Roden, E.E., and Muehe, E.M. (2015)
635 Geomicrobiology of Iron. In, Ehrlich,H.L. and Newman,D.K. (eds), *Ehrlich's*

636 *Geomicrobiology*. Boca Raton, US, pp. 343–399.

637 Kato, S., Ohkuma, M., Powell, D.H., Krepski, S.T., Oshima, K., Hattori, M., et al. (2015)
638 Comparative genomic insights into ecophysiology of neutrophilic, microaerophilic iron
639 oxidizing bacteria. *Front Microbiol* **6**: 1265.

640 Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015) The Phyre2
641 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845–858.

642 Kupka, D., Liljeqvist, M., Nurmi, P., Puhakka, J.A., Tuovinen, O.H., and Dopson, M. (2009)
643 Oxidation of elemental sulfur, tetrathionate and ferrous iron by the psychrotolerant
644 Acidithiobacillus strain SS3. *Research in Microbiology* **160**: 767–774.

645 Laczny, C.C., Sternal, T., Plugaru, V., Gawron, P., Atashpendar, A., Margossian, H.H., et al.
646 (2015) VizBin - an application for reference-independent visualization and human-
647 augmented binning of metagenomic data. *Microbiome 2017 5:1* **3**: 1.

648 Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat*
649 *Methods* **9**: 357–359.

650 Laufer, K., Nordhoff, M., Halama, M., Martinez, R.E., Obst, M., Nowak, M., et al. (2017)
651 Microaerophilic Fe(II)-oxidizing Zetaproteobacteria isolated from low-Fe marine coastal
652 sediments: Physiology and characterization of their twisted stalks. *Appl. Environ.*
653 *Microbiol.* **83**: e03118-16.

654 Li, W. and Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of
655 protein or nucleotide sequences. *Bioinformatics* **22**: 1658–1659.

656 Liu, J., Wang, Z., Belchik, S.M., Edwards, M.J., Liu, C., Kennedy, D.W., et al. (2012)
657 Identification and characterization of MtoA: A decaheme c-type cytochrome of the
658 neutrophilic Fe(II)-oxidizing bacterium *Sideroxydans lithotrophicus* ES-1. *Front Microbiol*
659 **3**: 37.

660 Majzlan, J. (2013) Minerals and aqueous species of iron and manganese as reactants and
661 products of microbial metal respiration. In, Gescher,J. and Kappler,A. (eds), *Microbial Metal*
662 *Respiration*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–28.

663 Makita, H., Tanaka, E., Mitsunobu, S., Miyazaki, M., Nunoura, T., Uematsu, K., et al. (2016)
664 Mariprofundus micogutta sp. nov., a novel iron-oxidizing zetaproteobacterium isolated from
665 a deep-sea hydrothermal field at the Bayonnaise knoll of the Izu-Ogasawara arc, and a
666 description of Mariprofundales ord. nov. and Zetaproteobacteria classis nov. *Arch Microbiol*
667 **199**: 335–346.

668 McAllister, S.M., Davis, R.E., McBeth, J.M., Tebo, B.M., Emerson, D., and Moyer, C.L. (2011)
669 Biodiversity and emerging biogeography of the neutrophilic iron-oxidizing
670 Zetaproteobacteria. *Appl. Environ. Microbiol.* **77**: 5445–5457.

671 McBeth, J.M., Little, B.J., Ray, R.I., Farrar, K.M., and Emerson, D. (2011) Neutrophilic iron-

672 oxidizing "Zetaproteobacteria" and mild steel corrosion in nearshore marine
673 environments. *Appl. Environ. Microbiol.* **77**: 1405–1412.

674 McGuffin, L.J., Bryson, K., and Jones, D.T. (2000) The PSIPRED protein structure prediction
675 server. *Bioinformatics* **16**: 404–405.

676 Merkley, E.D., Wrighton, K.C., Castelle, C.J., Anderson, B.J., Wilkins, M.J., Shah, V., et al.
677 (2015) Changes in protein expression across laboratory and field experiments in *Geobacter*
678 *bemidjiensis*. *J. Proteome Res.* **14**: 1361–1375.

679 Modi, N., Ganguly, S., Bárcena-Uribarri, I., Benz, R., van den Berg, B., and Kleinekathöfer, U.
680 (2015) Structure, dynamics, and substrate specificity of the OprO Porin from *Pseudomonas*
681 *aeruginosa*. *Biophys J* **109**: 1429–1438.

682 Moore, R.M., Harrison, A.O., McAllister, S.M., Marine, R.L., Chan, C.S., and Wommack, K.E.
683 (2017) Iroki: automatic customization for phylogenetic trees. *bioRxiv* 106138.

684 Moraes, T.F., Bains, M., Hancock, R.E.W., and Strynadka, N.C.J. (2006) An arginine ladder in
685 OprP mediates phosphate-specific transfer across the outer membrane. *Nat Struct Mol Biol*
686 **14**: 85–87.

687 Mori, J.F., Scott, J.J., Hager, K.W., Moyer, C.L., Küsel, K., and Emerson, D. (2017)
688 Physiological and ecological implications of an iron- or hydrogen-oxidizing member of the
689 Zetaproteobacteria, Ghiorsea bivora, gen. nov., sp. nov. *ISME J* **11**: 2624–2636.

690 Moya-Beltrán, A., Cárdenas, J.P., Covarrubias, P.C., Issotta, F., Ossandon, F.J., Grail, B.M., et
691 al. (2014) Draft genome sequence of the nominated type strain of "*Ferrovum myxofaciens,*"
692 an acidophilic, iron-oxidizing Betaproteobacterium. *Genome Announc.* **2**: e00834–14.

693 Mumford, A.C., Adaktylou, I.J., and Emerson, D. (2016) Peeking under the Iron Curtain:
694 Development of a microcosm for imaging colonization of steel surfaces by *Mariprofundus*
695 sp. DIS-1, an oxygen tolerant Fe-oxidizing bacterium. AEM.01990–16.

696 Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited.
697 *Microbiol Mol Biol R* **67**: 593–656.

698 Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017) metaSPAdes: a new
699 versatile metagenomic assembler. *Genome Res.* **27**: 824–834.

700 Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM:
701 assessing the quality of microbial genomes recovered from isolates, single cells, and
702 metagenomes. *Genome Res.* **25**: 1043–1055.

703 Petersen, T.N., Brunak, S., Heijne, von, G., and Nielsen, H. (2011) SignalP 4.0: discriminating
704 signal peptides from transmembrane regions. *Nat Methods* **8**: 785–786.

705 Quaiser, A., Bodi, X., Dufresne, A., Naquin, D., Francez, A.-J., Dheilly, A., et al. (2014)
706 Unraveling the stratification of an iron-oxidizing microbial mat by metatranscriptomics.

PLoS ONE **9**: e102561.

Salinero, K., Keller, K., Feil, W.S., Feil, H., Trong, S., Di Bartolo, G., and Lapidus, A. (2009) Metabolic analysis of the soil microbe Dechloromonas aromatica str. RCB: indications of a surprisingly complex life-style and cryptic anaerobic pathways for aromatic degradation. *BMC Genomics* **10**: 351.

Schulz, G.E. (2004) The structures of general porins. In *Bacterial and Eukaryotic Porins: Structure, Function, Mechanism.* Benz, R. (ed). Weinheim, Germany: Wiley-VCH, pp. 25-40.

Sieber, C.M.K., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G., and Banfield, J.F. (2017) Recovery of genomes from metagenomes via a dereplication, aggregation, and scoring strategy. *bioarXiv* **107789**: 1–24.

Singer, E., Emerson, D., Webb, E.A., Barco, R.A., Kuenen, J.G., Nelson, W.C., et al. (2011) *Mariprofundus ferrooxydans* PV-1 the first genome of a marine Fe(II) oxidizing Zetaproteobacterium. *PLoS ONE* **6**: e25386.

Söding, J., Biegert, A., and Lupas, A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* **33**: W244–W248.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Stookey, L.L. (1970) Ferrozine---a new spectrophotometric reagent for iron. *Anal Chem* **42**: 779–781.

Summers, Z.M., Gralnick, J.A., and Bond, D.R. (2013) Cultivation of an obligate Fe(II)-oxidizing lithoautotrophic bacterium using electrodes. *MBio* **4**: e00420–12.

Ullrich, S.R., González, C., Poehlein, A., Tischler, J.S., Daniel, R., Schlömann, M., et al. (2016) Gene loss and horizontal gene transfer contributed to the genome evolution of the extreme acidophile "Ferrovum." *Front Microbiol* **7**: e78237.

Ullrich, S.R., Poehlein, A., Tischler, J.S., González, C., Ossandon, F.J., Daniel, R., et al. (2016) Genome analysis of the biotechnologically relevant acidophilic iron oxidising strain JA12 indicates phylogenetic and metabolic diversity within the novel genus "*Ferrovum*." *PLoS ONE* **11**: e0146832.

Webb, B. and Sali, A. (2016) Comparative protein structure modeling using MODELLER. *Curr Protoc Bioinformatics* **54**: 5.6.1–5.6.37.

Wickham, H. (2009) ggplot2: Elegant Graphics for Data Analysis Springer, New York, NY.

Williams, K.H., Kemna, A., Wilkins, M.J., Druhan, J., Arntzen, E., N'Guessan, A.L., et al. (2009) Geophysical monitoring of coupled microbial and geochemical processes during stimulated subsurface bioremediation. *Environ Sci Technol* **43**: 6717–6723.

742 Williams, K.H., Long, P.E., Davis, J.A., Wilkins, M.J., N'Guessan, A.L., Steefel, C.I., et al.
743 (2011) Acetate availability and its influence on sustainable bioremediation of uranium-
744 contaminated groundwater. *Geomicrobiol J* **28**: 519–539.

745 Wu, Y.-W., Simmons, B.A., and Singer, S.W. (2016) MaxBin 2.0: an automated binning
746 algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**: 605–
747 607.

748 Yarzabal, A., Brasseur, G., Ratouchniak, J., Lund, K., Lemesle-Meunier, D., DeMoss, J., and
749 Bonnefoy, V. (2002) The high-molecular-weight cytochrome c Cyc2 of *Acidithiobacillus*
750 *ferrooxidans* is an outer membrane protein. *Journal of Bacteriology* **184**: 313–317.

751 Yoshikawa, S., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., Yamashita, E., Inoue, N., et al.
752 (1998) Redox-coupled crystal structural changes in bovine heart cytochrome c oxidase.
753 *Science* **280**: 1723–1729.

754 Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*
755 *2016 17:1* **9**: 40.

756

757

758 **Figure legends**

759 **Figure 1.** Cyc2 protein phylogenetic tree, with all FeOB groups labeled, including the
760 functionally characterized *At. ferrooxidans* Cyc2 and *L. ferriphilum* Cyt_{572}, which have both
761 been shown to oxidize Fe(II) *in vitro*. Note that the neutrophilic FeOB (Zetaproteobacteria,
762 Gallionellaceae, *Chlorobi*) fall into a well-supported cluster (bootstrap value 93%). Also closely
763 related is the *Chlorobi* photoferrotroph Cyc2. Bootstrap values (%) corresponding to nodes
764 highlighted by circles: Zetaproteobacteria-most sequences (92), Gallionellaceae (98/75),
765 Chlorobi (100), Acidithiobacillus ferrooxidans (100), Ferrovum (100), Zetaproteobacteria-4
766 sequences (100), Leptospirillum spp. (100). See Supplemental File 1 for full tree with all taxa
767 and nodes labeled.

768

769 **Figure 2.** Alignment of cytochrome-containing section for Cyc2 from representative neutrophilic
770 and acidophilic FeOB. Red=CXXCH heme-binding motif. Blue=other motifs discussed in the
771 text. This section is preceded by a signal sequence, as detected by SignalP, and followed by a
772 beta strand predicted by PSIPRED. See Supplemental Figure 2 for full Cyc2 alignment.
773
774 **Figure 3.** Two views of the Cyc2 porin model, generated using iTasser. The Cyc2 cytochrome
775 (orange sphere=hypothesized location) is connected to the N terminal end of the porin (blue
776 strand).
777
778 **Figure 4.** Heterologous expression of Cyc2 and Fe oxidation assay results. (A) Western blot
779 showing expression of Cyc2 (lane 4) and $Cyc2_{porin}$ (lane 6); U=uninduced, I=induced with 0.5
780 mM IPTG. See corresponding Coomassie-stained gel in Supplemental Fig. 6B. (B) Heme stain
781 showing successful insertion of heme into Cyc2. (C) Fe oxidation assay results. Dissolved Fe^{2+}
782 monitored over time after Fe^{2+} addition to a suspension of *E. coli* cells expressing Cyc2,
783 $Cyc2_{porin}$, or neither (empty vector). Fe^{2+} was added to Cyc2-expressing cells a second time.
784 Error bars represent 1 SD of triplicate experimental runs.
785

Table 1. Distribution of putative Fe oxidation genes in FeOB genomes

Organism	*mtoAB*	*cyc2*	Reference
Freshwater neutrophilic microaerophilic FeOB (Gallionellaceae)			
Sideroxydans lithotrophicus ES-1	**X**	x	Liu *et al.*, 2012; Emerson *et al.*, 2013
Gallionella capsiferriformans ES-2*	x	x	Emerson *et al.*, 2013
Gallionellaceae *sp.* KS	x	x	He *et al.*, 2016
Ferriphaselus amnicola OYT1*		x	Kato *et al.*, 2015
Ferriphaselus globulitus R-1*		x	Kato *et al.*, 2015
Marine neutrophilic microaerophilic FeOB (Zetaproteobacteria)			
Mariprofundus ferrooxydans PV-1*, JV-1*, M34*, EKF-M39*		x	Singer *et al.*, 2011; Barco *et al.*, 2015; Field *et al.*, 2015
Mariprofundus aestuarium CP-5*		x	Chiu *et al.*, 2017
Mariprofundus ferrinatatus CP-8*		x	Chiu *et al.*, 2017
Mariprofundus micogutta ET2*		x	Makita *et al.*, 2016
Mariprofundus sp. DIS-1*		x	Mumford *et al.*, 2016
Ghiorsea bivora TAG-1, SV108		x	Mori *et al.*, 2017
Acidophilic FeOB			
Acidithiobacillus ferrooxidans ATCC23270		**X**	Yarzabal *et al.*, 2002; Castelle *et al.*, 2008
Acidithiobacillus ferrivorans CF27, SS3		x	Hallberg *et al.*, 2009; Kupka *et al.*, 2009
Leptospirillum sp., Iron Mountain Mine community		**X** (*cyt_{572}*)	Jeans *et al.*, 2008
Leptospirillum ferriphilum DSM14647/P3A*		x	Cardenas *et al.*, 2014
Ferrovum myxofaciens P3G*		x	Johnson *et al.*, 2014
Ferrovum spp. JA12, PN-J185, Z-31		x	Ullrich, Poehlein, *et al.*, 2016; Ullrich, González, *et al.*, 2016
Thiomonas spp. FB-6, FB-Cd	x	x	Fabisch *et al.*, 2011
Phototrophic FeOB			
Chlorobium ferrooxidans KoFox		x	Heising *et al.*, 1999
Chlorobium phaeoferrooxidans KB01		x	Crowe *et al.*, 2017
Rhodopseudomonas palustris TIE-1	**X** (*pioAB*)		Jiao and Newman, 2007

x=present in genome

X=Fe oxidation function confirmed

*obligate FeOB, noted for pure cultures only

Table 2. *cyc2* and *mtoA* expression in Jewell et al. Rifle aquifer metatranscriptome

Gallionellaceae bin	*cyc2* max (RPKM)	*mtoA* max (RPKM)	*cyc2*:*mtoA* ratio
b22.1	136	0.714	190
b22.2	271	0.333	812
b22.3	173		
b22.5	1211		
b22.6	696	42.1	16.5
b22.7		5.09	
b22.8	176	0.166	1520
b22.9	171	0.317	540
all bins, median (n=11)	172	0.73	236

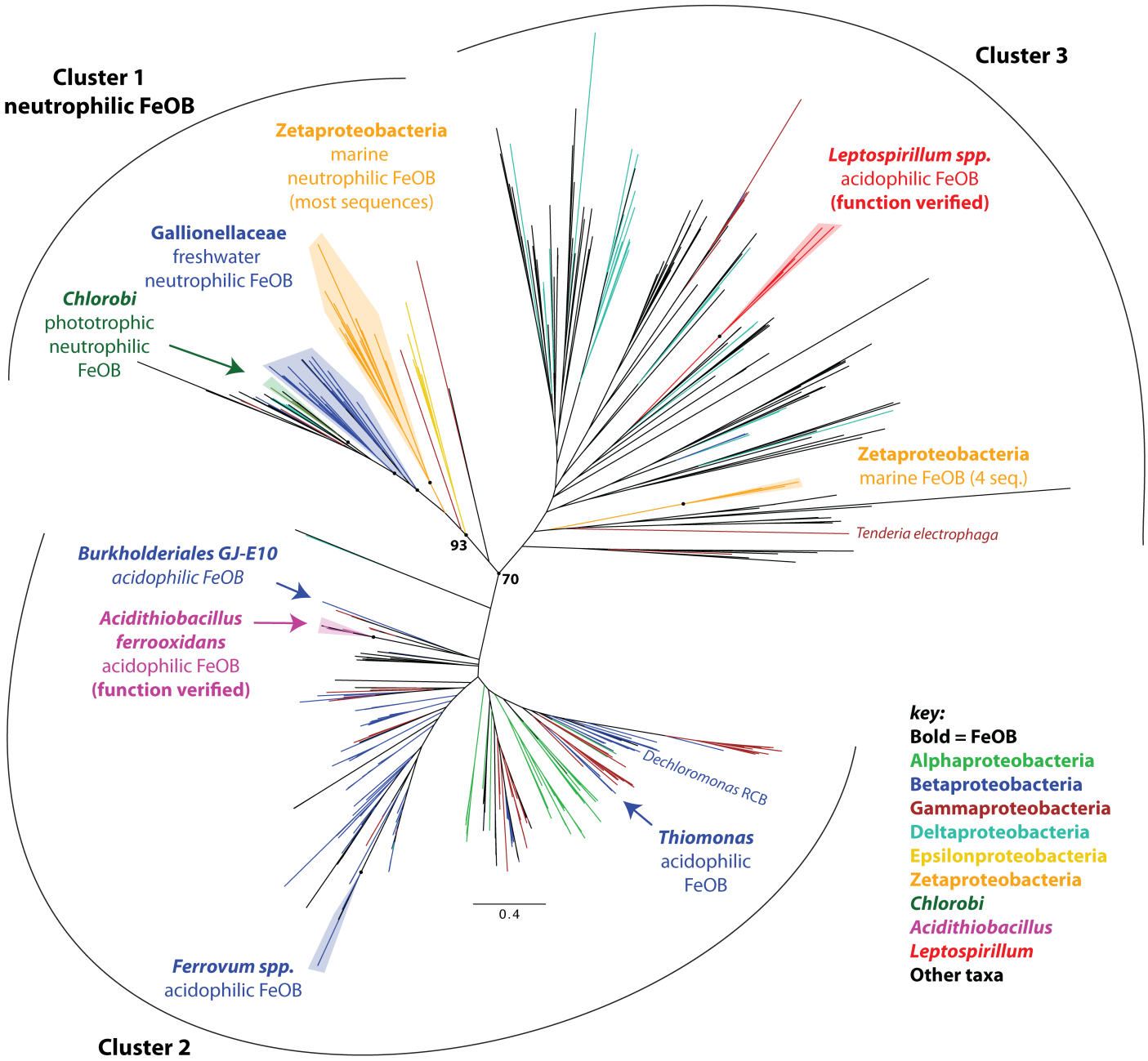

Figure 1. Cyc2 protein phylogenetic tree, with all FeOB groups labeled, including the functionally characterized *At. ferrooxidans* Cyc2 and *L. ferriphilum* Cyt_{572}, which have both been shown to oxidize Fe(II) *in vitro*. Note that the neutrophilic FeOB (Zetaproteobacteria, Gallionellaceae, *Chlorobi*) fall into a well-supported cluster (bootstrap value 93%). Also closely related is the *Chlorobi* photoferrotroph Cyc2. Bootstrap values (%) corresponding to nodes highlighted by circles: Zetaproteobacteria-most sequences (92), Gallionellaceae (98/75), Chlorobi (100), Acidithiobacillus ferrooxidans (100), Ferrovum (100), Zetaproteobacteria-4 sequences (100), Leptospirillum spp. (100). See Supplemental File 1 for full tree with all taxa and nodes labeled.

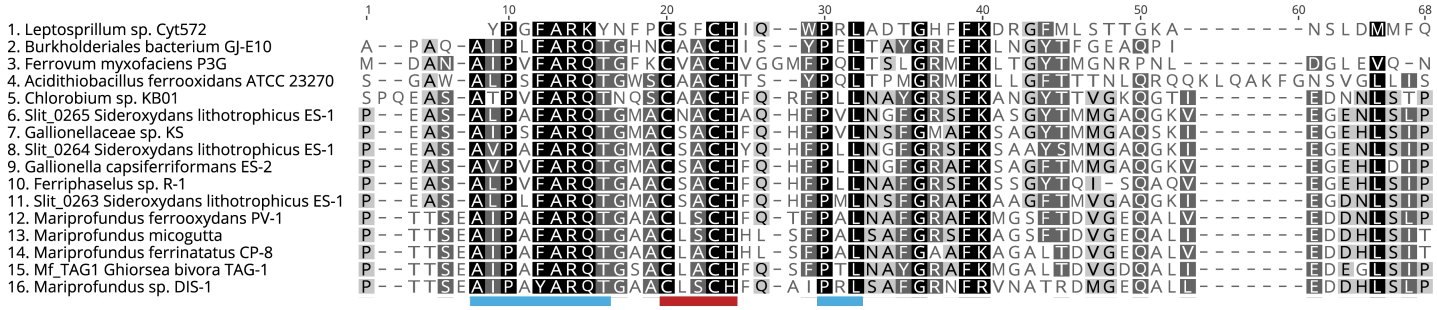

Figure 2. Alignment of cytochrome-containing section for Cyc2 from representative neutrophilic and acidophilic FeOB. Red=CXXCH heme-binding motif. Blue=other motifs discussed in the text. This section is preceded by a signal sequence, as detected by SignalP, and followed by a beta strand predicted by PSIPRED. See Supplemental Figure 2 for full Cyc2 alignment.

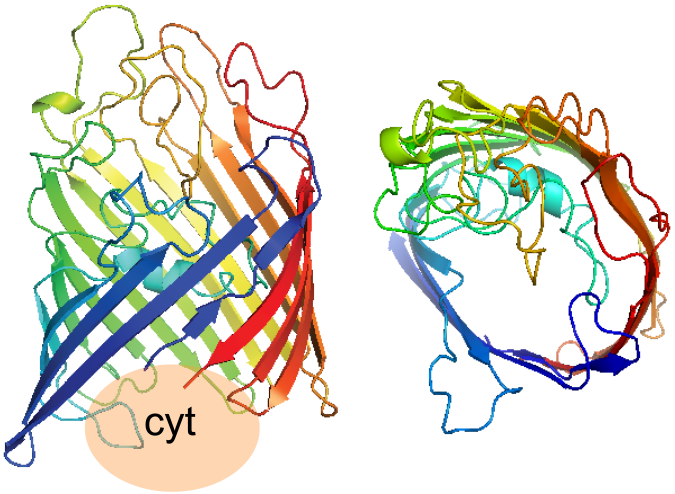

Figure 3. Two views of the Cyc2 porin model, generated using iTasser. The Cyc2 cytochrome (orange sphere=hypothesized location) is connected to the N terminal end of the porin (blue strand).

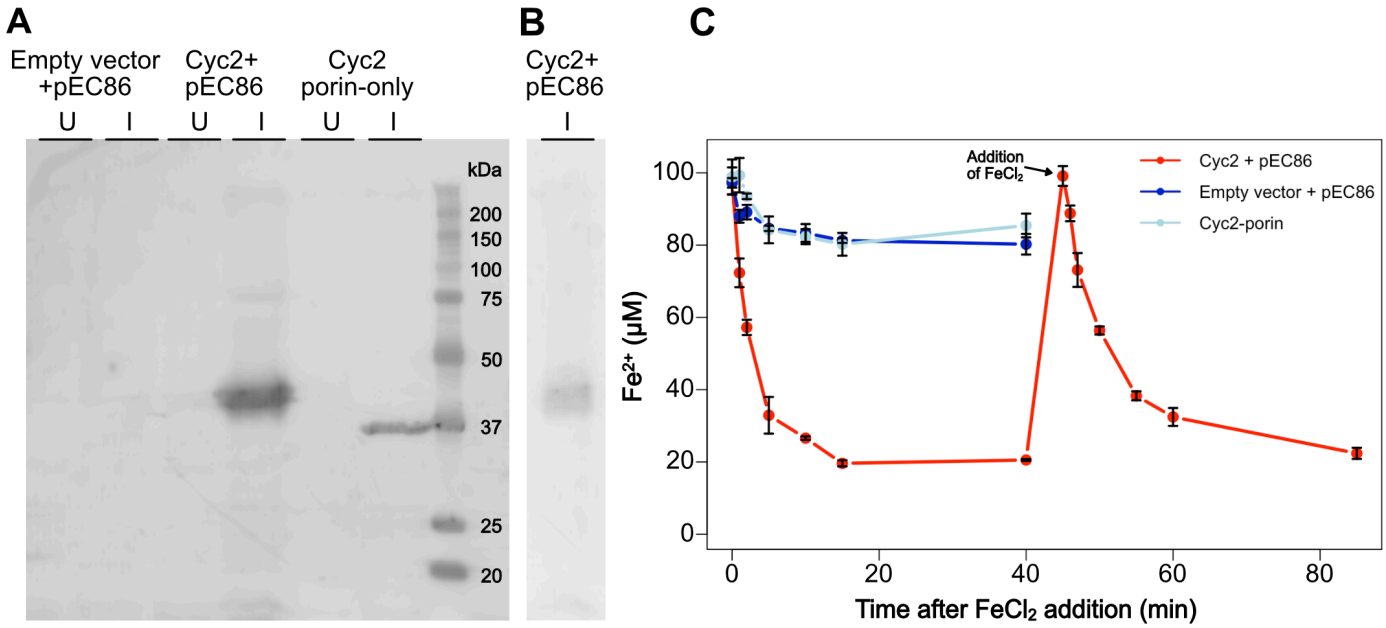

Figure 4. Heterologous expression of Cyc2 and Fe oxidation assay results. (A) Western blot showing expression of Cyc2 (lane 4) and $Cyc2_{porin}$ (lane 6); U=uninduced, I=induced with 0.5 mM IPTG. See corresponding Coomassie-stained gel in Supplemental Fig. 6B. (B) Heme stain showing successful insertion of heme into Cyc2. (C) Fe oxidation assay results. Dissolved Fe^{2+} monitored over time after Fe^{2+} addition to a suspension of *E. coli* cells expressing Cyc2, $Cyc2_{porin}$, or neither (empty vector). Fe^{2+} was added to Cyc2-expressing cells a second time. Error bars represent 1 SD of triplicate experimental runs.