1 Title: Unique metabolic strategies in Hadean analogues reveal hints for primordial

2 physiology

3 Authors:

4 Masaru Konishi Nobu[1†*], Ryosuke Nakai[1,2†], Satoshi Tamazawa[1,3], Hiroshi Mori[4], Atsushi

5 Toyoda[4], Akira Ijiri[5], Shino Suzuki[6,7], Ken Kurokawa[4], Yoichi Kamagata[1], and Hideyuki

6 Tamaki[1*]

7

8 Affiliation:

9 [1] Bioproduction Research Institute, National Institute of Advanced Industrial Science and

10 Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan

11 [2] Bioproduction Research Institute, National Institute of Advanced Industrial Science and

12 Technology (AIST), 2-17-2-1, Tsukisamu-Higashi, Sapporo, 062-8517, Japan

13 [3] Horonobe Research Institute for the Subsurface Environment (H-RISE), Northern

14 Advancement Center for Science & Technology, 5-3 Sakaemachi, Horonobe, Teshio,

15 Hokkaido, 098-3221, Japan

16 [4] National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan

17 [5] Kochi Institute for Core Sample Research, Japan Agency for Marine-Earth Science and

18 Technology (JAMSTEC), 200 Monobe Otsu, Nankoku, Kochi, Japan

19 [6] Institute for Extra-cutting-edge Science and Technology Avant-garde Research (X-star),

20 JAMSTEC, Natsushima 2-15, Yokosuka, Kanagawa 237-0061, Japan

21 [7] Institute of Space and Astronautical Science (ISAS), Japan Aerospace Exploration Agency

22 (JAXA), 3-1-1 Yoshinodai, Chuo-ku, Sagamihara, Kanagawa 252-5210, Japan

23

24 [†] These authors contributed equally.

25 * Corresponding authors: m.nobu@aist.go.jp and tamaki-hideyuki@aist.go.jp

Abstract

Primordial microorganisms are postulated to have emerged in H_2-rich alkaline Hadean serpentinite-hosted environments with homoacetogenesis as a core metabolism. However, investigation of two modern serpentinization-active analogues of early Earth reveals that conventional H_2-/CO_2-dependent homoacetogenesis is thermodynamically unfavorable *in situ* due to picomolar CO_2 levels. Through metagenomics and thermodynamics, we discover unique taxa capable of metabolism adapted to the habitat. This included a novel deep-branching phylum, "*Ca.* Lithoacetigenota", that exclusively inhabits Hadean analogues and harbors genes encoding alternative modes of H_2-utilizing lithotrophy. Rather than CO_2, these metabolisms utilize reduced carbon compounds detected *in situ* presumably serpentinization-derived: formate and glycine. The former employs a partial homoacetogenesis pathway and the latter a distinct pathway mediated by a rare selenoprotein – the glycine reductase. A survey of serpentinite-hosted system microbiomes shows that glycine reductases are diverse and nearly ubiquitous in Hadean analogues. "*Ca.* Lithoacetigenota" glycine reductases represent a basal lineage, suggesting that catabolic glycine reduction is an ancient bacterial innovation for gaining energy from geogenic H_2 even under serpentinization-associated hyperalkaline, CO_2-poor conditions. This draws remarkable parallels with ancestral archaeal H_2-driven methyl-reducing methanogenesis recently proposed. Unique non-CO_2-reducing metabolic strategies presented here may provide a new view into metabolisms that supported primordial life and the diversification of LUCA towards *Archaea* and *Bacteria*.

46   During the Hadean eon (~4.6-4.0 Ga), H_2-rich hyperalkaline fluids generated from widespread

47   serpentinization of ultramafic rocks are thought to have been conducive for the evolution of

48   primordial life [1-7]. Early microbial life is theorized to have catabolized H_2 through

49   homoacetogenesis [4,8,9], and recent studies point towards evolutionary antiquity of the central

50   enzyme of the pathway, the bifunctional CO dehydrogenase/acetyl-CoA synthase or

51   CODH/ACS [10-12]. Our recent study also shows that an acetyl-CoA pathway-like chain of

52   reactions can proceed in the presence of hydrothermal iron minerals [13], suggesting the pathway

53   preceded life and life simply encapsulated this into cells [14]. Protocells and the last universal

54   common ancestor (LUCA) are hypothesized to have evolved within alkaline hydrothermal

55   mineral deposits at the interface of serpentinization-derived fluid and ambient water (*e.g.*,

56   Hadean weakly acidic seawater) [15-17]. Although such interfaces no longer exist (*i.e.*, the Hadean

57   Earth lacked O_2 but most water bodies contain O_2 on modern Earth), some anaerobic terrestrial

58   and oceanic ecosystems harboring active serpentinization (*e.g.*, Lost City hydrothermal field)

59   [18-25] have been identified as modern analogues of ancient serpentinization-associated alkaline

60   fluids. These are ideal ecosystems for investigating what kind of H_2 catabolism may have

61   supported early post-LUCA life venture away from the interface deeper into the hyperalkaline

62   hydrothermal systems. Gaining independence from the gradient at the interface was likely a

63   critical step in the evolution of life, and life likely headed towards the hyperalkaline fluids due

64   to their reliance on serpentinization-derived energy sources and their cellular machinery being

65   alkaliphilic (*i.e.*, moving towards the weakly acidic seawater would have been an unfavorable

66   transition) [5,15-17]. However, microbiologists have yet to provide insight into such metabolic

67   strategies from extant organisms inhabiting modern analogues. In this study, we pair

68   metagenomics and thermodynamics to characterize uncultured putative anaerobic H_2 utilizers

69   inhabiting alkaline H_2-rich serpentinite-hosted systems (Hakuba Happo hot springs in Hakuba,

70   Japan and The Cedars springs in California, USA; pH ~10.9 and ~11.9, respectively) [26-28] and

71   elucidate lithotrophic catabolism that may have been relevant to early life.

72

73   **Thermogeochemistry**

74   To evaluate whether homoacetogenesis is viable *in situ*, we examined the *in situ* geochemical

75   environment and the thermodynamics of H_2/formate utilization and homoacetogenesis. The

76   spring waters of both Hakuba and The Cedars contained H_2 (*e.g.*, 201–664 μM in Hakuba [28]).

77   Formate, another compound thought to be abiotically generated through serpentinization, was

78   also detected in Hakuba (8 μM in drilling well #3 [29]) and The Cedars (6.9 μM in GPS1). Acetate

79   has also been detected *in situ* (4 μM in Hakuba [29] and 69.3 μM in The Cedars GPS1), suggesting

80   these ecosystems may host novel H_2- and/or formate-utilizing homoacetogens.

81   Thermodynamic calculations confirm that H_2 and formate are reductants *in situ* (*i.e.*, $H_2 = 2H^+$

82    $+ 2e^-$ / Formate$^- = H^+ + CO_2 + 2e^-$): the Gibbs free energy yields (ΔG) for oxidation (coupled

83   with physiological electron carriers $NADP^+$, NAD^+, and ferredoxin) are less than –4.78 kJ per

84   mol H_2 and –24.92 kJ per mol formate in Hakuba, and –10.73 and –22.03 in The Cedars

85   respectively (see Supplementary Results). However, serpentinite-hosted systems impose a

86   unique challenge to homoacetogenesis – a key substrate, CO_2, is at extremely low

87   concentrations due to the high alkalinity. We estimate that the aqueous CO_2 concentration is

88   below 0.0004 nM in Hakuba (pH 10.7 and <0.1 μM TIC) and 0.003 nM in The Cedars (pH

89   11.9 and 35 μM TIC) [26,28]. In Hakuba, H_2/CO_2-driven acetogenesis (ΔG of -1.68 kJ per mol

90   acetate) cannot support microbial energy generation ($\Delta G \leq$ -20 kJ per mol is necessary [30]; Fig.

91   S1). Moreover, in both Hakuba and The Cedars, one of the first steps in CO_2-reducing

92   homoacetogenesis, reduction of CO_2 to formate, is unfavorable based on the thermodynamics

93   presented above (ΔG > +24.92 or +22.02 kJ per mol formate). Thus, catabolic reduction of CO_2

94   to acetate is thermodynamically challenging *in situ* and may only run if investing ATP (*e.g.*,

95   Calvin-Benson-Bassham cycle [–6 ATP; ΔG of -361.68 kJ per mol acetate in Hakuba] or

96   reductive tricarboxylic acid [–1 ATP; -61.68 kJ per mol]). Under CO_2 limitation, autotrophs

97   are known to accelerate CO_2 uptake through HCO_3^- dehydration to CO_2 (carbonic anhydrase)

98   or carbonate mineral dissolution, but both only modify kinetics and are not effective in

99   changing the maximum CO_2 concentration (determined by equilibrium with carbonate species).

100   In addition, in Hakuba, the CO_3^{2-} concentration is too low (87.5 nM CO_3^{2-}) for carbonate

101   mineral precipitation (*e.g.*, $[CO_3^{2-}]$ must exceed 38.5 µM given K_s of 5×10^{-9} for $CaCO_3$ and

102   $[Ca^{2+}]$ of 0.13 mM).

103

104   Based on thermodynamic calculations, the energy obtainable from H_2/CO_2-driven

105   homoacetogenesis is too small to support life in many Hadean analogues (Fig. S2a), yet acetate

106   is detected in some of these ecosystems (Fig. S2b). Thus, CO_2-independent electron-disposing

107   metabolism may have been necessary for early life to gain energy from H_2 in the hyperalkaline

108   fluids of hydrothermal systems. Here, we explore the metabolic capacities of extant organisms

109   living in CO_2-limited Hadean analogues to gain insight into potential metabolic strategies that

110   LUCA or early post-LUCA organisms beginning diversification towards *Bacteria* and *Archaea*

111   may have utilized to thrive in alkaline serpentinite-active ecosystems.

112

113   **Diverse putative H_2- and formate-utilizing organisms**

114   Through metagenomic exploration of the two serpentinite-hosted systems (Table S1), we

115   discover a plethora of phylogenetically novel organisms encoding genes for H_2 and formate

116   metabolism (19 bins with 73.2-94.8% completeness and 0.0-8.1% contamination [86.1% and

117   3.8% on average respectively]; available under NCBI BioProject PRJNA453100) despite

118   challenges in acquisition of genomic DNA (15.7 and 18.9 ng of DNA from 233 and 720 L of

119   filtered Hakuba Happo spring water, respectively; RNA was below the detection limit). We

120   find metagenome-assembled genomes (MAGs) affiliated with lineages of *Firmicutes* (*e.g.*,

121   *Syntrophomonadaceae* and uncultured family SRB2), *Actinobacteria*, and candidate division

122   NPL-UPA2 [31] (Fig. S3). We also recovered MAGs for a novel candidate phylum, herein

123   referred to as "*Ca*. Lithoacetigenota", that inhabits both Hakuba and The Cedars and, to our

124   knowledge, no other ecosystems (Fig. 1a and S3). These genomes encode enzymes for

125   oxidizing H_2 and formate (*i.e.*, hydrogenases and formate dehydrogenases [32-39]; see

126   Supplementary Results), suggesting that organisms *in situ* can employ H_2 and formate as

127   electron donors.

128

129   **"*Ca*. Lithoacetigenota" has unique site-adapted metabolism**

130   Inspection of the serpentinite-hosted environment-exclusive phylum "*Ca*. Lithoacetigenota"

131   reveals specialization to H_2-driven lithotrophy potentially suitable for the low-CO_2 *in situ*

132   conditions (Fig. 1b). We discover that The Cedars-inhabiting population (*e.g.*, MAG BS5B28,

133   94.8% completeness and 2.9% contamination) harbors genes for H_2 oxidation ([NiFe]

134   hydrogenase Hox) and a nearly complete Wood-Ljungdahl pathway and an oxidoreductase

135   often associated with acetogenesis – NADH:ferredoxin oxidoreductase Rnf [40,41] (Table S2 and

136   S3). One critical enzyme, the formate dehydrogenase, is missing from all three "*Ca*.

137   Lithoacetigenota" MAGs from The Cedars (and unbinned contigs), indicating that these

138   bacteria can neither perform H_2/CO_2-driven nor formate-oxidizing acetogenesis (Fig. 1b).

139   However, even without the formate dehydrogenase, the genes present can form a coherent

140   pathway that uses formate rather than CO_2 as a starting point for the "methyl branch" of the

141   Wood-Ljungdahl pathway (*i.e.*, formate serves as an electron acceptor; Fig. 1b). This is a

142   simple yet potentially effective strategy for performing homoacetogenesis while circumventing

143   the unfavorable reduction of CO_2 to formate. Coupling H_2 oxidation with this formate-reducing

144   pathway is thermodynamically viable as it halves the usage of CO_2 ($3H_2$ + Formate$^-$ + CO_2 =

145   Acetate$^-$ + 2H$_2$O; ΔG of -29.62 kJ per mol acetate) and, as a pathway, is simply an intersection

146   between the conventional H$_2$/CO$_2$-driven and formate-disproportionating acetogenesis (Fig.

147   1b). Although use of formate as an electron acceptor for formate-oxidizing acetogenesis is

148   quite common, no previous homoacetogens have been observed to couple H$_2$ oxidation with

149   acetogenesis from formate, likely because CO$_2$ has a much higher availability than formate in

150   most ecosystems.

151

152   Interestingly, the Hakuba-inhabiting "*Ca*. Lithoacetigenota" (HKB210 and HKB111) also

153   encodes Hox for H$_2$ oxidation but lacks genes for homoacetogenesis (no homologs closely

154   related to The Cedars population genes were detected even in unbinned metagenomic contigs).

155   Perhaps this population forgoes the above H$_2$/formate-driven homoacetogenesis because the

156   estimated energy yield of the net reaction *in situ* (ΔG of –19.94 kJ per mol acetate) is extremely

157   close to the thermodynamic threshold of microbial catabolism (slightly above –20 kJ per mol)

158   and, depending on the actual threshold for "*Ca*. Lithoacetigenota" and/or even slight changes

159   in the surrounding conditions (*e.g.*, ΔG increases by 1 kJ per mol if H$_2$ decreases by 20 μM

160   decreases in Hakuba), the metabolism may be unable to recover energy. Through searching the

161   physicochemical environment for alternative exogenous electron acceptors and MAGs for

162   electron-disposing pathways, we detected a low concentration of glycine *in situ* (5.4 $\pm$ 1.6 nM;

163   Table S4) and found genes specific to catabolic glycine reduction (see next paragraph). We

164   suspect that some portion of this glycine is likely geochemically generated *in situ*, given that

165   (a) glycine is often detected as the most abundant amino acid produced by both natural and

166   laboratory-based serpentinization (*e.g.*, H$_2$ + Formate = Formaldehyde $\Rightarrow$ Formaldehyde +

167   NH$_3$ = Glycine) [1,7,42-47] and (b) no other amino acid was consistently detectable (if glycine was

168   cell-derived, other amino acids ought to also be consistently detected).

169

170   For utilization of the putatively abiotic glycine, the Hakuba "*Ca*. Lithoacetigenota" encode

171   glycine reductases (Grd; Fig. 2 and S4; Tables S2 and S3) – a unidirectional selenoprotein for

172   catabolic glycine reduction [48,49]. Based on the genes available, this population likely specializes

173   in coupling H_2 oxidation and glycine reduction (Fig. 1b). Firstly, the genomes encode NADP-

174   linked thioredoxin reductases (NADPH + Thioredoxin$_{ox}$ $\rightarrow$ NADP$^+$ + Thioredoxin$_{red}$) that can

175   bridge electron transfer from H_2 oxidation (H_2 + NADP $\rightarrow$ NADPH + H$^+$) to glycine reduction

176   (Glycine$^-$ + Thioredoxin$_{red}$ $\rightarrow$ Acetyl-P$_i$ + NH$_3$ + Thioredoxin$_{ox}$). Secondly, though glycine

177   reduction is typically coupled with amino acid oxidation (*i.e.*, Stickland reaction in *Firmicutes*

178   and *Synergistetes* [48,50]), similar metabolic couplings have been reported for some organisms

179   (*i.e.*, formate-oxidizing glycine reduction [via Grd] [51] and H_2-oxidizing trimethylglycine

180   reduction [via Grd-related betaine reductase] [52]). Thirdly, Grd is a rare catabolic enzyme, so

181   far found in organisms that specialize in amino acid (or peptide) catabolism, many of which

182   are reported to use glycine for the Stickland reaction (*e.g.*, *Peptoclostridium* of *Firmicutes* and

183   *Aminobacterium* of *Synergistetes* [53]). Lastly, the population lacks any discernable fermentative

184   and respiratory electron disposal pathways and oxidative organotrophy (Table S2 and S3).

185   Moreover, reflecting the lack of other catabolic pathways, the Hakuba "*Ca*. Lithoacetigenota"

186   MAGs display extensive genome streamlining, comparable to that of *Aurantimicrobium* [54,55],

187   "*Ca*. Pelagibacter" [56], and *Rhodoluna* [57] (Fig. S6). Thermodynamic calculations show that H_2-

188   oxidizing glycine reduction is thermodynamically favorable *in situ* ($\Delta G°'$ of -70.37 kJ per mol

189   glycine [ΔG of -85.84 in Hakuba]; Fig. S1). Further, based on the pathway identified, this

190   metabolism is >10 times more efficient in recovering energy from H_2 (1 mol ATP per mol H_2)

191   than acetogenesis utilizing H_2/CO_2 (0.075 mol ATP per mol H_2 based on the pathway

192   *Acetobacterium woodii* utilizes) or H_2/formate (0.075 mol ATP per mol H_2, assuming no

193   energy recovery associated with the formate dehydrogenase). We also detect glycine reductases

194   in The Cedars "*Ca*. Lithoacetigenota", indicating that it may also perform this metabolism (ΔG

195   of -76.87 in The Cedars, assuming 201 μM H_2). Thus, we propose glycine as an overlooked

196   thermodynamically and energetically favorable electron acceptor for H_2 oxidation in

197   serpentinite-hosted systems.

198

199   Given the phylogenetic and metabolic uniqueness of these populations, we report

200   provisional taxonomic assignment to "*Ca.* Lithoacetigenota" phylum. nov., "*Ca.*

201   Thermoacetigena glycinireducens" gen. nov., sp. nov. (HKB111 and HKB210), and "*Ca.*

202   Psychroacetigena formicireducens" gen. nov., sp. nov. (BS525, BS5B28, and GPS1B18) (see

203   Supplementary Results). Based on a concatenated ribosomal protein tree, this serpentinite-

204   hosted ecosystem-associated candidate phylum is closely related to a deep-branching group of

205   bacterial phyla (Fig. 1a) (*e.g.*, *Caldiserica* [Caldisericota in GTDB phylogeny],

206   *Coprothermobacterota,* and *Dictyoglomi* [Dictyoglomota] [58]), many of which have extremely

207   limited phylogenetic diversity (only 8, 1, and 2 genus-level lineages identified respectively via

208   cultivation and metagenomics [based on GTDB release 95]) and ecological distribution on

209   modern Earth. Comparative genomics shows that "*Ca*. Lithoacetigenota" shares 623 core

210   functions (based on Bacteria-level COGs/NOGs predicted by eggnog-mapper shared by the

211   two highest quality Hakuba and The Cedars MAGs HKB210 and BS5B28; Fig. 1c). When

212   compared with the core functions of two closest related phyla (*Caldiserica* and

213   *Coprothermobacterota*), 176 functions were unique to "*Ca*. Lithoacetigenota", including those

214   for NiFe hydrogenases (and their maturation proteins), selenocysteine utilization (essential for

215   Grd), and sodium:proton antiporter for alkaliphily. With *Coprothermobacterota*, 232 functions

216   were shared, including Grd, thioredoxin oxidoreductase (essential for electron transfer to Grd),

217   and additional proteins for NiFe hydrogenases and selenocysteine utilization, pointing towards

218   importance of H_2 metabolism and glycine reduction for these closely related phyla. More

219   importantly, among bacterial phyla in the deep-branching group, "*Ca*. Lithoacetigenota"

220   represents the first lineage inhabiting hyperalkaliphilic serpentinite-hosted Hadean analogue

221   ecosystems, suggesting that these organisms may be valuable extant windows into potential

222   physiologies LUCA and/or early post-LUCA organisms may have taken (albeit with 4 billion

223   years of evolution in between; see discussion regarding Grd below).

224

225   **Widespread glycine reduction in serpentinite-hosted systems**

226   Uncultured members of *Chloroflexi* (Chloroflexota) class *Dehalococcoidia* inhabiting The

227   Cedars and *Firmicutes* (Firmicutes_D) class SRB2 in Hakuba and The Cedars also possess

228   glycine reductases (Table S3). In addition, these populations encode hydrogenases and formate

229   dehydrogenases, suggesting that they may also link H_2 and formate metabolism to glycine

230   reduction. Moreover, we further discover closely related glycine reductases in other studied

231   serpentinite-hosted systems (47~94% amino acid similarity in Tablelands, Voltri Massif, and

232   Coast Range Ophiolite) [18,19,24]. Phylogenetic analysis of the glycine-binding "protein B"

233   subunit GrdB reveals close evolutionary relationships between glycine reductases from

234   distant/remote sites (Fig. 2, S4, and S5). (Tablelands spring glycine reductase sequences were

235   not included in the analysis as they were only detected in the unassembled metagenomic reads;

236   4460690.3; 69.7~82.2% similarity to Hakuba SRB2). Overall, "*Ca*. Lithoacetigenota",

237   *Dehalococcoidia*, and SRB2 glycine reductases are all detected in at least two out of the seven

238   metagenomically investigated systems despite the diverse environmental conditions (*e.g.*,

239   temperature), highlighting the potential importance of glycine metabolism by these clades in

240   serpentinite-hosted systems. We suspect that glycine reduction may be a valuable catabolic

241   strategy as the pathway requires few genes/proteins (a hydrogenase, Grd, acetate kinase, and

242   thioredoxin oxidoreductase) and conveniently provides acetate, ammonia, and ATP as basic

243   forms of carbon, nitrogen, and energy.

244

245   Comparison of the topology of the GrdB (Fig. 2, S4, and S5) and ribosomal protein (Fig.

246   1a) trees hints towards vertical transfer of GrdB dating back to one of the deepest divisions in

247   the bacterial domain. Given the ecosystem specificity and deep phylogeny of the newly

248   discovered phylum and Grd, catabolic glycine reduction may be a relatively ancient

249   metabolism viable in serpentinization-related habitats, rapidly lost due to its low utility in

250   modern ecosystems (*e.g.*, no severe nutrient/electron acceptor limitation and no excess glycine

251   via abiotic generation), but repurposed by some anaerobes for the fermentative Stickland

252   reaction in organic-rich ecosystems (*e.g.*, faeces and biodigesters) where excess amino acids

253   are available but no favorable electron acceptors are accessible (*i.e.*, anaerobic). We identified

254   the first archaeal GRD (in Miscellaneous Crenarchaeota Group [MCG] or *Ca.* Bathyarchaeota

255   member BA-1; Fig. 2, S4, and S5), but phylogenetic analysis shows that the gene was gained

256   through horizontal transfer and whether this gene truly belongs to this clade remains to be

257   verified given that the source is a metagenome-assembled genome. Thus, the currently

258   available data suggests that Grd (and catabolic glycine reduction) is an ancient bacterial

259   innovation (*i.e.*, originated in *Bacteria*) developed post-LUCA and largely exclusive to the

260   domain *Bacteria*. Further investigation of Hadean analogues is necessary to verify the history

261   of Grd (still unclear based on Bayesian inference of GrdB phylogeny; Fig. S5), uncover other

262   basal Grd and explore whether glycine reduction (or Grd family proteins) can be dated back

263   further (more ancient *Bacteria* or LUCA; *i.e.*, necessary to find more basal bacterial and/or

264   archaeal homologs).

265

266   **Other characteristics of putative indigenous homoacetogens**

267   In contrast with members of "*Ca.* Lithoacetigenota", several other putative homoacetogenic

268   populations encode the complete Wood-Ljundgahl pathway (Table S2 and S3), indicating that

269   other forms of acetogenesis may also be viable *in situ*. One putative homoacetogen in The

270   Cedars, NPL-UPA2, lacks hydrogenases but encodes formate dehydrogenases. Although the

271   NPL-UPA2 population cannot perform H_2/formate-driven acetogenesis, it may couple formate

272   oxidation with formate-reducing acetogenesis – another thermodynamically viable metabolism

273   (ΔG of –50.90 kJ per mol acetate in The Cedars; Fig. 1b). The pathway uses CO_2 as a substrate

274   but has lower CO_2 consumption compared to H_2/CO_2 homoacetogenesis and can produce

275   intracellular CO_2 from formate. In Hakuba, an *Actinobacteria* population affiliated with the

276   uncultured class UBA1414 (MAG HKB206) encodes hydrogenases and a complete Wood-

277   Ljungdahl pathway (Table S3) and, thus, may be capable of H_2/formate or the above formate-

278   disproportionating acetogenesis (Fig. 1b). Indeed, the UBA1414 population was enriched in

279   Hakuba-derived cultures aiming to enrich acetogens using the H_2 generated by the metallic

280   iron–water reaction [59] (Fig. S7). Many populations encoding a complete Wood-Ljungdahl

281   pathway possess monomeric CO dehydrogenases (CooS unassociated with CODH/ACS

282   subunits; NPL-UPA2, *Actinobacteria*, *Syntrophomonadaceae* [Hakuba and The Cedars], and

283   *Dehalococcoidia* [The Cedars]; Table S2). Although CO is below the detection limit in Hakuba

284   (personal communication with permission from Dr. Konomi Suda), another study shows that

285   CO metabolism takes place in an actively serpentinizing system with no detectable CO [60].

286   Given that CO is a known product of serpentinization [24,60], it may be an important substrate for

287   thermodynamically favorable acetogenesis *in situ*. However, further investigation is necessary

288   to verify this.

289

290   Another interesting adaptation observed for all putative homoacetogens detected in Hakuba

291   and The Cedars was possession of an unusual CODH/ACS complex. Although *Bacteria* and

292   *Archaea* are known to encode structurally distinct forms of CODH/ACS (designated as Acs

293   and Cdh respectively for this study), all studied Hakuba/The Cedars putative homoacetogens

294   encode genes for a hybrid CODH/ACS that integrate archaeal subunits for the CO

295   dehydrogenase (AcsA replaced with CdhAB) and acetyl-CoA synthase (AcsB replaced with

296   CdhC) and bacterial subunits for the corrinoid protein and methyltransferase components

297   (AcsCDE) (Table S2). The *Firmicutes* lineages also additionally encode the conventional

298   bacterial AcsABCDE. Given that all of the identified putative homoacetogens encode this

299   peculiar hybrid complex, we suspect that such CODH/ACS's may have features that adapted

300   to the high-*p*H low-CO_2 conditions (*e.g.*, high affinity for CO_2 and/or CO). In agreement, a

301   similar hybrid CODH/ACS has also been found in "*Ca.* Desulforudis audaxviator" inhabiting

302   an alkaline (*p*H 9.3) deep subsurface environment with a low CO_2 concentration (below

303   detection limit [61,62]).

304

305   **Implications for primordial biology**

306   Based on one of the most plausible theories of the origin of life, LUCA inhabited alkaline

307   geochemically active sites [63], like serpentinite-hosted systems widespread across Earth during

308   the Hadean to early Archean. Although reconstruction of the LUCA's physiology is

309   challenging [64], energy acquisition through H_2-oxidation-driven CO_2 reduction to acetyl-CoA

310   (*i.e.*, H_2/CO_2 homoacetogenesis) is theorized to be a core feature of primordial metabolism [11].

311   However, the free energy yield of such homoacetogenesis decreases with CO_2 limitation and

312   increasing temperature (Fig. S2a), suggesting that LUCA and/or early post-LUCA organisms

313   may have encountered CO_2-related thermodynamic challenges. Another study also points out

314   that CO_2 speciation towards a less bioavailable form (carbonate) under hyperalkaline pH may

315   have been a potential problem for primordial life [65]. At the interface of the alkaline fluids and

316   seawater where life is thought to have originated [7,15-17], H_2 and CO_2 from the respective fluids

317   could have come in contact and allow CO_2-dependent H_2 utilization; however, CO_2-dependent

318   metabolism would have become thermodynamically, and perhaps kinetically [66], challenging as

319   early post-LUCA organisms ventured away from the interface and deeper into serpentinite-

320   hosted systems. Through the investigation of CO_2-poor Hadean analogues, we discover a novel

321   Hadean analogue-exclusive alkaliphilic phylum that belongs to a deep-branching group of

322   bacterial phyla and possesses unconventional thermodynamically favorable less CO_2-

323   dependent H_2-oxidizing metabolic strategies (*e.g.*, coupled with formate or glycine reduction)

324   that may be compelling candidates for early H_2-dependent lithotrophy.

325

326   The thermodynamic and energetic favorability of catabolic H_2-oxidizing glycine reduction

327   mediated by GRD makes it a competitive metabolism for primordial *Bacteria*. Catabolic H_2-

328   utilizing glycine reduction also draws remarkable parallels with homoacetogenesis and

329   ancestral archaeal metabolism. Like homoacetogenesis, glycine reduction has a rare feature

330   thought to be essential for primordial metabolism – the ability to reductively synthesize acyl-

331   phosphate for substrate-level phosphorylation in the cytosol [67-70] – indicating high utility for

332   ancient organisms. Glycine reduction also shares features with a metabolism recently proposed

333   to be ancestral in *Archaea* – H_2-oxidizing methyl-reducing methanogenesis [71,72]. Both are

334   (nearly) domain-exclusive, potentially primordial, CO_2-independent, and energetically more

335   efficient than H_2/CO_2-driven acetogenesis (in terms of energy recovered per mol H_2). Moreover,

336   both utilize reduced serpentinization-derived carbon compounds, a thiol/disulfide electron

337   carrier as an electron donor, and NiFe hydrogenases as an upstream electron source (Fig. 3). It

338   is tempting to speculate that early post-LUCA *Bacteria* and *Archaea* both integrated geogenic

339   thermodynamically favorable electron acceptors to adapt to thrive under the extreme conditions

340   and interestingly occupied non-competing niches by using different acceptors (glycine and

341   methyl compounds respectively) and energy acquisition strategies (substrate-level

342   phosphorylation and proton motive force respectively), whereby allowing the two to co-exist

343   with little competition. If true, we may begin to speculate that, as *Bacteria* and *Archaea* weaned

344   themselves from serpentinite-hosted systems and explored the non-alkaline habitats of Earth,

345   GRD and MCR were connected to the methyl branch of the Wood-Ljungdahl pathway and

346   repurposed to facilitate CO_2 fixation [73] and CO_2-reducing methanogenesis respectively, both

347   CO_2-dependent pathways relevant today. However, further investigations are necessary to

348   verify because, unlike MCR, information on GRD is limited: glycine-mediated autotrophy was

349   only discovered recently, organisms still using GRD for catabolic glycine reduction are low in

350   diversity on modern Earth given the abundant availability of other electron acceptors, and the

351   geobiology of glycine-generating serpentinite-hosted systems have not been studied

352   extensively.

353

354   As was the case for MCR, metagenomic exploration of the uncharted rare biosphere

355   provided key data for genes and core pathways that provide a glimpse into how primordial life

356   or early bacteria may have proliferated (though we must still take caution in interpretation

357   given the uncertainty associated with metagenomics). The thermodynamic (i.e., more favorable

358   than the CO_2-reducing counterpart) and evolutionary (i.e., ancestral) parallels between archaeal

359   methyl-reducing and bacterial glycine-reducing H_2-dependent lithotrophy warrant

360   investigation of whether these metabolisms did indeed play a role in the formation or fixation

361   of the two domains. Further studies on the evolutionary origin, antiquity, and history of glycine

362   reduction will shed light on primordial way of life as well as bacterial/prokaryotic

363   diversification.

364

365   **Author contribution**

366   MKN and RN designed and performed metagenomic, phylogenetic, and thermodynamic

367   analyses, interpreted the data, and wrote the manuscript. ST performed sampling, DNA

368   extraction, and cultivation. HM, AT, and KK performed metagenomic sequencing and

369   metagenomic analysis. AI performed chemical measurements. SS supported chemical data

370   collection. HT and YK designed the project, supervised sampling and cultivation, and

371   supported manuscript preparation.

372

380

381   **Competing interests**

382   The authors declare no competing interests.

References

1   Schrenk, M. O., Brazelton, W. J. & Lang, S. Q. Serpentinization, Carbon, and Deep Life. *Reviews in Mineralogy and Geochemistry* **75**, 575-606, doi:10.2138/rmg.2013.75.18 (2013).

2   Komiya, T., Maruyama, S., Hirata, T., Yurimoto, H. & Nohda, S. Geochemistry of the oldest MORB and OIB in the Isua Supracrustal Belt, southern West Greenland: Implications for the composition and temperature of early Archean upper mantle. *Island Arc* **13**, 47-72, doi:10.1111/j.1440-1738.2003.00416.x (2004).

3   Sleep, N. H., Bird, D. K. & Pope, E. C. Serpentinite and the dawn of life. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 2857 (2011).

4   Martin, W. F. & Sousa, F. L. Early Microbial Evolution: The Age of Anaerobes. *Cold Spring Harbor Perspectives in Biology* **8**, doi:10.1101/cshperspect.a018127 (2016).

5   Lane, N. & Martin, William F. The Origin of Membrane Bioenergetics. *Cell* **151**, 1406-1416, doi:https://doi.org/10.1016/j.cell.2012.11.050 (2012).

6   Battistuzzi, F. U., Feijao, A. & Hedges, S. B. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evolutionary Biology* **4**, 44, doi:10.1186/1471-2148-4-44 (2004).

7   Shibuya, T. *et al.* Hydrogen-rich hydrothermal environments in the Hadean ocean inferred from serpentinization of komatiites at 300 °C and 500 bar. *Progress in Earth and Planetary Science* **2**, 46, doi:10.1186/s40645-015-0076-z (2015).

8   Camprubi, E., Jordan, S. F., Vasiliadou, R. & Lane, N. Iron catalysis at the origin of life. *IUBMB life* **69**, 373-381, doi:10.1002/iub.1632 (2017).

9   Decker, K., Jungermann, K. & Thauer, R. K. Energy Production in Anaerobic Organisms. *Angewandte Chemie International Edition in English* **9**, 138-158, doi:10.1002/anie.197001381 (1970).

10   Martin, W. F., Weiss, M. C., Neukirchen, S., Nelson-Sathi, S. & Sousa, F. L. Physiology, phylogeny, and LUCA. *Microbial Cell* **3**, 582-587, doi:10.15698/mic2016.12.545 (2016).

11   Weiss, M. C. *et al.* The physiology and habitat of the last universal common ancestor. *Nature Microbiology* **1**, 16116, doi:10.1038/nmicrobiol.2016.116 https://www.nature.com/articles/nmicrobiol2016116#supplementary-information (2016).

12   Adam, P. S., Borrel, G. & Gribaldo, S. Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proceedings of the National Academy of Sciences* **115**, E1166, doi:10.1073/pnas.1716667115 (2018).

13   Preiner, M. *et al.* A hydrogen-dependent geochemical analogue of primordial carbon and energy metabolism. *Nat Ecol Evol* **4**, 534-542, doi:10.1038/s41559-020-1125-6 (2020).

14   Martin, W. F. Older Than Genes: The Acetyl CoA Pathway and Origins. *Frontiers in microbiology* **11**, 817-817, doi:10.3389/fmicb.2020.00817 (2020).

15   Lane, N. Proton gradients at the origin of life. *Bioessays* **39**, 1600217, doi:https://doi.org/10.1002/bies.201600217 (2017).

16   Shibuya, T., Russell, M. J. & Takai, K. Free energy distribution and hydrothermal mineral precipitation in Hadean submarine alkaline vent systems: Importance of iron redox reactions under anoxic conditions. *Geochim Cosmochim Ac* **175**, 1-19, doi:https://doi.org/10.1016/j.gca.2015.11.021 (2016).

431   17   Sojo, V., Herschy, B., Whicher, A., Camprubí, E. & Lane, N. The Origin of Life in
432   Alkaline Hydrothermal Vents. *Astrobiology* **16**, 181-197, doi:10.1089/ast.2015.1406
433   (2016).

434   18   Brazelton, W. J. *et al.* Metagenomic identification of active methanogens and
435   methanotrophs in serpentinite springs of the Voltri Massif, Italy. *PeerJ* **5**, e2945,
436   doi:10.7717/peerj.2945 (2017).

437   19   Brazelton, W. J., Nelson, B. & Schrenk, M. O. Metagenomic Evidence for H(2)
438   Oxidation and H(2) Production by Serpentinite-Hosted Subsurface Microbial
439   Communities. *Frontiers in Microbiology* **2**, 268, doi:10.3389/fmicb.2011.00268
440   (2011).

441   20   Suzuki, S. *et al.* Unusual metabolic diversity of hyperalkaliphilic microbial
442   communities associated with subterranean serpentinization at The Cedars. *Isme J* **11**,
443   2584-2598, doi:10.1038/ismej.2017.111 (2017).

444   21   Kelley, D. S. *et al.* A Serpentinite-Hosted Ecosystem: The Lost City Hydrothermal
445   Field. *Science* **307**, 1428-1434, doi:10.1126/science.1102556 (2005).

446   22   Tiago, I. & Verissimo, A. Microbial and functional diversity of a subterrestrial high
447   pH groundwater associated to serpentinization. *Environ Microbiol* **15**, 1687-1706,
448   doi:10.1111/1462-2920.12034 (2013).

449   23   Crespo-Medina, M. *et al.* Methane Dynamics in a Tropical Serpentinizing
450   Environment: The Santa Elena Ophiolite, Costa Rica. *Frontiers in Microbiology* **8**,
451   916, doi:10.3389/fmicb.2017.00916 (2017).

452   24   Twing, K. I. *et al.* Serpentinization-Influenced Groundwater Harbors Extremely Low
453   Diversity Microbial Communities Adapted to High pH. *Frontiers in Microbiology* **8**,
454   308, doi:10.3389/fmicb.2017.00308 (2017).

455   25   Neubeck, A. *et al.* Microbial Community Structure in a Serpentine-Hosted Abiotic
456   Gas Seepage at the Chimaera Ophiolite, Turkey. *Appl Environ Microb* **83**,
457   doi:10.1128/aem.03430-16 (2017).

458   26   Morrill, P. L. *et al.* Geochemistry and geobiology of a present-day serpentinization
459   site in California: The Cedars. *Geochim Cosmochim Ac* **109**, 222-240,
460   doi:https://doi.org/10.1016/j.gca.2013.01.043 (2013).

461   27   Suda, K., Gilbert, A., Yamada, K., Yoshida, N. & Ueno, Y. Compound– and
462   position–specific carbon isotopic signatures of abiogenic hydrocarbons from on–land
463   serpentinite–hosted Hakuba Happo hot spring in Japan. *Geochim Cosmochim Ac* **206**,
464   201-215, doi:https://doi.org/10.1016/j.gca.2017.03.008 (2017).

465   28   Suda, K. *et al.* Origin of methane in serpentinite-hosted hydrothermal systems: The
466   CH4–H2–H2O hydrogen isotope systematics of the Hakuba Happo hot spring. *Earth*
467   *and Planetary Science Letters* **386**, 112-125,
468   doi:https://doi.org/10.1016/j.epsl.2013.11.001 (2014).

469   29   Suda, K. *Origins of Hydrocarbons in On-land Serpentinization Fields and Insights*
470   *into Hadean Hydrothermal Systems: Systematic Study using Stable Isotopes* Ph.D.
471   thesis, Tokyo Institute of Technology, (2016).

472   30   Schink, B. Energetics of syntrophic cooperation in methanogenic degradation.
473   *Microbiology and Molecular Biology Reviews* **61**, 262-280 (1997).

474   31   Suzuki, S., Nealson, K. H. & Ishii, S. i. Genomic and in-situ Transcriptomic
475   Characterization of the Candidate Phylum NPL-UPL2 From Highly Alkaline Highly
476   Reducing Serpentinized Groundwater. **9**, doi:10.3389/fmicb.2018.03141 (2018).

477   32   Morandi, P., Valzasina, B., Colombo, C., Curti, B. & Vanoni, M. A. Glutamate
478   synthase: identification of the NADPH-binding site by site-directed mutagenesis.
479   *Biochemistry-Us* **39**, 727-735, doi:10.1021/bi9920329 (2000).

480 33 Schneider, K. & Schlegel, H. G. Purification and properties of soluble hydrogenase
481   from Alcaligenes eutrophus H 16. *Biochimica et biophysica acta* **452**, 66-80 (1976).
482 34 Burgdorf, T. *et al.* The Soluble NAD(+)-Reducing [NiFe]-Hydrogenase from
483   Ralstonia eutropha H16 Consists of Six Subunits and Can Be Specifically Activated
484   by NADPH. *J Bacteriol* **187**, 3122-3132, doi:10.1128/JB.187.9.3122-3132.2005
485   (2005).
486 35 de Luca, G., de Philip, P., Rousset, M., Belaich, J. P. & Dermoun, Z. The NADP-
487   reducing hydrogenase of Desulfovibrio fructosovorans: evidence for a native complex
488   with hydrogen-dependent methyl-viologen-reducing activity. *Biochemical and*
489   *biophysical research communications* **248**, 591-596, doi:10.1006/bbrc.1998.9022
490   (1998).
491 36 Schut, G. J. & Adams, M. W. The iron-hydrogenase of *Thermotoga maritima* utilizes
492   ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen
493   production. *J Bacteriol* **191**, 4451-4457, doi:10.1128/JB.01582-08 (2009).
494 37 de Bok, F. A. *et al.* Two W-containing formate dehydrogenases (CO_2-reductases)
495   involved in syntrophic propionate oxidation by *Syntrophobacter fumaroxidans*.
496   *European journal of biochemistry / FEBS* **270**, 2476-2485 (2003).
497 38 Yamamoto, I., Saiki, T., Liu, S. M. & Ljungdahl, L. G. Purification and properties of
498   NADP-dependent formate dehydrogenase from Clostridium thermoaceticum, a
499   tungsten-selenium-iron protein. *J Biol Chem* **258**, 1826-1832 (1983).
500 39 Hidalgo-Ahumada, C. A. P. *et al.* Novel energy conservation strategies and behaviour
501   of *Pelotomaculum schinkii* driving syntrophic propionate catabolism. *Environ*
502   *Microbiol* **20**, 4503-4511, doi:10.1111/1462-2920.14388 (2018).
503 40 Nobu, M. K. *et al.* Microbial dark matter ecogenomics reveals complex synergistic
504   networks in a methanogenic bioreactor. *Isme J* **9**, 1710-1722,
505   doi:10.1038/ismej.2014.256 (2015).
506 41 Poehlein, A. *et al.* An ancient pathway combining carbon dioxide fixation with the
507   generation and utilization of a sodium ion gradient for ATP synthesis. *PLoS One* **7**,
508   e33439, doi:10.1371/journal.pone.0033439 (2012).
509 42 Fuchida, S., Mizuno, Y., Masuda, H., Toki, T. & Makita, H. Concentrations and
510   distributions of amino acids in black and white smoker fluids at temperatures over
511   200°C. *Org Geochem* **66**, 98-106,
512   doi:https://doi.org/10.1016/j.orggeochem.2013.11.008 (2014).
513 43 Haberstroh, P. R. & Karl, D. M. Dissolved free amino acids in hydrothermal vent
514   habitats of the Guaymas Basin. *Geochim Cosmochim Ac* **53**, 2937-2945,
515   doi:https://doi.org/10.1016/0016-7037(89)90170-1 (1989).
516 44 Svensson, E., Skoog, A. & Amend, J. P. Concentration and distribution of dissolved
517   amino acids in a shallow hydrothermal system, Vulcano Island (Italy). *Org Geochem*
518   **35**, 1001-1014, doi:https://doi.org/10.1016/j.orggeochem.2004.05.005 (2004).
519 45 Fox, S. W. & Windsor, C. R. Synthesis of amino acids by the heating of
520   formaldehyde and ammonia. *Science* **170**, 984-986 (1970).
521 46 Islam, M. N., Kaneko, T. & Kobayashi, K. in *Analytical Sciences/Supplements*
522   *Proceedings of IUPAC International Congress on Analytical Sciences 2001 (ICAS*
523   *2001).* i1631-i1634 (The Japan Society for Analytical Chemistry).
524 47 Inaba, S. Primary Formation Path of Formaldehyde in Hydrothermal Vents. *Origins*
525   *of Life and Evolution of Biospheres* **48**, 1-22, doi:10.1007/s11084-017-9550-5 (2018).
526 48 Andreesen, J. R. Glycine reductase mechanism. *Curr Opin Chem Biol* **8**, 454-461,
527   doi:10.1016/j.cbpa.2004.08.002 (2004).
528 49 Andreesen, J. R. in *Acetogenesis* (ed Harold L. Drake) 568-629 (Springer US,
529   1994).

50   Nisman, B. The Stickland reaction. *Bacteriological Reviews* **18**, 16-42 (1954).

51   Hormann, K. & Andreesen, J. R. Reductive cleavage of sarcosine and betaine by Eubacterium acidaminophilum via enzyme systems different from glycine reductase. *Archives of Microbiology* **153**, 50-59, doi:10.1007/bf00277541 (1989).

52   Moune, S. *et al. Haloanaerobacter salinarius* sp. nov., a novel halophilic fermentative bacterium that reduces glycine-betaine to trimethylamine with hydrogen or serine as electron donors; emendation of the genus *Haloanaerobacter. Int J Syst Bacteriol* **49 Pt 1**, 103-112, doi:10.1099/00207713-49-1-103 (1999).

53   Hamdi, O. *et al.* Aminobacterium thunnarium sp. nov., a mesophilic, amino acid-degrading bacterium isolated from an anaerobic sludge digester, pertaining to the phylum Synergistetes. *Int J Syst Evol Microbiol* **65**, 609-614, doi:10.1099/ijs.0.068965-0 (2015).

54   Nakai, R., Baba, T., Niki, H., Nishijima, M. & Naganuma, T. Aurantimicrobium minutum gen. nov., sp. nov., a novel ultramicrobacterium of the family Microbacteriaceae, isolated from river water. *Int J Syst Evol Microbiol* **65**, 4072-4079, doi:10.1099/ijsem.0.000541 (2015).

55   Nakai, R. *et al.* Complete Genome Sequence of Aurantimicrobium minutum Type Strain KNCT, a Planktonic Ultramicrobacterium Isolated from River Water. *Genome announcements* **4**, doi:10.1128/genomeA.00616-16 (2016).

56   Giovannoni, S. J. *et al.* Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**, 1242-1245, doi:10.1126/science.1114057 (2005).

57   Hahn, M. W., Schmidt, J., Taipale, S. J., Doolittle, W. F. & Koll, U. Rhodoluna lacicola gen. nov., sp. nov., a planktonic freshwater bacterium with stream-lined genome. *Int J Syst Evol Microbiol* **64**, 3254-3263, doi:10.1099/ijs.0.065292-0 (2014).

58   Betts, H. C. *et al.* Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nature Ecology & Evolution* **2**, 1556-1562, doi:10.1038/s41559-018-0644-x (2018).

59   Kato, S., Yumoto, I. & Kamagata, Y. Isolation of Acetogenic Bacteria That Induce Biocorrosion by Utilizing Metallic Iron as the Sole Electron Donor. *Appl Environ Microb* **81**, 67-73, doi:10.1128/aem.02767-14 (2015).

60   Morrill, P. L. *et al.* Investigations of potential microbial methanogenic and carbon monoxide utilization pathways in ultra-basic reducing springs associated with present-day continental serpentinization: the Tablelands, NL, CAN. *Frontiers in Microbiology* **5**, doi:10.3389/fmicb.2014.00613 (2014).

61   Chivian, D. *et al.* Environmental genomics reveals a single-species ecosystem deep within Earth. *Science* **322**, 275-278, doi:10.1126/science.1155495 (2008).

62   Lin, L.-H. *et al.* Long-Term Sustainability of a High-Energy, Low-Diversity Crustal Biome. *Science* **314**, 479 (2006).

63   Martin, W. & Russell, M. J. On the origin of biochemistry at an alkaline hydrothermal vent. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**, 1887 (2007).

64   Berkemer, S. J. & McGlynn, S. E. A New Analysis of Archaea–Bacteria Domain Separation: Variable Phylogenetic Distance and the Tempo of Early Evolution. *Molecular Biology and Evolution* **37**, 2332-2340, doi:10.1093/molbev/msaa089 (2020).

65   Boyd, E. S., Amenabar, M. J., Poudel, S. & Templeton, A. S. Bioenergetic constraints on the origin of autotrophic metabolism. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **378**, 20190151, doi:doi:10.1098/rsta.2019.0151 (2020).

579 66 Gonzalez-Cabaleiro, R., Lema, J. M., Rodriguez, J. & Kleerebezem, R. Linking
580   thermodynamics and kinetics to assess pathway reversibility in anaerobic
581   bioprocesses. *Energy & Environmental Science* **6**, 3780-3789,
582   doi:10.1039/C3EE42754D (2013).
583 67 Martin, W. F. & Thauer, R. K. Energy in Ancient Metabolism. *Cell* **168**, 953-955,
584   doi:10.1016/j.cell.2017.02.032 (2017).
585 68 Lipmann, F. in *The Origins of Prebiological Systems and of their Molecular Matrices*
586   259-280 (Elsevier, 1965).
587 69 Decker, K., Jungermann, K. & Thauer, R. K. Energy production in anaerobic
588   organisms. *Angew Chem Int Ed Engl* **9**, 138-158, doi:10.1002/anie.197001381 (1970).
589 70 Ferry, J. G. & House, C. H. The Stepwise Evolution of Early Life Driven by Energy
590   Conservation. *Molecular Biology and Evolution* **23**, 1286-1292,
591   doi:10.1093/molbev/msk014 (2006).
592 71 Baker, B. J. *et al.* Diversity, ecology and evolution of Archaea. *Nature Microbiology*
593   **5**, 887-900, doi:10.1038/s41564-020-0715-z (2020).
594 72 Borrel, G. *et al.* Wide diversity of methane and short-chain alkane metabolisms in
595   uncultured archaea. *Nature Microbiology* **4**, 603-613, doi:10.1038/s41564-019-0363-
596   3 (2019).
597 73 Sánchez-Andrea, I. *et al.* The reductive glycine pathway allows autotrophic growth of
598   Desulfovibrio desulfuricans. *Nature Communications* **11**, 5090, doi:10.1038/s41467-
599   020-18906-7 (2020).
600 74 Nobu, M. K., Narihiro, T., Kuroda, K., Mei, R. & Liu, W. T. Chasing the elusive
601   *Euryarchaeota* class WSA2: genomes reveal a uniquely fastidious methyl-reducing
602   methanogen. *Isme J* **10**, 2478-2487, doi:10.1038/ismej.2016.33 (2016).
603 75 Heuer, V. *et al.* Online δ13C analysis of volatile fatty acids in sediment/porewater
604   systems by liquid chromatography‐isotope ratio mass spectrometry. *Limnology and
605   Oceanography: Methods* **4**, 346-357, doi:doi:10.4319/lom.2006.4.346 (2006).
606 76 Ijiri, A. *et al.* Biogeochemical processes involving acetate in sub-seafloor sediments
607   from the Bering Sea shelf break. *Org Geochem* **48**, 47-55,
608   doi:https://doi.org/10.1016/j.orggeochem.2012.04.004 (2012).
609 77 Hanselmann, K. W. Microbial energetics applied to waste repositories. *Experientia*
610   **47**, 645-687, doi:10.1007/BF01958816 (1991).
611 78 Wang, G., Spivack, A. J. & D'Hondt, S. Gibbs energies of reaction and microbial
612   mutualism in anaerobic deep subseafloor sediments of ODP Site 1226. *Geochim
613   Cosmochim Ac* **74**, 3938-3947, doi:https://doi.org/10.1016/j.gca.2010.03.034 (2010).
614 79 Mason, O. U. *et al.* Metagenome, metatranscriptome and single-cell sequencing
615   reveal microbial response to Deepwater Horizon oil spill. *Isme J* **6**, 1715-1727,
616   doi:10.1038/ismej.2012.59 (2012).
617 80 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina
618   sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170
619   (2014).
620 81 Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications
621   to single-cell sequencing. *Journal of computational biology* **19**, 455-477,
622   doi:10.1089/cmb.2012.0021 (2012).
623 82 Wu, Y.-W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: an automated binning
624   algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*
625   **32**, 605-607, doi:10.1093/bioinformatics/btv638 (2016).
626 83 Wu, Y. W., Tang, Y. H., Tringe, S. G., Simmons, B. A. & Singer, S. W. MaxBin: an
627   automated binning method to recover individual genomes from metagenomes using

628　　　　an expectation-maximization algorithm. *Microbiome* **2**, 26, doi:10.1186/2049-2618-2-
629　　　　26 (2014).
630　84　Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W.
631　　　　CheckM: assessing the quality of microbial genomes recovered from isolates, single
632　　　　cells, and metagenomes. *Genome Res* **25**, 1043-1055, doi:10.1101/gr.186072.114
633　　　　(2015).
634　85　Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-
635　　　　2069, doi:10.1093/bioinformatics/btu153 (2014).
636　86　Huerta-Cepas, J. *et al.* eggNOG 4.5: a hierarchical orthology framework with
637　　　　improved functional annotations for eukaryotic, prokaryotic and viral sequences.
638　　　　*Nucleic Acids Res* **44**, D286-D293, doi:10.1093/nar/gkv1248 (2016).
639　87　Katoh, K., Kuma, K.-i., Toh, H. & Miyata, T. MAFFT version 5: improvement in
640　　　　accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**, 511-518,
641　　　　doi:10.1093/nar/gki198 (2005).
642　88　Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for
643　　　　automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*
644　　　　**25**, 1972-1973, doi:10.1093/bioinformatics/btp348 (2009).
645　89　Imachi, H. *et al.* Isolation of an archaeon at the prokaryote-eukaryote interface.
646　　　　*Nature* **577**, 519-525, doi:10.1038/s41586-019-1916-6 (2020).
647　90　Parks, D. H. *et al.* A complete domain-to-species taxonomy for Bacteria and Archaea.
648　　　　*Nat Biotechnol*, doi:10.1038/s41587-020-0501-8 (2020).
649　91　Guindon, S. *et al.* New Algorithms and Methods to Estimate Maximum-Likelihood
650　　　　Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology* **59**, 307-
651　　　　321, doi:10.1093/sysbio/syq010 (2010).
652　92　Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol
653　　　　Biol Evol* **25**, 1307-1320, doi:10.1093/molbev/msn067 (2008).
654　93　Lemoine, F. *et al.* Renewing Felsenstein's phylogenetic bootstrap in the era of big
655　　　　data. *Nature* **556**, 452-456, doi:10.1038/s41586-018-0043-0 (2018).
656　94　Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-
657　　　　generation sequencing data. *Bioinformatics* **28**, 3150-3152,
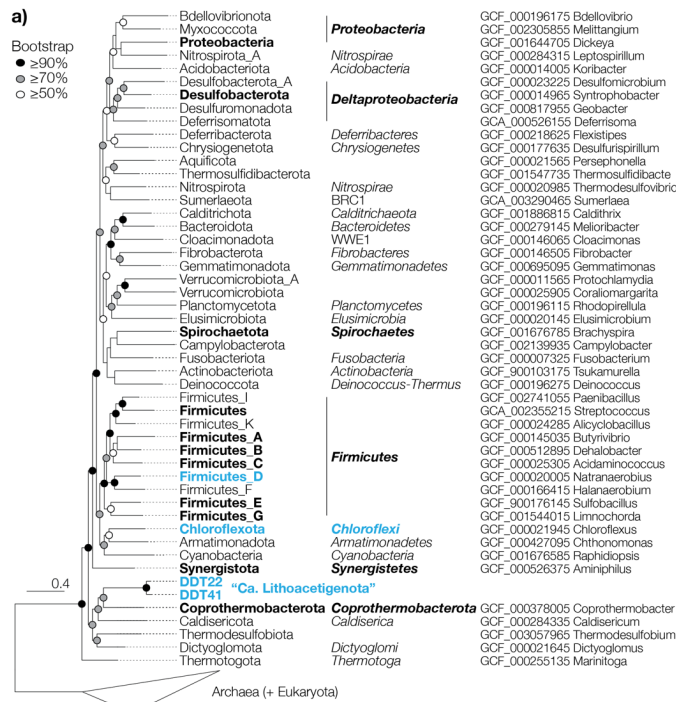658　　　　doi:10.1093/bioinformatics/bts565 (2012).
659

660 **Figures**

Figure 1. "*Ca.* Lithoacetigenota" phylogeny, lithotrophic acetate generation pathways, and comparative genomics with neighboring phyla. (a) A maximum likelihood tree was calculated using PhyML using the LG model, 4 gamma categories, and 100 bootstrap replicates from a concatenated alignment of universally conserved ribosomal protein sequences from representative genomes of individual phyla aligned with MAFFT (default parameters) and trimmed with trimAl prior to concatenation. Bootstrap values were recalculated using BOOSTER. Phyla that possess glycine reductases (black+bolded) and phyla for which glycine reductases were detected in serpentinite-hosted systems are indicated (blue+bolded). Phylum names are shown for both NCBI taxonomy (italicized) and GTDB classification. (b) Metabolic pathways potentially adapted to the CO_2-limited hyperalkaline conditions encoded by "*Ca.* Lithoacetigenota" members and others: formate- and glycine-reducing acetate generation. Arrow colors indicate oxidative (pink), reductive (blue), ATP-yielding (orange), and ATP-consuming (green) steps. (c) Venn diagram of COGs/NOGs (as predicted by eggnog-mapper) fully conserved across all members of each phylum (genomes included in GTDB release 95 with completeness ≥85% and contamination ≤5%). COGs/NOGs related to lithotrophy and alkaliphily are highlighted. * "COG" abbreviated.
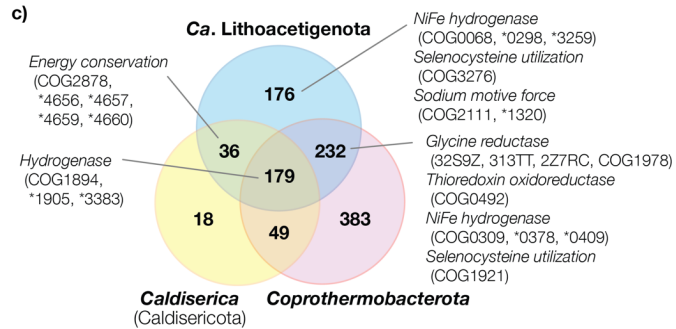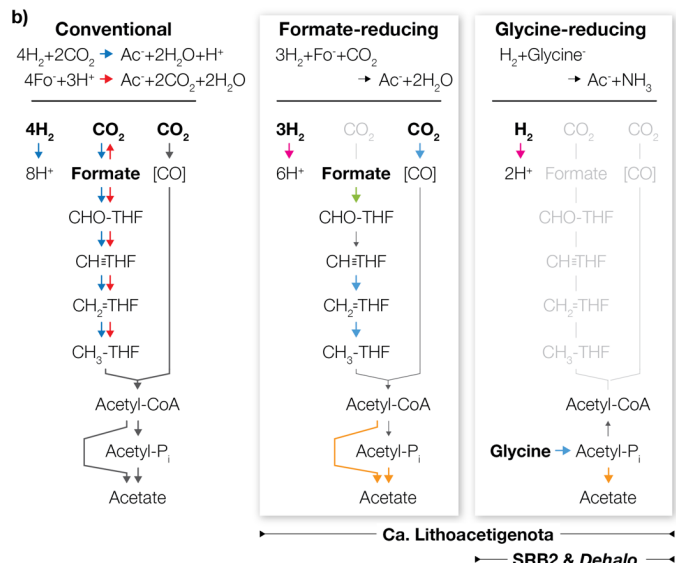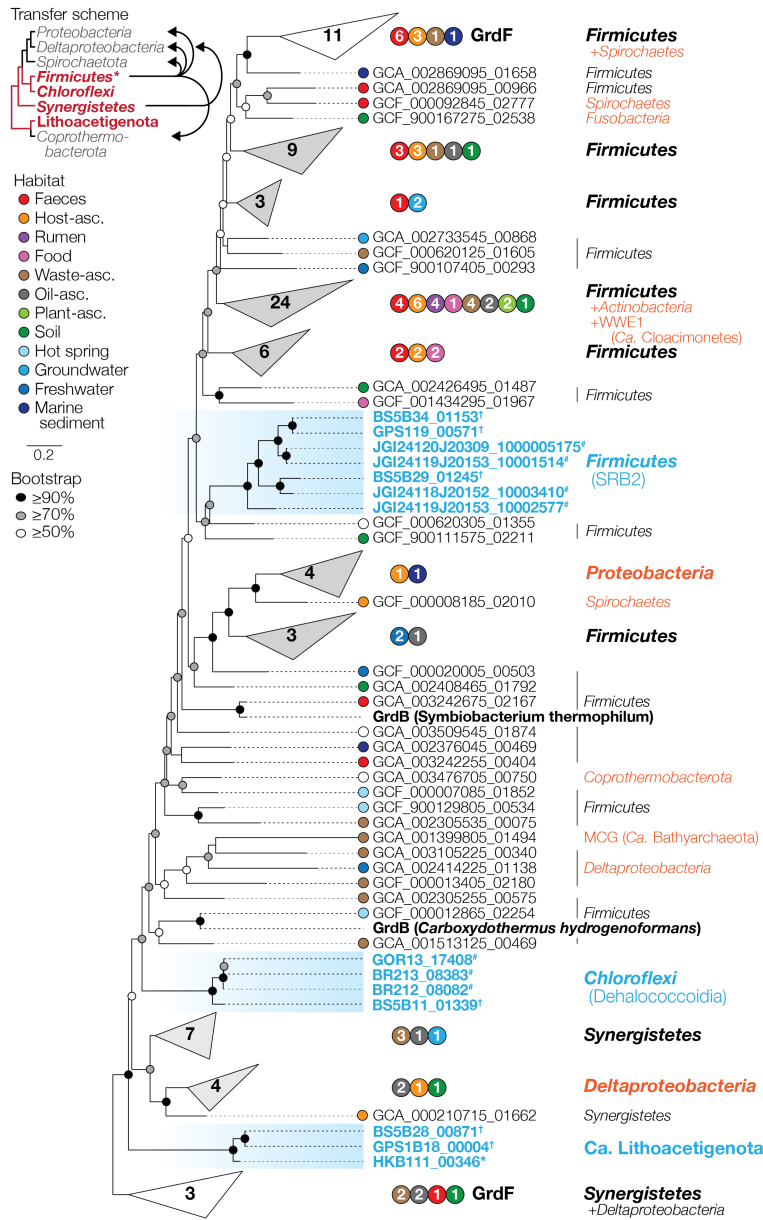
Figure 2. Phylogeny of serpentinite-hosted microbiome glycine reductase subunit GrdB homologs (Hakuba Happo hot spring*, The Cedars springs[†], and other serpentinite-hosted system metagenomes[#]) and a brief scheme for evolutionary history of Grd. COG1978 homologs were collected from the representative species genomes in GTDB, filtered using a GrdB motif conserved across members of phyla known to perform glycine-reducing Stickland reaction and a GrdF motif conserved across sequences that form a distinct cluster around the biochemically characterized *Peptoclostridium acidaminophilum* sarcosine reductase subunit GrdF (see Methods and Supplementary Fig. S4), and clustered with 75% amino acid sequence similarity using CD-HIT (-c 0.75). A maximum likelihood tree was calculated as described in Fig. 1. For each sequence, the original habitat the isolate or MAG was obtained from is shown (white circle = unknown). Large sequence clusters were grouped (number of representative sequences included are shown). Note that the outgroup is a cluster of uncharacterized *Synergistetes* and *Deltaproteobacteria* sequences that was inferred to function as GrdF given that it shares the motif found in the *Firmicutes* GrdF. Taxa thought to have gained GrdB through horizontal transfer are shown in orange. See Supplementary Fig. S4 for complete tree. In the brief scheme of Grd evolution, the cladogram topology is based on Fig. 1a. Vertical transfer (red lines in cladogram) and horizontal transfer (black arrows) inferred from tree structures are shown. Phyla that may have acquired Grd vertically (red) and horizontally (gray) are indicated. GTDB phyla belonging to *Firmicutes* were grouped together.
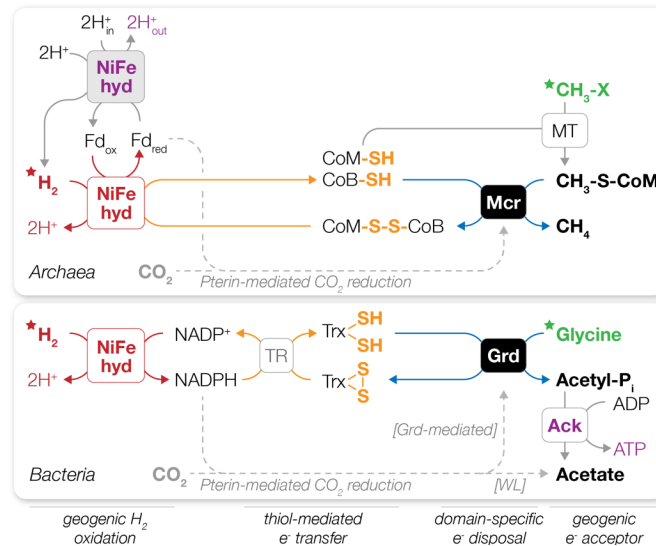
Figure 3. Proposed non-CO$_2$-reducing metabolic strategies of ancestral archaea and bacteria inhabiting serpentinite-hosted systems. Geogenic carbon compounds, in particular methyl groups for archaea and glycine for bacteria (marked by a green star), can serve as electron acceptors for the oxidation of geogenic H$_2$ (red star). The bacterial glycine-reducing H$_2$-dependent lithotrophy is strategically similar to the ancestral archaeal methyl-reducing methanogenesis recently suggested [71,72]. Both reactions can be mediated by nickel-iron [NiFe] hydrogenase H$_2$ oxidation (red) and reduction of thiol/disulfide electron carriers (orange) and forgo pterin-mediated CO$_2$ reduction (gray dotted line). Energy recovery (purple) is contrasting between the two metabolisms: archaea employing proton motive force (membrane-bound proteins gray) and bacteria substrate level phosphorylation. Reactions are not balanced with H$^+$, H$_2$O, and P$_i$. The pathways shown are based on "*Ca.* Methanofastidiosum" [74] and "*Ca.* Thermolithoacetigena glycinireducens". NiFe hyd, NiFe hydrogenase; Fd, ferredoxin; CoM-SH, coenzyme M thiol; CoB-SH, coenzyme B thiol; CoM-S-S-CoB, heterodisulfide of coenzyme M and coenzyme B; Mcr, methyl-coenzyme M reductase; CH$_3$-X, methyl compounds; MT, methyltransferase; CH$_3$-S-CoM, methyl-coenzyme M; TR, thioredoxin reductase; Trx-(SH)$_2$, thioredoxin dithiol; Trx-S$_2$, thioredoxin disulfide; Grd, glycine reductase; acetyl-P$_i$, acetyl-phosphate; Ack, acetate kinase.

764 **Methods**

765 *Sampling site and sample collection*

766 The Hakuba Happo samples for geochemical and microbiological analysis were artificially

767 pumped from a drilling well (700 m in depth), which was previously described and named

768 Happo #3 (36° 42' N 137° 48' E [28]). For microbiological analysis, two spring water samples

769 were taken at different time points, 233 L taken in July 2016 (labelled HKB701) and 720 L

770 taken in October 2016 (labelled HKB702), respectively. To collect cells, samples were filtered

771 through a 0.1-μm Omnipore™ membrane filter (Merck Millipore) using a 90 mm diameter

772 stainless-steel filter holder (Merck Millipore) attached to FDA Viton® tubing (Masterflex) at a

773 sampling site. After filtration, filters were immediately transferred to sterile tubes and frozen

774 in a dry ice-ethanol bath, transported in dry ice, and stored at −80°C until DNA extraction. For

775 NH_3 and amino acid analysis, water samples were collected in October 2017, filtered as

776 described above, transferred to dry-heat-sterilized nitrogen-purged 100-ml glass vials, and

777 stored at 4°C.

778

779 *Geochemical analysis*

780 The water temperature of hot spring water was measured using a thermometer (CT-430WP,

781 Custom Ltd.) at a site. The pH, oxidation reduction potential (ORP), electrical conductivity

782 (EC), and dissolved oxygen (DO) level were determined with portable devices, including a *p*H

783 meter (D-23, Horiba), an ORP meter (RM-30P, TOA-DKK), an EC meter (CM-31P, TOA-

784 DKK), and a DO meter (DO-31P, TOA-DKK), correspondingly. The ion concentrations of

785 Na^+, K^+, Ca^{2+}, and NO_3^- were determined using portable sensors (LAQUAtwin™ series,

786 Horiba). The *in situ* NH_3 concentration was determined by measuring aqueous NH_4^+ and

787 gaseous NH_3 (purged with N_2 gas, gas dissolved into deionized water, and measured dissolved

788 NH_4^+) of a sample stored as described above using high-performance liquid chromatography

789   (HPLC; Prominence; Shimadzu), then adding the two together. For amino acid quantification,

790   the sample was concentrated under a stream of nitrogen gas and then analyzed following

791   Shimadzu protocol no. L323 (https://www.ssi.shimadzu.com/products/literature/lc/L323.pdf)

792   using HPLC with minor modifications (fluorescence detector RF-20Axs; sodium hypochlorite

793   solution was not added for detection of proline). The Cedars spring concentrations of formate

794   and acetate were determined by Isotope-Ratio-Monitoring Liquid Chromatography Mass

795   Spectrometry (IRM-LCMS); Thermo- Finnigan Delta Plus XP isotope-ratio mass spectrometer

796   connected to LC IsoLink, as described by Heuer et al. [75] and Ijiri et al. [76].

797

798   *Thermodynamic calculations*

799   The Hakuba and Cedars calculations Gibbs free energy yield (ΔG) are based on $\Delta G°_f$ and $\Delta H°_f$

800   values at 298 K, respective pH (10.8 and 11.9), and adjustment to the *in situ* temperatures (48

801   and 17°C) through the Gibbs-Helmholtz equation [77]. The effect of pressure was approximated

802   as described by Wang *et al* [78]. For both Hakuba and The Cedars calculations, the glycine

803   concentration (5.4 nM) was based on measurements from Hakuba. Formate, acetate, and NH_3

804   concentrations were based on respective measurements from Hakuba (8 μM formate, 4 μM

805   acetate, and 2.9 μM NH_3) and The Cedars (6.9 μM formate [average of 6.777 and 7.079 μM

806   measured on September 2017], 69.3 μM acetate [average of 69.601 and 68.967 μM measured

807   on September 2017], and 1 μM NH_3 [below detection limit]). For Hakuba, the H_2 concentration

808   measured in Hakuba drilling well #3 (DNA source) was used (201 μM H_2). For The Cedars,

809   the highest detected H_2 concentration in Hakuba was used (664 μM H_2 in drilling well #1).

810

811   *Metagenome sequencing, assembly and binning*

812   The filter was aseptically cut into 16 equal pieces using sterilized tweezers, and each piece was

813   placed in the bead-beating tube (Lysing Matrix E tube; MP Biomedicals). After DNA

814   extraction following the bead-beating method described previously [79], the 16 DNA samples

815   were mixed and then stored at −80°C until used. Sequence libraries were prepared with Nextera

816   XT DNA Library Preparation kit (Illumina) with a genomic DNA fragment size ranging from

817   200 to 2,000 bp. These libraries sequenced on HiSeq2500 sequencing platform (Illumina) with

818   HiSeq Rapid SBS kit v2 (Illumina), generating paired-end reads up to 250 bp. The generated

819   sequences were trimmed using Trimmomatic v0.33 [80] with a quality cutoff of 30, sliding

820   window of 6 bp, and minimum length cutoff of 78 bp. The trimmed sequences were assembled

821   using SPAdes v3.10.1 [81] with the "-meta" option and k-mer values of 21, 31, 41, 53, 65, and

822   77. The assembled contigs were binned using MaxBin2.2.1 [82,83]. The completeness and

823   contamination of each bin was checked using CheckM [84]. These bins were manually curated

824   as described in our metagenomics study [40]. Genes were then annotated using Prokka v1.12 [85]

825   and eggnog-mapper [86]. For interpretation and comparison of microbial metabolism, bin

826   genomes were also constructed from public metagenomic data generated from The Cedars [20]

827   (trimmed with sliding window of 6, quality cutoff of 20, and minimum length of 68 bp through

828   Trimmomatic v0.33, normalized using BBMap 36.99 (https://jgi.doe.gov/data-and-

829   tools/bbtools/) with target and minimum coverages of 40 and 2, assembled using SPAdes

830   v3.10.1 with the "-meta" option and k-mer values of 21, 33, 45, 55, 67, and binned through

831   MaxBin2.2.1) and were then analyzed collectively.

832

833   *Phylogenomic and phylogenetic analysis*

834   For tree construction, sequences were aligned with MAFFT [87] v7.453 (default parameters) and

835   trimmed using trimAl [88] v1.2rev59 (-gt 0.9). For ribosomal protein trees, a concataenated

836   alignment of universally conserved ribosomal proteins [89] was used. Protein sequences were

837   retrieved by downloading the GTDB [90] database and predicting protein sequences using Prokka

838   [85] 1.14 (--kingdom Bacteria/Archaea --rnammer --addgenes --mincontiglen 200). Maximum

839   likelihood trees were calculated using PhyML [91] v3.3.20190909 using the LG [92] model and 100

840   bootstrap replicates (-b 100 -d aa -m LG -v e). Bootstrap values were recalculated using

841   BOOSTER [93]. Sequence clustering was performing through CD-HIT [94] v4.8.1. For glycine

842   reductase GrdB and sarcosine reductase GrdF, conserved motifs were predicted by first

843   identifying fully conserved residues in the sequence cluster including the biochemically

844   characterized *Peptoclostridium acidaminophilum* GrdF (see Supplementary Fig. S4;

845   YxNx(6)GGE   x(34,38)   CGD   x(27,35)   GPxF[NF]AGRYG   x(150,181)

846   IHGGYDRx(6)[IP]x(4)PxD x(19,20) TTGTGTx(7)F x(12) [HILV]), then identifying fully

847   conserved residues in the phylogenetic clusters (see Supplementary Fig. S4) that include GrdB

848   from phyla known to perform the Stickland reaction (*Firmicutes*, *Spirochaetes*, and

849   *Synergistetes*) subtracting any sequences clusters that contain the GrdF motif above

850   (YxNx(6)GGE x(34,38) CGD x(27,35) GPxF[NF]AGRYG x(157,178) AHGGxD[QTAP] x(8)

851   RV[IL]PxD x(19,20) TxGNxTxV). Bayesian trees were calculated using Phylobayes (MPI)

852   v1.8 using the LG model with 4 gamma categories (-lg -ncat 1 -dgam 4) with four chains (chain

853   length 10000; -x 1 10000) of which three chains that converged (maxdiff=0.1859,

854   meandiff=0.014) were used for determination of a consensus tree with a burn-in of 1000 (nodes

855   with posterior probability less than 0.8 were collapsed; bpcomp -c 0.8 -x 1000 1).

856

857   *Data availability*

858   The datasets generated during and/or analyzed during the current study are available in the

859   National Center for Biotechnology Information (NCBI) under BioProject PRJNA453100 and

860   BioSamples SAMN08978938-SAMN08978962.