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

Authors:

Masaru Konishi Nobu^{1†*}, Ryosuke Nakai^{1,2†}, Satoshi Tamazawa^{1,3}, Hiroshi Mori⁴, Atsushi Toyoda⁴, Akira Ijiri⁵, Shino Suzuki^{6,7}, Ken Kurokawa⁴, Yoichi Kamagata¹, and Hideyuki Tamaki^{1*}

Affiliation:

¹ Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan

² Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 2-17-2-1, Tsukisamu-Higashi, Sapporo, 062-8517, Japan

³ Horonobe Research Institute for the Subsurface Environment (H-RISE), Northern Advancement Center for Science & Technology, 5-3 Sakaemachi, Horonobe, Teshio, Hokkaido, 098-3221, Japan

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

⁵ Kochi Institute for Core Sample Research, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 200 Monobe Otsu, Nankoku, Kochi, Japan

⁶ Institute for Extra-cutting-edge Science and Technology Avant-garde Research (X-star), JAMSTEC, Natsushima 2-15, Yokosuka, Kanagawa 237-0061, Japan

⁷ Institute of Space and Astronautical Science (ISAS), Japan Aerospace Exploration Agency (JAXA), 3-1-1 Yoshinodai, Chuo-ku, Sagami-hara, Kanagawa 252-5210, Japan

[†] These authors contributed equally.

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

Abstract

Primordial microorganisms are postulated to have emerged in H₂-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 H₂-/CO₂-dependent homoacetogenesis is thermodynamically unfavorable *in situ* due to picomolar CO₂ 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 H₂-utilizing lithotrophy. Rather than CO₂, 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 H₂ even under serpentinization-associated hyperalkaline, CO₂-poor conditions. This draws remarkable parallels with ancestral archaeal H₂-driven methyl-reducing methanogenesis recently proposed. Unique non-CO₂-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), H₂-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 H₂ 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 O₂ but most water bodies contain O₂ 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 H₂ 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 H₂ utilizers
 69 inhabiting alkaline H₂-rich serpentinite-hosted systems (Hakuba Happo hot springs in Hakuba,

Japan and The Cedars springs in California, USA; $pH \sim 10.9$ and ~ 11.9 , respectively)²⁶⁻²⁸ and elucidate lithotrophic catabolism that may have been relevant to early life.

Thermogeochemistry

To evaluate whether homoacetogenesis is viable *in situ*, we examined the *in situ* geochemical environment and the thermodynamics of H_2 /formate utilization and homoacetogenesis. The spring waters of both Hakuba and The Cedars contained H_2 (*e.g.*, 201–664 μM in Hakuba²⁸). Formate, another compound thought to be abiotically generated through serpentinization, was also detected in Hakuba (8 μM in drilling well #3²⁹) and The Cedars (6.9 μM in GPS1). Acetate has also been detected *in situ* (4 μM in Hakuba²⁹ and 69.3 μM in The Cedars GPS1), suggesting these ecosystems may host novel H_2 - and/or formate-utilizing homoacetogens. Thermodynamic calculations confirm that H_2 and formate are reductants *in situ* (*i.e.*, $H_2 = 2H^+ + 2e^-$ / $Formate^- = H^+ + CO_2 + 2e^-$): the Gibbs free energy yields (ΔG) for oxidation (coupled with physiological electron carriers $NADP^+$, NAD^+ , and ferredoxin) are less than -4.78 kJ per mol H_2 and -24.92 kJ per mol formate in Hakuba, and -10.73 and -22.03 in The Cedars respectively (see Supplementary Results). However, serpentinite-hosted systems impose a unique challenge to homoacetogenesis – a key substrate, CO_2 , is at extremely low concentrations due to the high alkalinity. We estimate that the aqueous CO_2 concentration is below 0.0004 nM in Hakuba (pH 10.7 and <0.1 μM TIC) and 0.003 nM in The Cedars (pH 11.9 and 35 μM TIC)^{26,28}. In Hakuba, H_2/CO_2 -driven acetogenesis (ΔG of -1.68 kJ per mol acetate) cannot support microbial energy generation ($\Delta G \leq -20$ kJ per mol is necessary³⁰; Fig. S1). Moreover, in both Hakuba and The Cedars, one of the first steps in CO_2 -reducing homoacetogenesis, reduction of CO_2 to formate, is unfavorable based on the thermodynamics presented above ($\Delta G > +24.92$ or $+22.02$ kJ per mol formate). Thus, catabolic reduction of CO_2 to acetate is thermodynamically challenging *in situ* and may only run if investing ATP (*e.g.*,

Calvin-Benson-Bassham cycle [-6 ATP; ΔG of -361.68 kJ per mol acetate in Hakuba] or reductive tricarboxylic acid [-1 ATP; -61.68 kJ per mol]). Under CO_2 limitation, autotrophs are known to accelerate CO_2 uptake through HCO_3^- dehydration to CO_2 (carbonic anhydrase) or carbonate mineral dissolution, but both only modify kinetics and are not effective in changing the maximum CO_2 concentration (determined by equilibrium with carbonate species). In addition, in Hakuba, the CO_3^{2-} concentration is too low (87.5 nM CO_3^{2-}) for carbonate mineral precipitation (*e.g.*, $[\text{CO}_3^{2-}]$ must exceed 38.5 μM given K_s of 5×10^{-9} for CaCO_3 and $[\text{Ca}^{2+}]$ of 0.13 mM).

Based on thermodynamic calculations, the energy obtainable from H_2/CO_2 -driven homoacetogenesis is too small to support life in many Hadean analogues (Fig. S2a), yet acetate is detected in some of these ecosystems (Fig. S2b). Thus, CO_2 -independent electron-disposing metabolism may have been necessary for early life to gain energy from H_2 in the hyperalkaline fluids of hydrothermal systems. Here, we explore the metabolic capacities of extant organisms living in CO_2 -limited Hadean analogues to gain insight into potential metabolic strategies that LUCA or early post-LUCA organisms beginning diversification towards *Bacteria* and *Archaea* may have utilized to thrive in alkaline serpentinite-active ecosystems.

Diverse putative H_2 - and formate-utilizing organisms

Through metagenomic exploration of the two serpentinite-hosted systems (Table S1), we discover a plethora of phylogenetically novel organisms encoding genes for H_2 and formate metabolism (19 bins with 73.2-94.8% completeness and 0.0-8.1% contamination [86.1% and 3.8% on average respectively]; available under NCBI BioProject PRJNA453100) despite challenges in acquisition of genomic DNA (15.7 and 18.9 ng of DNA from 233 and 720 L of filtered Hakuba Happo spring water, respectively; RNA was below the detection limit). We

find metagenome-assembled genomes (MAGs) affiliated with lineages of *Firmicutes* (e.g., *Syntrophomonadaceae* and uncultured family SRB2), *Actinobacteria*, and candidate division NPL-UPA2³¹ (Fig. S3). We also recovered MAGs for a novel candidate phylum, herein referred to as “*Ca. Lithoacetigenota*”, that inhabits both Hakuba and The Cedars and, to our knowledge, no other ecosystems (Fig. 1a and S3). These genomes encode enzymes for oxidizing H₂ and formate (i.e., hydrogenases and formate dehydrogenases³²⁻³⁹; see Supplementary Results), suggesting that organisms *in situ* can employ H₂ and formate as electron donors.

“*Ca. Lithoacetigenota*” has unique site-adapted metabolism

Inspection of the serpentinite-hosted environment-exclusive phylum “*Ca. Lithoacetigenota*” reveals specialization to H₂-driven lithotrophy potentially suitable for the low-CO₂ *in situ* conditions (Fig. 1b). We discover that The Cedars-inhabiting population (e.g., MAG BS5B28, 94.8% completeness and 2.9% contamination) harbors genes for H₂ oxidation ([NiFe] hydrogenase Hox) and a nearly complete Wood-Ljungdahl pathway and an oxidoreductase often associated with acetogenesis – NADH:ferredoxin oxidoreductase Rnf^{40,41} (Table S2 and S3). One critical enzyme, the formate dehydrogenase, is missing from all three “*Ca. Lithoacetigenota*” MAGs from The Cedars (and unbinned contigs), indicating that these bacteria can neither perform H₂/CO₂-driven nor formate-oxidizing acetogenesis (Fig. 1b). However, even without the formate dehydrogenase, the genes present can form a coherent pathway that uses formate rather than CO₂ as a starting point for the “methyl branch” of the Wood-Ljungdahl pathway (i.e., formate serves as an electron acceptor; Fig. 1b). This is a simple yet potentially effective strategy for performing homoacetogenesis while circumventing the unfavorable reduction of CO₂ to formate. Coupling H₂ oxidation with this formate-reducing pathway is thermodynamically viable as it halves the usage of CO₂ ($3\text{H}_2 + \text{Formate}^- + \text{CO}_2 =$

Acetate⁻ + 2H₂O; ΔG of -29.62 kJ per mol acetate) and, as a pathway, is simply an intersection between the conventional H₂/CO₂-driven and formate-disproportionating acetogenesis (Fig. 1b). Although use of formate as an electron acceptor for formate-oxidizing acetogenesis is quite common, no previous homoacetogens have been observed to couple H₂ oxidation with acetogenesis from formate, likely because CO₂ has a much higher availability than formate in most ecosystems.

Interestingly, the Hakuba-inhabiting “*Ca. Lithoacetigenota*” (HKB210 and HKB111) also encodes Hox for H₂ oxidation but lacks genes for homoacetogenesis (no homologs closely related to The Cedars population genes were detected even in unbinned metagenomic contigs). Perhaps this population forgoes the above H₂/formate-driven homoacetogenesis because the estimated energy yield of the net reaction *in situ* (ΔG of -19.94 kJ per mol acetate) is extremely close to the thermodynamic threshold of microbial catabolism (slightly above -20 kJ per mol) and, depending on the actual threshold for “*Ca. Lithoacetigenota*” and/or even slight changes in the surrounding conditions (*e.g.*, ΔG increases by 1 kJ per mol if H₂ decreases by 20 μM decreases in Hakuba), the metabolism may be unable to recover energy. Through searching the physicochemical environment for alternative exogenous electron acceptors and MAGs for electron-disposing pathways, we detected a low concentration of glycine *in situ* (5.4 ± 1.6 nM; Table S4) and found genes specific to catabolic glycine reduction (see next paragraph). We suspect that some portion of this glycine is likely geochemically generated *in situ*, given that (a) glycine is often detected as the most abundant amino acid produced by both natural and laboratory-based serpentinization (*e.g.*, H₂ + Formate = Formaldehyde ⇒ Formaldehyde + NH₃ = Glycine)^{1,7,42-47} and (b) no other amino acid was consistently detectable (if glycine was cell-derived, other amino acids ought to also be consistently detected).

For utilization of the putatively abiotic glycine, the Hakuba “*Ca. Lithoacetigenota*” encode glycine reductases (Grd; Fig. 2 and S4; Tables S2 and S3) – a unidirectional selenoprotein for catabolic glycine reduction^{48,49}. Based on the genes available, this population likely specializes in coupling H₂ oxidation and glycine reduction (Fig. 1b). Firstly, the genomes encode NADP-linked thioredoxin reductases ($\text{NADPH} + \text{Thioredoxin}_{\text{ox}} \rightarrow \text{NADP}^+ + \text{Thioredoxin}_{\text{red}}$) that can bridge electron transfer from H₂ oxidation ($\text{H}_2 + \text{NADP} \rightarrow \text{NADPH} + \text{H}^+$) to glycine reduction ($\text{Glycine}^- + \text{Thioredoxin}_{\text{red}} \rightarrow \text{Acetyl-P}_i + \text{NH}_3 + \text{Thioredoxin}_{\text{ox}}$). Secondly, though glycine reduction is typically coupled with amino acid oxidation (*i.e.*, Stickland reaction in *Firmicutes* and *Synergistetes*^{48,50}), similar metabolic couplings have been reported for some organisms (*i.e.*, formate-oxidizing glycine reduction [via Grd]⁵¹ and H₂-oxidizing trimethylglycine reduction [via Grd-related betaine reductase]⁵²). Thirdly, Grd is a rare catabolic enzyme, so far found in organisms that specialize in amino acid (or peptide) catabolism, many of which are reported to use glycine for the Stickland reaction (*e.g.*, *Peptoclostridium* of *Firmicutes* and *Aminobacterium* of *Synergistetes*⁵³). Lastly, the population lacks any discernable fermentative and respiratory electron disposal pathways and oxidative organotrophy (Table S2 and S3). Moreover, reflecting the lack of other catabolic pathways, the Hakuba “*Ca. Lithoacetigenota*” MAGs display extensive genome streamlining, comparable to that of *Aurantimicrobium*^{54,55}, “*Ca. Pelagibacter*”⁵⁶, and *Rhodoluna*⁵⁷ (Fig. S6). Thermodynamic calculations show that H₂-oxidizing glycine reduction is thermodynamically favorable *in situ* (ΔG° of -70.37 kJ per mol glycine [ΔG of -85.84 in Hakuba]; Fig. S1). Further, based on the pathway identified, this metabolism is >10 times more efficient in recovering energy from H₂ (1 mol ATP per mol H₂) than acetogenesis utilizing H₂/CO₂ (0.075 mol ATP per mol H₂ based on the pathway *Acetobacterium woodii* utilizes) or H₂/formate (0.075 mol ATP per mol H₂, assuming no energy recovery associated with the formate dehydrogenase). We also detect glycine reductases in The Cedars “*Ca. Lithoacetigenota*”, indicating that it may also perform this metabolism (ΔG

of -76.87 in The Cedars, assuming $201 \mu\text{M H}_2$). Thus, we propose glycine as an overlooked thermodynamically and energetically favorable electron acceptor for H_2 oxidation in serpentinite-hosted systems.

Given the phylogenetic and metabolic uniqueness of these populations, we report provisional taxonomic assignment to “*Ca. Lithoacetigenota*” phylum. nov., “*Ca. Thermoacetigena glycinireducens*” gen. nov., sp. nov. (HKB111 and HKB210), and “*Ca. Psychroacetigena formicireducens*” gen. nov., sp. nov. (BS525, BS5B28, and GPS1B18) (see Supplementary Results). Based on a concatenated ribosomal protein tree, this serpentinite-hosted ecosystem-associated candidate phylum is closely related to a deep-branching group of bacterial phyla (Fig. 1a) (e.g., *Caldiserica* [Caldisericota in GTDB phylogeny], *Coprothermobacterota*, and *Dictyoglomi* [Dictyoglomota]⁵⁸), many of which have extremely limited phylogenetic diversity (only 8, 1, and 2 genus-level lineages identified respectively via cultivation and metagenomics [based on GTDB release 95]) and ecological distribution on modern Earth. Comparative genomics shows that “*Ca. Lithoacetigenota*” shares 623 core functions (based on Bacteria-level COGs/NOGs predicted by eggno-mapper shared by the two highest quality Hakuba and The Cedars MAGs HKB210 and BS5B28; Fig. 1c). When compared with the core functions of two closest related phyla (*Caldiserica* and *Coprothermobacterota*), 176 functions were unique to “*Ca. Lithoacetigenota*”, including those for NiFe hydrogenases (and their maturation proteins), selenocysteine utilization (essential for Grd), and sodium:proton antiporter for alkaliphily. With *Coprothermobacterota*, 232 functions were shared, including Grd, thioredoxin oxidoreductase (essential for electron transfer to Grd), and additional proteins for NiFe hydrogenases and selenocysteine utilization, pointing towards importance of H_2 metabolism and glycine reduction for these closely related phyla. More importantly, among bacterial phyla in the deep-branching group, “*Ca. Lithoacetigenota*”

represents the first lineage inhabiting hyperalkaliphilic serpentinite-hosted Hadean analogue ecosystems, suggesting that these organisms may be valuable extant windows into potential physiologies LUCA and/or early post-LUCA organisms may have taken (albeit with 4 billion years of evolution in between; see discussion regarding Grd below).

Widespread glycine reduction in serpentinite-hosted systems

Uncultured members of *Chloroflexi* (Chloroflexota) class *Dehalococcoidia* inhabiting The Cedars and *Firmicutes* (Firmicutes_D) class SRB2 in Hakuba and The Cedars also possess glycine reductases (Table S3). In addition, these populations encode hydrogenases and formate dehydrogenases, suggesting that they may also link H₂ and formate metabolism to glycine reduction. Moreover, we further discover closely related glycine reductases in other studied serpentinite-hosted systems (47~94% amino acid similarity in Tablelands, Voltri Massif, and Coast Range Ophiolite)^{18,19,24}. Phylogenetic analysis of the glycine-binding "protein B" subunit GrdB reveals close evolutionary relationships between glycine reductases from distant/remote sites (Fig. 2, S4, and S5). (Tablelands spring glycine reductase sequences were not included in the analysis as they were only detected in the unassembled metagenomic reads; 4460690.3; 69.7~82.2% similarity to Hakuba SRB2). Overall, "*Ca. Lithoacetigenota*", *Dehalococcoidia*, and SRB2 glycine reductases are all detected in at least two out of the seven metagenomically investigated systems despite the diverse environmental conditions (*e.g.*, temperature), highlighting the potential importance of glycine metabolism by these clades in serpentinite-hosted systems. We suspect that glycine reduction may be a valuable catabolic strategy as the pathway requires few genes/proteins (a hydrogenase, Grd, acetate kinase, and thioredoxin oxidoreductase) and conveniently provides acetate, ammonia, and ATP as basic forms of carbon, nitrogen, and energy.

Comparison of the topology of the GrdB (Fig. 2, S4, and S5) and ribosomal protein (Fig. 1a) trees hints towards vertical transfer of GrdB dating back to one of the deepest divisions in the bacterial domain. Given the ecosystem specificity and deep phylogeny of the newly discovered phylum and Grd, catabolic glycine reduction may be a relatively ancient metabolism viable in serpentinization-related habitats, rapidly lost due to its low utility in modern ecosystems (*e.g.*, no severe nutrient/electron acceptor limitation and no excess glycine via abiotic generation), but repurposed by some anaerobes for the fermentative Stickland reaction in organic-rich ecosystems (*e.g.*, faeces and biodigesters) where excess amino acids are available but no favorable electron acceptors are accessible (*i.e.*, anaerobic). We identified the first archaeal GRD (in Miscellaneous Crenarchaeota Group [MCG] or *Ca.* Bathyarchaeota member BA-1; Fig. 2, S4, and S5), but phylogenetic analysis shows that the gene was gained through horizontal transfer and whether this gene truly belongs to this clade remains to be verified given that the source is a metagenome-assembled genome. Thus, the currently available data suggests that Grd (and catabolic glycine reduction) is an ancient bacterial innovation (*i.e.*, originated in *Bacteria*) developed post-LUCA and largely exclusive to the domain *Bacteria*. Further investigation of Hadean analogues is necessary to verify the history of Grd (still unclear based on Bayesian inference of GrdB phylogeny; Fig. S5), uncover other basal Grd and explore whether glycine reduction (or Grd family proteins) can be dated back further (more ancient *Bacteria* or LUCA; *i.e.*, necessary to find more basal bacterial and/or archaeal homologs).

Other characteristics of putative indigenous homoacetogens

In contrast with members of “*Ca.* Lithoacetigenota”, several other putative homoacetogenic populations encode the complete Wood-Ljungdahl pathway (Table S2 and S3), indicating that other forms of acetogenesis may also be viable *in situ*. One putative homoacetogen in The

Cedars, NPL-UPA2, lacks hydrogenases but encodes formate dehydrogenases. Although the NPL-UPA2 population cannot perform H₂/formate-driven acetogenesis, it may couple formate oxidation with formate-reducing acetogenesis – another thermodynamically viable metabolism (ΔG of –50.90 kJ per mol acetate in The Cedars; Fig. 1b). The pathway uses CO₂ as a substrate but has lower CO₂ consumption compared to H₂/CO₂ homoacetogenesis and can produce intracellular CO₂ from formate. In Hakuba, an *Actinobacteria* population affiliated with the uncultured class UBA1414 (MAG HKB206) encodes hydrogenases and a complete Wood-Ljungdahl pathway (Table S3) and, thus, may be capable of H₂/formate or the above formate-disproportionating acetogenesis (Fig. 1b). Indeed, the UBA1414 population was enriched in Hakuba-derived cultures aiming to enrich acetogens using the H₂ generated by the metallic iron–water reaction ⁵⁹ (Fig. S7). Many populations encoding a complete Wood-Ljungdahl pathway possess monomeric CO dehydrogenases (CooS unassociated with CODH/ACS subunits; NPL-UPA2, *Actinobacteria*, *Syntrophomonadaceae* [Hakuba and The Cedars], and *Dehalococcoidia* [The Cedars]; Table S2). Although CO is below the detection limit in Hakuba (personal communication with permission from Dr. Konomi Suda), another study shows that CO metabolism takes place in an actively serpentinizing system with no detectable CO ⁶⁰. Given that CO is a known product of serpentinization ^{24,60}, it may be an important substrate for thermodynamically favorable acetogenesis *in situ*. However, further investigation is necessary to verify this.

Another interesting adaptation observed for all putative homoacetogens detected in Hakuba and The Cedars was possession of an unusual CODH/ACS complex. Although *Bacteria* and *Archaea* are known to encode structurally distinct forms of CODH/ACS (designated as Acs and Cdh respectively for this study), all studied Hakuba/The Cedars putative homoacetogens encode genes for a hybrid CODH/ACS that integrate archaeal subunits for the CO

dehydrogenase (AcsA replaced with CdhAB) and acetyl-CoA synthase (AcsB replaced with CdhC) and bacterial subunits for the corrinoid protein and methyltransferase components (AcsCDE) (Table S2). The *Firmicutes* lineages also additionally encode the conventional bacterial AcsABCDE. Given that all of the identified putative homoacetogens encode this peculiar hybrid complex, we suspect that such CODH/ACS's may have features that adapted to the high-pH low-CO₂ conditions (*e.g.*, high affinity for CO₂ and/or CO). In agreement, a similar hybrid CODH/ACS has also been found in "*Ca. Desulforudis audaxviator*" inhabiting an alkaline (pH 9.3) deep subsurface environment with a low CO₂ concentration (below detection limit ^{61,62}).

Implications for primordial biology

Based on one of the most plausible theories of the origin of life, LUCA inhabited alkaline geochemically active sites ⁶³, like serpentinite-hosted systems widespread across Earth during the Hadean to early Archean. Although reconstruction of the LUCA's physiology is challenging ⁶⁴, energy acquisition through H₂-oxidation-driven CO₂ reduction to acetyl-CoA (*i.e.*, H₂/CO₂ homoacetogenesis) is theorized to be a core feature of primordial metabolism ¹¹. However, the free energy yield of such homoacetogenesis decreases with CO₂ limitation and increasing temperature (Fig. S2a), suggesting that LUCA and/or early post-LUCA organisms may have encountered CO₂-related thermodynamic challenges. Another study also points out that CO₂ speciation towards a less bioavailable form (carbonate) under hyperalkaline pH may have been a potential problem for primordial life ⁶⁵. At the interface of the alkaline fluids and seawater where life is thought to have originated ^{7,15-17}, H₂ and CO₂ from the respective fluids could have come in contact and allow CO₂-dependent H₂ utilization; however, CO₂-dependent metabolism would have become thermodynamically, and perhaps kinetically ⁶⁶, challenging as early post-LUCA organisms ventured away from the interface and deeper into serpentinite-

hosted systems. Through the investigation of CO₂-poor Hadean analogues, we discover a novel Hadean analogue-exclusive alkaliphilic phylum that belongs to a deep-branching group of bacterial phyla and possesses unconventional thermodynamically favorable less CO₂-dependent H₂-oxidizing metabolic strategies (*e.g.*, coupled with formate or glycine reduction) that may be compelling candidates for early H₂-dependent lithotrophy.

The thermodynamic and energetic favorability of catabolic H₂-oxidizing glycine reduction mediated by GRD makes it a competitive metabolism for primordial *Bacteria*. Catabolic H₂-utilizing glycine reduction also draws remarkable parallels with homoacetogenesis and ancestral archaeal metabolism. Like homoacetogenesis, glycine reduction has a rare feature thought to be essential for primordial metabolism – the ability to reductively synthesize acyl-phosphate for substrate-level phosphorylation in the cytosol⁶⁷⁻⁷⁰ – indicating high utility for ancient organisms. Glycine reduction also shares features with a metabolism recently proposed to be ancestral in *Archaea* – H₂-oxidizing methyl-reducing methanogenesis^{71,72}. Both are (nearly) domain-exclusive, potentially primordial, CO₂-independent, and energetically more efficient than H₂/CO₂-driven acetogenesis (in terms of energy recovered per mol H₂). Moreover, both utilize reduced serpentinization-derived carbon compounds, a thiol/disulfide electron carrier as an electron donor, and NiFe hydrogenases as an upstream electron source (Fig. 3). It is tempting to speculate that early post-LUCA *Bacteria* and *Archaea* both integrated geogenic thermodynamically favorable electron acceptors to adapt to thrive under the extreme conditions and interestingly occupied non-competing niches by using different acceptors (glycine and methyl compounds respectively) and energy acquisition strategies (substrate-level phosphorylation and proton motive force respectively), whereby allowing the two to co-exist with little competition. If true, we may begin to speculate that, as *Bacteria* and *Archaea* weaned themselves from serpentinite-hosted systems and explored the non-alkaline habitats of Earth,

GRD and MCR were connected to the methyl branch of the Wood-Ljungdahl pathway and repurposed to facilitate CO₂ fixation⁷³ and CO₂-reducing methanogenesis respectively, both CO₂-dependent pathways relevant today. However, further investigations are necessary to verify because, unlike MCR, information on GRD is limited: glycine-mediated autotrophy was only discovered recently, organisms still using GRD for catabolic glycine reduction are low in diversity on modern Earth given the abundant availability of other electron acceptors, and the geobiology of glycine-generating serpentinite-hosted systems have not been studied extensively.

As was the case for MCR, metagenomic exploration of the uncharted rare biosphere provided key data for genes and core pathways that provide a glimpse into how primordial life or early bacteria may have proliferated (though we must still take caution in interpretation given the uncertainty associated with metagenomics). The thermodynamic (i.e., more favorable than the CO₂-reducing counterpart) and evolutionary (i.e., ancestral) parallels between archaeal methyl-reducing and bacterial glycine-reducing H₂-dependent lithotrophy warrant investigation of whether these metabolisms did indeed play a role in the formation or fixation of the two domains. Further studies on the evolutionary origin, antiquity, and history of glycine reduction will shed light on primordial way of life as well as bacterial/prokaryotic diversification.

Author contribution

MKN and RN designed and performed metagenomic, phylogenetic, and thermodynamic analyses, interpreted the data, and wrote the manuscript. ST performed sampling, DNA extraction, and cultivation. HM, AT, and KK performed metagenomic sequencing and metagenomic analysis. AI performed chemical measurements. SS supported chemical data

collection. HT and YK designed the project, supervised sampling and cultivation, and supported manuscript preparation.

Acknowledgements

The authors greatly appreciate Mr. Sejima and others of Happo-one Development Co., Ltd. for their kind permission and cooperation in conducting field studies at Hakuba Happo hot springs. This work was supported by the JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas, “Hadean Bioscience” project (no. JP26106006). This work was partly funded by the JSPS KAKENHI Grant-in-aid for Scientific Research nos. JP26710012 and JP18H02426 to H. Tamaki, JP15H05620 to R. Nakai, and JP18H03367 to M.K. Nobu.

Competing interests

The authors declare no competing interests.

383 References

- 384 1 Schrenk, M. O., Brazelton, W. J. & Lang, S. Q. Serpentinization, Carbon, and Deep
385 Life. *Reviews in Mineralogy and Geochemistry* **75**, 575-606,
386 doi:10.2138/rmg.2013.75.18 (2013).
- 387 2 Komiya, T., Maruyama, S., Hirata, T., Yurimoto, H. & Nohda, S. Geochemistry of
388 the oldest MORB and OIB in the Isua Supracrustal Belt, southern West Greenland:
389 Implications for the composition and temperature of early Archean upper mantle.
390 *Island Arc* **13**, 47-72, doi:10.1111/j.1440-1738.2003.00416.x (2004).
- 391 3 Sleep, N. H., Bird, D. K. & Pope, E. C. Serpentinite and the dawn of life.
392 *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 2857
393 (2011).
- 394 4 Martin, W. F. & Sousa, F. L. Early Microbial Evolution: The Age of Anaerobes. *Cold*
395 *Spring Harbor Perspectives in Biology* **8**, doi:10.1101/cshperspect.a018127 (2016).
- 396 5 Lane, N. & Martin, William F. The Origin of Membrane Bioenergetics. *Cell* **151**,
397 1406-1416, doi:<https://doi.org/10.1016/j.cell.2012.11.050> (2012).
- 398 6 Battistuzzi, F. U., Feijao, A. & Hedges, S. B. A genomic timescale of prokaryote
399 evolution: insights into the origin of methanogenesis, phototrophy, and the
400 colonization of land. *BMC Evolutionary Biology* **4**, 44, doi:10.1186/1471-2148-4-44
401 (2004).
- 402 7 Shibuya, T. *et al.* Hydrogen-rich hydrothermal environments in the Hadean ocean
403 inferred from serpentinization of komatiites at 300 °C and 500 bar. *Progress in Earth*
404 *and Planetary Science* **2**, 46, doi:10.1186/s40645-015-0076-z (2015).
- 405 8 Camprubi, E., Jordan, S. F., Vasiliadou, R. & Lane, N. Iron catalysis at the origin of
406 life. *IUBMB life* **69**, 373-381, doi:10.1002/iub.1632 (2017).
- 407 9 Decker, K., Jungermann, K. & Thauer, R. K. Energy Production in Anaerobic
408 Organisms. *Angewandte Chemie International Edition in English* **9**, 138-158,
409 doi:10.1002/anie.197001381 (1970).
- 410 10 Martin, W. F., Weiss, M. C., Neukirchen, S., Nelson-Sathi, S. & Sousa, F. L.
411 Physiology, phylogeny, and LUCA. *Microbial Cell* **3**, 582-587,
412 doi:10.15698/mic2016.12.545 (2016).
- 413 11 Weiss, M. C. *et al.* The physiology and habitat of the last universal common ancestor.
414 *Nature Microbiology* **1**, 16116, doi:10.1038/nmicrobiol.2016.116
415 <https://www.nature.com/articles/nmicrobiol2016116#supplementary-information> (2016).
- 416 12 Adam, P. S., Borrel, G. & Gribaldo, S. Evolutionary history of carbon monoxide
417 dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes.
418 *Proceedings of the National Academy of Sciences* **115**, E1166,
419 doi:10.1073/pnas.1716667115 (2018).
- 420 13 Preiner, M. *et al.* A hydrogen-dependent geochemical analogue of primordial carbon
421 and energy metabolism. *Nat Ecol Evol* **4**, 534-542, doi:10.1038/s41559-020-1125-6
422 (2020).
- 423 14 Martin, W. F. Older Than Genes: The Acetyl CoA Pathway and Origins. *Frontiers in*
424 *microbiology* **11**, 817-817, doi:10.3389/fmicb.2020.00817 (2020).
- 425 15 Lane, N. Proton gradients at the origin of life. *Bioessays* **39**, 1600217,
426 doi:<https://doi.org/10.1002/bies.201600217> (2017).
- 427 16 Shibuya, T., Russell, M. J. & Takai, K. Free energy distribution and hydrothermal
428 mineral precipitation in Hadean submarine alkaline vent systems: Importance of iron
429 redox reactions under anoxic conditions. *Geochim Cosmochim Acta* **175**, 1-19,
430 doi:<https://doi.org/10.1016/j.gca.2015.11.021> (2016).

- 17 Sojo, V., Herschy, B., Whicher, A., Camprubí, E. & Lane, N. The Origin of Life in Alkaline Hydrothermal Vents. *Astrobiology* **16**, 181-197, doi:10.1089/ast.2015.1406 (2016).
- 18 Brazelton, W. J. *et al.* Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy. *PeerJ* **5**, e2945, doi:10.7717/peerj.2945 (2017).
- 19 Brazelton, W. J., Nelson, B. & Schrenk, M. O. Metagenomic Evidence for H₂ Oxidation and H₂ Production by Serpentinite-Hosted Subsurface Microbial Communities. *Frontiers in Microbiology* **2**, 268, doi:10.3389/fmicb.2011.00268 (2011).
- 20 Suzuki, S. *et al.* Unusual metabolic diversity of hyperalkaliphilic microbial communities associated with subterranean serpentinitization at The Cedars. *Isme J* **11**, 2584-2598, doi:10.1038/ismej.2017.111 (2017).
- 21 Kelley, D. S. *et al.* A Serpentinite-Hosted Ecosystem: The Lost City Hydrothermal Field. *Science* **307**, 1428-1434, doi:10.1126/science.1102556 (2005).
- 22 Tiago, I. & Verissimo, A. Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinitization. *Environ Microbiol* **15**, 1687-1706, doi:10.1111/1462-2920.12034 (2013).
- 23 Crespo-Medina, M. *et al.* Methane Dynamics in a Tropical Serpentinizing Environment: The Santa Elena Ophiolite, Costa Rica. *Frontiers in Microbiology* **8**, 916, doi:10.3389/fmicb.2017.00916 (2017).
- 24 Twing, K. I. *et al.* Serpentinization-Influenced Groundwater Harbors Extremely Low Diversity Microbial Communities Adapted to High pH. *Frontiers in Microbiology* **8**, 308, doi:10.3389/fmicb.2017.00308 (2017).
- 25 Neubeck, A. *et al.* Microbial Community Structure in a Serpentine-Hosted Abiotic Gas Seepage at the Chimaera Ophiolite, Turkey. *Appl Environ Microb* **83**, doi:10.1128/aem.03430-16 (2017).
- 26 Morrill, P. L. *et al.* Geochemistry and geobiology of a present-day serpentinitization site in California: The Cedars. *Geochim Cosmochim Ac* **109**, 222-240, doi:<https://doi.org/10.1016/j.gca.2013.01.043> (2013).
- 27 Suda, K., Gilbert, A., Yamada, K., Yoshida, N. & Ueno, Y. Compound- and position-specific carbon isotopic signatures of abiogenic hydrocarbons from on-land serpentinite-hosted Hakuba Happo hot spring in Japan. *Geochim Cosmochim Ac* **206**, 201-215, doi:<https://doi.org/10.1016/j.gca.2017.03.008> (2017).
- 28 Suda, K. *et al.* Origin of methane in serpentinite-hosted hydrothermal systems: The CH₄-H₂-H₂O hydrogen isotope systematics of the Hakuba Happo hot spring. *Earth and Planetary Science Letters* **386**, 112-125, doi:<https://doi.org/10.1016/j.epsl.2013.11.001> (2014).
- 29 Suda, K. *Origins of Hydrocarbons in On-land Serpentinization Fields and Insights into Hadean Hydrothermal Systems: Systematic Study using Stable Isotopes* Ph.D. thesis, Tokyo Institute of Technology, (2016).
- 30 Schink, B. Energetics of syntrophic cooperation in methanogenic degradation. *Microbiology and Molecular Biology Reviews* **61**, 262-280 (1997).
- 31 Suzuki, S., Nealson, K. H. & Ishii, S. i. Genomic and in-situ Transcriptomic Characterization of the Candidate Phylum NPL-UPL2 From Highly Alkaline Highly Reducing Serpentinized Groundwater. **9**, doi:10.3389/fmicb.2018.03141 (2018).
- 32 Morandi, P., Valzasina, B., Colombo, C., Curti, B. & Vanoni, M. A. Glutamate synthase: identification of the NADPH-binding site by site-directed mutagenesis. *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 (CO₂-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 Acta* **53**, 2937-2945,
515 doi:[https://doi.org/10.1016/0016-7037\(89\)90170-1](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).

- 530 50 Nisman, B. The Stickland reaction. *Bacteriological Reviews* **18**, 16-42 (1954).
- 531 51 Hormann, K. & Andreesen, J. R. Reductive cleavage of sarcosine and betaine by
- 532 Eubacterium acidaminophilum via enzyme systems different from glycine reductase.
- 533 *Archives of Microbiology* **153**, 50-59, doi:10.1007/bf00277541 (1989).
- 534 52 Moune, S. *et al.* *Haloanaerobacter salinaris* sp. nov., a novel halophilic
- 535 fermentative bacterium that reduces glycine-betaine to trimethylamine with hydrogen
- 536 or serine as electron donors; emendation of the genus *Haloanaerobacter*. *Int J Syst*
- 537 *Bacteriol* **49 Pt 1**, 103-112, doi:10.1099/00207713-49-1-103 (1999).
- 538 53 Hamdi, O. *et al.* *Aminobacterium thunnarium* sp. nov., a mesophilic, amino acid-
- 539 degrading bacterium isolated from an anaerobic sludge digester, pertaining to the
- 540 phylum Synergistetes. *Int J Syst Evol Microbiol* **65**, 609-614,
- 541 doi:10.1099/ijs.0.068965-0 (2015).
- 542 54 Nakai, R., Baba, T., Niki, H., Nishijima, M. & Naganuma, T. *Aurantimicrobium*
- 543 *minutum* gen. nov., sp. nov., a novel ultramicrobacterium of the family
- 544 Microbacteriaceae, isolated from river water. *Int J Syst Evol Microbiol* **65**, 4072-
- 545 4079, doi:10.1099/ijsem.0.000541 (2015).
- 546 55 Nakai, R. *et al.* Complete Genome Sequence of *Aurantimicrobium minutum* Type
- 547 Strain KNCT, a Planktonic Ultramicrobacterium Isolated from River Water. *Genome*
- 548 *announcements* **4**, doi:10.1128/genomeA.00616-16 (2016).
- 549 56 Giovannoni, S. J. *et al.* Genome streamlining in a cosmopolitan oceanic bacterium.
- 550 *Science* **309**, 1242-1245, doi:10.1126/science.1114057 (2005).
- 551 57 Hahn, M. W., Schmidt, J., Taipale, S. J., Doolittle, W. F. & Koll, U. *Rhodoluna*
- 552 *lacidicola* gen. nov., sp. nov., a planktonic freshwater bacterium with stream-lined
- 553 genome. *Int J Syst Evol Microbiol* **64**, 3254-3263, doi:10.1099/ijs.0.065292-0 (2014).
- 554 58 Betts, H. C. *et al.* Integrated genomic and fossil evidence illuminates life's early
- 555 evolution and eukaryote origin. *Nature Ecology & Evolution* **2**, 1556-1562,
- 556 doi:10.1038/s41559-018-0644-x (2018).
- 557 59 Kato, S., Yumoto, I. & Kamagata, Y. Isolation of Acetogenic Bacteria That Induce
- 558 Biocorrosion by Utilizing Metallic Iron as the Sole Electron Donor. *Appl Environ*
- 559 *Microb* **81**, 67-73, doi:10.1128/aem.02767-14 (2015).
- 560 60 Morrill, P. L. *et al.* Investigations of potential microbial methanogenic and carbon
- 561 monoxide utilization pathways in ultra-basic reducing springs associated with present-
- 562 day continental serpentinization: the Tablelands, NL, CAN. *Frontiers in*
- 563 *Microbiology* **5**, doi:10.3389/fmicb.2014.00613 (2014).
- 564 61 Chivian, D. *et al.* Environmental genomics reveals a single-species ecosystem deep
- 565 within Earth. *Science* **322**, 275-278, doi:10.1126/science.1155495 (2008).
- 566 62 Lin, L.-H. *et al.* Long-Term Sustainability of a High-Energy, Low-Diversity Crustal
- 567 Biome. *Science* **314**, 479 (2006).
- 568 63 Martin, W. & Russell, M. J. On the origin of biochemistry at an alkaline hydrothermal
- 569 vent. *Philosophical Transactions of the Royal Society B: Biological Sciences* **362**,
- 570 1887 (2007).
- 571 64 Berkemer, S. J. & McGlynn, S. E. A New Analysis of Archaea-Bacteria Domain
- 572 Separation: Variable Phylogenetic Distance and the Tempo of Early Evolution.
- 573 *Molecular Biology and Evolution* **37**, 2332-2340, doi:10.1093/molbev/msaa089
- 574 (2020).
- 575 65 Boyd, E. S., Amenabar, M. J., Poudel, S. & Templeton, A. S. Bioenergetic constraints
- 576 on the origin of autotrophic metabolism. *Philosophical Transactions of the Royal*
- 577 *Society A: Mathematical, Physical and Engineering Sciences* **378**, 20190151,
- 578 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 $\delta^{13}\text{C}$ 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

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

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 CO₂-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 eggno-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 3. Proposed non-CO₂-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 H₂ (red star). The bacterial glycine-reducing H₂-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 H₂ oxidation (red) and reduction of thiol/disulfide electron carriers (orange) and forgo pterin-mediated CO₂ 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⁺, H₂O, 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; CH₃-X, methyl compounds; MT, methyltransferase; CH₃-S-CoM, methyl-coenzyme M; TR, thioredoxin reductase; Trx-(SH)₂, thioredoxin dithiol; Trx-S₂, thioredoxin disulfide; Grd, glycine reductase; acetyl-P_i, acetyl-phosphate; Ack, acetate kinase.

Methods

Sampling site and sample collection

The Hakuba Happo samples for geochemical and microbiological analysis were artificially pumped from a drilling well (700 m in depth), which was previously described and named Happo #3 (36° 42' N 137° 48' E ²⁸). For microbiological analysis, two spring water samples were taken at different time points, 233 L taken in July 2016 (labelled HKB701) and 720 L taken in October 2016 (labelled HKB702), respectively. To collect cells, samples were filtered through a 0.1-µm Omnipore™ membrane filter (Merck Millipore) using a 90 mm diameter stainless-steel filter holder (Merck Millipore) attached to FDA Viton® tubing (Masterflex) at a sampling site. After filtration, filters were immediately transferred to sterile tubes and frozen in a dry ice-ethanol bath, transported in dry ice, and stored at -80°C until DNA extraction. For NH₃ and amino acid analysis, water samples were collected in October 2017, filtered as described above, transferred to dry-heat-sterilized nitrogen-purged 100-ml glass vials, and stored at 4°C.

Geochemical analysis

The water temperature of hot spring water was measured using a thermometer (CT-430WP, Custom Ltd.) at a site. The pH, oxidation reduction potential (ORP), electrical conductivity (EC), and dissolved oxygen (DO) level were determined with portable devices, including a pH meter (D-23, Horiba), an ORP meter (RM-30P, TOA-DKK), an EC meter (CM-31P, TOA-DKK), and a DO meter (DO-31P, TOA-DKK), correspondingly. The ion concentrations of Na⁺, K⁺, Ca²⁺, and NO₃⁻ were determined using portable sensors (LAQUAtwin™ series, Horiba). The *in situ* NH₃ concentration was determined by measuring aqueous NH₄⁺ and gaseous NH₃ (purged with N₂ gas, gas dissolved into deionized water, and measured dissolved NH₄⁺) of a sample stored as described above using high-performance liquid chromatography

(HPLC; Prominence; Shimadzu), then adding the two together. For amino acid quantification, the sample was concentrated under a stream of nitrogen gas and then analyzed following Shimadzu protocol no. L323 (<https://www.ssi.shimadzu.com/products/literature/lc/L323.pdf>) using HPLC with minor modifications (fluorescence detector RF-20AXs; sodium hypochlorite solution was not added for detection of proline). The Cedars spring concentrations of formate and acetate were determined by Isotope-Ratio-Monitoring Liquid Chromatography Mass Spectrometry (IRM-LCMS); Thermo- Finnigan Delta Plus XP isotope-ratio mass spectrometer connected to LC IsoLink, as described by Heuer et al. ⁷⁵ and Ijiri et al. ⁷⁶.

Thermodynamic calculations

The Hakuba and Cedars calculations Gibbs free energy yield (ΔG) are based on ΔG°_f and ΔH°_f values at 298 K, respective *pH* (10.8 and 11.9), and adjustment to the *in situ* temperatures (48 and 17°C) through the Gibbs-Helmholtz equation ⁷⁷. The effect of pressure was approximated as described by Wang *et al* ⁷⁸. For both Hakuba and The Cedars calculations, the glycine concentration (5.4 nM) was based on measurements from Hakuba. Formate, acetate, and NH₃ concentrations were based on respective measurements from Hakuba (8 μM formate, 4 μM acetate, and 2.9 μM NH₃) and The Cedars (6.9 μM formate [average of 6.777 and 7.079 μM measured on September 2017], 69.3 μM acetate [average of 69.601 and 68.967 μM measured on September 2017], and 1 μM NH₃ [below detection limit]). For Hakuba, the H₂ concentration measured in Hakuba drilling well #3 (DNA source) was used (201 μM H₂). For The Cedars, the highest detected H₂ concentration in Hakuba was used (664 μM H₂ in drilling well #1).

Metagenome sequencing, assembly and binning

The filter was aseptically cut into 16 equal pieces using sterilized tweezers, and each piece was placed in the bead-beating tube (Lysing Matrix E tube; MP Biomedicals). After DNA

extraction following the bead-beating method described previously ⁷⁹, the 16 DNA samples were mixed and then stored at -80°C until used. Sequence libraries were prepared with Nextera XT DNA Library Preparation kit (Illumina) with a genomic DNA fragment size ranging from 200 to 2,000 bp. These libraries sequenced on HiSeq2500 sequencing platform (Illumina) with HiSeq Rapid SBS kit v2 (Illumina), generating paired-end reads up to 250 bp. The generated sequences were trimmed using Trimmomatic v0.33 ⁸⁰ with a quality cutoff of 30, sliding window of 6 bp, and minimum length cutoff of 78 bp. The trimmed sequences were assembled using SPAdes v3.10.1 ⁸¹ with the “-meta” option and k-mer values of 21, 31, 41, 53, 65, and 77. The assembled contigs were binned using MaxBin2.2.1 ^{82,83}. The completeness and contamination of each bin was checked using CheckM ⁸⁴. These bins were manually curated as described in our metagenomics study ⁴⁰. Genes were then annotated using Prokka v1.12 ⁸⁵ and eggno-mapper ⁸⁶. For interpretation and comparison of microbial metabolism, bin genomes were also constructed from public metagenomic data generated from The Cedars ²⁰ (trimmed with sliding window of 6, quality cutoff of 20, and minimum length of 68 bp through Trimmomatic v0.33, normalized using BBMap 36.99 (<https://jgi.doe.gov/data-and-tools/bbtools/>) with target and minimum coverages of 40 and 2, assembled using SPAdes v3.10.1 with the “-meta” option and k-mer values of 21, 33, 45, 55, 67, and binned through MaxBin2.2.1) and were then analyzed collectively.

Phylogenomic and phylogenetic analysis

For tree construction, sequences were aligned with MAFFT ⁸⁷ v7.453 (default parameters) and trimmed using trimAl ⁸⁸ v1.2rev59 (-gt 0.9). For ribosomal protein trees, a concatenated alignment of universally conserved ribosomal proteins ⁸⁹ was used. Protein sequences were retrieved by downloading the GTDB ⁹⁰ database and predicting protein sequences using Prokka ⁸⁵ 1.14 (--kingdom Bacteria/Archaea --rnammer --addgenes --mincontiglen 200). Maximum

likelihood trees were calculated using PhyML⁹¹ v3.3.20190909 using the LG⁹² model and 100 bootstrap replicates (-b 100 -d aa -m LG -v e). Bootstrap values were recalculated using BOOSTER⁹³. Sequence clustering was performing through CD-HIT⁹⁴ v4.8.1. For glycine reductase GrdB and sarcosine reductase GrdF, conserved motifs were predicted by first identifying fully conserved residues in the sequence cluster including the biochemically characterized *Peptoclostridium acidaminophilum* GrdF (see Supplementary Fig. S4; YxNx(6)GGE x(34,38) CGD x(27,35) GPxF[NF]AGRYG x(150,181) IHGGYDRx(6)[IP]x(4)Px D x(19,20) TTGTGTx(7)F x(12) [HILV]), then identifying fully conserved residues in the phylogenetic clusters (see Supplementary Fig. S4) that include GrdB from phyla known to perform the Stickland reaction (*Firmicutes*, *Spirochaetes*, and *Synergistetes*) subtracting any sequences clusters that contain the GrdF motif above (YxNx(6)GGE x(34,38) CGD x(27,35) GPxF[NF]AGRYG x(157,178) AHGGxD[QTAP] x(8) RV[IL]Px D x(19,20) TxGNxTxV). Bayesian trees were calculated using Phylobayes (MPI) v1.8 using the LG model with 4 gamma categories (-lg -ncat 1 -dgam 4) with four chains (chain length 10000; -x 1 10000) of which three chains that converged (maxdiff=0.1859, meandiff=0.014) were used for determination of a consensus tree with a burn-in of 1000 (nodes with posterior probability less than 0.8 were collapsed; bpcomp -c 0.8 -x 1000 1).

Data availability

The datasets generated during and/or analyzed during the current study are available in the National Center for Biotechnology Information (NCBI) under BioProject PRJNA453100 and BioSamples SAMN08978938-SAMN08978962.