# Sulfur reduction coupled with anaerobic ammonium oxidation drove prebiotic proto-anabolic networks

Peng Bao [1, 2,*], Guo-Xiang Li [1, 2, 3], Jun-Yi Zhao [1, 2], Kun Wu [1, 2], Juan Wang [1, 2], Xiao-Yu Jia[1, 2], Hui-En Zhang [4], Yu-Qin He [1, 2, 5], Hu Li [1], Ke-Qing Xiao [6]

[1] Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, P. R. China

[2] Ningbo Urban Environment Observation and Station, Chinese Academy of Sciences, Ningbo 315800, P. R. China

[3] Center for Applied Geosciences (ZAG), Eberhard Karls University Tuebingen, Sigwartstrasse 10, Tuebingen 72076, Germany

[4] College of Biological and Environmental Sciences, Zhejiang Wanli University, Ningbo 315100, P. R. China

[5] University of Chinese Academy of Sciences, Beijing 100049, P. R. China

[6] School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK

***Corresponding Author**

Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, P. R. China; E-mail: pbao@iue.ac.cn

## SUMMARY

The geochemical energy that drove the transition from geochemistry to biochemistry on early Earth remains unknown. Here, we show that the reduction of sulfurous species, such as thiosulfate, sulfite, elemental sulfur, and sulfate, coupled with anaerobic ammonium oxidation (Sammox), could have provided the primordial redox equivalents and proton potential for prebiotic proto-anabolic networks consisting of the reductive acetyl-CoA pathway combined with the incomplete reductive tricarboxylic acid (rTCA) cycle under mild hydrothermal conditions. Sammox-driven prebiotic proto-anabolic networks (SPPN) include $CO_2$ reduction, esterification, reductive amination, pyrrole synthesis, and peptides synthesis, in one geochemical setting. Iron-sulfur (FeS) minerals, as the proto-catalysts, enhanced the efficiency of SPPN. Thiols/thioesters were used as the energy currency in non-enzymatic phosphate-independent metabolism and accelerated SPPN. Peptides that consisted of 15 proteinogenic amino acids were the end products of SPPN with bicarbonate as the only source of carbon. Most peptides shared high similarity with the truly minimal protein content (TMPC) of the last universal common ancestor (LUCA). The peptides and/or proteinogenic amino acids might have endowed SPPN with autocatalysis and homochirality. Thus, Sammox drove the coupling transformation of carbon, hydrogen, oxygen, nitrogen, sulfur, and/or iron simultaneously in the far-from-equilibrium environment, thereby initiating the emergence of biochemistry. The existing Sammox microorganisms might belong to the phylum of *Planctomycetes*, and might be transitional forms between the three

42 domains of life.

43

44

## INTRODUCTION

The chemistry of life is based on reduction-oxidation (redox) reactions that comprise successive transfers of electrons and protons from the six major elements—carbon (C), hydrogen (H), oxygen (O), nitrogen (N), sulfur (S), and phosphorous (P) [1-3]. The $H_2/CO_2$ redox couple has been proposed as the first energy source to drive the reductive acetyl-CoA pathway, an ancient metabolic route [4-7]; however, the primordial energy source of this redox couple suffers from the difficulty that the exergonic reaction competes with the endergonic reaction for available $H_2$ [8]. The subsequent theories on surface metabolism and thioesters could not explain how the required reduced carbon compounds were synthesized with $CO_2$ as a substrate [9, 10], but did emphasize the important roles of thioesters and metal sulfide (FeS) catalysts in driving the primordial reductive tricarboxylic acid (rTCA) cycle, which is a central anabolic biochemical pathway whose origins have been proposed to trace back to geochemistry [9, 11, 12]. It might be more plausible that proto-anabolic networks, consisting of the reductive acetyl-CoA pathway, together with the incomplete rTCA cycle, which might be catalyzed by FeS and/or thioesters, were spontaneously driven into existence as a mechanism by which to dissipate geochemical redox gradients [3, 13-17]. We realize that there could have been a redox reaction which involved all major elements, and it makes sense that this redox reaction should have been catalyzed by FeS minerals and/or facilitated by thioesters. The roles played by the geochemical transformations of nitrogen and sulfur in the origin of life have been largely ignored because a computational analysis has suggested that nitrogen and sulfur were essential

for thermodynamically feasible phosphate-independent metabolism before the rise of the last universal common ancestor (LUCA) [18]. Phylogenetic distribution and functional grouping of sulfite reductase clusters suggest that a sulfite reductase, containing coupled siroheme-[$Fe_4$-$S_4$] cluster, was most likely present in the LUCA [2, 19-21]. Sulfite reductases from some sources can catalyze the reduction of both sulfite and nitrite [22], implying that S and N biochemistry may have a common evolutionary origin. Thus, the proto-anabolic networks might have been driven by carbon, hydrogen, oxygen, nitrogen, sulfur, and/or iron coupling transformations in the far-from-equilibrium environments. We speculated that thermodynamically feasible reduction coupling of sulfurous species with anaerobic ammonium oxidation reaction [23-27] (Eqs. 1, 2, 3, and 4; at 70 and 100°C), with or without the FeS and thioester catalyst, might have been the prebiotically relevant reaction for the creation of prebiotic proto-anabolic networks. The redox reaction between elemental sulfur and ammonium is not possible according to the value of $\Delta G^0$ (Eq. 3). However, if the concentration of elemental sulfur and ammonium reached 1 mM and the concentrations of reaction products were 0.01 mM, $\Delta G^0_r$ will be −22.78 kJ mol$^{-1}$ at 70°C and −31.99 kJ mol$^{-1}$ at 100 °C.

$$8NH_4^+ + 3S_2O_3^{2-} \rightarrow 4N_{2(g)} + 6H_2S_{(aq)} + 9H_2O + 2H^+$$

(70°C, $\Delta G^0$ = −167.2 kJ mol$^{-1}$; 100°C, $\Delta G^0$ = −194.35 kJ mol$^{-1}$) (Equation 1)

$$6NH_4^+ + 3SO_3^{2-} \rightarrow 3N_{2(g)} + 3H_2S_{(aq)} + 9H_2O$$

(70°C, $\Delta G^0$ = −332.04 kJ mol$^{-1}$; 100°C, $\Delta G^0$ = −370.77 kJ mol$^{-1}$)(Equation 2)

89 (70°C, $\Delta G^0$ = 62.81 kJ mol$^{-1}$; 100°C, $\Delta G^0$ = 53.6 kJ mol$^{-1}$) (Equation 3)

90 $$8NH_4^+ + 3SO_4^{2-} \rightarrow 4N_{2(g)} + 3H_2S_{(aq)} + 12H_2O + 2H^+$$

91 (70°C, $\Delta G^0$ = −113.38 kJ mol$^{-1}$; 100°C, $\Delta G^0$ = −152.98 kJ mol$^{-1}$) (Equation 4)

92 On early Earth, usually referred to as Earth in its first one billion years, sulfite,
93 elemental sulfur, and thiosulfate were abundantly produced from volcanic and
94 hydrothermal $SO_2$ or from the oxidation of $H_2S$ by iron oxides in the sulfide-rich
95 hydrothermal fluid [2, 13, 28]. The nitrogen species in these fluids released from the
96 mantle of the reduced young Earth into the early oceans might have comprised mostly
97 $NH_3$ [29, 30]. The $CO_2$ concentration in the oceans on early Earth would have been much
98 higher than that in the oceans today because there was perhaps up to 1,000 times more
99 $CO_2$ in the atmosphere [31]. Hence, when the prebiotically plausible sulfurous species
100 and $NH_3$ in the Hadean hydrothermal systems contacted $CO_2$, spontaneous redox
101 reaction transfers for energy generation and organic molecule synthesis via the
102 reduction of sulfurous species coupled with ammonium oxidation occurred. As sulfate
103 would have been severely limited in ancient oceans [2, 3, 32], we termed this process
104 Sammox, which more likely used thiosulfate, sulfite, and elemental sulfur as electron
105 acceptors instead of sulfate, particularly in the Hadean hydrothermal systems.

106 The phosphate-independent prebiotic proto-anabolic networks should have at
107 least started with carbon fixation and continued until peptides synthesis, as peptides
108 are at the heart of life and the best-known biocatalysts in the cell. Therefore, a
109 prebiotic reaction that led to the origin of life should have contributed to the
110 autocatalysis and evolution of the prebiotic proto-anabolic networks. A robust idea on

the origin of life should be able to explain the origin of homochirality of biological molecules (the use of only left-handed amino acids and only right-handed sugars), as it is a basic feature of life and should have come along with the origin of life. Here, we present prebiotic chemical evidences of Sammox-driven, prebiotic, reductive acetyl-CoA pathway combined with the incomplete rTCA cycle under hydrothermal conditions. The scenario of $CO_2$ fixation, bio-molecule synthesis, peptides formation, and the origin of autocatalysis and homochirality in one geochemical setting will provide profound insights into the earliest origins of life and fill in the gap in the understanding of the emergence from geochemistry to biochemistry.

## RESULTS AND DISCUSSION

### Sammox drives the emergence and combination of the prebiotic reductive acetyl-CoA pathway with incomplete rTCA cycle

The aim of this study was to first verify the feasibility of a Sammox-driven prebiotic reductive acetyl-CoA pathway, an incomplete rTCA cycle, and off-cycle reactions (esterification, reductive amination, and co-factor and peptide synthesis) in thiosulfate/sulfite/elemental sulfur/sulfate-fueled Sammox-driven prebiotic proto-anabolic networks (SPPN). Formate and acetate were the products of the Sammox-driven prebiotic reductive acetyl-CoA pathway with bicarbonate as the carbon source (Fig. 1), although formate was represented only in thiosulfate- and sulfate-fueled SPPN (Fig. 1a, d). We have not determined the presence of methanol but have qualitatively identified methyl acetate as a product in all sulfur-fueled SPPN

(Extended Data Fig. 1), which implies that methanol should be an intermediate of the Sammox-driven prebiotic reductive acetyl-CoA pathway. More importantly, this result confirmed the possibility of Sammox-powered esterification, which is critical for the synthesis of biopolymers [33]. We did not detect pyruvate as the end-product of the Sammox-driven prebiotic reductive acetyl-CoA pathway. When pyruvate was added to the SPPN, it quickly entered the incomplete rTCA cycle and was either consumed as a reaction substrate or was reductively aminated to amino acids (Figs. 2 and 3).

We quantitatively identified succinate and α-ketoglutarate as the products of the Sammox-driven prebiotic incomplete rTCA cycle (Figs. 1 and 2; Extended Data Fig. 2). The presence of polypyrroles implied the generation of pyrrole from succinate, as the off-cycle reaction products of the Sammox-driven prebiotic incomplete rTCA cycle (Fig. 3; Extended Data Fig. 3). A few peptides were found to be the end products of SPPN, especially in thiosulfate and sulfite-fueled SPPN (Figs. 1 and 2). The yield of peptides was relatively low in sulfate and elemental sulfur-fueled SPPN (Figs. 1 and 2). In this study, peptides might be the final products of SPPN (Fig. 3).

To further prove the feasibility of the Sammox-driven prebiotic, incomplete rTCA cycle, we used oxaloacetate, malate, fumarate, succinate, and α-ketoglutarate, as substrates in all sulfur-fueled SPPNs, and all eventually produced α-ketoglutarate and/or peptides (Fig. 4; Extended Data Fig. 4a–d). Interestingly, we found fumarate, succinate, α-ketoglutarate, and peptides as the products of malate amendment treatments in all sulfur-fueled SPPNs (Extended Data Fig. 4b), which suggested that

the generation of the intermediates was in accordance with the rTCA direction. Our results suggest that incomplete rTCA metabolites specifically inter-converted among the incomplete rTCA intermediates driven by Sammox. The Sammox-driven prebiotic reductive acetyl-CoA pathway and prebiotic, incomplete rTCA cycle integrated as prebiotic proto-anabolic networks.

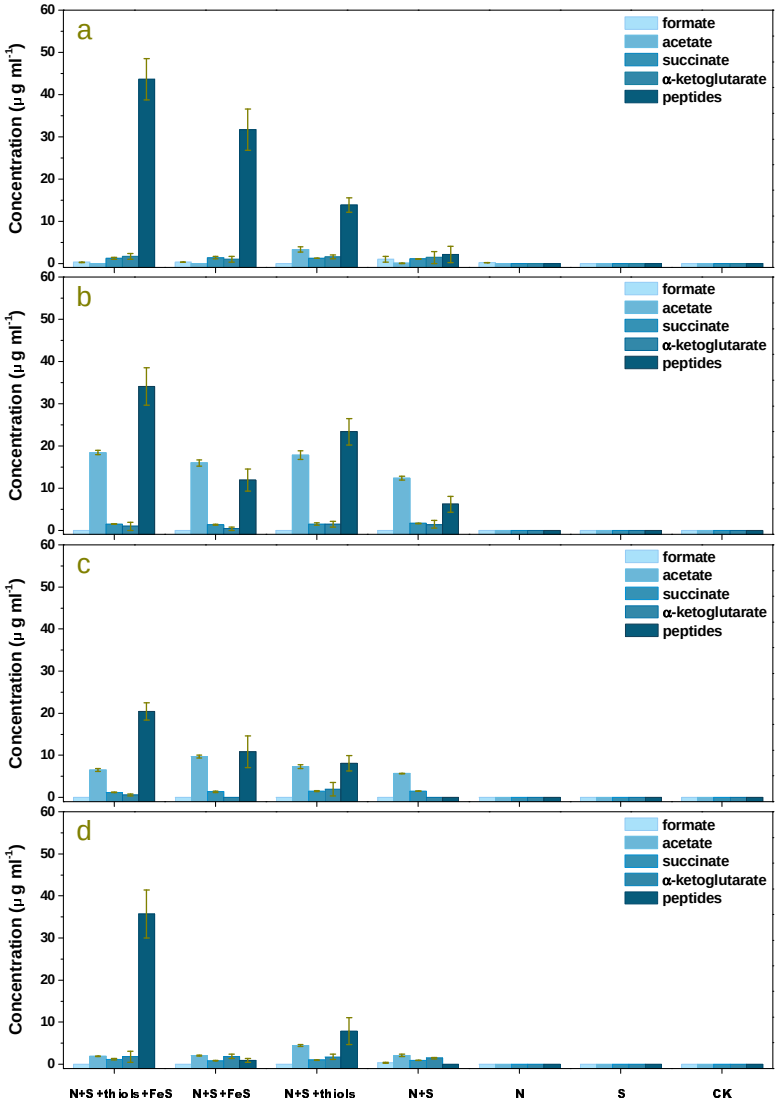
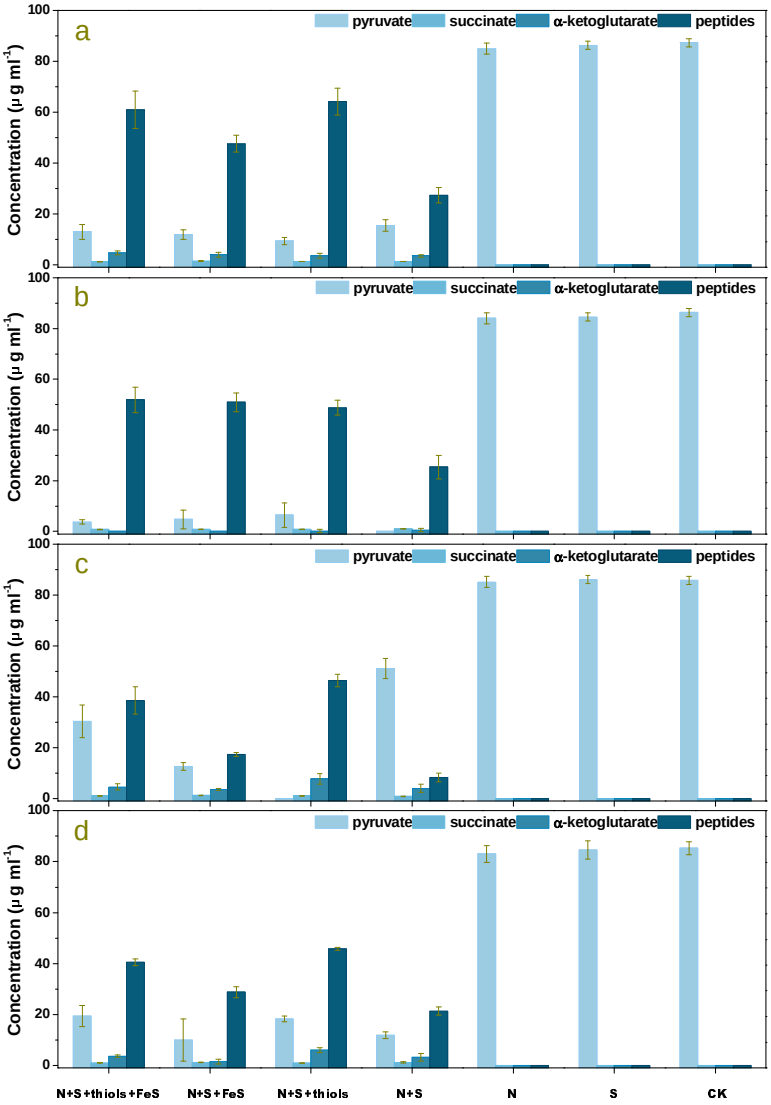



**Figure 1. Organic products generated from Sammox-driven prebiotic proto-anabolic networks (SPPN), with bicarbonate as the sole carbon source,**

**under hydrothermal conditions. Treatments were as follows (a, thiosulfate-fueled SPPN; b, sulfite-fueled SPPN; c, elemental sulfur-fueled SPPN; d, sulfate-fueled SPPN) from left to right: (i) sulfurous species, ammonium, FeS minerals, and methanethiol; (ii) sulfurous species, ammonium, and FeS minerals; (iii) sulfurous species, ammonium, and methanethiol; (iv) sulfurous species and ammonium; (v) ammonium; (vi) sulfurous species; (vii) CK. The bar chart shows the yields of formate, acetate, succinate, α-ketoglutarate, and peptides in each treatment group. Error bars represent standard deviations of three replicates.**




171

172 **Figure 2. Organic products of Sammox-driven prebiotic, incomplete rTCA cycle**

173 **with pyruvate as substrate under hydrothermal conditions. Treatments were as**

174 **follows (a, thiosulfate-fueled Sammox; b, sulfite-fueled Sammox; c, elemental**

175 **sulfur-fueled Sammox; d, sulfate-fueled Sammox) from left to right: (i) sulfurous**

176 **species, ammonium, FeS minerals, and methanethiol; (ii) sulfurous species,**

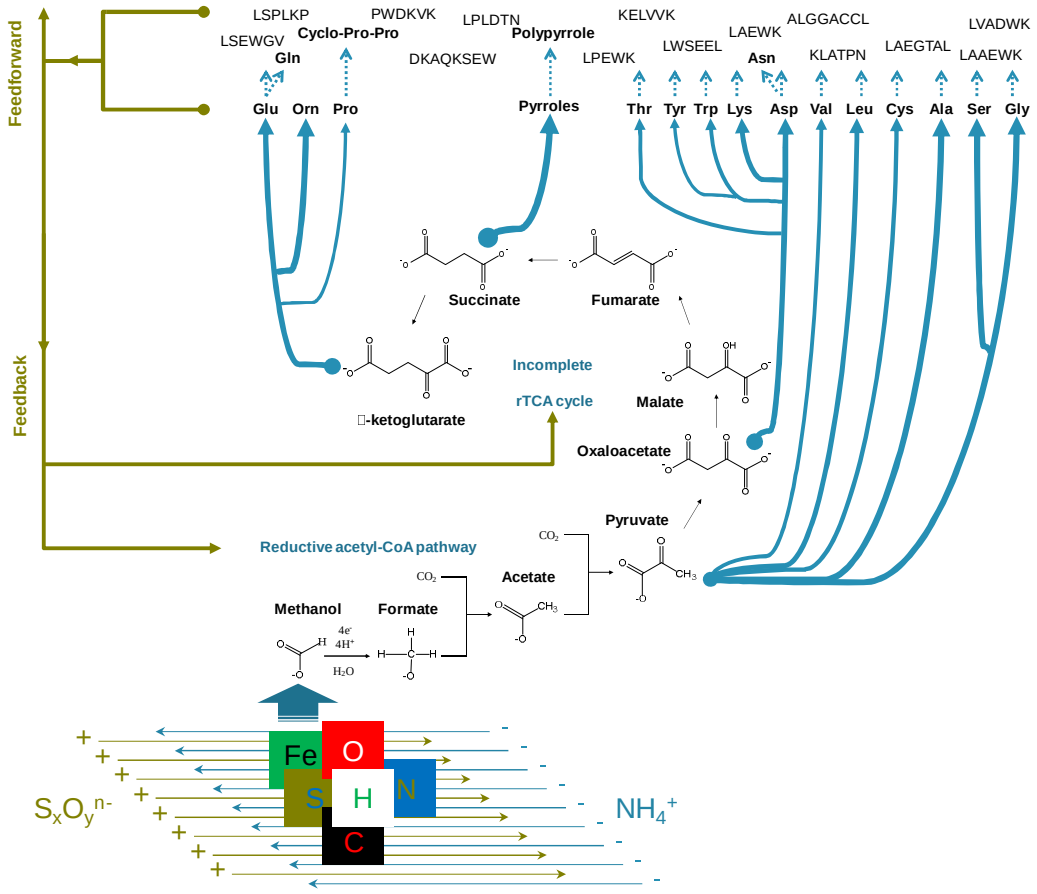
177 **ammonium, and FeS minerals; (iii) sulfurous species, ammonium, and**

178 **methanethiol; (iv) sulfurous species and ammonium; (v) ammonium; (vi)**

179 **sulfurous species; (vii) CK. The bar chart shows the concentrations of pyruvate,**

180 **succinate, α-ketoglutarate, and peptides in each treatment group. Error bars**

181 **represent standard deviations of three replicates.**




182

183 **Figure 3. The conceptual model of Sammox-driven coupling transformation of**

184 **carbon, hydrogen, oxygen, nitrogen, sulfur, and/or iron simultaneously in the**

185 **far-from-equilibrium environments, initiating the emergence of prebiotic**

186 **proto-anabolic networks, the combination of the reductive acetyl-CoA pathway,**

187 **and the incomplete rTCA cycle. Feedback and feedforward effects stem from the**

188 **products of SPPN, amino acids, and peptides. x = 1, 2; y = 0, 3, or 4; n = 0, 2.**

189

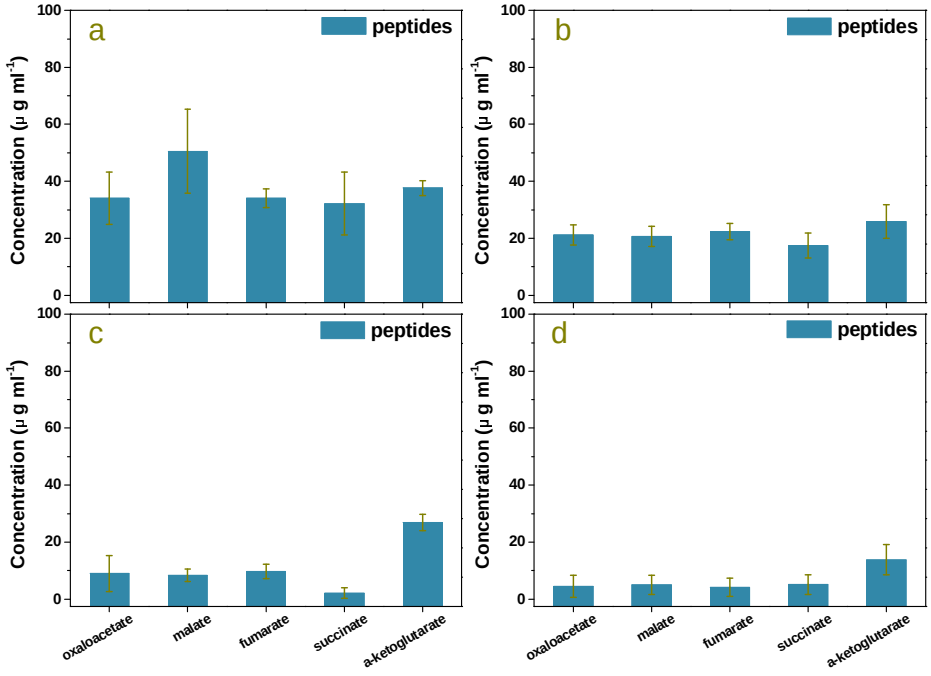



**Figure 4. Peptides generated from Sammox-driven prebiotic, incomplete rTCA cycle with oxaloacetate, malate, fumarate, succinate, and α-ketoglutarate as substrates. Treatments were as follows (a, thiosulfate-fueled Sammox; b, sulfite-fueled Sammox; c, elemental sulfur-fueled Sammox; d, sulfate-fueled Sammox). The bar chart shows the concentrations of peptides in each treatment group. Error bars represent standard deviations of three replicates.**

## The specificity and reasonability of SPPN

As expected, SPPN should have adequate specificity to function in a sustained way. It is also critical that side-reactions that would disrupt SPPN are avoided [14]. To find out whether side-reactions could move materials irreversibly out of the incomplete rTCA cycle, we evaluated the most problematic reactions—the reductions of pyruvate and α-ketoglutarate [14]. Our results showed that there were no reductions

of pyruvate to lactate or of α-ketoglutarate to α-hydroxyglutarate in all sulfur-fueled SPPNs (Figs. 1 and 2; Extended Data Figs. 2 and 4). Indeed, we did not find any unnecessary products in the SPPN. This suggests that SPPN is an efficient cycle, with naturally high specificity.

Both FeS minerals and methanethiol (used to synthesize thioesters) could enhance the efficiency of sulfur-fueled SPPN (Figs. 1 and 2). FeS cluster proteins are believed to be evolutionarily ancient and fundamental to central metabolism [34]. The physical and chemical principles are unchanged between biological and geological settings. Hence, it is logical that SPPNs could be catalyzed by FeS minerals as a proto-catalyst. Similarly, SPPN could be facilitated by thioesters as they were the energy currency before the emergence of adenosine triphosphate (ATP) [18, 34]. Methanethiol incorporates into the biochemical transformations of carbon, hydrogen, oxygen, and nitrogen through a thiol-thioester exchange [9], transforming itself and other organic reaction products into new organic products, resulting in the coupling biochemical transformation of carbon, hydrogen, oxygen, nitrogen, and sulfur and the expansion of SPPN. FeS minerals and thioesters were all accessible for SPPN in ancient oceans. Thiosulfate and elemental sulfur were the main products of sulfide oxidation by metallic oxide under anaerobic conditions in ancient oceans [35, 36]. When thiosulfate and elemental sulfur-fueled SPPN occurred, FeS minerals and methanethiol were found to be present in the same geological setting. As FeS minerals, thiosulfate, and elemental sulfur are products of sulfide oxidation by ferric oxide, and methanethiol is the product when FeS reacts with $H_2S$ and $CO_2$ [37].

Thiosulfate and sulfite-fueled SPPNs are the most efficient, generating a considerable number of organic products even without the facilitation of FeS minerals and thioesters (Figs. 1 and 2). Sulfite, thiosulfate, and elemental sulfur-fueled Sammox reactions could be the primordial power source for the creation of SPPN, instead of sulfate-fueled Sammox reactions. Consistent with our results, the metabolism of ancestral sulfur was found to be likely by sulfite reduction, sulfur disproportionation, and disulfide disproportionation rather than sulfate reduction [38].

**Conceivable emergence of autocatalysis and homochirality in SPPN**

Reductive amination occurred in SPPN as an off-cycle reaction of incomplete rTCA, with peptides as the end products. Moreover, we found that the peptide mixtures consisted of 15 proteinogenic amino acids, including alanine, glycine, valine, cystine, leucine, serine, aspartate, asparagine, lysine, glutamate, glutamine, tyrosine, threonine, tryptophan, and proline (Table 1 and 2; Fig. 5; Extended data Table 1; Extended data Fig. 5). The 15 proteinogenic amino acids required only a few steps in their metabolism from the incomplete rTCA (Fig. 3) [39]. Alanine, glycine, valine, cystine, leucine, and serine are closely linked to pyruvate (Fig. 3). Aspartate, asparagine, lysine, tyrosine, threonine, and tryptophan are linked to oxaloacetate (Fig. 3). Glutamate, glutamine, and proline are linked to α-ketoglutarate (Fig. 3). L-ornithine was also identified in our study (Table 2), which might have been derived from L-glutamate semialdehyde. Notably, compared to other amino acids, the frequency of the most ancient amino acids—glycine, alanine, aspartate, and glutamate—was relatively high (Table 2), which can be explained by many theories [40,

248 [41]. The frequency of serine was also significantly high, as it is conserved in ancestral
249 ferredoxin [34]. More importantly, the frequency of glycine, alanine, aspartate,
250 glutamate, and serine as conserved amino acids is consistent with the stage of genetic
251 code evolution [42]. Our results suggest that Sammox might be the right condition for
252 the origin of life.

253 The combined results from determination of amino acid content and peptide
254 identification showed that 56 SPPN-generated peptides shared similarity with truly
255 minimal protein content (TMPC) of LUCA, with only one exception, covering all
256 categories of TMPC, including electron transport, metabolism,
257 replication/recombination/repair/modification, transcription/regulation,
258 translation/ribosome, RNA processing, cellular processes, and transport/membrane [43,
259 44] (Table 1 and 2). A total of 55 peptides had high sequence similarity with more than
260 one category of TMPC (Table 1). SPPN-generated primordial peptides should be
261 multifunctional peptides with low substrate specificity; thus, related mechanistically
262 and evolutionarily. It has been suggested that any sufficiently complex set of
263 polypeptides will inevitably derive reflexively autocatalytic sets of peptides and
264 polypeptides [45]. In 1996, Lee et al. demonstrated that a rationally designed 32-residue
265 α-helical peptide could act autocatalytically in templating its own synthesis by
266 accelerating thioester-promoted amide-bond condensation in neutral aqueous
267 solutions, indicating that the peptide has the possibility of self-replication [46]. Other
268 studies have suggested that not only do some dipeptides and short peptide have
269 catalytic activities, even a single proline can have aldolase activity [47, 48, 49]. Therefore,

the SPPN-generated primordial peptides could be complex enough to facilitate the emergence of reflexive autocatalysis, and thus make SPPN autocatalytic (Fig. 3). Moreover, the molecules of the autocatalytic SPPN might promote the synthesis of more complex molecules via a feedforward effect (Fig. 3), thus, ensuring autocatalytic SPPN self-promotion.

In our study, as expected, 22 peptides shared similarity with oxidoreductase, and 11 shared similarity with ferredoxin-dependent proteins (Table 1). Four peptides shared similarity with thiosulfate sulfurtransferase as TMPC (Table 1). Thiosulfate sulfurtransferase (multifunctional rhodanese) catalyzes thiosulfate cleavage to sulfite. It has also been proposed to have an assimilatory role, using dithiol dihydrolipoate as the sulfur acceptor and acting as a sulfur insertase, involved in the formation of prosthetic groups in iron-sulfur proteins, such as ferredoxin [50, 51]. On the other hand, the ferredoxin polypeptide can accelerate the assembly of its own iron-sulfur clusters *in vitro* [51]. This suggests that suitable conditions might have been present for the formation of primordial iron-sulfur proteins in SPPN if with the involvement of Fe-S minerals. In addition to TMPC, we found six peptides which shared similarity with sulfite/nitrite reductase (Table 1). Dissimilatory sulfite reductase (DsrAB) is closely related to the assimilatory enzyme present in all domains of life and is an enzyme of ancient origin [52]. In fact, the functional divergence of assimilatory and dissimilatory sulfite reductases precedes the divergence of the bacterial and archaeal domains [19-21]; therefore, it has been concluded that a primordial siroheme-containing sulfite reductase was most likely present in LUCA [2]. In the Sammox reaction scenario, the

generation of pyrrole might have provided the possibility of the emergence of a primordial siroheme-containing sulfite reductase, as pyrrole is the building block of siroheme. Similar to sulfite reductase, nitrite reductase is widely distributed in all domains of life; sulfite reductases from some sources can catalyze the reduction of both sulfite and nitrite [22], suggesting that S and N biochemistry might have a common evolutionary origin derived from SPPN. Previous study reported a sulfate-fueled Sammox (also named as suramox) organism *Anammoxoglobus sulfate*, that belong to the phylum *Planctomycetes* [53]. In fact, so far, all anaerobic ammonium-oxidizing (anammox) organisms belong to a monophyletic group, deeply branching inside the phylum *Planctomycetes* [54]. Some of anammox *Planctomycetes* organisms have been suggested containing dissimilatory sulfite reductase [55, 56] (Fig. 6), implied that some *Planctomycetes* organisms may use sulfurous species as the electron acceptor instead of nitrite during anaerobic ammonium oxidation process. Figure 6 shows that anammox organisms dissimilatory sulfite reductases, thiosulfate reductase (eg. sulfurtransferase or rhodanese), elemental sulfur reductase (eg. polysulfide reductase), and sulfate adenylyltransferase are taxonomically related to that of two typical sulfate/sulfite reducing bacterium, *Desulfurispora thermophila* and *Moorella thermoacetica*. It is of great significance to study whether anammox organisms have Sammox function. More interestingly, *Planctomycetes* are one of the deepest lines of divergence in the Bacteria domain, and might be transitional forms between the three domains of life [57], implying a planctobacterial origin of neomura (eukaryotes, archaebacteria) [58]. Thus confirm a very ancient role of Sammox metabolism.

Sammox-generated primordial peptides, similar to the enzymes of carbon fixation metabolism, including the reductive acetyl-CoA pathway and incomplete rTCA cycle, such as carbon monoxide dehydrogenase, formate acetyltransferase, acetyl-coenzyme A synthetase, acetoacetate-CoA ligase, succinate-CoA ligase, and fumarate hydratase, may have the feedforward effect for the formation of primordial enzymes to catalyze the reductive acetyl-CoA pathway and incomplete rTCA cycle.

The origin of homochirality of biological molecules (the use of only L-amino acids and only D-sugars) is an inevitable question in the process of explaining the origin of life. During the origin of life origin, proline induced rapid and stereoselective conversions during organic syntheses, resulting in pronounced peptide growth in an RNA-dependent fashion [59]. Results from previous studies show that cyclo-Pro-Pro, a cyclic dipeptide, can catalyze the chiral selection of reactions and is a peptide precursor [60, 61, 62]. Cyclo-Pro-Pro can be formed directly from unprotected proline in aqueous trimetaphosphate solution under mild conditions, with an yield up to 97%, whereas other amino acids were found to form proline-containing cyclic dipeptides with lower yield under the same conditions [62]. In this study, cyclo-Pro-Pro could be found in sulfite-, elemental sulfur-, and sulfate-fueled SPPNs (Fig. 7). We did not find cyclo-Pro-Pro in thiosulfate-fueled SPPN because it quickly returns to the peptide form (Fig. 7a). This result might suggest the emergence of homochirality in SPPN-generated peptides. In this study, the Sammox reaction scenario was a far-from-equilibrium environment; accordingly, symmetry breaking and the asymmetric amplification of SPPN-generated L-peptides might result from

336 autocatalytic far-from-equilibrium chiral amplification reactions, such as Soai
337 autocatalytic reaction [63]. A very recent mechanistic study highlighted the crucial role
338 of isotope chiraliy in symmetry breaking and asymmetric amplification via the Soai
339 autocatalytic reaction [64] (Fig. 8).

340 The origin of the D-sugars in life becomes realizable under Sammox reaction
341 conditions because of the emergence of L-amino acids. Breslow and Cheng found that
342 D-glyceraldehyde, the basic unit from which all other D-sugars are built, synthesized
343 by the reaction of formaldehyde with glycolaldehyde, is catalyzed under prebiotic
344 conditions to D/L ratios greater than 1, to as much as 60/40, by L-serine, L-alanine,
345 L-phenylalanine, L-valine, L-leucine, and L-glutamic acid [66]. Moreover, catalysis
346 mixtures of natural D-sugars leads to enantio-enrichment of natural L-amino acid
347 precursors [66], thus showing the complementary nature of these two classes of
348 molecules in the origin of biological homochirality. Further effort is necessary to
349 explore the emergence of D-sugars in SPPN.

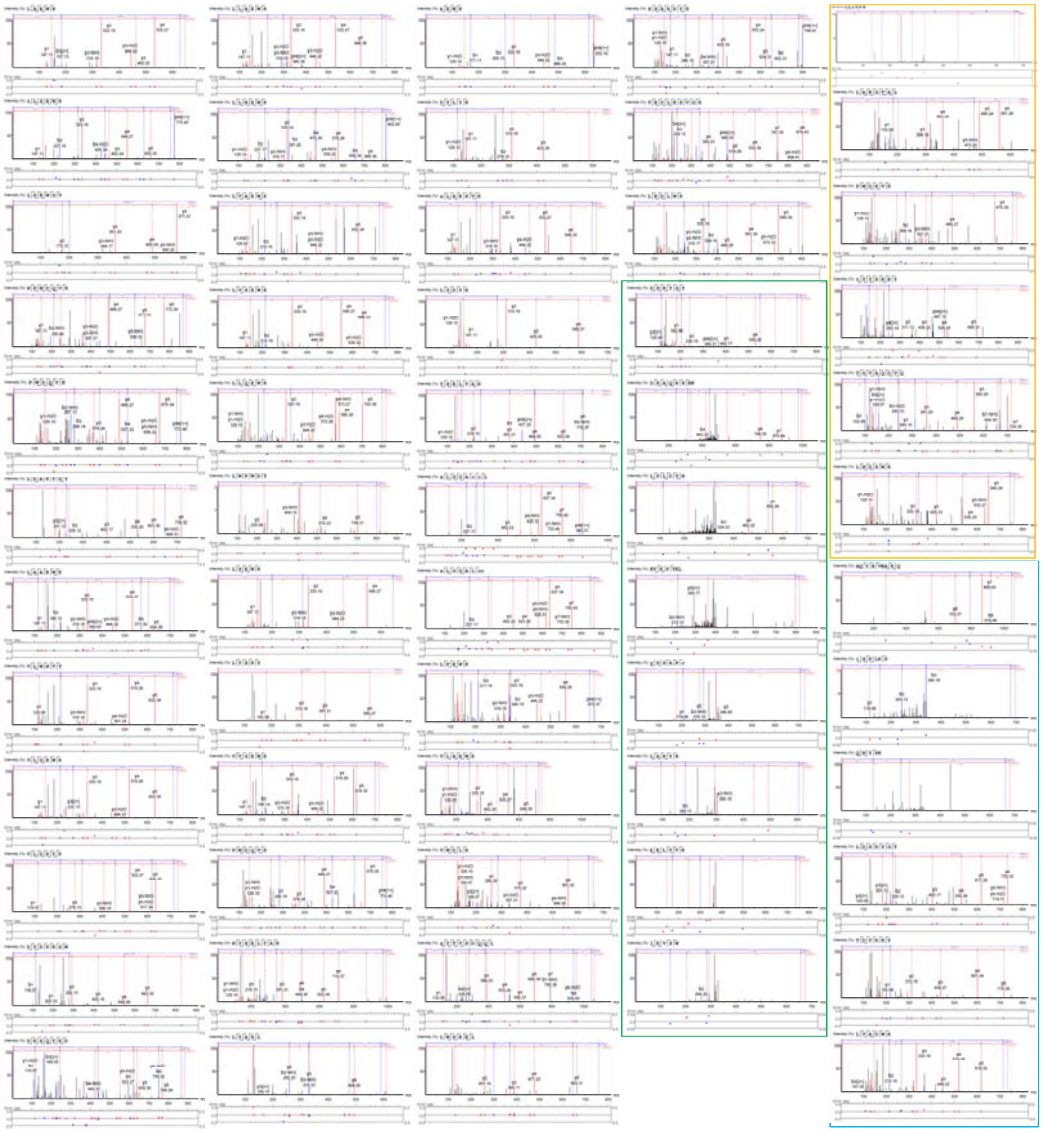


**Figure 5. MS/MS spectra of SPPN-generated peptides. The spectra of the eight peptides from sulfite-fueled SPPN are shown in the green box. The spectra of the six peptides from elemental sulfur-fueled SPPN are shown in the yellow box. The spectra of the six peptides from sulfate-fueled SPPN are shown in the blue box. The spectra of the other 37 peptides from thiosulfate-fueled SPPN are shown.**

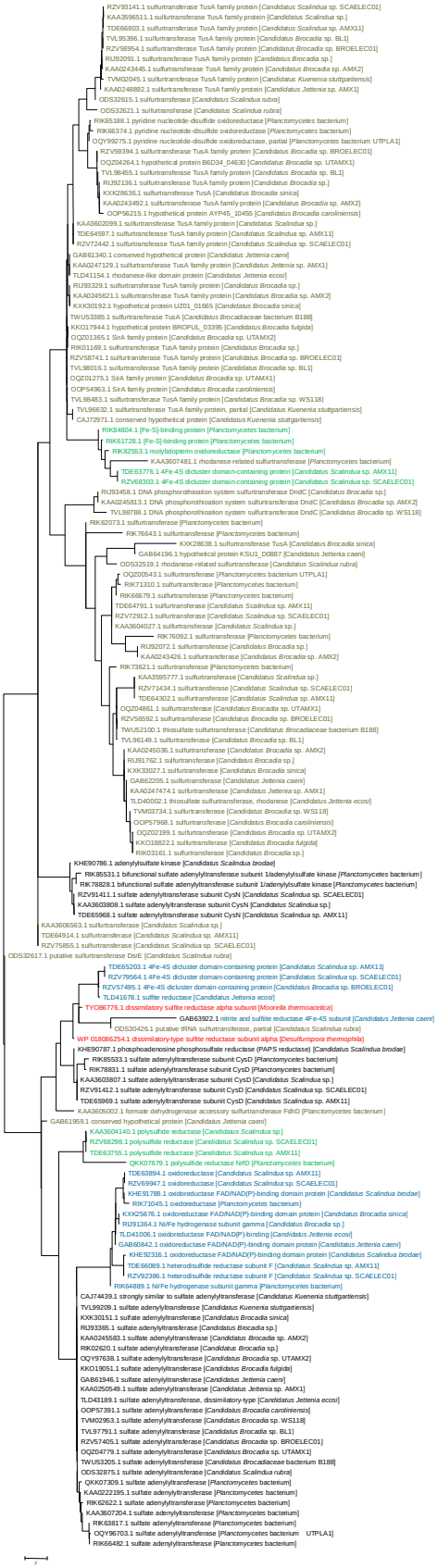


**Figure 6. Phylogenetic dendrogram based on the results of sequences comparisons. Phylogeny was inferred using the maximum-likelihood method. The percentages of replicate trees in which the associated taxa clustered together were obtained using bootstrap test with 1000 replications. Olive colour; putitive thiosulfate reductase. Blue; putitive sulfite reductase. Green; putitive elemental sulfur reductase. Black; sulfate reduction related reductase. Dissimilatory sulfite reductase alpha subunit of *Moorella thermoacetica* and *Desulfurispora thermophila* were included to show their position.**

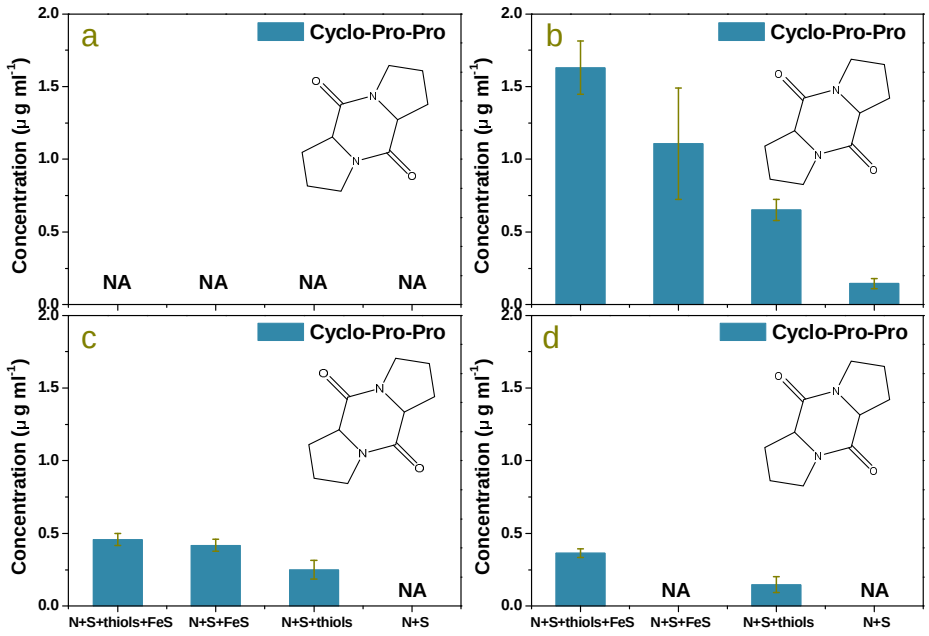



**Figure 7. Cyclo-Pro-Pro generated from SPPN. Treatments were as follows (a, thiosulfate-fueled SPPN; b, sulfite-fueled SPPN; c, elemental sulfur-fueled SPPN; d, sulfate-fueled SPPN). The bar chart shows the concentrations of cyclo-Pro-Pro in each treatment group. Error bars represent standard deviations of three replicates. NA: Not detected.**

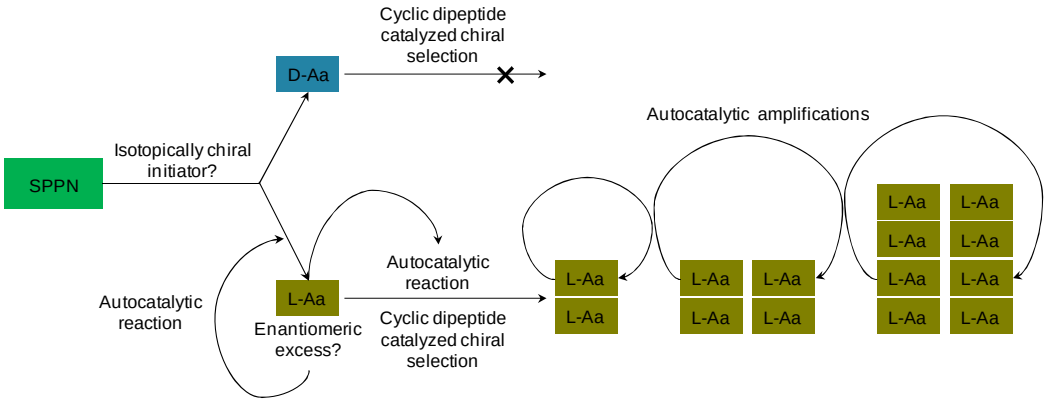



**Figure 8. Schematic diagram of asymmetric amplification of L-peptides in Soai autocatalysis initiated by isotopically chiral molecules in SPPN.**

## CONCLUSIONS

In this study, we provided convincing evidence of the presence of SPPN—the combination of the acetyl-CoA pathway and incomplete rTCA cycle under mild hydrothermal conditions. SPPN consists of $CO_2$ fixation and sulfur and nitrogen transformation, which are derived from the simplest substances—$CO_2$, sulfurous species (thiosulfate, sulfite, elemental sulfur, or sulfate), and ammonium to carboxylic acids, amino acids, pyrrole, and peptides in one geological setting. Peptides with 15 proteinogenic amino acids shared high similarity with TMPC of LUCA, which might endow the SPPN with autocatalysis and homochirality. FeS minerals as proto-catalyst and thiols/thioesters as primordial energy currency enhanced the efficiency of SPPN. Sammox organisms might belong to the phylum of *Planctomycetes*, as one of the deepest lines of divergence in the Bacteria domain, and might be transitional forms between the three domains of life. We suggest that the proto-anabolic networks might arise from simultaneously coupling transformation of carbon, hydrogen, oxygen,

nitrogen, sulfur, and/or iron in the far-from-equilibrium environments. The coupling transformation of sulfur and nitrogen were vital for driving the origin of life in planetary systems.

## MATERIAL AND METHODS

### Chemicals

All reagents and organic solvents were purchased from Alfa Aesar, Ark Pharm, J&K Scientific, and Sigma-Aldrich and used without further purification, unless otherwise noted. Ultrapure water was prepared using the Millipore purification system (Billerica, MA, USA).

### General procedure for prebiotic Sammox-driven $CO_2$ fixation with or without FeS catalysis and thiol/thioester promotion

A total of 100 mL ultrapure water was transferred into 120 mL serum bottles and sealed with butyl rubber stoppers and aluminum crimp caps. The solution in the serum bottles was autoclaved and cooled at 25$^{o}$C after being flushed with helium (He) gas (purity = 99.999%). Additional sulfur (thiosulfate, sulfite, elemental sulfur, and sulfate), ammonium ($^{14}NH_4Cl$), and bicarbonate were added into the serum bottles as the "Sammox reaction system." The above-mentioned ingredients were aseptically added to the serum bottles as follows: sodium thiosulfate, sulfite, and sulfate (1 mL, 3 mM final concentration), elemental sulfur (96 mg $L^{-1}$, equivalent to 3 mM final concentration), ammonium solution (0.5 mL, 1 mM final concentration), bicarbonate

412 solution (1 mL, 20 mM final concentration), and 1 mM freshly precipitated FeS
413 mineral were added into the serum bottles, as required. A solution of methanethiol
414 (0.8 mM final concentration), used to synthesize thioesters, was added to the Sammox
415 reaction system if necessary. The reaction systems were heated at 100°C in an oil bath
416 in the dark for 24 h, maintained at 70°C in the dark for 24 h, and removed from the oil
417 bath and allowed to cool to room temperature before analysis were conducted.

418 This study was performed using the following series of experiments:

419 (i) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 1 mM $NH_4Cl$ + 20 mM
420 $HCO_3^-$,

421 (ii) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 1 mM $NH_4Cl$ + 20 mM
422 $HCO_3^-$ + 1 mM FeS,

423 (iii) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 1 mM $NH_4Cl$ + 20 mM
424 $HCO_3^-$ + 0.8 mM methanethiol,

425 (iv) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 1 mM $NH_4Cl$ + 20 mM
426 $HCO_3^-$ + 1 mM FeS + 0.8 mM methanethiol,

427 (v) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 20 mM $HCO_3^-$,

428 (vi) 1 mM $NH_4Cl$ + 20 mM $HCO_3^-$, and

429 (vii) ultrapure water

430

431 **Experimental procedure for verifying the Sammox-driven combination of**
432 **non-enzymatic reductive acetyl-CoA pathway and incomplete rTCA cycle**

433 We designed this experimental set-up to verify whether the Sammox-driven

434 reductive acetyl-CoA pathway could go into the Sammox-driven incomplete rTCA
435 cycle. The experimental procedure was the same as that noted above. Pyruvate (1 mM
436 final concentration) was added into the Sammox reaction system as the substrate.
437 Serum bottles were heated at 100°C in a water bath in the dark for 24 h, maintained at
438 70°C in the dark for another 24 h, and allowed to cool to room temperature before
439 analysis.

440 This study was performed using the following series of experiments:

441 (i) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 1 mM $NH_4Cl$ + 20 mM
442 $HCO_3^-$ + 1 mM pyruvate + 1 mM FeS + 0.8 mM methanethiol,

443 (ii) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 1 mM $NH_4Cl$ + 20 mM
444 $HCO_3^-$ + 1 mM FeS + 1 mM pyruvate,

445 (iii) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 1 mM $NH_4Cl$ + 20 mM
446 $HCO_3^-$ + 1 mM pyruvate,

447 (iv) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 20 mM $HCO_3^-$ + 1 mM
448 pyruvate,

449 (v) 1 mM $NH_4Cl$ + 20 mM $HCO_3^-$ + 1 mM pyruvate, and

450 (vi) Ultrapure water + 1 mM pyruvate.

451

452 **General procedure for verifying the Sammox-driven prebiotic incomplete rTCA**
453 **cycle**

454 To further prove the feasibility of the Sammox-driven prebiotic incomplete
455 rTCA cycle, we used oxaloacetate, malate, fumarate, succinate, and α-ketoglutarate (1

456 mM final concentration), as substrates in Sammox systems. Serum bottles were
457 heated at 100°C in a water bath in the dark for 24 h, maintained at 70°C in the dark
458 for another 24 h, and allowed to cool to room temperature before analysis. This study
459 was performed using the following series of experiments:
460 (i) 3 mM thiosulfate/sulfite/elemental sulfur/sulfate + 1 mM $NH_4Cl$ + 20 mM
461 $HCO_3^-$ + 1 mM oxaloacetate/malate/fumarate/succinate/α-ketoglutarate.
462

463 **Sampling analytical methods**

464 ***Ion chromatographic analysis of the products of the Sammox-driven prebiotic***
465 ***reductive acetyl-CoA pathway***

466 To determine formate and acetate, 0.5 mL of each sample was filtered through
467 0.22-μm pores to remove particulates that could interfere with ion chromatography
468 (IC). The IC system comprised an ICS-5000$^+$ SP pump (Thermo Fisher Scientific Inc.
469 Sunnyvale, CA, USA), an ICS-5000$^+$ DC column oven, and a DC-5 electrochemical
470 detector [67]. The Dionex Ionpac AS11-HC column was used for IC. The operating
471 conditions were an eluent of 30 mM KOH at a flow rate of 1 mL $min^{-1}$.

472 ***Derivatization procedure and identification of carboxylic acids of the***
473 ***Sammox-driven prebiotic incomplete rTCA by GC-MS***

474 For optimal GC-MS resolution, the carboxylic acids were converted to ethyl
475 esters using a mixture of ethanol/ethyl chloroformate (EtOH/ECF). The carboxylic
476 acids were derivatized into esters as published before [68]. The products of this reaction
477 were identified by comparing the mass spectra and retention times against

478 analogously derivatized authentic samples. ECF derivatization was preferred for small
479 molecule substrates (pyruvate, lactate, malate, fumarate, succinate, α-ketoglutarate,
480 and α-hydroxyglutarate). To confirm a Sammox-driven prebiotic complete rTCA
481 cycle, $C_5$ and $C_6$ carboxylic acids of complete rTCA, such as *cis*-aconitate,
482 tricarballylate, iso-citrate, and citrate, were converted to methyl esters, using a
483 mixture of MeOH/MCF and following the same procedure as that used for ECF
484 derivatization [68].

485 GC-MS analysis was conducted on a 7890B GC System, connected to an MSD
486 block 5977A using the Agilent high-resolution GC column as follows: PN
487 19091S-433, HP-5MS, 28 m × 0.25 mm, 0.25 μm, SN USN 462366H. Samples were
488 prepared in 200 μL (sample volume) ethyl acetate and were analyzed on a splitless 1
489 μL injection volume with an injection port temperature of 250°C. The column oven
490 temperature protocol was as follows: 60°C for 1 min, ramped at 30°C $min^{-1}$ to 310°C
491 with a 3-min hold, and a total running time of 12.33 min. The mass spectrometer was
492 turned on after 3 min and was operated in the electron ionization mode with a
493 quadrupole temperature of 150°C. Data were acquired in the full-scan mode (50–500).
494 Hydrogen (99.999 % purity) was used as the carrier gas at a constant flow rate of 1.5
495 mL $min^{-1}$.

496 ***Liquid chromatography-MS method to determine polypyrrole***

497 For qualitative detection of polypyrrole, the Waters ACQUITY UPLC system
498 with an online coupled SYNAPT G2 mass spectrometer Q-TOF was used. Samples
499 were separated using the ACQUITY UPLC HSS T3 column (1.8 μm; 2.1 × 100 mm;

column temperature, 30°C). Solvent A contained 2.5% methanol, 0.2% formic acid in UPLC-grade water, and solvent B was 97.5% UPLC-grade water with 0.2% formic acid. Injection of 2 μL sample into the column at 0.2 mL $min^{-1}$ was followed by gradient elution. The mass spectrometric, qualitative detection of polypyrrole was conducted in resolution mode. Polypyrrole was identified by comparing the mass spectra and retention times against pure commercially available polypyrrole (100 μL $mL^{-1}$) that was heated at 100°C in a water bath in the dark for 24 h and maintained at 70°C in the dark for another 24 h. Tandem MS data were analyzed using MassLynx v 4.1.

***BCA protein assay***

The BCA protein assay was performed according to the instructions of the commercial kit. The detection solution was prepared using the BCA solution and the $Cu^{2+}$ solution at a ratio of 50:1. Thereafter, standard BSA solutions were added into each hole from 0–20 μL and the volume in each hole was made up to 20 μL with PBS. Thereafter, 20 μL sample solutions, followed by 200 μL of the detection solution were added to each hole. The microplate was incubated at 37°C in the analyzer for 20 min. The absorbance of the colorimetric solution was measured at 560 nm using a microplate reader (Infinite 200 PRO, TECAN).

***Nano LC-MS/MS analysis***

The sample solution was reduced using 10 mM DTT at 56°C for 1 h and alkylated with 20 mM IAA at room temperature, in dark for 1 h. Thereafter, the extracted peptides were lyophilized to near dryness and resuspended in 2–20 μL of

0.1% formic acid before LC-MS/MS analysis. LC-MS/MS analysis was performed on the UltiMate 3000 system (Thermo Fisher Scientific, USA) coupled to a Q Exactive™ Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Fisher Scientific, USA). The chromatographic separation of peptides was achieved using a nanocolumn—a 150 μm × 15 cm column—made in-house and packed with the reversed-phase ReproSil-Pur C18-AQ resin (1.9 μm, 100 A, Dr. Maisch GmbH, Germany). A binary mobile phase and gradient were used at a flow rate of 600 nL $min^{-1}$, directed into the mass spectrometer. Mobile phase A was 0.1% formic acid in water, and mobile phase B was 0.1% formic acid in acetonitrile. LC linear gradient: from 6–9% B for 5 min, from 9–50% B for 45 min, from 50–95% B for 2 min, and from 95–95% B for 4 min. The injection volume was 5 μL. MS parameters were set as follows: resolution at 70,000; AGC target at 3e6; maximum IT at 60 ms; number of scan ranges at 1; scan range at 300 to 1,400 m/z; and spectrum data type was set to profile. MS/MS parameters were set as follows: resolution was set at 17,500; AGC target at 5e4; maximum IT at 80 ms; loop count at 20; MSX count at 1; TopN at 20; isolation window at 3 m/z; isolation offset at 0.0 m/z; scan range at 200 to 2,000 m/z; fixed first mass at 100 m/z; stepped NCE at 27; spectrum data type at profile; intensity threshold at 3.1e4; and dynamic exclusion at 15 s. The raw MS files were analyzed and searched against target protein databases, based on the species of the samples using Peaks studio and MaxQuant (1.6.2.10), combined with manual comparison in the UniProt and NCBI databases. The parameters were set as follows: protein modifications were carbamidomethylation (C) (fixed), oxidation (M)

(variable), and acetylation (N-term) (variable); enzyme was set to unspecific; the maximum missed cleavages were set to 2; the precursor ion mass tolerance was set to 20 ppm, and MS/MS tolerance was 20 ppm. Only peptides identified with high confidence were chosen for downstream protein identification analysis.

For the analysis of amino acids content, reaction solution of SPPN was condensed 20 times, followed by acid hydrolysis. A total of 10 μL acid hydrolysate was mixed with 30 μL acetonitrile, vortexed for 1 min, and centrifuged for 5 min at 13,200 r $min^{-1}$ at 4°C. Thereafter, 10 μL of supernatant was added to 10 μL water and vortexed for 1 min. Subsequently, 10 μL of the mixture was added to 70 μL of borate buffer (from AccQTag kit) and vortexed for 1 min. A total of 20 μL of AccQ Tag reagent (from AccQTag kit) was added to the sample, vortexed for 1 min, and the sample was allowed to stand at ambient temperature for 1 min. Finally, the solution was heated for 10 min at 55°C, and centrifuged for 2 min at 13,200 r $min^{-1}$ and 4°C.

Multiple reaction monitoring analysis was performed by using a Xevo TQ-S mass spectrometer. All experiments were performed in positive electrospray ionization (ESI+) mode. The ion source temperature and capillary voltage were kept constant and set to 150°C and 2 kV, respectively. The cone gas flow rate was 150 L $h^{-1}$ and desolvation temperature was 600°C. The desolvation gas flow was 1,000 bar. The system was controlled using the analysis software.

***High performance liquid chromatography analysis of cyclo-Pro-Pro***

The Phenomenex Luna CN 5u column, which is a non-porous analytical column, packed with 5 μm particles (250 mm × 4.6 mm inner diameter, Phenomenex Inc, USA)

566 was used. Mobile phase A contained 0.05 M sodium acetate, while solvent B was 20%
567 methanol–60% acetonitrile–20% ultrapure water. Gradient profiling involved a linear
568 gradient elution from A/B (95:5) to A/B (52:48) for 39 min, a linear gradient elution
569 from A/B (52:48) to A/B (0:100) for 1 min, a linear gradient elution from A/B (0:100)
570 for 5 min, a linear gradient elution from A/B (0:100) to A/B (95:5) for 1 min, and A/B
571 (95:5), maintained for 4 min. The flow rates of the mobile phase and the column
572 temperature were set at 1 mL $min^{-1}$ and 35°C, respectively. The detection wave was
573 UV-220 nm. Cyclo-Pro-Pro was identified by comparing the retention times against
574 pure commercially available cyclo-Pro-Pro.

575 **Phylogenetic analyses of dissimilatory sulfite reductase, thiosulfate reductase,**
576 **elemental sulfur reductase, and sulfate adenylyltransferase**

577 Proteins sequences and the taxonomy of species classification information will
578 be downloaded from NCBI GenBank database (https://www.ncbi.nlm.nih.gov). The
579 evolutionary history was inferred by using the Maximum Likelihood method based on
580 the JTT matrix-based model [69]. The tree with the highest log likelihood (-3868.8161)
581 is shown. Initial tree for the heuristic search was obtained automatically as follows.
582 When the number of common sites was < 100 or less than one fourth of the total
583 number of sites, the maximum parsimony method was used; otherwise BIONJ method
584 with MCL distance matrix was used. The tree is drawn to scale, with branch lengths
585 measured in the number of substitutions per site. The analysis involved 150 amino
586 acid sequences. Evolutionary analyses were conducted in MEGA5 [70].

587

## ACKNOWLEDGMENTS

This research was financially supported by the National Natural Science Foundation of China (General Program No. 41571240).

## AUTHOR CONTRIBUTIONS

Peng Bao conceived the study, designed and carried out the experiment, and wrote the manuscript. Guo-Xiang Li carried out experiments and analysis. Peng Bao, Guo-Xiang Li, and Ke-Qing Xiao contributed to interpreting the data. Jun-Yi Zhao, Kun Wu, Juan Wang, Xiao-Yu Jia, Hui-En Zhang, Yu-Qin He, and Hu Li carried out sample analysis.

## COMPETING INTERESTS

The authors declare no competing interests.

## REFERENCES

(1) Falkowski, P. G.; Fenchel, T.; Delong, E. F. The Microbial Engines that Drive Earth's Biogeochemical Cycles. *Science*. **2008**, *320*, 1034–1039.

(2) Sousa, F. L.; Thiergart, T.; Landan, G.; Nelson-Sathi, S.; Pereira, I. A. C.; Allen, J. F.; Lane, N.; Martin, W. F. Early Bioenergetic Evolution. *Philos. Trans. R. Soc. B-Biol. Sci.* **2013**, *368*.

(3) Moore, E. K.; Jelen, B. I.; Giovannelli, D.; Raanan, H.; Falkowski, P. G. Metal Availability and the Expanding Network of Microbial Metabolisms in the Archaean Eon. *Nat. Geosci.* **2017**, *10*, 629–636.

(4) Weiss, M. C.; Sousa, F. L.; Mrnjavac, N.; Neukirchen, S.; Roettger, M.;

610 Nelson-Sathi, S.; Martin, W. F. The Physiology and Habitat of the Last Universal
611 Common Ancestor. *Nat. Microbiol.* **2016**, *1*.
612 (5) Ljungdah, L.; Irion, E.; Wood, H. G. Total Synthesis of Acetate from $CO_2$. I.
613 Co-Methylcobyric Acid and Co-(methyl)-5-Methoxybenzimidazolylcobamide as
614 Intermediates with *Clostridium thermoaceticum*. *Biochemistry* **1965**, *4*,
615 2771–2780.
616 (6) Kelley, D. S.; Karson, J. A.; Fruh-Green, G. L.; Yoerger, D. R.; Shank, T. M.;
617 Butterfield, D. A.; Hayes, J. M.; Schrenk, M. O.; Olson, E. J.; et al. A
618 Serpentinite-Hosted Ecosystem: the Lost City Hydrothermal Field. *Science* **2005**,
619 *307*, 1428–1434.
620 (7) Lane, N.; Martin, W. F. The Origin of Membrane Bioenergetics. *Cell* **2012**, *151*,
621 1406–1416.
622 (8) Wächtershäuser, G. Pyrite Formation, the First Energy Source for Life: a
623 Hypothesis. *Syst. Appl. Microbiol.* **1988**, *10*, 207–210.
624 (9) Wächtershäuser, G. Evolution of the First Metabolic Cycles. *Proc. Natl. Acad. Sci.*
625 *U. S. A.* **1990**, *1*, 200–204.
626 (10) De Duve, C. Blueprint for a Cell: the Nature and Origin of Life. Neil Patterson,
627 **1991**,275.
628 (11) Evans, M. C. W.; Buchanan, B. B.; Arnon, D. I. A New Ferredoxin-Dependent
629 Carbon Reduction Cycle in a Photosynthetic Bacterium. *Proc. Natl. Acad. Sci. U.*
630 *S. A.* **1966**, *55*, 928–934.
631 (12) Smith, E.; Morowitz, H. J. Universality in Intermediary Metabolism. *Proc. Natl.*

*Acad. Sci. U. S. A.* **2004**, *101*, 13168–13173.

(13) Martin, W.; Russell, M. J. On the Origin of Biochemistry at an Alkaline Hydrothermal Vent. *Philos. Trans. R. Soc. B-Biol. Sci.* **2007**, *362*, 1887–1925.

(14) Orgel, L. E. The Implausibility of Metabolic Cycles on the Prebiotic Earth. *PLoS. Biol.* **2008**, *6*, 5–13.

(15) Braakman, R.; Smith, E. The Emergence and Early Evolution of Biological Carbon-Fixation. *PLoS Comput. Biol.* **2012**, *8*.

(16) Camprubi, E.; Jordan, S. F.; Vasiliadou, R.; Lane, N. Iron Catalysis at the Origin of Life. *IUBMB Life* **2017**, *69*, 373–381.

(17) Russell, M. J.; Hall, A. J.; Mellersh, A. R. Natural and Laboratory Simulated Thermal Geochemical Processes. *Springer*, **2003**; pp 325–388.

(18) Goldford, J. E.; Hartman, H.; Smith, T. F.; Segre, D. Remnants of an Ancient Metabolism without Phosphate. *Cell* **2017**, *168*, 1126–1134.

(19) Crane, B. R.; Siegel, L. M.; Getzoff, E. D. Sulfite Reductase Structure at 1.6 Angstrom—Evolution and Catalysis for Reduction of Inorganic Anions. *Science* **1995**, *270*, 59–67.

(20) Molitor, M.; Dahl, C.; Molitor, I.; Schafer, U.; Speich, N.; Huber, R.; Deutzmann, R.; Truper, H. G. A Dissimilatory Sirohaem-Sulfite-Reductase-Type Protein from the Hyperthermophilic Archaeon *Pyrobaculum islandicum*, *Microbiology* **1998**, *144*, 529–541.

(21) Dhillon, A.; Goswami, S.; Riley, M.; Teske, A.; Sogin, M. Domain Evolution and Functional Diversification of Sulfite Reductases. *Astrobiology* **2005**, *5*,

18–29.

(22) Crane, B. R.; Getzoff, E. D. The Relationship between Structure and Function for the Sulfite Reductases. *Curr. Opin. Struct. Biol.* **1996**, *6*, 744–756.

(23) Fdz-Polanco, F.; Fdz-Polanco, M.; Fernandez, N.; Uruena, M. A.; Garcia, P. A.; Villaverde, S. Combining the Biological Nitrogen and Sulfur Cycles in Anaerobic Conditions. *Water Sci. Technol.* **2001**, *44*, 77–84.

(24) Amend, J. P.; Rogers, K. L.; Shock, E. L.; Gurrieri, S.; Inguaggiato, S. Energetics of Chemolithoautotrophy in the Hydrothermal System of Vulcano Island, Southern Italy. *Geobiology* **2003**, *1*, 37–58.

(25) Schrum, H. N.; Spivack, A. J.; Kastner, M.; D'Hondt, S. Sulfate-Reducing Ammonium Oxidation: A Thermodynamically Feasible Metabolic Pathway in Subseafloor Sediment. *Geology* **2009**, *37*, 939–942.

(26) Rios-Del Toro, E. E.; Valenzuela, E. I.; Lopez-Lozano, N. E.; Cortes-Martinez, M. G.; Sanchez-Rodriguez, M. A.; Calvario-Martinez, O.; Sanchez-Carrillo, S.; Cervantes, F. J. Anaerobic Ammonium Oxidation Linked to Sulfate and Ferric Iron Reduction Fuels Nitrogen Loss in Marine Sediments. *Biodegradation* **2018**, *29*, 429–442.

(27) Amend, J. P.; Shock, E. L. Energetics of Overall Metabolic Reactions of Thermophilic and Hyperthermophilic Archaea and Bacteria. *FEMS Microbiol Rev* **2001**, 25, 175–243.

(28) Philippot, P.; Zuilen, M. V.; Lepot, K.; Thomazo, C.; Farquhar, J.; Van Kranendonk, M. J. Early Archaean Microorganisms Preferred Elemental Sulfur,

not Sulfate. *Science* **2007**, *317*, 1534–1537.

(29) Li, Y.; Keppler, H. Nitrogen Speciation in Mantle and Crustal Fluids. *Geochim. Cosmochim. Acta*. **2014**, *129*, 13–32.

(30) Mikhail, S.; Sverjensky, D. A. Nitrogen Speciation in upper Mantle Fluids and the Origin of Earth's Nitrogen-Rich Atmosphere. *Nat. Geosci.* **2014**, *7*, 816–819.

(31) Sleep, N. H. The Hadean-Archaean Environment. *Cold Spring Harbor Perspect. Biol*. **2010**, *2*.

(32) Crowe, S. A.; Paris, G.; Katsev, S.; Jones, C.; Kim, S. T.; Zerkle, A. L.; Nomosatryo, S.; Fowle, D. A.; Adkins, J. F.; et al. Sulfate was a Trace Constituent of Archean Seawater. *Science* **2014**, *346*, 735–739.

(33) Eigen, M. Self-Organization of Matter and the Evolution of Biological Macromolecules. *Naturwissenschaften* **1971**, *58*, 465–562.

(34) Eck, R. V.; Dayhoff, M. O. Evolution of the Structure of Ferredoxin Based on Living Relics of Primitive Amino Acid Sequences. *Science* **1966**, *152*, 363–366.

(35) Jorgensen, B. B. A Thiosulfate Shunt in the Sulfur Cycle of Marine-Sediments. *Science* **1990**, *249*.

(36) Lohmayer, R.; Kappler, A.; Losekann-Behrens, T.; Planer-Friedrich, B. Sulfur Species as Redox Partners and Electron Shuttles for Ferrihydrite Reduction by *Sulfurospirillum deleyianum*. *Appl. Environ. Microbiol*. **2014**, *80*, 3141−3149.

(37) Heinen, W.; Lauwers, A. M. Organic Sulfur Compounds Resulting from the Interaction of Iron Sulfide, Hydrogen Sulfide and Carbon Dioxide in an Anaerobic Aqueous Environment. *Orig. Life Evol. Biosph.* **1996**, *26*, 131–150.

(38) Milucka, J.; Ferdelman, T. G.; Polerecky, L.; Franzke, D.; Wegener, G.; Schmid, M.; Lieberwirth, I.; Wagner, M.; Widdel, F.; Kuypers, M. M. M. Zero-Valent Sulphur is a Key Intermediate in Marine Methane Oxidation. *Nature* **2012**, *491*, 541–546.

(39) Hartman, H. Speculations on the Origin and Evolution of Metabolism. *J. Mol. Evol.* **1975**, *4*, 359–370.

(40) Wong, J. T. F. Coevolutlon Theory of Genetic Code at Age Thirty. *Bioessays*. **2005**, *27*, 416–425.

(41) Trifonov, E. N.; Volkovich, Z.; Frenkel, Z. M. Multiple Levels of Meaning in DNA Sequences, and One More. *Ann. NY Acad. Sci.* **2012**, *1267*, 35–38.

(42) Davis, B. K. Molecular Evolution Before the Origin of Species. *Prog. Biophys. Mol. Biol.* **2002**, *79*, 77–133.

(43) Koonin, E. V. Comparative Genomics, Minimal Gene-Sets and the Last Universal Common Ancestor, *Nat. Rev. Microbiol*. **2003**, *1*, 127–136.

(44) Ouzounis, C. A.; Kunin, V.; Darzentas, N.; Goldovsky, L. A Minimal Estimate for the Gene Content of the Last Universal Common Ancestor-Exobiology From a Terrestrial Perspective. *Res. Microbiol*. **2006**, *157*(1):57–68.

(45) Kauffman, S. A. Autocatalytic Sets of Proteins. *J. theor. Biol.* **1986**, *119*, 1–24.

(46) Lee, D. H.; Granja, J. R.; Martinez, J. A.; Severin, K.; Ghadiri, M. R. A Self-replicating Peptide. *Nature* **1996**, *382*, 525–528.

(47) Sakthivel, K.; Notz, W.; Bui, T.; Barbas, C. F., III. Amino Acid Catalyzed Direct Asymmetric Aldol Reactions: A Bioorganic Approach to Catalytic Asymmetric

720 Carbon-Carbon Bond Formation. *J. Am. Chem. Soc*. **2001**, *123*, 5260.

721 (48) Eilzabeth, R. J.; Miller, S. J. Amino Acids and Peptides As Asymmetric

722 Organocatalysis. *Tetrahedron*. **2002**, *58* (13): 2481–2495.

723 (49) Kroiss, D.; Ashkenasy, G.; Braunschweig, A. B.; Tuttle, T.; Ulijn, R. V. Catalyst:

724 can systems chemistry unravel the mysteries of the chemical origins of life?

725 *Chem.* **2019**, *5*, 1917–1920.

726 (50) Lefaou, A.; Rajagopal, B. S.; Daniels, L.; Fauque, G. Thiosulfate, polythionates

727 and elemental sulfur assimilation and reduction in the bacterial world. *FEMS*

728 *Microbiol Lett*. **1990**, *75*, 351–82.

729 (51) Bonomi, F.; Werth, M. T.; Kurtz, D. M. Assembly of [FenSn(SR)]2-(n=2, 4) in

730 aqueous media from iron salts, thiols and sulfur, sulfide, thiosulfide plus

731 rhodonase. *Inorg Chem.* **1985**, *24*, 4331–5.

732 (52) Grein, F.; Ramos, A. R.; Venceslau, S. S.; Pereira, I. A. C. Unifying Concepts in

733 Anaerobic Respiration: Insights from Dissimilatory Sulfur Metabolism.

734 *BBA-Bioenerg*. **2013**, *1827*, 145–160.

735 (53) Liu, S.; Yang, F.; Gong, Z.; Meng, F.; Chen, H.; Xue, Y.; Furukawa, K.

736 Application of anaerobic ammonium-oxidizing consortium to achieve

737 completely autotrophic ammonium and sulfate removal. *Bioresour. Technol.*

738 **2008**, *99* (15), 6817–6825.

739 (54) van Niftrik, L.; Jetten, M. S. M. Anaerobic ammonium-oxidizing bacteria:

740 unique microorganisms with exceptional properties. *Microbiol. Mol. Biol. R.*

741 **2012**, *76*, 585–596.

742 (55) Kartal, B.; Keltjens, J. T. Anammox biochemistry: a tale of heme *c* proteins.
743 *Trends. Biochem. Sci*. **2016**, *41*(12), 998–1011.
744 (56) Anantharaman, K.; Hausmann, B.; Jungbluth, S. P.; Kantor, R. S.; Lavy, A.;
745 Warren, L. A.; Rappé, M. S.; Pester, M.; Loy, A.; Thomas, B. C.; Banfield, J. F.
746 Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle.
747 *ISME J*. **2018**, *12*, 1715–1728.
748 (57) Reynaud, E. G.; Devos, D. P. Transitional forms between the three domains of
749 life and evolutionary implications. *Proc. Biol. Sci.* **2011**, *278*, 3321–3328.
750 (58) Cavalier-Smith, T.; Chao, E. E. Multidomain ribosomal protein trees and the
751 planctobacterial origin of neomura (eukaryotes, archaebacteria). *Protoplasma.*
752 **2020**, *257*, 621–753.
753 (59) Griesser, H.; Bechthold, M.; Tremmel, P.; Kervio, E.; Richert, C. Amino
754 acid-Specific, Ribonucleotide-Promoted Peptide Formation in the Absence of
755 Enzymes. *Angew. Chem. Int. Ed*. **2017**, *56*, 1224–1228.
756 (60) Borthwick, A. D. 2, 5-Diketopiperazines: Synthesis, Reactions, Medicinal
757 Chemistry, and Bioactive Natural Products. *Chemical. reviews.* **2012**, *112*,
758 3641–3716.
759 (61) Danger, G.; Plasson, R.; Pascal, R. Pathways For the Formation and Evolution of
760 Peptides In Prebiotic Environments. *Chem. Soc. Rev.* **2012**, *41*, 5416–5429.
761 (62) Ying, J.; Lin, R.; Xu, P.; Wu, Y.; Liu, Y.; Zhao, Y. Prebiotic Formation of Cyclic
762 Dipeptides Under Potentially Early Earth Conditions. *Sci. Rep*. **2018**, *8*, 936.
763 (63) Soai, K.; Shibata, T.; Morioka, H.; Choji, K. Asymmetric Autocatalysis and

764 Amplification of Enantiomeric Excess of a Chiral Molecule. *Nature*. **1995**, *378*,
765 767–768.
766 (64) Hawbaker, N.; Blackmond, D. G. Rationalization of Asymmetric Amplification
767 via Autocatalysis Triggered by Isotopically Chiral Molecules. *ACS Cent. Sci*.
768 **2018**, *4*, 776–780.
769 (65) Breslow, R.; Cheng, Z. L. L-Amino acids Catalyze the Formation of An Excess
770 of D-glyceraldehyde, and Thus of Other D Sugars, Under Credible Prebiotic
771 Conditions. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 5723−5725.
772 (66) Wagner, A. J.; Zubarev, D. Y.; Aspuru-Guzik, A.; Blackmond, D. G. Chiral
773 Sugars Drive Enantioenrichment in Prebiotic Amino Acid Synthesis. *ACS Cent.*
774 *Sci*. **2017**, *3*, 322−328.
775 (67) Bao, P.; Li, G. X. Sulfur-Driven Iron Reduction Coupled to Anaerobic
776 Ammonium Oxidation. *Environ. Sci. Technol*. **2017**, *51*, 6691−6698.
777 (68) Muchowska, K. B.; Varma, S. J.; Chevallot-Beroux, E.; Lethuillier-Karl, L.; Li,
778 G.; Moran, J. Metals Promote Sequences of the Reverse Krebs Cycle. *Nat. Ecol.*
779 *Evol.* **2017**, *1*, 1716–1721.
780 (69) Jones, D. T.; Taylor, W. R.; Thornton. J. M. The rapid generation of mutation
781 data matrices from protein sequences. *Comput. Appl. Biosci*. **1992**, *8*, 275–282.
782 (70) Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5:
783 molecular evolutionary genetics analysis using maximum likelihood, evolutionary
784 distance, and maximum parsimony methods. *Mol. Biol. Evol*. **2011**, *28*, 2731–9.
785

**Table 1. Sammox-driven prebiotic proto-anabolic networks (SPPN)-generated peptides compared within the UniProt and NCBI databases. Thios, peptide from thiosulfate-fueled SPPN; Sulfi, peptide from sulfite-fueled SPPN; ES, peptide from elemental sulfur-fueled SPPN; Sulfa, peptide from sulfate-fueled SPPN. Summary of proteins identified from the truly minimal protein content (TMPC) of the last universal common ancestor (LUCA), iron-sulfur protein (in blue), proposed Sammox metabolic protein (in blue), including function of proteins, and organismal source.**

| Peptide ID | Denovo peptide | Function, and organismal source of matched protein |
|---|---|---|
| Thios-1 | LAEWK | Putative thiosulfate sulfurtransferase, *Methanothermobacter thermautotrophicus*/4Fe-4S ferredoxin, Deltaproteobacteria bacterium/acyl-CoA synthetase, *Kandeliimicrobium roseum*/Acetyl-CoA acetyltransferase, *Weissella oryzae*/ATP-dependent DNA helicase II subunit 1, *Chaetomium globosum*/Acetolactate synthase large subunit, *Bacillus subtilis*/Glycerol-3-phosphate dehydrogenase [NAD(P)+], *Corynebacterium glutamicum*/Lysine--tRNA ligase, *Pseudarthrobacter* sp. AG30/Branched-chain-amino-acid aminotransferase-like protein 1, *Arabidopsis thaliana*/Histidine ammonia-lyase, *Caulobacter vibrioides*/Phenylalanine--tRNA ligase alpha subunit, *Roseiflexus* sp. RS-1/ Flavin-dependent thymidylate synthase, *Thermococcus kodakarensis* |
| Thios-2 | LLSEWK | Glutamine-dependent NAD(+) synthetase, *Wallemia ichthyophaga*/thioester reductase family protein, *Photorhabdus luminescens*/Mediator of RNA polymerase II transcription subunit 16, *Aedes aegypti* (Yellowfever mosquito) /Diaminopimelate decarboxylase, *Aureibacillus halotolerans*/Sulfite reductase, ferredoxin dependent, *filamentous* cyanobacterium/Thiamine-monophosphate kinase, *Candidatus Omnitrophica* bacterium/DNA helicase, *Candidatus Rokubacteria* bacterium/DNA mismatch repair protein MutS, *Desulfovibrio* sp. An276/CTP synthase, *Candidatus Poseidoniales* archaeon |
| Thios-3 | LSEWGV | Oxidored_nitro domain-containing protein, *Firmicutes* bacterium/NADH-quinone oxidoreductase subunit I, *Rhodopseudomonas palustris*/Cysteine--tRNA ligase, *Methylocella silvestris*/CTP synthase, *Chloroflexi* bacterium/ 3-isopropylmalate dehydratase large subunit, *Fictibacillus phosphorivorans*/Orotate phosphoribosyltransferase, *Bacillus campisalis*/Probable cysteine desulfurase, *Chlamydia trachomatis* |
| Thios-4 | NPWDQVK | Serine/threonine protein kinase, putative, *Talaromyces stipitatus*/Amidase domain-containing protein, *Chloroflexi* bacterium/ABC transporter, *Verrucomicrobiales* bacterium/long-chain fatty acid--CoA ligase, *Corynebacterium* sp. CNJ-954 /phosphoribosylformylglycinamidine synthase, *Candidatus Accumulibacter phosphatis* |
| Thios-5 | PWDQVK | Acetyl-coenzyme A synthetase, *Methanomicrobiales* archaeon/Cluster: His-tRNA synthetase (Fragment), *Teinoptila guttella*/Cluster: NAD(P)/ FAD-dependent oxidoreductase, bacterium M00.F/Cluster: Translation elongation factor G, Firmicutes bacterium/Cluster: Amino acid permease family protein, *Acinetobacter nosocomialis*/Methyltransf_2 domain-containing protein, *Hypholoma sublateritium*/Aminotransferase, Bacteroidetes bacterium |
| Thios-6 | LDAKYGY | Magnesium and cobalt transporter CorA, *Coprococcus* sp. HPP0048/ Peptidase, *Staphylococcus*.OJ82 |
| Thios-7 | LAAEWK | Adenylosuccinate lyase, *Candidatus Moduliflexus flocculans*/Acetyl-CoA synthetase, *Prevotella* sp. kh1p2/Carbon monoxide dehydrogenase, *Clostridium formicaceticum*/NADH-quinone oxidoreductase subunit N, *Acidobacteria* bacterium/Shikimate dehydrogenase (NADP(+)) aroE, *Paenibacillus rhizosphaerae*/Lysine--tRNA ligase, *Candidatus Campbellbacteria* bacterium/DNA helicase, *Baekduia soli* |
| Thios-8 | VLWNVT; VLGEWK | Acetoacetate--CoA ligase, Deltaproteobacteria bacterium/dehydratase, *Aliikangiella* marina/Multifunctional fusion protein [Includes: ADP-dependent (S)-NAD(P)H-hydrate dehydratase, (ADP-dependent NAD(P)HX dehydratase); NAD(P)H-hydrate epimerase, (NAD(P)HX epimerase)], Bacteroidales bacterium/Aldehyde ferredoxin oxidoreductase, *Candidatus Bathyarchaeota* archaeon B24-2/Branched-chain amino acid ABC transporter permease, *Streptomyces* sp. FR-008/Acetylornithine aminotransferase, *Koribacter versatilis*/Carbamoyl-phosphate synthase large chain, *Chryseobacterium* sp. 1,059 |
| Thios-9 | VLDKYP | Aspartate--tRNA ligase, cytoplasmic, *Mus musculus* (Mouse)/Adenosylhomocysteinase, *Vibrionimonas magnilacihabitans*/Peptidase T, *Chitinophaga cymbidii*/ABC transporter, ATP-binding protein, *Collinsella stercoris*/ Cysteine--tRNA ligase, *Olsenella* sp. oral taxon/Sulfotransferase family protein, *Gemmatimonadetes* bacterium/Sulfate adenylyltransferase, *Paenimaribius caenipelagi* |
| Thios-10 | DYDKKSW | Formate dehydrogenase-N subunit alpha fdnG, *Paracoccus* sp. JC/phosphate ABC transporter permease PstA, *Comamonas testosteroni* |
| Thios-11 | DKSQYKD | Methyltransf_25 domain-containing protein, *Gammaproteobacteria* bacterium/Thiol oxidoreductase, Alphaproteobacteria bacterium/3-oxoacyl-[acyl-carrier-protein] synthase, *Chiloscyllium punctatum*/Electron transfer flavoprotein alpha, *Odoribacter splanchnicus*/beta-subunit, *Pelosinus* fermentans B4/Aspartyl-tRNA synthetase, *Trichoderma harzianum* |
| Thios-12 | LLAEWK | Flavin-dependent thymidylate synthase, *Thermococcus kodakarensis*/Indolepyruvate ferredoxin oxidoreductase family protein, *Luteimonas* sp. H23/N-acetyltransferase domain-containing protein, *Saccharopolyspora rectivirgula*/Alcohol dehydrogenase, *Drosophila erecta*/ATPase, *Actinomyces oris*/DNA polymerase III subunit alpha, *Cobetia crustatorum*/ Chemotaxis protein methyltransferase, *Vibrio alginolyticus* |
| Thios-13 | LLNEWK | Nitrogen assimilation regulatory protein NtrX, *Wohlfahrtiimonas chitiniclastica*/Esterase/lipase/thioesterase family protein, hydrothermal vent metagenome/NAD(P)H-quinone oxidoreductase subunit 2, *Frullania nodulosa* /Peptidase_MA_2 domain-containing protein, *Novibacillus thermophilus*/Phenylalanine--tRNA ligase alpha subunit, *Marinithermus hydrothermalis*/DNA mismatch repair protein MutS, *Desulfovibrio vulgaris* |

| | | |
|---|---|---|
| Thios-14 | LVAEWK | FAD/NAD(P)-binding oxidoreductase family protein, *Artemisia annua*/Adenylosuccinate lyase, *Capnocytophaga* sp. oral taxon/Leucine--tRNA ligase, *Pseudoflavonifractor* sp. An176/Acetolactate synthase, *Rhizobiales* bacterium 24-66-13/ Peptidase_M23 domain-containing protein, *Marinobacter flavimaris* |
| Thios-15 | LVSEWK | Thiol: disulfide interchange protein, *Ottowia* sp./FMN-dependent NADH-azoreductase, *Erwinia persicina*/DNA polymerase, *Candidatus Thermochlorobacteriaceae* bacterium GBChlB/DNA helicase, candidate division TM6 bacterium/Lysine--tRNA ligase, *Amnibacterium* sp. M8JJ-5/Adenylosuccinate lyase, *Thermaurantimonas aggregans*/Glycine dehydrogenase (decarboxylating), *Rhodococcus* sp. LHW51113/Anthranilate phosphoribosyltransferase, *Magnetospirillum gryphiswaldense* |
| Thios-16 | LLQEWK | Orotate phosphoribosyltransferase, *Enterococcus faecalis*/Imidazole glycerol phosphate synthase subunit HisH, *Exiguobacterium* sp. AB2/Aspartate aminotransferase, *Smittium megazygosporum*/Oxidoreductase, aldo/keto reductase family, *Pseudomonas yamanorum*/Methyltransferase, *Cyanobacteria* bacterium UBA11370/Orotate phosphoribosyltransferase, *Enterococcus faecalis*/Acetolactate synthase, *Chloroflexi* bacterium/Glycerol kinase, *Sphingobacterium* sp. SSI9/DNA mismatch repair protein MutS, *Spirochaetes* bacterium/Argininosuccinate lyase, uncultured archaeon |
| Thios-17 | LWPWDT | SDR family NAD(P)-dependent oxidoreductase, *Micromonospora* sp. PPF5-17/Hist_deacetyl domain-containing protein, *Coccomyxa subellipsoidea* strain C-169 |
| Thios-18 | LSEWK | 4Fe-4S ferredoxin-type domain-containing protein, marine sediment metagenome/DNA ligase, *Yarrowia lipolytica*/N-acetyl-gamma-glutamyl-phosphate reductase, *Bacillus anthracis* |
| Thios-19 | LYSKY | Aminomethyltransferase, *Prochlorococcus marinus* subsp. *pastoris*/NADH-quinone oxidoreductase subunit J, *Bacillus thuringiensis* subsp. *Darmstadiensis*/Ferredoxin-dependent glutamate synthase, *Antithamnion* sp. (Red alga)/NADH-ubiquinone oxidoreductase chain 1, *Citrullus lanatus* (Watermelon) (Citrullus vulgaris)/NADH-quinone oxidoreductase subunit J, *Bacillus cereus* Rock4-2/Pyruvate dehydrogenase E1 component subunit beta, '*Echinacea purpurea*' witches'-broom phytoplasma/Thiol-disulfide oxidoreductase, *Halobacillus salinus*/Ferrochelatase, (Heme synthase) (Protoheme ferro-lyase) hemH, *Brevibacillus nitrificans*/Citrate synthase, *Eubacterium* sp./ATP-dependent RNA helicase RhlE, *Marinobacter* sp. C1S70/DNA helicase, *Thermovibrio guaymasensis* |
| Thios-20 | VVGEWK | Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex, *Anaerotignum neopropionicum*/Sulfotransfer_1 domain-containing protein, *Actinobacteria* bacterium/Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex, *Anaerotignum neopropionicum*/Succinate--CoA ligase [ADP-forming] subunit beta, *Roseovarius indicus*/Isoleucine--tRNA ligase, *Parabacteroides distasonis*/ATPase, *Pseudobutyrivibrio* sp. AR14/Electron transfer flavoprotein, alpha subunit, *Caballeronia sordidicola*/ATP-dependent RNA helicase, *Erinaceus europaeus*/DNA topoisomerase 1, *Flavobacteriales* bacterium |
| Thios-21 | PWDQVK | Acetyl-coenzyme A synthetase, *Methanomicrobiales* archaeon HGW-Methanomicrobiales-6/Amino acid permease family protein, *Acinetobacter baumannii* 1571545/His-tRNA synthetase, *Teinoptila guttella* |
| *Thios-22* | *NYSELYAK* | *dicarboxylate transporter, DctP subunit, Enterococcus sp. 9E7_DIV0242* |
| Thios-23 | LYAKL | Chorismate synthase, *Helicobacter pylori* strain G27/DNA-directed RNA polymerase subunit beta, *Deinococcus geothermalis* strain DSM 11300/Proline--tRNA ligase, *Ehrlichia ruminantium* strain Gardel |
| Thios-24 | K(+42.01)EWK | NADH-quinone oxidoreductase subunit N, *Azobacteroides pseudotrichonymphae genomovar.* CFP2/Periplasmic nitrate reductase, *Nautilia profundicola* strain BAA-1463/30S ribosomal protein S15, *Sulfurisphaera tokodaii* strain 16993//Putative aldehyde dehydrogenase-like protein, *Schizosaccharomyces pombe* strain 972 |
| Thios-25 | DYLYK | Argininosuccinate synthase, *Clostridium botulinum* strain 19397/Glucose-1-phosphate adenylyltransferase, *Haemophilus influenzae* strain ATCC 51907/Dipeptidyl peptidase 4, *Felis catus* (Cat) (Felis silvestris catus)/DNA-directed RNA polymerase subunit beta', RNAP subunit beta', *Ureaplasma parvum* serovar 3 strain ATCC 700970/Oligoendopeptidase, *Lactobacillus casei*/Carbon monoxide dehydrogenase, *Desulfopila* sp. IMCC35006/Aspartate--tRNA ligase, *Firmicutes* bacterium |
| Thios-26 | A(+42.01)LAESVK | Ferrochelatase, (Heme synthase) (Protoheme ferro-lyase) hemH, *Acidithiobacillus* sp. SH/Alcohol dehydrogenase, *Scaptomyza albovittata* (Fruit fly)/Peptidyl-prolyl cis-trans isomerase, *Aliivibrio fischeri* strain MJ11/Peptidyl-prolyl cis-trans isomerase, *Aliivibrio salmonicida* strain LFI1238 |
| Thios-27 | LYDYK | Argininosuccinate synthase, *Bifidobacterium longum* strain NCC 2705/ATP-dependent helicase/nuclease subunit A, *Enterococcus faecalis* strain ATCC 700802 |
| Thios-28 | YSELYAK | Peptidase_M3 domain-containing protein, *Clostridiales* bacterium/Acetyl-coenzyme A synthetase, *Rhodoblastus acidophilus*/Oxidoreductase, 2OG-Fe(II) oxygenase family, *Penicillium digitatum* strain Pd1 |
| Thios-29 | A(+42.01)LGGAC(+57.02)C(+57.02)L; A(+42.01)LGGALC(+57.02)C(+57.02) | Radical SAM protein (iron-sulfur cluster binding), *Azospirillum* sp. TSO22-1/DsrB, partial, uncultured bacterium (Marine deep biosphere microbial communities assemble in near-surface sediments)/amino acid permease, *Planctomycetes* bacterium Poly41/NADH-quinone oxidoreductase subunit M, *Lysobacter lycopersici*/DsrB, partial, uncultured bacterium/ respiratory nitrate reductase subunit gamma, *Streptomyces* sp. LHW50302 |
| Thios-30 | LPEWK | 2Fe-2S ferredoxin-type domain-containing protein, *Actinobacteria* bacterium/Fumarate hydratase class II, *Streptomyces* sp. CB01580/N-acetyltransferase domain-containing protein, *Treponema* sp./Glucose-6-phosphate isomerase, *Candida albicans* strain SC5314/Glutamate--tRNA ligase, *Cupriavidus pinatubonensis* strain JMP/Na_H_Exchanger domain-containing protein, *Streptomyces klenkii*/Probable glycine dehydrogenase (decarboxylating) subunit 1, *Thermus thermophilus* strain SG0.5JP17-16/Succinyl-CoA synthetase, beta subunit, *Chloracidobacterium thermophilum* strain B/ Fumarate hydratase class II, *Thiocapsa rosea*/Sulfurtransferase, *Bacillus* bacterium/Ferredoxin-NADP reductase, *Arthrobacter crystallopoietes*/Adenylosuccinate lyase, *Pasteurellaceae* bacterium/Heme chaperone HemW, *Rhodobacterales* bacterium/Thiosulfate sulfurtransferase SseA, *Flavobacteriales* bacterium ALC-1/Sulfite reductase, ferredoxin dependen, *Oscillatoriales cyanobacterium* USR001/ArsR family transcriptional regulator, *Mesorhizobium* sp. SEMIA 3007 |
| Thios-31 | VLAEWK | Bifunctional ligase/repressor BirA (Biotin--[acetyl-CoA-carboxylase] ligase, *Desulfocucumis palustris*/Ferredoxin--nitrite reductase, *Synechococcus* sp. Lanier/Branched-chain-amino-acid aminotransferase-like protein 1, *Arabidopsis thaliana*/Acetyl-CoA acetyltransferase, *Mycolicibacterium tokaiense*/Ammonium transporter, *Acetobacter aceti* |
| Thios-32 | VWDLK | DNA ligase, *Aquifex aeolicus* strain VF5/Oxidoreductase, *Stenotrophomonas maltophilia*/SDR family oxidoreductase, *Stenotrophomonas maltophilia*/NAD(P)-dependent oxidoreductase, *Variovorax* sp. T529/Methyltransferase, *Leptospira haakeii*/ABC transporter permease, *Anaerobacillus arseniciselenatis*/NAD(P)-dependent dehydrogenase (Short-subunit alcohol dehydrogenase family), *Stenotrophomonas* sp. AG209/DNA helicase, *Hungateiclostridium straminisolvens*/Phenylalanine--tRNA ligase, *Gemmatimonadetes* bacterium/Thiamine-monophosphate kinase, *Blastocatellia* bacterium |
| Thios-33 | A(+42.01)YTVSDQQL | acetyl-CoA synthetase-like protein, *Wolfiporia cocos* MD-104 SS10/aminopeptidase N, *Idiomarina seosinensis*/peptidylprolyl isomerase, *Novimethylophilus kurashikiensis*/aminopeptidase N, *Idiomarina seosinensis* |
| Thios-34 | LWSEEL | Fe-S oxidoreductase, *Marinilabiliales* bacterium/Formate acetyltransferase, *Escherichia coli* DEC6A/PFL-like enzyme TdcE (Keto-acid formate acetyltransferase) (Keto-acid formate-lyase) (Ketobutyrate formate-lyase, KFL, EC 2.3.1.-) (Pyruvate formate-lyase, PFL, EC 2.3.1.54) , *Escherichia coli* strain K12 |
| Thios-35 | SSAKDYK | ArsR family transcriptional regulator, *Virgibacillus soli*/Rhodanese-related sulfurtransferase, *Bacillus simplex* |
| Thios-36 | VEDLESVGK | SDR family NAD(P)-dependent oxidoreductase, *Salinigranum* sp. YJ-53/methyltransferase, *Leptotrichia* sp. oral taxon 847 |
| Thios-37 | LEDLWK | Methionine--tRNA ligase, *Sulfurospirillum* sp. UBA12182/Thioredoxin domain-containing protein, *Clastoptera* |

| | | |
|---|---|---|
| | | *arizonana*/Adenylosuccinate synthetase, AMPSase, *Gemella* sp. WT2a/Threonine--tRNA ligase, *Candidatus Aegiribacteria* bacterium MLS_C/Diaminopimelate decarboxylase, *Idiomarina salinarum* |
| Sulfi-1 | DSKYGY | Methionine aminopeptidase, *Acetivibrio ethanolgignens*/Peptidylprolyl isomerase, *Plasmodium ovale* |
| Sulfi-2 | DKAQKSEW | Sulfurtransferase, bacterium GenBank: TNE68199.1/ferrochelatase, *Microbulbifer aggregans*/acyl-CoA dehydrogenase, *Marivita* sp. XM-24bin2 |
| Sulfi-3 | LPLDTN | Dihydrolipoamide dehydrogenase, *Brevibacterium antiquum*/Mannose-1-phosphate guanylyltransferase (GDP), *Phyllobacterium sophorae*/mannose-6-phosphate isomerase, type 2, *Phyllobacterium* sp. OV277/Putative oxidoreductase ferredoxin-type protein, hydrothermal vent metagenome/Oxidoreductase, short chain dehydrogenase/reductase family, hydrothermal vent metagenome/Aspartate--tRNA ligase, *Moranella endobia* strain PCIT/NAD(P)-dependent oxidoreductase, *Burkholderia* sp. AU18528/Glycerol-3-phosphate dehydrogenase, *Helicobacter pylori*/(Fe-S)-binding protein, *Helicobacter pylori*/ATP-dependent DNA helicase II subunit 2, *Hyalomma excavatum*/Long-chain-fatty-acid--CoA ligase, *Pseudomonas syringae* pv. aptata 1/Peptidase_M28 domain-containing protein, *Dibothriocephalus latus* |
| Sulfi-4 | EVKPTKL | Lysine--tRNA ligase, *Saccharicrinis carchari*/Fructose-1,6-bisphosphatase class 1, *Rhodovulum* sp. PH10/30S ribosomal protein S9, *Candidatus Peregrinibacteria* bacterium CG1/Ferrochelatase hemH, *Marivivens niveibacter*/dehydratase family protein, *Actinobacteria* bacterium TMED270/Alanine--tRNA ligase, *Euryarchaeota* archaeon |
| Sulfi-5 | Y(+42.01)SSAP C(+57.02) | Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta, *Anthoceros angustus* (Hornwort) (Anthoceros formosae)/Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta, *Deparia lancea* (False spleenwort) (Asplenium lanceum)/Na(+)/H(+) antiporter NhaA, *Komagataeibacter sucrofermentans*/Formate dehydrogenase, *Methanoculleus bourgensis* |
| Sulfi-6 | LAKYK | 30S ribosomal protein S4, *Sulfolobus acidocaldarius*/Peptidyl-prolyl cis-trans isomerase D, *Candida glabrata*/ Nitrogenase molybdenum-iron protein alpha chain, nifD, *Rhodobacter capsulatus*/Adenylosuccinate synthetase, AMPSase, *Komagataella phaffii*/Aspartate carbamoyltransferase, *Buchnera aphidicola* subsp. *Schizaphis graminum* strain Sg/Sulfate adenylyltransferase, *Debaryomyces hansenii*/Aminotransferase, *Oscillibacter* sp./Aspartokinase, *Chryseobacterium* sp. 36-9 |
| Sulfi-7 | K(+42.01)ELVV K | Nitric oxide reductase transcriptional regulator NorR, *Providencia heimbachae*/Nitrite reductase (cytochrome; ammonia-forming), *Moorea producens*/4Fe-4S ferredoxin, *Nitrospirae* bacterium GWD2/NADPH-dependent 2,4-dienoyl-CoA reductase, sulfur reductase, *Paramaledivibacter caminithermalis*/Nitric oxide reductase transcriptional regulator NorR, *Proteus mirabilis*/Iron-sulfur cluster repair di-iron protein, *Niastella koreensis*/50S ribosomal protein L13, *Acholeplasma laidlawii* strain PG-8A/Thioredoxin domain-containing protein, marine sediment metagenome/Pyruvate kinase, *Lactobacillus parabrevis*/4Fe-4S ferredoxin, *Nitrospirae* bacterium GWD2/Peptidyl-prolyl cis-trans isomerase, *Streptosporangium roseum*/Aconitate hydratase, Aconitase, *Gemmatimonadetes* bacterium |
| Sulfi-8 | L(+42.01)KVSW | NADH-quinone oxidoreductase subunit, *Buchnera aphidicola* subsp. *Baizongia pistaciae* strain Bp/Thiol: disulfide interchange protein DsbD, *Shewanella sediminis* strain HAW-EB3/Thiol: disulfide interchange protein, *Arcobacter ebronensis*/DNA repair protein, *Ashbya gossypii*/Aminopeptidase, *Teladorsagia circumcincta*/NADH-ubiquinone oxidoreductase chain 5, *Agarophyton chilensis*/Cytochrome P450, *Nonomuraea* sp. KC333 |
| ES-1 | K(+42.01)LATP N | Ferredoxin-type protein NapG, *Helicobacter* sp./Proline--tRNA ligase, *Pleurocapsa* sp. CCALA 161/Sulfite reductase, *Methylomonas lenta*/Sulfite reductase (NADPH), *Methylobacter tundripaludum*/Acetylornithine aminotransferase, *Phenylobacterium hankyongense*/50S ribosomal protein L14, *Pricia antarctica*/UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase, *Flavobacteriaceae* bacterium CG02/ATP synthase subunit beta, *Alteromonas* sp./A Carbamoyl-phosphate synthase large chain, *Desulfococcus oleovorans* |
| ES-2 | LAEGTAL | DNA topoisomerase 1, *Halomonas elongata*/DNA mismatch repair protein MutS, *Tepidamorphus gemmatus*/Glycerol-3-phosphate dehydrogenase (NAD(P)(+)), *Martelella* sp. AD-3/NifU-like protein, *Methylorubrum extorquens*/Phosphoribosylaminoimidazole-succinocarboxamide synthase, *Hwanghaeicola aestuarii*/UDP-glucose 6-dehydrogenase, Actinobacteria bacterium 69-20 |
| ES-3 | PWDKVK | Peptidyl-prolyl cis-trans isomerase, *Crocinitomicaceae* bacterium/Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha, *Clostridium vincentii*/Methionine--tRNA ligase, *Metallosphaera sedula*/Leucine--tRNA ligase, *Candidatus Marinimicrobia* bacterium/Sulfate adenylyltransferase, Deltaproteobacteria bacterium/NADH dehydrogenase (Ubiquinone) 1 alpha subcomplex 4, *Cryptococcus depauperatus* CBS 7855 |
| ES-4 | T(+42.01)TYSK EY | Adenine phosphoribosyltransferase, termite gut metagenome/Methylenetetrahydrofolate reductase, *Paenibacillus mucilaginosus* K02 |
| ES-5 | TGVAQDVQ | Sugar ABC transporter permease, *Mesorhizobium* sp. LSHC420B00 |
| ES-6 | LNDSWK | Acetylornithine aminotransferase, *Pseudomonas* sp. HUK17/FAD-binding oxidoreductase, *Amycolatopsis alba* |
| Sulfa-1 | AQYEVEAKQ | Sugar transporter, *Sparassis crispa*/SDR family NAD(P)-dependent oxidoreductase, *Paenibacillus durus*/long-chain-fatty-acid--CoA ligase, *Paraburkholderia* sp. UYCP14C |
| Sulfa-2 | L(+42.01)SPLK P | 3-isopropylmalate dehydrogenase, *Streptococcus thermophilus*/Acetyl-CoA acetyltransferase, *Fusobacterium nucleatum* subsp. *polymorphum*/NH(3)-dependent NAD(+) synthetase, *Pusillimonas* sp. YR330/Dipeptide epimerase, *Sphingomonas* sp. SRS2/NH(3)-dependent NAD(+) synthetase, *Legionella busanensis*/Sulfite reductase [NADPH] flavoprotein alpha-component, *Serratia symbiotica* str. 'Cinara cedri'/DNA helicase, *Mycoplasma haemocanis* strain Illinois/Cytidylate kinase, *Xanthomonadaceae* bacterium NML95-0200/Phenylalanine--tRNA ligase beta subunit, *Comamonas* sp. SCN 65-56/Cytidylate kinase, *Streptococcus pyogenes*/ABC transporter permease, *Anaerotruncus* sp. G3/Malic enzyme, *Ashbya aceri*/DNA helicase, *Candidatus Raymondbacteria* bacterium RIFOXYD12/Amino acid ABC transporter substrate-binding protein, *Shewanella fodinae*/Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial, *Oryzias latipes* (Japanese rice fish) |
| Sulfa-3 | QNYKK | Orotate phosphoribosyltransferase, Deltaproteobacteria bacterium/Sugar ABC transporter permease, *Vagococcus elongatus*/ Aconitate hydratase, *Hydrogenothermus* sp. |
| Sulfa-4 | LDSKYGY | Peptidylprolyl isomerase, *Plasmodium malariae*/Sugar ABC transporter permease, *Atopobacter* sp. AH10 |
| Sulfa-5 | TDYKKY | Methionine--tRNA ligase, *Epulopiscium* sp./Acetyltransf_6 domain-containing protein, *Thiothrix lacustris* |
| Sulfa-6 | LVADWK | Thiol: disulfide interchange protein, Gammaproteobacteria bacterium/N-acetyltransferase, *Pseudomonas frederiksbergensis*/FAD_binding_3 domain-containing protein, *Archangium* sp./Methyltranfer_dom domain-containing protein, Parcubacteria group bacterium |

794

795

796

797

798 **Table 2. Amino acid relative abundance in Sammox-driven prebiotic**
799 **proto-anabolic networks (SPPN). Summary of the average amino acid relative**
800 **abundance (molar ratio) of thiosulfate, sulfite, elemental sulfur, and**
801 **sulfate-fueled SPPN. The associated errors are standard deviations of three**
802 **replicates.**

| Amino acids | Molar ratio (%) | | | | |
|---|---|---|---|---|---|
| | Thiosulfate-fueled SPPN | Sulfite-fueled SPPN | Elemental sulfur-fueled SPPN | Sulfate-fueled SPPN | Average |
| L-Serine | 7.04 ± 2.47 | 10.43 ± 1.24 | 17.69 ± 1.60 | 27.28 ± 1.71 | 15.61 ± 5.5 |
| Glycine | 15.32 ± 3.73 | 14.99 ± 1.22 | 16.04 ± 2.39 | 25.96 ± 1.23 | 18.08 ± 3.15 |
| L-Aspartic acid | 6.02 ± 0.79 | 8.12 ± 0.78 | 8.64 ± 0.73 | 10.12 ± 0.46 | 8.23 ± 0.92 |
| L-Glutamic acid | 19.71 ± 2.81 | 3.20 ± 0.48 | 6.43 ± 0.15 | 3.96 ± 0.12 | 8.32 ± 4.55 |
| L-Threonine | 2.35 ± 0.43 | 3.41 ± 0.89 | 5.06 ± 1.60 | 8.36 ± 0.16 | 4.79 ± 1.53 |
| L-Alanine | 10.26 ± 1.11 | 13.95 ± 2.78 | 11.89 ± 1.97 | 17.16 ± 1.61 | 13.32 ± 1.79 |
| L-Proline | 3.81 ± 0.53 | 5.18 ± 1.13 | 2.26 ± 0.15 | 5.28 ± 0.33 | 4.13 ± 0.88 |
| L-Ornithine | 10.56 ± 1.76 | 12.86 ± 2.76 | 11.22 ± 0.50 | 14.08 ± 0.44 | 12.18 ± 1.03 |
| L-Cysteine | 3.77 ± 0.53 | 4.78 ± 1.52 | 2.94 ± 0.13 | 7.04 ± 0.32 | 4.63 ± 1.02 |
| L-Lysine | 7.61 ± 0.13 | 6.64 ± 1.86 | 5.35 ± 0.97 | 7.92 ± 0.28 | 6.88 ± 0.71 |
| L-Tyrosine | 4.14 ± 0.44 | 5.55 ± 1.24 | 4.42 ± 0.14 | 10.56 ± 0.45 | 6.17 ± 1.76 |
| L-Valine | 4.65 ± 0.31 | 4.81 ± 0.99 | 3.89 ± 0.23 | 6.60 ± 0.22 | 5.35 ± 0.62 |
| L-Leucine | 4.39 ± 0.61 | 5.08 ± 1.06 | 4.16 ± 0.83 | 7.01 ± 0.26 | 5.16 ± 0.74 |

803