Origin of first cells at terrestrial, anoxic geothermal fields

Armen Y. Mulkidjian, Andrew Yu. Bychkov, Daria V. Dibrova, Michael Y. Galperin, and Eugene V. Koonin

School of Physics, University of Osnabrück, D-49069 Osnabrück, Germany; A. N. Belozersky Institute of Physico-Chemical Biology and Schools of *Geology and Bioengineering and Bioinformatics, Moscow State University, Moscow 119992, Russia; and *National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894

All cells contain much more potassium, phosphate, and transition metals than modern (or reconstituted primeval) oceans, lakes, or rivers. Cells maintain ion gradients by using sophisticated, energy-dependent membrane enzymes (membrane pumps) that are embedded in elaborate ion-tight membranes. The first cells could possess neither ion-tight membranes nor membrane pumps, so the concentrations of small inorganic molecules and ions within protocells and in their environment would equilibrate. Hence, the ion composition of modern cells might reflect the inorganic composition of the habitats of protocells. We attempted to reconstitute the "hatcheries" of the first cells by combining geochemical analysis with phylogenomic scrutiny of the inorganic requirements of universal components of modern cells. These ubiquitous, and by inference primordial, proteins and functional systems show affinity to and functional requirement for K⁺, Zn²⁺, Mn²⁺, and phosphate. Thus, protocells must have evolved in habitats with a high K⁺/Na⁺ ratio and relatively high concentrations of Zn, Mn, and phosphorous compounds. Geochemical reconstruction shows that the ionic composition conducive to the origin of cells could not have existed in marine settings but is compatible with emissions of vapor-dominated zones of inland geothermal systems. Under the anoxic, CO₂-dominated primordial atmosphere, the chemistry of basins at geothermal fields would resemble the internal milieu of modern cells. The precellular stages of evolution might have transpired in shallow ponds of condensed and cooled geothermal vapor that were lined with porous silicate minerals mixed with metal sulfides and enriched in K⁺, Zn²⁺, and phosphorous compounds.

prebiotic chemistry | abiotic photosynthesis | hydrothermal alteration | origin of life | Na⁺/K⁺ gradient

The utility of the geological record for reconstruction of the habitats of the earliest life forms is limited. Because of the heavy impact bombardment, the Earth surface underwent major changes approximately 3.8 to 3.9 Gigayears (Gyr) ago, so that only few rock samples are older than 4.0 Gyr (1, 2). Diverse recent data indicate that life might be older than the oldest known rocks (2). If life originated in the Hadean, finding any geological traces of the first life forms is unlikely. In 1926, Archibald Macallum noted that, although similarities between seawater and organismal fluids, such as blood and lymph, indicate that the first animals emerged in the sea, the inorganic composition of the cell cytosol dramatically differs from that of modern sea water (3). Macallum insightfully pointed out that "the cell... has endomembranes transmitted from a past almost as remote as the origin of life on earth." Thus, in our inference of the features of the primordial organisms and their environment, we are left with the biological record which, given the evolutionary continuity, is as old as life itself. The ideas of Macallum (3) can be generalized in a "chemistry conservation principle" (4): the chemical traits of organisms are more conservative than the changing environment and hence retain information about ancient environmental conditions. Chemistry conservation is manifest, for example, in the highly reduced state of the cell interior even in those organisms that dwell in oxygenated habitats (4). The reduced state of the cytoplasm indicates that the major biochemical pathways were fixed before the atmosphere became oxygenated as a result of the activity of cyanobacteria approximately 2.4 Gyr ago (5), so that substantial modification of these pathways in response to the oxygenation of the atmosphere was impossible. Instead, cellular life forms have evolved numerous energy-requiring membrane transport systems to sustain redox and (electro)chemical gradients between their interior and the environment.

It stands to reason that simultaneous consideration of various boundary conditions has the potential to eliminate most of the vast number of scenarios for the early evolution of life that appear possible in principle (4). Under this premise, we have previously addressed diverse facets of the early life problem from the viewpoint of photochemistry (6), comparative genomics (7–9), and energetics (10, 11). The principle of chemistry conservation can be used as an additional major constraint for reconstructing primordial environmental conditions in the absence of reliable geological record. For example, ancient, ubiquitous proteins often use Zn and Mn, but not Fe, as transition metal cofactors; this preference is retained across the three domains of life (12). The abundance of Zn- and Mn-dependent enzymes during the earliest steps of evolution and the later recruitment of Fe has been inferred also from a global phylogenomic reconstruction (13). The prevalence of Zn-dependent ancestral enzymes is particularly remarkable given the low estimated concentration of Zn in the anoxic ocean of 10⁻¹² to 10⁻¹⁰ M (14, 15) and indicates that the first organisms might have dwelled in specific, Zn-enriched habitats (12, 16).

Here we combine geochemical evidence with the data on the overall ionic composition of the modern cells, with a particular emphasis on their universal preference for K⁺ ions over Na⁺ ions. Geochemical analysis shows that, contrary to the common belief that associates the origin of life with marine environments, the first cells could have emerged at inland geothermal fields within ponds of condensed and cooled geothermal vapor. Conceptually, this scenario of early evolution resembles Darwin’s “warm little pond” vision (17). Under this scenario, the ocean...
was invaded by life at a later stage, following the emergence of ion-tight phospholipid membranes.

**Results and Discussion**

**Inorganic Ion Requirements of Ubiquitous Cellular Systems.** The total intracellular content of an ion reflects the ability of the cell to accumulate this ion against the concentration gradient. In particular, Table 1 shows that concentrations of K⁺, Zn²⁺, phosphate, and several other inorganic ions in all cells are orders of magnitude higher than the levels of these ions in modern sea water, as well as in the primordial, anoxic ocean. Conversely, the magnitude lower than the levels of these ions in modern sea water, phosphate, and several other inorganic ions in all cells are orders of magnitude higher than the levels of these ions in modern sea water, as well as in the primordial, anoxic ocean. Conversely, the content of Na⁺ ions in the cells is much lower than it is in the sea water. Many halophiles that can tolerate high external levels of NaCl increase the internal K⁺ concentration up to approximately 1.0 M, to keep the internal K⁺/Na⁺ ratio high (18). Apparently, it is not so much the actual concentrations of K⁺ and Na⁺ but the K⁺/Na⁺ ratio of at least 1 that is critical for the proper functioning of the cell.

Modern cells can maintain the ionic disequilibria because their membranes are ion-tight and contain a plethora of membrane-embedded, energy-dependent ion-translocating protein complexes (i.e., ion pumps). Accordingly, cells invest large amounts of energy into sustaining the respective ion gradients. For example, neurons, even in the resting state, use approximately 20% of their ATP to maintain the K⁺/Na⁺ gradient across the membrane (19).

Under the chemistry conservation principle, the striking difference between the intracellular inorganic chemistry and the composition of sea water suggests that the first cellular organisms dwelled in specific habitats that were enriched for the elements that are prevalent in modern cells (3, 4, 12, 16, 20). A potential alternative to this explanation is that the chemical differences between the intracellular milieu and the environment are unrelated to the conditions under which the first cells evolved (21). Then, the dramatic enrichment of modern cells for K⁺, Zn²⁺, and phosphate could be viewed as a relatively late shift that came after the emergence of powerful ion-translocating membrane pumps and was driven by the growing demand of the newly evolving enzymes for particular inorganic ions as catalysts or substrates.

To distinguish between these two explanations, we turned to the proteins that are shared by (nearly) all cellular organisms related to the conditions under which the first cells. Indeed, the ion-tight membranes of modern cells are extremely complex energy conversion and transport systems that obviously are products of long evolution and could not possibly exist in the first protocells. According to the available reconstructions, the first lipids were simple and single-tailed (28–31). The experiments with such lipids compounds have shown that vesicles made of fatty acids (28, 32) or of phosphorylated isoprenoids (33) can reliably encapsulate biological polymers and even facilitate their transmembrane localization but could not prevent (almost) free exchange of small molecules and ions with the environment. Furthermore, before the emergence of diverse membrane translocators, the exchange of small molecules via leaky membranes should have been of vital importance for the first cells, which also implies that their interior was equilibrated with the surroundings, at least with respect to small molecules and ions (30, 32, 34–38).

**SI Appendix, Table S1,** lists the ion requirements and affinities of the ubiquitous proteins that represent the heritage of the LUCA and probably of protocells (7, 27). Besides the preference for Zn and Mn, which has been discussed previously (12, 16), several proteins and functional systems that can be traced back to the LUCA—and probably beyond—require K⁺, whereas none of the surveyed ancestral proteins specifically requires Na⁺. The majority of the (nearly) universal proteins that can be confidently traced to the LUCA are involved in translation, which is potassium-dependent both in bacteria (39) and in archaea (40, 41). Potassium seems to be required for at least two essential ribosomal reactions. First, K⁺ ions are needed for the peptidyl transferase center to assume its functional conformation (42). Second, our sequence and structure comparisons indicate that the key translation factors are K⁺-dependent GTPases (**SI Appendix, Figs. S1–S4 and Table S2** provide further details).

Phylogenetic analysis of GTPases shows that extensive diversification of GTPase domains antedated the LUCA (43). The K⁺-binding sites are highly conserved in diverse GTPases, indicating that they were already present in the primordial GTPase domains (**SI Appendix**). Perhaps even more telling are reconstructions showing that the peptidyl transferase center is the core, ancestral part of the ribosome (44, 45). Thus, the K⁺-dependent components of the translation system appear to stem from the protocell (or even earlier) stage of evolution. Apparently, the dominance of K⁺ over Na⁺ in modern cells, which is reverse to the case in sea water, was important also for the protocells.

The concentration of phosphate in the cytosol is at least four orders of magnitude greater than in the sea water (Table 1). Not surprisingly, the energetics of the protocells, which can be inferred from the inspection of the ubiquitous protein set, must have been based on phosphate transfer reactions and specifically

<table>
<thead>
<tr>
<th>Ion, mol/L</th>
<th>Modern sea water</th>
<th>Anoxic water of primordial ocean</th>
<th>Cell cytoplasm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na⁺</td>
<td>0.4</td>
<td>&gt;0.4</td>
<td>0.01</td>
</tr>
<tr>
<td>K⁺</td>
<td>0.01</td>
<td>−0.01</td>
<td>0.1</td>
</tr>
<tr>
<td>Ca²⁺</td>
<td>0.01</td>
<td>−0.01</td>
<td>0.001</td>
</tr>
<tr>
<td>Mg²⁺</td>
<td>0.05</td>
<td>−0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Fe⁺⁺</td>
<td>10⁻⁸ (mostly Fe⁺⁺)</td>
<td>10⁻⁴</td>
<td>10⁻⁶ to 10⁻⁴</td>
</tr>
<tr>
<td>Mn²⁺</td>
<td>10⁻⁸</td>
<td>10⁻⁶ to 10⁻⁸</td>
<td>10⁻⁶</td>
</tr>
<tr>
<td>Zn²⁺</td>
<td>10⁻⁹</td>
<td>&lt;10⁻¹²</td>
<td>10⁻³ to 10⁻⁴</td>
</tr>
<tr>
<td>Cu⁺⁺</td>
<td>10⁻⁴ (Cu⁺⁺)</td>
<td>&lt;10⁻²⁰ (Cu⁺⁺)</td>
<td>10⁻⁵</td>
</tr>
<tr>
<td>Cl⁻</td>
<td>0.5</td>
<td>&gt;0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>PO₄³⁻</td>
<td>10⁻⁶ to 10⁻⁹</td>
<td>&lt;10⁻⁵</td>
<td>~10⁻² (mostly bound)</td>
</tr>
</tbody>
</table>

The intracellular concentration is defined here as the total content of a particular element divided by the cell volume and should be discriminated from the much lower free ion concentration, which does not account for the ions that are bound to biological molecules. The reconstructed chemical composition of the anoxic ocean includes data from refs. 14, 15, 58, 141. The data on intracellular concentrations of different chemical elements are based on refs. 14, 142–145.
on hydrolysis of nucleoside triphosphates (SI Appendix, Table S1). That phosphate-based metabolism is ancestral in cellular life follows also from the results of the recent global phylogenomic analysis (13). Given that the backbones of nucleic acids contain phosphate groups, there is no doubt that phosphate was a central component of life from its inception.

However, the concentration of phosphate ions in natural aqueous systems, such as lakes or seas, could never be as high as it is inside cells because of the poor solubility of Ca and Mg phosphates. Thus, although the requirement for a high phosphate concentration in the protocols is indisputable, it remains unclear how the protocols could accumulate phosphate without tight membranes and phosphate-scavenging pumps. It has been argued that more reduced phosphorous compounds such as hypophosphate (PO₃⁻) and/or phosphate (PO₄³⁻), which are approximately 1,000 times more soluble than phosphate, could have been abundant under primordial reduced conditions (46–49).

Hence a major conundrum:

a) Intracellular concentrations of key ions, in particular K⁺, Zn²⁺, and phosphate, are several orders of magnitude higher compared with sea water, both extant and that of Hadean ocean (according to the available reconstruction; Table 1);

b) (Nearly) universal, and by inference primordial, proteins and functional systems show affinity to and functional requirement for K⁺, Mg²⁺, Zn²⁺, Mn²⁺, and phosphate, but not Na⁺ (SI Appendix, Table S1); and
c) It is extremely unlikely that protocols possessed ion-tight membranes with built-in ion pumps.

Given these observations and inferences, it appears most likely that protocols evolved in habitats characterized by a high K⁺/Na⁺ ratio and relatively high concentrations of Zn²⁺, Mn²⁺ and phosphorous compounds.

Vapor-Dominated Zones of Terrestrial Geothermal Systems as Possible Hatcheries of First Cells. Is it possible to envision any natural habitats with high levels of transition metals and phosphorous compounds, as well as a K⁺/Na⁺ ratio substantially greater than 1?

As argued previously (10–12), high concentrations of transition metals, such as Zn and Mn, are found only where extremely hot hydrothermal fluids leach metal ions from the crust and bring them to the surface. Such thermal systems operate either on the sea floor (50, 51), or at sites of continental (i.e., terrestrial) geothermal activity where the metal ions are carried not only by hot fluids, but also by steam (52, 53).

Phosphate concentrations are low both in the sea water (Table 1) and in the fluids of the deep sea hydrothermal vents (~0.5 μM) (50). The content of phosphorous compounds is higher in terrestrial thermal springs, where it varies within a broad range, reaching 60 to 70 μM in some Yellowstone springs (54) and as much as 1 mM in the acidic mud pots of Kamchatka (55). In an attempt to discriminate phosphate from phosphate in field samples, Pech et al. have found comparable amounts of phosphate and phosphate in a pristine geothermal pool at Hot Creek Gorge near Mammoth Lakes, CA, which is fed by hot, bicarbonate-rich geothermal waters (56). The discovery of highly soluble phosphate in a modern geothermal pool can at least partly account for high amounts of phosphorous in the discharges of terrestrial geothermal systems. Furthermore, this finding could explain why diverse prokaryotes possess systems of hypophosphite and phosphate oxidation (57).

The high K⁺/Na⁺ ratio should be taken as the key search criterion because accumulation of transitional metals or phosphorous compounds is conceivable in primordial evaporating water basins; evaporation, however, cannot affect the K⁺/Na⁺ ratio. No marine environment with a K⁺/Na⁺ ratio greater than 1 has ever been described or reconstructed to our knowledge. In trapped samples of Archaean seawater, the K⁺/Na⁺ ratio is approximately 0.025 and is similar to that in modern oceans (58). Arguably, this low K⁺/Na⁺ ratio was established in the ocean shortly after its formation, when it was still too hot to be compatible with life (2, 58). The K⁺/Na⁺ ratio is similarly low in hydrothermal fluids of marine hot vents because these vents are fed predominantly by sea water (50).

Terrestrial aqueous systems, which are mostly fed by water from rain and snow, are more variable with respect to the K⁺/Na⁺ ratio. Generally, the concentrations of K⁺ and Na⁺ in rivers and lakes are much less than 1 mM, and the K⁺/Na⁺ ratio is in the range of 0.1 to 1.0, although in streams that interact with potassium-rich igneous rocks, this ratio can reach 2 or 3 (59, 60). At sites of inland geothermal activity, the levels of K⁺ and Na⁺ are higher as a result of extensive leaching of metals from rocks by hot, carbonate-enriched waters, and the K⁺/Na⁺ ratio varies within a broad range (54, 55) owing to the intrinsic heterogeneity of such systems. The heterogeneity is a result of the boiling of the ascending hot hydrothermal fluids at shallower depths followed by separation of the vapor phase from the liquid phase (Fig. 1). Upon separation, gaseous compounds, such as H₂S, CO₂, and NH₃, redistribute into vapor that rises upward toward the surface. The subsurface area in which steam and gas prevail in open fractures is called the vapor-dominated zone (Fig. 1). The exhalations from vapor-dominated zones, which are enriched in

![Diagram](https://example.com/diagram.png)
H₂S, CO₂, NH₃ and metal cations, discharge as steam (i.e., fumaroles) or, after condensation, as mud pots (SI Appendix, Fig. S5) because of the silica that is also carried by the vapor (52–55, 61, 62). Numerous fumaroles and mud pots overlaying a vapor-dominated zone make a geothermal field.

The emissions from the vapor-dominated zones of inland geothermal systems are K⁺-enriched, unlike the discharges from the liquid-dominated zones, which contain much more Na⁺ than K⁺ (54, 55). To our knowledge, the causes of this enrichment have not been explicitly addressed. Comparison of the concentrations of some essential elements in the fluids of thermal springs and in the vapor of the same springs (Table 2 shows data from Kamchatka volcanic system) sheds light on the probable thermodynamics of vapor-dominated geothermal fields. As follows from the data in Table 2, the K⁺/Na⁺ ratio is, on average, higher in the vapor condensate than in the liquid. A similar dependence can be inferred from data on the two largest vapor-dominated geothermal systems simultaneously show K⁺/Na⁺ ratios much greater than 1, which ascend with the vapor, are weak acids, and their acidity is usually compensated by the interaction with basic rocks.

Methods

The pH of the discharges from the vapor-dominated zones should have been closer to neutral because both H₂S and CO₂, which ascend with the vapor, are weak acids, and their acidity is usually compensated by the interaction with basic rocks.

At neutral pH, silica would precipitate at the outlets of thermal springs and around them not as amorphous kaolinite/mud, as it does now (61), but as porous, ordered silicate minerals. Thus, the formation of clays such as smectite/montmorillonite and illite, as well as zeolites such as laumontite and natrolite, should be expected.

In the absence of oxygen on the primordial Earth, the geochemistry of vapor-dominated geothermal fields should have been quite different:

1) The pH of the discharges from the vapor-dominated zones should have been closer to neutral because both H₂S and CO₂, which ascend with the vapor, are weak acids, and their acidity is usually compensated by the interaction with basic rocks;

2) At neutral pH, silica would precipitate at the outlets of thermal springs and around them not as amorphous kaolinite/mud, as it does now (61), but as porous, ordered silicate minerals. Thus, the formation of clays such as smectite/montmorillonite and illite, as well as zeolites such as laumontite and natrolite, should be expected;

3) In the absence of oxygen, sulfide ions would cause precipitation of metal sulfides, as is the case at modern deep-sea hydrothermal systems, where slowly precipitating ZnS particles form halos around the vent throats which are built of fast-precipitating sulfides of iron and copper (50, 51). At ancient geothermal fields, because of the high silica content in the

<table>
<thead>
<tr>
<th>Element</th>
<th>S6–14</th>
<th>S6–15</th>
<th>S6–16</th>
<th>S6–17</th>
<th>S6–18</th>
<th>S6–19</th>
</tr>
</thead>
<tbody>
<tr>
<td>t, °C</td>
<td>94.00</td>
<td>93.00</td>
<td>89.00</td>
<td>93.00</td>
<td>96.00</td>
<td>96.00</td>
</tr>
<tr>
<td>pH</td>
<td>0.50</td>
<td>−0.28</td>
<td>0.25</td>
<td>−0.58</td>
<td>−0.09</td>
<td>−0.30</td>
</tr>
<tr>
<td>B</td>
<td>95.109</td>
<td>54.142</td>
<td>35.927</td>
<td>72.639</td>
<td>83.813</td>
<td>133,910</td>
</tr>
<tr>
<td>Ca</td>
<td>279,893</td>
<td>121,911</td>
<td>455,703</td>
<td>213,657</td>
<td>334,430</td>
<td>168,640</td>
</tr>
<tr>
<td>Fe</td>
<td>384,075</td>
<td>174,308</td>
<td>245,163</td>
<td>258,688</td>
<td>446,416</td>
<td>250,982</td>
</tr>
<tr>
<td>K</td>
<td>89,606</td>
<td>138,879</td>
<td>22,881</td>
<td>882,720</td>
<td>86,835</td>
<td>155,190</td>
</tr>
<tr>
<td>Mg</td>
<td>168,491</td>
<td>68,883</td>
<td>118,968</td>
<td>76,868</td>
<td>202,059</td>
<td>98,071</td>
</tr>
<tr>
<td>Mn</td>
<td>7,355</td>
<td>2,909</td>
<td>3,358</td>
<td>3,942</td>
<td>9,424</td>
<td>4,325</td>
</tr>
<tr>
<td>Na</td>
<td>128,609</td>
<td>100,599</td>
<td>79,224</td>
<td>479,027</td>
<td>143,699</td>
<td>121,597</td>
</tr>
<tr>
<td>Ni</td>
<td>140</td>
<td>89</td>
<td>82</td>
<td>96</td>
<td>593</td>
<td>67</td>
</tr>
<tr>
<td>P</td>
<td>7,399</td>
<td>8,615</td>
<td>6,434</td>
<td>33,689</td>
<td>7,568</td>
<td>9,163</td>
</tr>
<tr>
<td>Ti</td>
<td>9,170</td>
<td>2,345</td>
<td>2,300</td>
<td>3,106</td>
<td>8,533</td>
<td>7,874</td>
</tr>
<tr>
<td>Zn</td>
<td>657</td>
<td>324</td>
<td>734</td>
<td>471</td>
<td>830</td>
<td>439</td>
</tr>
</tbody>
</table>

For Mutnovsky volcano, Kamchatka peninsula, see Methods and refs. 62, 95.
exhalations of the vapor-dominated systems, the formation of metal-sulfide–contaminated clays and zeolites rather than pure metal-sulfide precipitates should be expected.

It is generally believed that the primordial atmosphere was CO₂-dominated and that the atmospheric pressure was higher than it is now (67, 68). Both these factors would boost the transportation of diverse ions by the ascending vapor. The high CO₂ concentration would enhance the leaking from the rock by carbonate ions, whereas the high atmospheric pressure would bring the boiling isotherm (Fig. 1) closer to the surface, shorten the distance that had to be covered by the ascending vapor, and thereby increase the amount of transported inorganic ions.

In summary, the operation of geothermal systems under anoxic, CO₂-dominated atmosphere would result in vigorous discharge of neutral geothermal fluids and steam from their vapor-dominated zones; the discharges would have a K⁺/Na⁺ ratio greater than 1 and would be enriched in NH₃, H₂S, CO₂, phosphorous compounds, and transition metals. These terrestrial zones; the discharges would have a K⁺/Na⁺ ratio greater than 1 and would be enriched in NH₃, H₂S, CO₂, phosphorous compounds, and transition metals. These terrestrial fields appear to provide the best environment on the primordial Earth for the origin of protocells.

**Evolution of Protocells at Anoxic Geothermal Fields.** Fig. 2 shows a scenario for the origin of protocells at anoxic geothermal fields overlaying the vapor-dominated zone of a primordial geothermal system (as detailed in the legend to Fig. 2). Such systems should have been typical of the first Earth continent(s) that are believed to have formed from Mg-, K-rich ultramafic rocks (2, 69). The analysis of the 4.02- to 4.19-Gyr-old inclusion-bearing zircons indicates an early presence of subduction zones and, hence, the overlaying geothermal fields (70). In the absence of oxygen, the transition metals would precipitate mostly as sulfides. While ZnS and MnS precipitate slowly, Cu₂S, PbS, and FeS₂ are promptly removed by precipitation at neutral pH and at temperatures lower than 300 °C (71–73). Therefore, Cu₂S, PbS, and FeS₂ cannot spread far away from points of discharge, especially taking into account the cooling of the geothermal fluids to the ambient temperatures. In addition, Zn is much more volatile than Fe, as could be judged from the analyses of geothermal springs (Table 2) and volcanic vapor (74). Hence, far-off ponds and puddles, fed by cooled geothermal fluids and condensed vapor, would have been particularly enriched in slowly precipitating Zn²⁺ and Mn²⁺ ions, with their beds covered by clays and zeolites contaminated by sulfides and carbonates of Zn and Mn (Fig. 2A).

We hypothesize that such loose, Zn- and Mn-enriched sediments served as the cradles for protocells (Fig. 2B). The affinity of many ubiquitous proteins for Zn²⁺ and, to a lesser extent, Mn²⁺ (SI Appendix, Table S1) implies that these proteins might have evolved in such environments.

The absence of any enzymes related to autotrophy in the ubiquitous protein set (SI Appendix, Table S1) suggests that the protocells were heterotrophs, i.e., their growth depended on the supply of abiotically produced organic compounds (32, 75–77). At least two continuous, abiotic sources of such compounds would exist in the described geothermal systems. First, even in modern vapor-dominated geothermal systems, exhalations carry organic molecules that are believed to be formed, at least partly, in the process of hydrothermal alteration of ultramafic rocks (78, 79). Hydrothermal alteration occurs when iron-containing rocks interact with water at temperatures of approximately 300 °C, which is typical of terrestrial geothermal systems. Under these conditions, part of the Fe²⁺ in the rock is oxidized to Fe³⁺ yielding magnetite (Fe₃O₄). The electrons released through this reaction are accepted by protons of water yielding H₂; in the presence of water-dissolved CO₂, diverse hydrocarbons are ultimately produced (78). It could be argued that the hydrothermal rock alteration might also account for the reduction of insoluble apatite to soluble phosphate (47), explaining the presence of phosphite in the geothermal fluids (56). Similar reactions could lead to the ammonia formation (80), which might account for the high ammonia content in the exhalations of geothermal fields [as much as 130 mg/L in the mud pot solutions of Kamchatka (55)].

In addition, diverse organic molecules could be produced by abiotic photosynthesis catalyzed by ZnS and MnS particles (81–84). Such crystals are semiconductors, which can trap quanta with a λ of less than 320 nm and transiently store their energy in a form of charge-separated states, capable of reducing diverse compounds at the surface (81). Thereby, crystals of ZnS are the most powerful photocatalysts known in nature (10). Particles of ZnS can catalyze photopolymerization reactions (85) and photoreduce carbonaceous compounds to diverse organic molecules, including intermediates of the tricarboxylic acid cycle (83, 84); the highest quantum yield of 80% was observed upon reduction of CO₂ to formate (81).

Generally, two types of environments relevant for the early stages of evolution can be discriminated at primordial geothermal fields: (i) periodically wetted and illuminated mineral surfaces that could serve as templates and catalysts for diverse abiotic syntheses and (ii) geothermal pools that could serve as hatcheries of first replicating life forms (Fig. 2). At mineral surfaces of primordial geothermal fields, ammonia, sulfide, phosphate, and phosphate ions would react with carbonaceous compounds, yielding aminated, sulfurated, and phosphorylated surfaces that could serve as templates and catalysts for diverse abiotic syntheses and (ii) geothermal pools that could serve as hatcheries of first replicating life forms (Fig. 2). At mineral surfaces of primordial geothermal fields, ammonia, sulfide, phosphate, and phosphate ions would react with carbonaceous compounds, yielding aminated, sulfurated, and phosphorylated surfaces that could serve as templates and catalysts for diverse abiotic syntheses and (ii) geothermal pools that could serve as hatcheries of first replicating life forms (Fig. 2). At mineral surfaces of primordial geothermal fields, ammonia, sulfide, phosphate, and phosphate ions would react with carbonaceous compounds, yielding aminated, sulfurated, and phosphorylated surfaces that could serve as templates and catalysts for diverse abiotic syntheses and (ii) geothermal pools that could serve as hatcheries of first replicating life forms (Fig. 2). At mineral surfaces of primordial geothermal fields, ammonia, sulfide, phosphate, and phosphate ions would react with carbonaceous compounds, yielding aminated, sulfurated, and phosphorylated surface...
for the protocells within the geothermal ponds. Each such pool would “harvest,” with the help of geothermal streams and rain water, substrates from its catchment area. Only water-soluble compounds or compounds that could be carried by water (e.g., as micelles of amphiphilic molecules) could reach such ponds. This harvesting mechanism essentially excludes the interference of “tar,” which would inevitably form under conditions of abiotic syntheses (4), with the chemistry within geothermal ponds.

In the absence of an ozone shield, the protocells would need protection from the UV component of solar light (86). Both ZnS and MnS crystals efficiently scavenge UV up to approximately 320 nm (81, 87). The molar absorption coefficient of ZnS particles is approximately 2 mM cm−1 at 260 nm, at which nucleotides absorb (88). It is easy to estimate that a thin, 5-μm layer of ZnS would attenuate the UV light by a factor of 1010. Thus, even conservatively assuming a 90% porosity of ZnS-containing sediments and a 1% ZnS content in the sediments, a 5-mm layer of ZnS-containing precipitates would give the same UV protection as a greater than 100 m water column (cf. ref. 86). This is a low bound estimate because other mineral constituents of siliceous sediments would also absorb UV and protect the pro-

mordial life forms (89). Hence, a stratified system could be established within geothermal ponds, where the illuminated upper layers would be involved in the “harvesting” and production of reduced organic compounds, whereas the deeper, less productive but better protected layers could provide shelter for the protocells (Fig. 2B). The porosity of the silica minerals would enable metabolite transport between the layers. Both the light gradient and the interlayer metabolic exchange are typical of modern stratified phototropic microbial communities (90).

Thus, Hadean anoxic geothermal fields would provide:

(a) Water basins with ionic composition compatible with that of modern cells, meeting the chemistry conservation criterion;
(b) A supply of organic molecules that could fuel biosynthetic reactions;
(c) Abundant, efficient, and versatile (photo)catalysts, above all ZnS and Zn2+;
(d) Microcompartments within porous, siliceous ZnS- and MnS-containing masses.

The proposed scenario is robust because its critical parameters, such as the K+/Na+ ratio greater than 1 and the continuous supply of reduced compounds, are sustained by multiple complementary mechanisms. In particular, the high K+ levels and the K+/Na+ ratio greater than 1 would have been maintained by the K-enrichment of the primordial igneous rocks (2), by the higher mobility of K+ ions in the vapor phase (Table 2), and by ability of 2:1 clay minerals, such as smectites and illite, to select potassium over sodium (91). The only vital parameter for the model is the absence of atmospheric oxygen, which is not disputed when it comes to the first eons of Earth history (5, 67).

Furthermore, geothermal fields have autonomous heat sources and good thermal isolation provided by the air, so the temperature and chemical composition of water basins in these habitats are defined primarily by the geothermal activity and are effectively independent of the climate, potentially allowing protocells to endure climate changes or even periods of early glaciations (67). Taken together, these considerations seem to make inland anoxic geothermal fields the best incubators for the protocells among all currently known habitats on Earth.

Terrestrial Anoxic Geothermal Fields as Cradles for Earliest Life Forms? So far, we have focused on the conditions under which the protocells might have evolved, without addressing the earlier steps of evolution. Comparison of extant genomes does not directly yield information on pre-LUCA life forms. However, features of these primordial organisms can be gleaned from the analysis of those protein families that were represented in the LUCA by multiple paralogues such as GTPases or aminoacyl-tRNA synthetases (92) (SI Appendix, Table S1). Most likely, the ancestors of these protein families shared the ionic requirements of the extant family members, such as those for K+ and Zn2+.

A similar preference for Zn2+, Mn2+, and ATP as substrate is shown by viral hallmark genes (SI Appendix, Table S3). These genes encode proteins which are present in many viral families but are absent from cellular organisms and could stem from organisms that preceded the LUCA (9, 93). Thus, extending the chemistry conservation principle, we hypothesize that terrestrial geothermal fields, similar to those illustrated in Fig. 24, might have also served as the cradles of life itself, sheltering the first, precellular life forms up to the stage of the LUCA. This scenario seems to be compatible with several lines of evidence:

a) Remaining almost independent of the ambient climate, inland geothermal fields could exist for millions of years, long enough to serve as incubators not only for the protocells but also for the preceding life forms.

b) The major biochemical building blocks are derivatives of those molecules that preferentially partition to the vapor phase upon the geothermal separation, namely simple carbonaceous and phosphorous compounds, ammonia, and sulfide. In addition, the vapor phase of geothermal systems is particularly enriched in borate, the concentration of which can reach 10 mM (Table 2) (54, 94, 95) and which seems to be important for the stabilization of ribose (96, 97).

c) Geothermal fields should have offered ample opportunity for the reagents to concentrate and interact upon evaporation. Specifically, the wetted surfaces would undergo continuous drying resulting in selective accumulation of the least volatile compounds, which, in this case, would be simple amides, with boiling points of approximately 200 °C due to their ability to form strong hydrogen bonds. Formamide, the likely key building block for abiotic synthesis of nucleotides and amino acids (98–108), could form via hydrolysis of hydrogen cyanide, which is found in volcanic gases and in exhalations of geothermal fields (109). In addition, elimination of a water molecule from ammonia salts of carboxylic acids could also yield amides, in particular, formamide from amino formate. As noted earlier, exhalations of geothermal fields contain high amounts of ammonia (55); part of this ammonia is of non-sedimentary origin (110) and could have been present already in the primordial geothermal vapor. Formate and other carboxylic acids would also have been produced at anoxic geothermal fields (as detailed earlier). Hence, anoxic geothermal fields could selectively accumulate simple amides, primarily formamide, most likely mixed with water and other simple molecules in different ratios. The yield of photochemical and thermal syntheses in amide-containing solutions could be further enhanced by catalytic action of mineral surfaces. Specifically, it has been shown that silica minerals catalyze the formation of adenine and cytosine from formamide (103, 111) and that TiO2, the main component of the mineral rutile, could catalyze the formation not only of purine derivatives but also of thymine, 5-hydroxymethyluracil, and even acyclo-

nucleosides (112). Even widespread iron oxides have been shown to catalyze the synthesis of nucleobases from formamid e (113).

d) Spontaneous polymerization events, which are thermodynamically unfavorable in the bulk water, would be favored at geothermal fields. Strikingly, a thermodynamic “window” at concentrations of formamide of greater than 30% has been identified, at which polynucleotides were more stable than mononucleotides (114, 115). In addition, condensation reactions would be favored by the wet/dry cycles driven by the intrinsic pulsation of thermal springs (66), daily oscillations
of temperature and light, and the capacity of silicate minerals to serve as apt templates (116–118).

e) The exceptional photostability of biological nucleotides suggests that they could have been selected under solar UV radiation from a plethora of diverse abiotically (photo) synthesized organic compounds (6, 119–122). Analogously, photoselection might have facilitated the transition from complex mixtures of small organic molecules to the “RNA world” (123) by favoring photostable RNA-like polymers with excitonically coupled, stacked nucleotides forming Watson–Crick pairs (6, 119, 124). In addition, solar UV radiation could support primeval syntheses not only by catalyzing photopolymerization, but also by breaking the less photostable organic molecules and thus supplying building blocks for new synthetic cycles (10).

f) Under the low luminosity of the young sun (67), the daily temperature oscillations could lead to periodic freezing events, favoring the concentration of reactants, the endurance of RNA-like oligomers, and their pairing (37).

g) The Zn2+ and Mn2+ ions could shape the primeval biochemistry as selective catalysts and as stabilizers of nascent biopolymers (10, 12). It has been shown that Zn2++, to a much greater extent than any other transition metal ion, favored the formation of naturally occurring 3′–5′ phosphodiester bonds during abiotic polymerization of activated nucleotides (125).

h) Last but not least, evolution of life from the very first RNA-like molecules to the stage of protocells in the same habitats is the most parsimonious scenario: otherwise, one would have to envision mechanisms for relocation of the first precellular organisms to geothermal fields from some other location and their accommodation in new habitats.

Protocells Could Not Emerge in Marine Habitat: Late Escape of Life to the Ocean. Apparently, no marine environment could ever provide a K+/Na+ ratio of greater than 1 or concentrate phosphate up to its level in the cells. Thus, our analysis argues against the widespread belief that the first cells evolved in marine habitats. Although early evolutionary scenarios usually considered shallow seawaters where solar light was available as an energy source (116, 126), deep-sea environments have been invoked later, initially because of the protection against the hazards of the solar UV that the water column would provide to the primordial life forms. In particular, it has been estimated that the UV component should have been attenuated by a factor as high as 1016 to avoid irreparable damage to the first organisms (86). Russell and coworkers have noticed that FeS/FeS2 precipitates around hydrothermal vents form expansive honeycomb-like structures and suggested that such iron-sulfide “bubbles” could encase and protect the first life forms before the emergence of cells with modern-type membranes (127, 128). Subsequently, attention has been drawn to low-temperature vents where the hydrothermal fluids are enriched in diverse organic compounds that are formed through serpentinization, a hydrothermal alteration process that is typical of the basaltic oceanic crust (129).

The terrestrial scenario outlined here incorporates all the features of the hydrothermal vents that favor the origin and early evolution of life, and adds more (Table P1 in Summary). Our scenario includes production of organic molecules from CO2 not only in reactions of hydrothermal alteration within the rocks but also via abiotic photosynthesis at the surface. The UV protection by ZnS, MnS, and silicate minerals is much more efficient than the protection by a water column. Continental geothermal fields are even more compartmentalized than marine hydrothermal systems. Not only do they include microcompartments, such as variably hydrated pores within ZnS and MnS-containing silicate minerals, but in addition, each pond or puddle can be itself considered a separately evolving macrocompartment; occasional exchange of genetic material between these macrocompartments could be triggered by rains or overflying of the geothermal fields.

Detailed analysis of the transition from the first biomolecules to the first cells is beyond the scope of this work; it is nevertheless clear that this transition should have been accompanied by selection for increasingly tighter cellular envelopes (36–38). Increasing sequestering of primordial life forms should have followed the evolution of their metabolic pathways (36, 130) and also would protect the informational systems from external hazards (10, 12).

The dramatic difference between the ionic compositions of the cytosol and seawater (Table 1) implies that cellular organisms could invade the ocean only after the emergence of ion-tight membranes. These membranes and the appropriate ion pumps were required to maintain the intracellular chemical environment similar to that in which the protocells evolved. Being encased by ion-tight membranes and endowed with ion pumps, the first cells could invade terrestrial water basins with low K+ / Na+ ratios and then, via rivers, reach the ocean, where they would have been severely challenged by the high sodium levels. Therefore, they would require ion pumps capable of ejecting Na+ ions out of the cell against large concentration backpressure. As argued previously on the basis of phylogenomic analysis of rotary ATPases, the interplay between several Na+ pumps might have led to the emergence of membrane bioenergetics, initially in its ancestral, Na+-using form (38, 131, 132).

The proposed terrestrial origin of the first cells implies that life started not as a planetary but as a local event, confined to a long-lasting inland geothermal field or to a network of such fields at a continental volcanic system. Only the invasion of the ocean by membrane-encased organisms transformed life into a planetary phenomenon.

Conclusions

Building on the geochemical data and the results of phylogenomic analysis, we argue here that anoxic geothermal fields overlaying the vapor-dominated zones of terrestrial hydrothermal/volcanic systems could be the most suitable hatcheries for the protocells and, most likely, the preceding replicator systems. These putative cradles of life share all of the advantages of the deep sea hydrothermal vents that have been previously proposed in the same capacity (127–129), including the presence of ion-tight compartments, high concentrations of catalysts, and sources of organic matter (Table P1 in Summary). In addition, and in contrast to deep sea vents, terrestrial geothermal fields are conducive to condensation reactions and enable the involvement of solar light as an energy source and a selective factor that would favor the accumulation of nucleotides, which are particularly photostable (6, 121, 124). Also in contrast to deep sea vents, the geothermal vapor is enriched in phosphorous and boron compounds (Table 2) that could be essential for the emergence of first RNA-like oligomers (96, 97).

Reconstruction of conditions under which the first life forms might have emerged is important for experimental modeling of the origin of life (32, 37). Some of the most successful attempts to simulate primitive abiotic reactions have been conducted under conditions that are compatible with reconstructed conditions at the geothermal fields of the anoxic Earth. These promising experiments include syntheses of biologically relevant compounds in formamide solutions (98–108, 111–115), photosynthesis/photoselection of natural nucleotides (120–122, 133), montmorillonite-catalyzed formation of long RNA oligomers (118) and membrane vesicles (134), RNA polymerization in the eutectic phase in water–ice (135), abiotic UV photosynthesis of the tricarboxylic acid cycle intermediates at ZnS (83, 84) and TiO2 crystals (136), as well as UV-triggered recharging of ADP to ATP (137). Further experimental exploration of models that mimic the conditions at anoxic geothermal fields are expected to shed more light on precellular evolution.
Methods

Steam samples were collected by using a specially constructed condensing device that aimed to minimize the possible contamination from the drops of liquid phase or incomplete condensation of vapors. The thermal spring (i.e., mud pot) was covered by a vapor collector that contained a refractor to prevent the eventual contamination by drops of liquid (SI Appendix, Fig. S6) (95). The temperature was controlled by a temperature sensor; the difference between the temperature in the vent and at the wall of the collector did not exceed 1 °C. The collector was connected to a glass Allihn condenser (i.e., bulb condenser). The condenser was continuously cooled by cold water from a condenser tank. The flow rate was regulated by changing the placement of the vapor collector. The sampling conditions were chosen in such a way that the temperature of the condensate outflow did not exceed 30 °C. Accordingly, if the vapor flow was too strong, the condenser was elevated so part of the steam could escape around the edges of the collector (SI Appendix, Fig. S6). After installation at a steam vent, the collector was equilibrated for 10 min. After that, the samples were gathered in several 50-mL vials (at least two per spring) during 2 h to ensure the reproducibility of results. When checked afterward, the concentration difference between samples obtained from the same spring did not exceed 10%, whereas the concentration differences between the samples taken from different springs could vary by orders of magnitude (Table 2). The samples of the liquid phase of the same thermal springs were filtered at the spot by using 0.45-μm membrane filters. All samples were preserved by the addition of HNO3 to ∼1% of the sample volume.

The samples were later analyzed by inductively coupled plasma MS by using an Element2 (Finngan) mass spectrometer.

ACKNOWLEDGMENTS. Valuable discussions with Drs. D. A. Cherepanov, M. Eigen, R. M. Haagen, G. F. Joyce, M. J. van Kranendonk, V. N. Kompanichenko, D.-H. Lankena, D. L. Pinti, M. J. Russell, V. P. Skulachev, H.-J. Steinhoff, Szostak, N. E. Voelckerbyorkova, R. P. Williams, Y. J. Wolf and A. Yonath are greatly appreciated. The authors are thankful to Drs. A. S. Karyagina and I. Y. Nikolaeva for providing photographs of boiling mud pots. This study was supported by Deutsche Forschungsgemeinschaft (DFG) Grants DFG-Mu 1285/1-10 and DFG-436-RUS 113/9630-1 to A.Y.M., Russian Government Grant 02.740.11.5228 (to A.Y.M.), the Volkswagen Foundation (A.Y.M.), EU COST CM0902 Action (A.Y.M.), Deutscher Akademischer Austausch Dienst (D.V.D.), Russian Foundation for Basic Research 10-05-03022a to A.Y.B. and D-04-91331 to D.V.D.), and the Intramural Research Program of the National Library of Medicine at the National Institutes of Health (M.Y.G. and E.V.K.).