Bayesian analysis of Enceladus' plume data to assess methanogenesis

Antonin Affholder, François Guyot, Boris Sauterey, Régis Ferrière, Stéphane Mazevet

Title

Bayesian analysis of Enceladus' plume data to assess methanogenesis

Author list

Antonin Affholder[1,2,†], François Guyot[3,4], Boris Sauterey[1], Régis Ferrière[1,5,6,*,‡] and Stéphane Mazevet[2,*]

Affiliations

[1] Institut de Biologie de l'École Normale Supérieure, ENS, Université Paris Sciences et Lettres, Paris, France. [2] Institut de Mécanique Céleste et de Calcul des Éphémérides, Observatoire de Paris, Université Paris Sciences et Lettres, Paris, France. [3] Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, Muséum National d'Histoire Naturelle, CNRS, Sorbonne Universités, Paris, France. [4] Institut Universitaire de France (IUF). [5] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, USA. [6] International Research Laboratory for Interdisciplinary Global Environmental Studies (iGLOBES), CNRS, ENS, Université Paris Sciences et Lettres, University of Arizona, Tucson, USA.
[*] These authors jointly supervised this work.
[†] e-mail : antonin.affholder@bio.ens.psl.eu. [‡] e-mail : regis.ferriere@bio.ens.psl.eu

Abstract

Observations from NASA's Cassini spacecraft established that Saturn's moon Enceladus has an internal liquid ocean. Analysis of a plume of ocean material ejected into space suggests alkaline hydrothermal vents on Enceladus' seafloor. On Earth, such deep-sea vents harbor microbial ecosystems rich in methanogenic archaea. Here, we use a Bayesian statistical approach to quantify the probability that methanogenesis (biotic methane production) might explain the escape rates of molecular hydrogen and methane in Enceladus' plume, as measured by Cassini instruments. We find that the observed escape rates (i) cannot be explained solely by the abiotic alteration of the rocky core by serpentinization; (ii) are compatible with the hypothesis of habitable conditions for methanogens; (iii) score the highest likelihood under the hypothesis of methanogenesis, assumed the probability of life emerging is high enough. If the probability of life emerging on Enceladus is low, the Cassini measurements are consistent with habitable yet uninhabited hydrothermal vents and point to unknown sources of methane (*e.g.*, primordial methane) awaiting to be discovered by future missions.

Main Text

The NASA-led Cassini space mission has provided unprecedented insights into the Saturnian system[1]. Observations from multiple fly-bys of the moon Enceladus confirmed the existence of a global ocean[2]. Sampling of the south-pole plume and analysis by Cassini revealed the presence of molecular hydrogen (H_2), methane (CH_4), and other gases. The detection of H_2 by Cassini's Ion and Neutral Mass Spectrometer (INMS) was confirmed and interpreted as a signature of hydrothermal activity[3]. These discoveries pointed to ocean worlds like Enceladus, but also Europa in the Jovian system and recently Ceres in the asteroid belt[4,5], as possibly harboring conditions that resemble Earth's habitable hydrothermal vents[6–8] (Fig. 1a). If Enceladus' plume sampled by Cassini has its origin in a global ocean harboring hydrothermal vents[9], the composition of the plume provides information on the environment of the deep-sea vents.

Hydrothermal circulation on Enceladus may resemble Earth's alkaline hydrothermal activity, producing H_2 and CH_4, among other gases[10,11]. If H_2 present in the hydrothermal fluid is produced in Enceladus' core[3], serpentinization is a plausible source[11], while other potential sources are likely insufficient to explain the observed H_2 flux[3]. Methane might originate from the (abiotic) conversion of H_2 and carbon dioxide (CO_2) to CH_4 in serpentinization waters according to reaction (1) (written for one mol of the electron donor H_2)[12]:

$$H_2 + \frac{1}{4}\,CO_2 \longrightarrow \frac{1}{4}\,CH_4 + \frac{1}{2}\,H_2O \qquad (1)$$

or from other abiotic sources such as a primordial stock of CH_4 [13], or a stock of organic molecules in the core being pyrolysed[14].

Hydrothermal activity might also sustain a biological origin of methane[3]. In Earth deep-sea hydrothermal vents, chemoautotrophic single-celled organisms harness thermodynamical disequilibrium generated by chemical gradients as their energy source, rather than sunlight[15]. Some hydrothermal vents are known to be habitats for abundant thermophilic and hyperthermophilic methanogenic microorganisms that convert H_2 and CO_2 (or its dissolved forms, dissolved inorganic carbon, or DIC) into methane[16,17] according to (1) as the net reaction. On Enceladus, hydrothermal vents shaped by internal heating might provide similar energy sources[11,18]. Considering the habitability of Earth's hydrothermal vents to methanogens and the putative role of methanogenic metabolism in early Earth biosphere function and evolution[19], Enceladus' habitability by hydrogenotrophic methanogens has been discussed[14,20,36]. Theoretical work evaluated the viability of methanogenesis based on observed thermodynamical disequilibrium[3], while the successful incubation of Earth methanogens in putative Enceladean environments was achieved experimentally[20].

Our goal is to take this approach a critical step further: to quantify the probability that biological methanogenesis might explain the Cassini observations of the plume that were reported in Waite et al. (2017)[3]. By integrating mathematical models of geophysical, geochemical and biological processes within a Bayesian statistical framework, we aim at quantifying the plausibility of alternate scenarios of habitability and biological activity, rather than assessing their mere possibility, and evaluate in probabilistic terms how distinguishable they are from one another.

Results

The difficulty in estimating the plausibility of alternate scenarios of habitability and inhabitation arises from dealing with information that is incomplete and comes from non-independent

sources. A Bayesian framework is well suited for quantifying inference uncertainty[21,22]. In general, the Bayesian approach aims at quantifying the probability that alternate hypotheses be true given actual observations. By translating the different hypotheses into mathematical models, simulations can be used to find the likelihoods of the data occurring under the alternate scenarios. *A priori* (prior) knowledge of factors that influence the probability of each hypothesis being true, including previous observations, is combined with the likelihoods to give the *a posteriori* (posterior) probability of each hypothesis being true.

Here we apply this framework and contrast a model for a biotic scenario, denoted by B, in which methanogenesis might occur at Enceladus' hydrothermal vents, to a null abiotic model, denoted by A, in which methane production is abiotic. Absence of methanogenesis may be because the hydrothermal environments are not habitable for organisms with a methanogenic metabolism, or because, even if they were, such metabolisms never evolved – due to life not emerging on Enceladus in the first place, or the evolutionary steps to methanogenesis not having taken place. The probability of the B model given our knowledge of the system follows from Bayes' theorem[21] (Extended Data Fig. 1):

$$P(B|x^0) = \frac{P(x^0|B)P_{prior}(B)}{P(x^0|A)P_{prior}(A) + P(x^0|B)P_{prior}(B)} \tag{2}$$

where x^0 is the vector of observed values of "observables", carefully chosen quantities that may include a biosignature; $P(B|x^0)$ is the *a posteriori* probability of scenario B given the observations; $P_{prior}(X)$ is the *a priori* probability of scenario X, where $X = A$ or B; and $P(x^0|X)$ is the likelihood of the data under scenario X. The observables that describe the composition of the plume are the inferred H_2 and CH_4 escape rates (Φ_{H_2} and Φ_{CH_4} in mol yr^{-1}) and the ratio ($H_2 : CH_4$) [3]. These escape rates are derived from mass spectrometry abundances relative to water and the estimated water ejection rate in the plume[3]. The escape rates are insightful because they and the production rates are expected to be of the same order of magnitude[3,23] (Methods).

To estimate each term in equation (2), we combine a process-based modeling approach with a summary of current knowledge of possible conditions in Enceladus' hydrothermal vents. We begin with constructing a mechanistic model of relevant interior processes. Although considerable uncertainty remains on Enceladus' interior and history[24], significant advances have been made regarding the structure of its core and ocean circulation. Based on these results, we build a 1D model of fluid and heat flow. Assuming a porous core with at least one heat source (tidal friction or radiogenic heating)[18,25,26], we carefully address the upward advection of water from hydrothermal vents to the ice ceiling in oceanic plumes shaped by buoyancy and Coriolis forces[11,27,28]. We then assemble a model of hydrothermal fluid (hereafter abbreviated HF) and oceanic water mixing in a mixing layer (ML) that sets a gradient of conditions around a hydrothermal vent.

Next we design an ecological model of methanogenesis constrained by our model of interior processes. The ecological model describes the biomass dynamics of hydrogenotrophic methanogenic metabolisms similar to thermophilic and hyperthermophilic archaea that have been found in Earth's hydrothermal vents[29]. We assume the putative ecosystem to be located at the core-ocean interface, in the ML surrounding a hydrothermal vent and in which ocean water and HF from the core are mixed (Methods). The biomass growth rate depends on the amount of energy obtained from catabolism (equation (1)). The catabolic rate competes with the rate at which energy is lost through processes of biological decay[12]. Chemical environmental

conditions determine the Gibbs free energy obtained for each mol of electron donor; environmental temperature influences both catabolic[30] and maintenance[31] energy rates, setting limits to the viability of the ecosystem (Fig. 1b, c). The model thus yields a quantitative criterion for habitability: a set of conditions on physical and chemical input parameters under which biomass production by a hydrogenotrophic methanogenic metabolism can reach a stationary state and persist (Fig. 1d). The biotic scenario corresponds to the case of a so-defined habitable environment that is actually inhabited. The model can then predict the composition of the surrounding water mass at biomass production steady-state.

Input parameters of the ecological model are the composition and temperature of the HF and composition of the ocean in species of interest (H_2, DIC, CH_4), gathered from the literature and forming our prior knowledge of the system (Extended Data Table 1). The distributions of these parameters thus define the prior space, θ. A well-supported estimate for the temperature of Enceladus' ocean is 275 K, close to thermal equilibrium with ice[3,9,10]. The HF temperature is estimated to be in the range of 300-620 K [11]. On Enceladus, the source of CO_2 on a geological timescale is unlikely to be magma outgassing, like it is on Earth. As a consequence, we assume that the ocean is the reservoir of CO_2, possibly as carbonates[10]. Using chemical data from the Cassini mission, the ocean DIC concentration was estimated between 10^{-1} and $4 \cdot 10^{-3}$ mol kg^{-1} [10]. The CH_4 concentration in the HF is also a critical parameter of methanogenesis, because its abundance may negatively influence the thermodynamical drive of reaction (1). We first consider methane having its origin in serpentinization, for which an upper bound of CH_4 concentration in the HF can be estimated. There is much debate about the abiotic rate of reaction (1) in serpentinization waters[32] but recent experimental studies set a maximum concentration of 10^{-4} mol kg^{-1} at the hydrostatic pressure of shallow hydrothermal vents ($\approx$ 10 MPa)[33], close to what is expected at Enceladus' hydrothermal systems[20]. Next we extend our analysis by raising the upper bound on the concentration of CH_4 in the HF to 0.1 mol kg^{-1}, to account for other possible abiotic sources for which mechanistic modeling remains beyond reach.

From these estimates, we define the prior distributions, capturing the knowledge and uncertainties of possible conditions in Enceladus' interior. From the prior distributions, the model input parameters are drawn to run 50,000 simulations, resulting in a distribution of so-called "pseudo-data" in the space of observables (Fig. 2 and Extended Data Fig. 2). The pseudo-data distribution can be split into the two classes: "habitable" (H), which correspond to the set of parameter values for which there is at least one methanogenic metabolism that can function steadily in the surroundings of the hydrothermal vent; and "uninhabitable" ($\bar{H}$) otherwise. Measuring the volume of the parameters space falling in the "habitable" class gives an *a priori* probability that Enceladus be habitable, $P_{prior}(H) \approx 0.32$. This value is obtained by integrating over the whole prior distributions, including compositions in the putative environment of methanogenic metabolisms that may be likely as well as unlikely. Thus, a rather large range of individually credible values of interior parameters allow methanogenic metabolisms to persist, according to our model.

Simulations accounting for methanogenesis in combination with serpentinization water chemistry consistently explain the observed rates of CH_4 and H_2 emissions and the ratio ($H_2 : CH_4$) (Fig. 2, green distributions). In contrast, simulations in which CH_4 only originates from abiotic reactions in serpentinization water, cannot explain all the observations (Fig. 2, blue and orange distributions). Only the class of biotic models yields a non-zero likelihood for every observable (Extended Data Fig. 3a), provided the *a priori* probability of life emergence, $P(B|H)$, is large enough (Extended Data Fig. 3c). In fact, the concentration of CH_4 in the HF must be at least approximately $10^{-3.5}$ mol kg^{-1} in order to explain the observed flux of $\approx 10^9$ mol yr^{-1} (Extended

Data Fig. 4a), whereas the maximum concentration of CH_4 in the HF allowed by serpentinization alone, 1 mmol kg^{-1} [33] (see Methods) translates into a maximal CH_4 escape rate of $\approx 10^8$ mol yr^{-1} (Extended Data Fig. 2a). The observed flux of CH_4 thus lies far from what can be expected from the abiotic conversion of H_2 and CO_2 through serpentinization; in contrast, it falls in the range of expected values when there is biological methanogenesis (Fig. 2b, d, e). Conversely, if the methane flux were re-measured at values ten-fold lower than Cassini's, serpentinization alone could explain the data, yet methanogenesis would still be possible (Extended Data Fig. 4b).

Besides serpentinization, CH_4 in Enceladus' plume might originate from other abiotic sources such as primordial outgassing[13] or organic pyrolysis[14]. These sources of CH_4 in Enceladus' ocean are too poorly constrained to be modeled mechanistically. Nevertheless, we can test the robustness of our results by considering an alternative scenario in which the maximal concentration of CH_4 in the HF is set to an arbitrary higher value of 0.1 mol kg^{-1}, possibly due in part to pyrolysis[3]. The high methane concentration, abiotic scenario yields a set of abiotic-habitable simulations that encompasses the empirical observations (Extended Data Fig. 4 and 5), and a slightly lower *a priori* probability that Enceladus be habitable $P_{prior}(H) \approx 0.27$. In this case, the observations may be explained either by the biotic or abiotic scenario. However, the estimated likelihoods still indicate that the observations are more likely under the B scenario (Extended Data Fig. 3b) provided that $P(B|H)$ is large enough (Extended Data Fig. 3d). To compute the posterior probabilities of the uninhabitable, abiotic-habitable, and biotic models, across the whole range $P(B|H) \in [0,1]$ (or conversely, $P_{prior}(B) \in [0, 0.27]$), we use a random forest classification algorithm (ABC-RF, see Methods)[34]. Results are shown in Fig. 3, except for the probability of uninhabitability, which remains the lowest of all three models for all values of $P(B|H)$. With the *a priori* belief (in the Bayesian terminology) that all values of $P(B|H)$ between 0 and 1 are equally likely, the preferred scenario is, on average, B with probability $P_{P(B|H)\in[0,1]}(B|x^0) = 0.59$. If $P(B|H)$ is less than 0.2, the preferred model predicts abiotic-habitable conditions, yet $P(B|x^0)$ is still higher than the *a priori* probability $P_{prior}(B)$. For values of $P(B|H)$ in the range $0.2 - 0.4$, the abiotic-habitable and biotic models are equally probable (which is also reflected in the ABC-RF classifier score dipping), and higher values of $P(B|H)$ lead to select methanogenesis as the preferred model.

Our findings shed light on the significance of H_2 and CH_4 as biosignatures. It has been argued that the detection of H_2 might signal environments that are potentially habitable yet uninhabited (inhabitation by hydrogenotrophs supposedly implying consumption and depletion of H_2). Assuming that the H_2 plume ejection rate matches the production rate in the core[3], our model shows that levels of H_2 outgassing in the plume do correlate with the energy source available to putative hydrogenotrophs around hydrothermal vents (Extended Data Fig. 2a). However, the escape rate of H_2 appears to be poorly informative of whether or not hydrogenotrophic metabolisms might actually be present (Fig. 2a). In fact, the observed escape rate of H_2 is compatible with active methanogenesis. This is because the change in H_2 escape rate that would be due to the activity of hydrogenotrophs is negligible in comparison with measurement uncertainty combined with prior uncertainty on H_2 molecular concentration in the HF. Thus, the H_2 escape rate may be a signature of habitability but it may not be used to infer inhabitation.

Another potential biosignature is the production of CH_4 [35,36]. Because increasing the CH_4 concentration reduces the available energy that reaction (1) yields to the cell (Methods), abiotic production may compete with biotic processes[32]. It appears that the abiotic conversion of H_2 and CO_2 to CH_4 in serpentinization waters is too weak to significantly affect habitability,

as a higher concentration of CH_4 in the hydrothermal fluid ($[CH_4]_f$) does not correlate with a significantly lower fraction of habitable simulations (Extended Data Fig. 2a). As a consequence, the observation of CH_4 in the plume is not informative on whether hydrothermal vents might be habitable to methanogens. On the other hand, in line with general considerations of CH_4 as a biosignature[35], quantifying CH_4 escape in the plume is very informative about whether or not models that include biological methanogenesis can explain the data (Extended Data Fig. 2b).

Discussion

In summary, models that combine abiotic processes involving production of H_2 in Enceladus' core and biological methanogenesis near hydrothermal vents can explain the Cassini data (Fig. 2). Given the observations, the posterior probability of the biotic scenario is the highest provided the prior probability of life emergence in habitable environments exceeds a threshold, *ca.* 0.35. Below this threshold, the highest posterior probability is scored by the subset of abiotic models that predict habitability (Fig. 3).

Considering the possibility that deep-sea vent conditions might be conducive to life emergence[37], better constraining the age and duration of such conditions of temperature and composition in Enceladus should help define ranges of credible values for the probability of life emergence, $P(B|H)$ (Fig. 3). Alternate hypotheses have been proposed for the age and formation of Enceladus, as well as for the evolution of its interior – ranging from a modern origin of the moon 500 Myr ago without an ocean, to a primordial origin at 4,500 Myr with the existence of a modern ocean[24]. Waite et al. (2017)[3] estimate that the observed production of H_2 may have been sustained by serpentinization in Enceladus for several billion years. Such insights suggest that habitable conditions on Enceladus may have existed around deep-sea vents long enough for life to emerge and methanogenesis to evolve. In the alternate scenario where life origin would require a surface environment with ultraviolet light and atmospheric sources of molecules to feed the prebiotic chemistry[38],the prior probability of life emergence in Enceladus' ocean, $P(B|H)$, might be very small. In this case, methane levels inferred from Cassini's observations can be explained in a large fraction of models that predict abiotic habitability, with an abiotic source of methane that remains to be identified.

We emphasize that methane in Enceladus' plume might originate from abiotic sources not included in our model. In our approach, we sampled interior parameters, including $[CH_4]_f$ in log-uniform, independent prior distributions. More accurate modeling of alternate abiotic sources will require to reconsider the structure of these distributions, *e.g.* correlating $[CH_4]_f$ with the H_2 concentration in the hydrothermal fluid ($[H_2]_f$). Our framework has the capacity to accommodate these changes and more generally to account for mechanisms of abiotic methane production other than serpentinization chemistry (outgassing of primordial methane accumulated in the core during formation or pyrolytic production[3]) when new insights into these processes become available for Enceladus.

Isotopic measurements such as the ^{14}N and ^{15}N of NH_3 as well as the ^{13}C to ^{12}C and deuterium-to-hydrogen ratios of CH_4 might shed new light on the nature of Enceladus' core and the origin (primordial or modern) of methane. There are challenges, however – instrumental (the Ion and Neutral Mass Spectrometer (INMS) onboard Cassini had an insufficient mass resolution to accurately measure these ratios[13]) and analytical (environmental conditions characteristic of hydrothermal vents –pressure and temperature– may alter isotope fractionation of carbon by methanogens, rendering the isotopic markers of life less clear[39]). Groundbreaking data

might come from observations of the plumes of the Jovian icy moon Europa (James Webb Space Telescope (JWST)[40], Clipper[41]) and from proposed missions targeting Enceladus itself (ELSAH[42], ELF[43] or E2T[44]). Such data are critically needed to identify abiotic sources of methane that might explain Cassini's observations without calling for methanogenesis. In a broader perspective, our work demonstrates how the integration of interior and ecological models can be achieved, so that a Bayesian inferential approach for the detection of habitability and biosignatures can be implemented – an approach that holds much promise for ocean worlds that future programs will discover and observe[21,22].

References

[1] Spilker, L. Cassini-Huygens' exploration of the Saturn system: 13 years of discovery. *Science* **364**, 1046–1051 (2019).

[2] Thomas, P. et al. Enceladus's measured physical libration requires a global subsurface ocean. *Icarus* **264**, 37–47 (2016).

[3] Waite, J. H. et al. Cassini finds molecular hydrogen in the Enceladus plume: evidence for hydrothermal processes. *Science* **356**, 155–159 (2017).

[4] Nathues, A. et al. Recent cryovolcanic activity at Occator crater on Ceres. *Nature Astronomy* **4**, 794–801 (2020).

[5] Schmidt, B. et al. Post-impact cryo-hydrologic formation of small mounds and hills in Ceres's Occator crater. *Nature Geoscience* **13**, 605–610 (2020).

[6] Reynolds, R. T., Squyres, S. W., Colburn, D. S. & McKay, C. P. On the habitability of Europa. *Icarus* **56**, 246–254 (1983).

[7] Martin, A. & McMinn, A. Sea ice, extremophiles and life on extra-terrestrial ocean worlds. *International Journal of Astrobiology* **17**, 1–16 (2018).

[8] McCollom, T. M. Methanogenesis as a potential source of chemical energy for primary biomass production by autotrophic organisms in hydrothermal systems on Europa. *Journal of Geophysical Research: Planets* **104**, 30729–30742 (1999).

[9] Hsu, H.-W. et al. Ongoing hydrothermal activities within Enceladus. *Nature* **519**, 207 (2015).

[10] Glein, C. R., Baross, J. A. & Waite Jr, J. H. The pH of Enceladus' ocean. *Geochimica et Cosmochimica Acta* **162**, 202–219 (2015).

[11] Choblet, G. et al. Powering prolonged hydrothermal activity inside Enceladus. *Nature Astronomy* **1**, 841 (2017).

[12] Kleerebezem, R. & Van Loosdrecht, M. C. A generalized method for thermodynamic state analysis of environmental systems. *Critical Reviews in Environmental Science and Technology* **40**, 1–54 (2010).

[13] Mousis, O. et al. Formation conditions of Enceladus and origin of its methane reservoir. *The Astrophysical Journal Letters* **701**, L39 (2009).

[14] McKay, C., Khare, B. N., Amin, R., Klasson, M. & Kral, T. A. Possible sources for methane and C2–C5 organics in the plume of Enceladus. *Planetary and Space Science* **71**, 73–79 (2012).

[15] Jannasch, H. W. & Mottl, M. J. Geomicrobiology of deep-sea hydrothermal vents. *Science* **229**, 717–725 (1985).

[16] Schrenk, M. O., Kelley, D. S., Bolton, S. A. & Baross, J. A. Low archaeal diversity linked to subseafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environmental Microbiology* **6**, 1086–1095 (2004).

[17] Hedderich, R. & Whitman, W. B. Physiology and biochemistry of the methane-producing Archaea. *The Prokaryotes: Prokaryotic Physiology and Biochemistry* , 635–662 (2013).

[18] Travis, B. & Schubert, G. Keeping Enceladus warm. *Icarus* **250**, 32–42 (2015).

[19] Martin, W., Baross, J., Kelley, D. & Russell, M. J. Hydrothermal vents and the origin of life. *Nature Reviews Microbiology* **6**, 805 (2008).

[20] Taubner, R.-S. et al. Biological methane production under putative Enceladus-like conditions. *Nature Communications* **9**, 748 (2018).

[21] Catling, D. C. et al. Exoplanet biosignatures: a framework for their assessment. *Astrobiology* **18**, 709–738 (2018).

[22] Lorenz, R. D. A Bayesian approach to biosignature detection on ocean worlds. *Nature Astronomy* **3**, 466–467 (2019).

[23] Bouquet, A., Mousis, O., Waite, J. H. & Picaud, S. Possible evidence for a methane source in Enceladus' ocean. *Geophysical Research Letters* **42**, 1334–1339 (2015).

[24] Neveu, M. & Rhoden, A. R. Evolution of Saturn's mid-sized moons. *Nature astronomy* **3**, 543–552 (2019).

[25] Prialnik, D. & Merk, R. Growth and evolution of small porous icy bodies with an adaptive-grid thermal evolution code: I. Application to Kuiper belt objects and Enceladus. *Icarus* **197**, 211–220 (2008).

[26] Roberts, J. H. The fluffy core of Enceladus. *Icarus* **258**, 54–66 (2015).

[27] Goodman, J. C., Collins, G. C., Marshall, J. & Pierrehumbert, R. T. Hydrothermal plume dynamics on Europa: Implications for chaos formation. *Journal of Geophysical Research: Planets* **109** (2004).

[28] Goodman, J. C. & Lenferink, E. Numerical simulations of marine hydrothermal plumes for Europa and other icy worlds. *Icarus* **221**, 970–983 (2012).

[29] Topçuoğlu, B. D. et al. Hydrogen limitation and syntrophic growth among natural assemblages of thermophilic methanogens at deep-sea hydrothermal vents. *Frontiers in microbiology* **7**, 1240 (2016).

[30] Daniel, R. M. et al. The molecular basis of the effect of temperature on enzyme activity. *Biochemical journal* **425**, 353–360 (2010).

[31] Tijhuis, L., Van Loosdrecht, M. C. & Heijnen, J. A thermodynamically based correlation for maintenance Gibbs energy requirements in aerobic and anaerobic chemotrophic growth. *Biotechnology and bioengineering* **42**, 509–519 (1993).

[32] Sleep, N., Meibom, A., Fridriksson, T., Coleman, R. & Bird, D. H2-rich fluids from serpentinization: geochemical and biotic implications. *Proceedings of the National Academy of Sciences* **101**, 12818–12823 (2004).

[33] McCollom, T. M. Abiotic methane formation during experimental serpentinization of olivine. *Proceedings of the National Academy of Sciences* **113**, 13965–13970 (2016).

[34] Pudlo, P. et al. Reliable ABC model choice via random forests. *Bioinformatics* **32**, 859–866 (2015).

[35] Krissansen-Totton, J., Olson, S. & Catling, D. C. Disequilibrium biosignatures over Earth history and implications for detecting exoplanet life. *Science advances* **4**, eaao5747 (2018).

[36] McKay, C. P., Porco, C. C., Altheide, T., Davis, W. L. & Kral, T. A. The possible origin and persistence of life on Enceladus and detection of biomarkers in the plume. *Astrobiology* **8**, 909–919 (2008).

[37] Russell, M. J. et al. The drive to life on wet and icy worlds. *Astrobiology* **14**, 308–343 (2014).

[38] Sasselov, D. D., Grotzinger, J. P. & Sutherland, J. D. The origin of life as a planetary phenomenon. *Science Advances* **6**, eaax3419 (2020).

[39] Takai, K. et al. Cell proliferation at 122 C and isotopically heavy CH4 production by a hyperthermophilic methanogen under high-pressure cultivation. *Proceedings of the National Academy of Sciences* **105**, 10949–10954 (2008).

[40] Kalirai, J. Scientific discovery with the James Webb Space Telescope. *Contemporary Physics* **59**, 251–290 (2018).

[41] Phillips, C. B. & Pappalardo, R. T. Europa Clipper mission concept: Exploring Jupiter's ocean moon. *Eos, Transactions American Geophysical Union* **95**, 165–167 (2014).

[42] Eigenbrode, J., Gold, R. E., McKay, C. P., Hurford, T. & Davila, A. Searching for Life in an Ocean World: The Enceladus Life Signatures and Habitability (ELSAH) mission concept. *cosp* **42**, F3–6 (2018).

[43] Cable, M. L. et al. Enceladus Life Finder: The search for life in a habitable moon (2016).

[44] Mitri, G. et al. Explorer of Enceladus and Titan (E2T): Investigating ocean worlds' evolution and habitability in the solar system. *Planetary and space science* **155**, 73–90 (2018).

Acknowledgements We are grateful for discussion with Daniel Apai, Alex Bixel, Benjamin Charnay, Zach Grochau-Wright, Betul Kacar, James Kasting, Charles Lineweaver, Valentin Thouzeau and members of the OCAV Project at PSL University and of NASA's Nexus for Exoplanet System Science (NExSS) research coordination network and its Earths in Other Solar Systems Project based at the University of Arizona. This work is supported by France *Investissements d'Avenir* program (ANR-10-LABX-54 MemoLife, ANR-10-IDEX-0001-02 PSL) through PSL IRIS OCAV and PSL - University of Arizona Mobility Program. R.F. acknowledges support from the U.S. National Science Foundation, Dimensions of Biodiversity (DEB-1831493), Biology Integration Institute-Implementation (DBI-2022070), and National Research Traineeship (DGE-2022055) programs; and from the United States National Aeronautics and Space Administration, Interdisciplinary Consortium for Astrobiology Research program.

Author contributions R.F., F.G. and S.M. conceived the study. R.F. designed the ecosystem model. F.G. and S.M. designed the geochemical model. A.A. and B.S. refined the models and developed the code. A.A. analyzed the results and wrote the first version of the manuscript. All authors finalized the paper.

Competing interests The authors declare no competing interests.

Supplementary information is available for this paper at xxx.doi.org

Figures

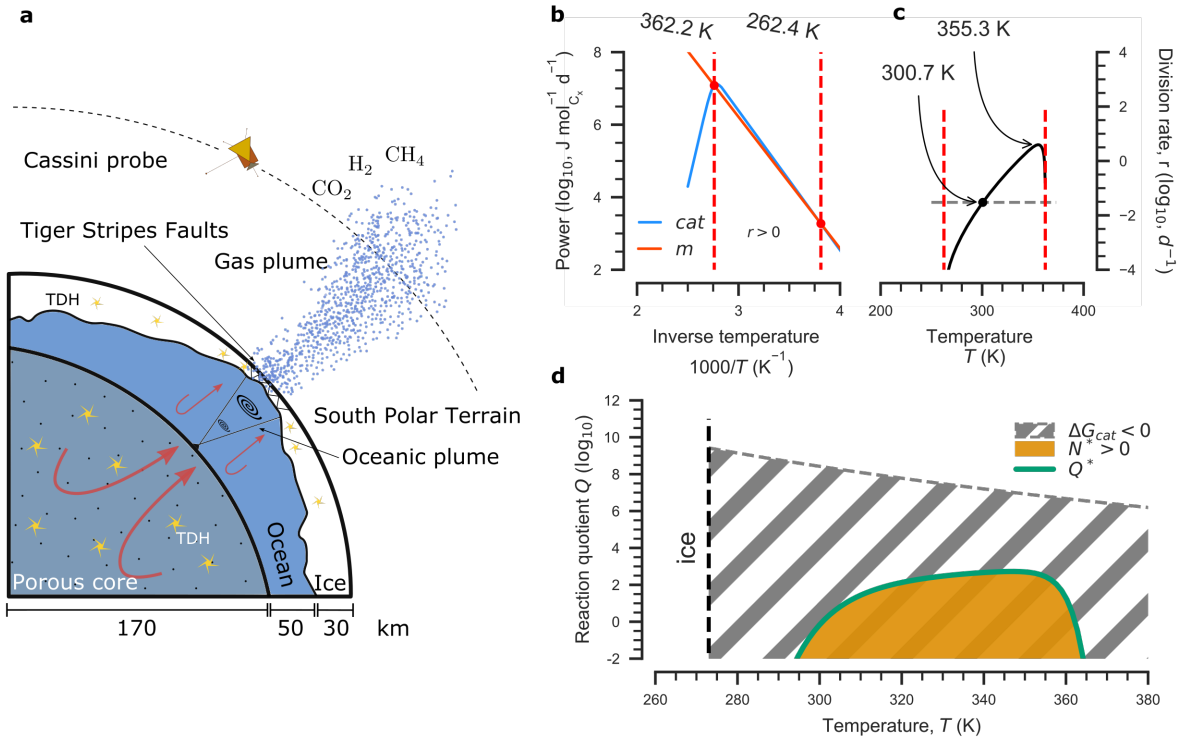

Figure 1 | General modeling framework. a Assumed interior structure of Enceladus. Yellow stars indicate Tidal Dissipative Heating (TDH). Red arrows denote water circulation, in the core [11] and in the ocean, where ocean plumes are shaped by buoyancy and Coriolis forces [11,28] (black spirals). The Cassini spacecraft is represented (not to scale) in one of its fly-bys of Enceladus, during which it crossed the gas plume of oceanic material escaping into space (dotted trajectory). **b** Catabolic (blue curve) and maintenance (red curve) energy of the modeled hyperthermophilic hydrogenotrophic methanogen in standard chemical conditions ($\Delta G = \Delta G_0$) as a function of temperature. Red dashed lines (vertical) delineate the temperature range over which the energy produced by catabolism is greater than maintenance requirements, resulting in positive growth rate, hence a positive division rate, r. **c** Division rate as a function of temperature. Red dashed lines (vertical) delineate the range of temperature over which the cell division rate is positive. The gray dashed line (horizontal) corresponds to a baseline death rate of 0.03 d^{-1}, which sets a minimum division rate needed for net population growth. At the black dot (300.7 K), the division rate and the death rate balance out. Between the lower thermal limit for $r > 0$ and the black dot, division occurs, but at a slower rate than death, thus leading the population to extinction. The temperature at which the division rate is highest (355.3 K) is indicated. This temperature is called the optimal temperature, T_{opt}. **d** Thermo-chemical (local) conditions conducive to a non-zero population size at steady state (orange). N^* is the number of individuals at population steady-state. The green line is the chemical quotient (Q) imposed by the population at steady-state, denoted by Q^*. The gray hatched region bounded by the gray dashed curve indicates the set of conditions under which methanogenesis is thermodynamically favored, *i.e.* $\Delta G_{cat} < 0$. The black dashed line (vertical) indicates the freezing temperature of water. Parameters used in b-d are given in Extended Data Table 2. See Methods for detail.

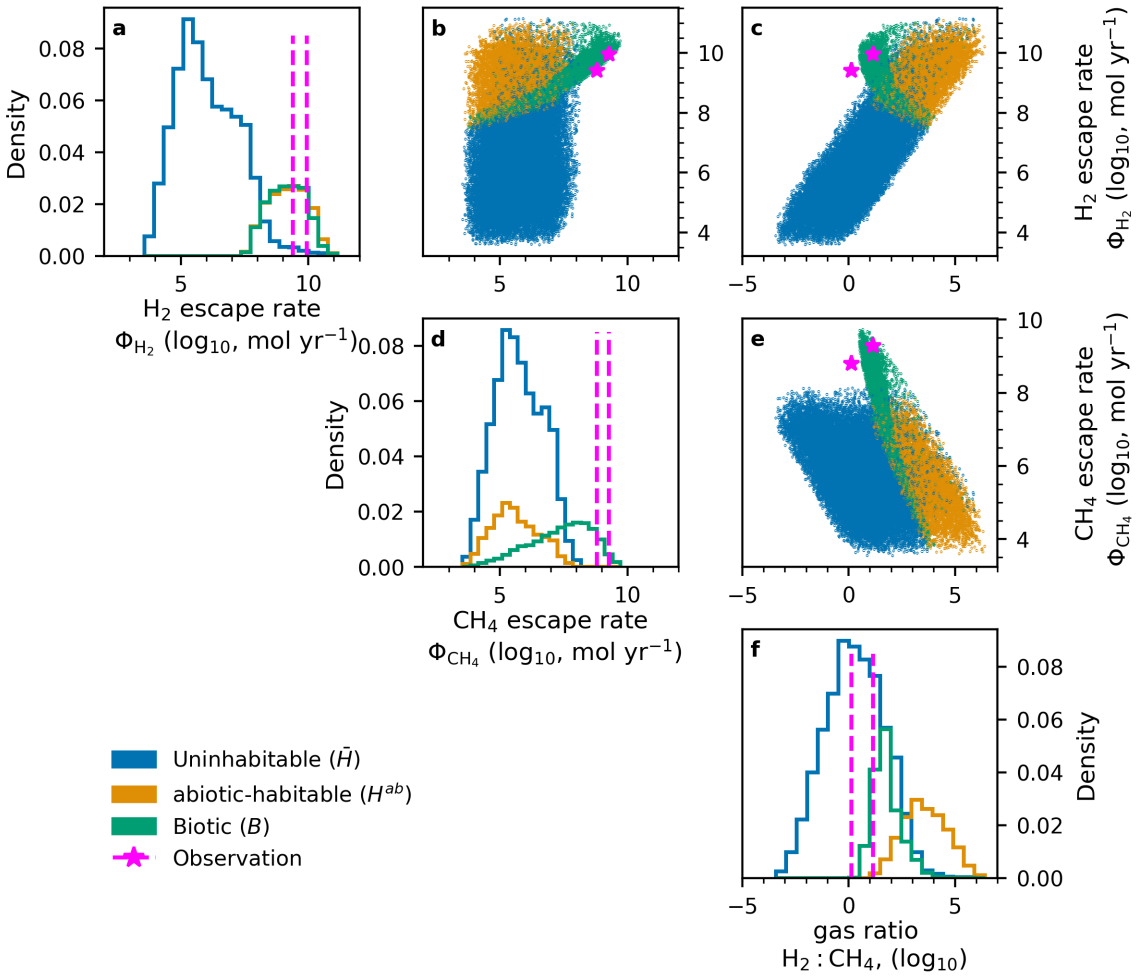

Figure 2 | Cassini observations and distributions of model outputs over the space of observables : inferred H_2 and CH_4 fluxes (Φ_{H_2} and Φ_{CH_4}), and gas ratio H_2 : CH_4. a, d, f, Density distributions of observables pseudo-data generated by model simulations. In blue, subset of simulations in which no population of methanogens could grow ('uninhabitable' scenario, $\bar{H}$). In orange, subset of simulations in which a population of methanogens is possible but their activity is not taken into account ('abiotic-habitable' scenario, H^{ab}). In green, subset of simulations in which methanogens can grow, and their activity is taken into account to produce pseudo-data ('biotic' scenario, B). Magenta dashed lines (vertical) indicate the actual Cassini observations[3]. Note the $\log_{10}$ scale on the horizontal axis. **b, c, e,** Joint distributions of observable pseudo-data generated by the model. Magenta stars indicate the Cassini observations. Note the $\log_{10}$ scale on both axes. This figure was generated from a set of 50,000 simulations using random values of internal parameters drawn from the distributions described in Extended Data Table 1. In about 32% of these simulations, the drawn set of parameters allowed the modeled population of methanogens to grow ('habitable', H), the remaining 68% were found to pertain to the 'uninhabitable' scenario ($\bar{H}$, blue). The 'habitable' subset was evenly split at random into two subsets of 8,008 simulations. In one subset, we did not simulate methanogenesis ('abiotic-habitable' scenario, H^{ab}, orange), while we simulated methanogenesis coupled with the geochemical model in the second subset ('biotic' scenario, B, green). See Methods and Extended Data Tables 2 and 3 for model equations and parameter values.

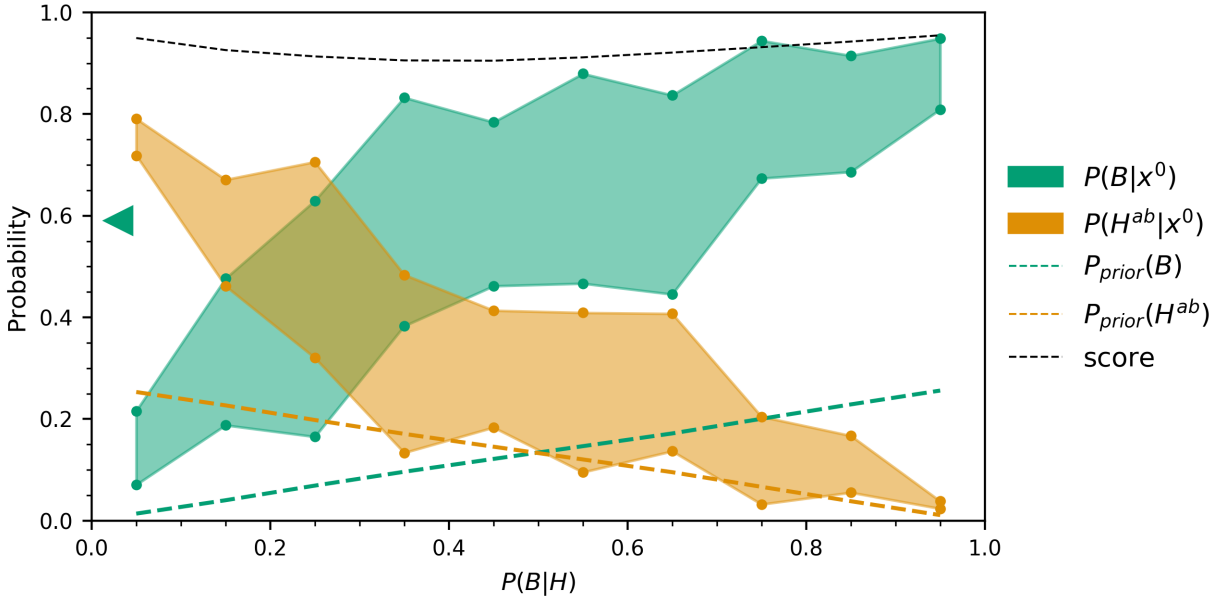

Figure 3 | Posterior probabilities of abiotic-habitable and biotic models, and classifier score. $P(B|H)$ is the prior probability of life emergence in habitable environments. Here the model accounts for unknown abiotic sources of methane in the hydrothermal fluid, by allowing for an *a priori* concentration higher than in the case of serpentinization only. This scenario yields $P_{prior}(H) \approx 0.27$. For each $P(B|H)$ value, the circles (green or orange) indicate the posterior probability of the biotic scenario (B, green) and abiotic-habitable scenario (H^{ab}, orange) for the lower and upper bounds of the observation range; the intervals between circles are colored for legibility. The green dashed line is $P_{prior}(B) = P_{prior}(H) \times P(B|H)$, the orange dashed line is $P_{prior}(H^{ab}) = P_{prior}(H) \times (1 - P(B|H))$. The black dotted line is the ABC-RF classifier score (see Methods). The green triangle on the vertical axis shows the mean value of the posterior probability $P(B|x^0)$ across $P(B|H) \sim U(0,1)$. Note that $P(\bar{H}|x^0)$ (not shown) is very low, and always less than $P(B|x^0)$ or $P(H^{ab}|x^0)$. See Methods for model equations, Extended Data Table 1 for prior internal parameter ranges, and Extended Data Tables 2, 3 for parameters values of model simulations and ABC-RF.

Methods

Mixing of core-ocean interface waters and plume composition

A hydrothermal hot spot dissipates thermal energy with power F (W) by advection of hydrothermal fluid (HF) at temperature T_f into seafloor oceanic waters (OW) at temperature T_o. We follow McCollom, (1999)[8] to model HF and OW mixing in a horizontally structured mixing layer (ML). ML waters buoyantly rise according to their temperature and escape the ML to be replaced by OW and HF. We first define the shape of J_f (kg s^{-1} m^{-2}), the mass flux density of HF into the ML as a function of the distance, u (m), to the hot spot center:

$$J_f(u) = J_{max} e^{-\left(\frac{u}{c}\right)^2} \tag{3}$$

where J_{max} (kg s^{-1} m^{-2}) and c (m) are derived by constraining the hot spot to dissipate power F and by forcing the hot spot center (where $J_f = J_{max}$) to be entirely composed of HF:

$$\int_S J_f dS = \int_0^\infty 2\pi r J_f(u) du = \frac{F}{C_p(T_f - T_o)} \tag{4}$$

where S (m^2) is the hot spot surface area, hence

$$c^2 = \frac{F}{\pi J_{max} C_p (T_f - T_o)} \tag{5}$$

where C_p is the specific heat capacity of water (4,200 J K^{-1} kg^{-1}).

The mass flux density J_c (kg s^{-1} m^{-2}) of buoyancy escape of ML waters is expressed as a function of the temperature anomaly $T' = T(r) - T_o$ following Goodman et al. (2004)[27]:

$$J_c = \frac{1}{u_s} \rho_o \sqrt{2\varepsilon g \alpha T'} \tag{6}$$

where $u_s = 1$ m^2 is a unitary surface (for readability, it is replaced with its value in following equations), ρ_o (1,000 kg m^{-3}) is seawater mass density, ε (1 m) is the ML thickness, α (3×10^{-4} K^{-1}) is the coefficient of thermal expansion of water and g (0.12 m s^{-2}) is the local gravitational acceleration (see values in Extended Data Table 2). Assuming a non-divergent fluid, we can write a system of partial differential equations to express the dimensionless mixing ratio $x = \frac{m_f}{m_f + m_o}$ (subscript f denotes the HF and o denotes the OW) and temperature $T = T_o + x(T_f - T_o)$ as a function of time:

$$\begin{aligned} \frac{\partial x}{\partial t} &= \frac{1}{m}(J_f - x J_c) \\ \frac{\partial T}{\partial t} &= \frac{\partial T}{\partial x}\frac{\partial x}{\partial t} = \frac{1}{m}\left(J_f(T_f - T_o) - J_c(T - T_o)\right) \end{aligned} \tag{7}$$

where $m = m_f + m_o$ is the ML mass density, which is constant. The steady state can be solved analytically. We can then derive the steady-state composition and temperature of the ML:

$$\begin{cases} T^*(u) &= \left[J_f(u)\frac{(T_f - T_o)}{\rho_o \sqrt{g 2\varepsilon\alpha}}\right]^{\frac{2}{3}} + T_o \\ J_c(T^*(u)) &= \left(2\rho_o^2 g \varepsilon \alpha J_f(u)(T_f - T_o)\right)^{\frac{1}{3}} \\ x^*(u) &= \left[\frac{J_f(u)^2}{2\rho_o^2 g \varepsilon \alpha (T_f - T_o)}\right]^{\frac{1}{3}} \end{cases} \tag{8}$$

where the asterisk symbol superscript denotes the steady-state value.

In the absence of chemical reactions taking place in the ML, we can write the concentration C_i^0 of compound i (mol kg^{-1}) in the steady-state ML as a function of the concentration of i in the HF (C_i^f) and in the ocean (C_i^o):

$$C_i^0(r) = x^* C_i^f + (1 - x^*) C_i^o \tag{9}$$

The derived set of concentrations then serves as the initial condition for our biological model. An example of a hot spot compositional structure is shown in Supplementary Fig. 1a.

The scaling relationships derived in Goodman et al. (2004)[27] and Goodman & Lenferink (2012)[28] for Europa were used in Choblet et al. (2017)[11] to show that the Coriolis effect was important for the most part of Enceladus' oceanic plumes (OPs) ascension, thus suggesting well-defined plumes that experience little dilution with oceanic waters. In Goodman et al. (2004)[27], an OP starts with warm water rising in a turbulent regime, forming a cone, until the Coriolis effect becomes important and the water mass forms a cylinder. When the ocean plume encounters the ice ceiling, it starts growing into a cone shape (Fig. 1a); a steady state is reached when the cone radius equals the Rossby radius of deformation. At this point, the plume begins to shed large eddies that travel horizontally. In our approach, we propose that the composition of the OP is the total mix of all convecting water masses:

$$C_i^{OP} = \frac{\int J_c C_i dS}{\int J_c dS} \tag{10}$$

Internal cell dynamics model

Here, we describe how we model the dependence of metabolism on environmental conditions through the thermodynamics of the catabolic reaction, upon which the bio-ecological model is built as a simplified variant of Sauterey et al. (2020)[45]. Catabolism is the set of reactions that a cell uses to harvest the energy needed for the synthesis of complex organic molecules to repair cell components or to grow. Catabolism can be written for 1 mol of electron donor (eD) as:

$$1eD + \{Y_i R_i\}_{i \neq eD} \longrightarrow \{Y_j P_j\} \tag{11}$$

where $\{R_i\}$ are the reactants, $\{P_j\}$ the products and $\{Y_{i,j}\}$ their stoichiometry (counted negative for reactants and positive for products). The Gibbs free energy of the catabolic reaction (J mol$_{eD}^{-1}$) is expressed as a function of temperature T (K) and reaction quotient Q_{cat} (dimensionless):

$$\Delta G_{cat} = \Delta G_{0,cat} + RT \log_e(Q_{cat}) \tag{12}$$

where

$$Q_{cat} = \prod_i a_i^{Y_i} \approx \prod_i C_i^{Y_i}. \tag{13}$$

Here a_i is the activity of species i, approximated by its concentration C_i. $\Delta G_{0,cat}$ is the Gibbs free energy of the reaction when all activities are unity, expressed as a function of temperature and standard values $\Delta G_{0,S}$ and $\Delta H_{0,S}$ found in chemistry tables:

$$\Delta G_{0,cat} = \Delta G_{0,S} \frac{T}{T_S} + \Delta H_{0,S} \frac{T_S - T}{T_S} \tag{14}$$

where $T_S = 298$ K is the standard temperature. Even spontaneous reactions ($\Delta G < 0$) may not occur, should their kinetics be infinitely slow; that is, when the activation energy needed to start them is very high. A fundamental property of living organisms is that they synthesize enzymes that significantly lower the activation energy of catabolic reactions so that they may occur at higher rates, making the potential energy they hold available for maintenance and anabolism (*i.e.* biomass production).

Our biological model is designed to track both the catabolic reaction rate and the energy that catabolism yields to the cell. The cell is able to run its catabolism at a certain specific rate, q_{cat} ($\mathrm{mol}_{eD}\ \mathrm{s}^{-1}\ \mathrm{mol}_{C_x}^{-1}$), which depends on temperature. A general model of enzymatic rate q_e ($\mathrm{mol}_{eD}\ \mathrm{s}^{-1}\ \mathrm{mol}_{enzyme}^{-1}$) as a function of temperature was derived in Daniel et al. (2010)[30] and validated experimentally. In their framework, the enzymatic rate is determined by the abundance of activated enzymes E_{act}. Higher temperature accelerates the catalyzed reaction, but also the conversion of activated enzymes to a deactivated state E_{inact}. The overall net enzyme catalysis rate q_e is modeled using an Arrhenius law for the rate of catalysis of the reaction by the activated enzyme (k_{cat}, $\mathrm{mol}_{eD}\ \mathrm{s}^{-1}\ \mathrm{mol}_{E_{act}}^{-1}$) and an equilibrium law for the ratio of inactivated to activated enzyme $K_{eq} = \frac{E_{inact}}{E_{act}}$ (dimensionless):

$$\begin{aligned} q_e &= \frac{k_{cat}}{1+K_{eq}} \\ k_{cat} &= \frac{k_b}{h} T e^{-\frac{\Delta G_{acat}}{RT}} \\ K_{eq} &= e^{\frac{\Delta H_{eq}}{R}\left(\frac{1}{T_{eq}}-\frac{1}{T}\right)} \end{aligned} \tag{15}$$

where k_b (1.38×10^{-23} J K^{-1}) is the Boltzmann constant, h (6.63×10^{-34} J s) is the Planck constant. The other kinetic parameters are enzyme-dependent. In our approach of modeling a generic hyperthermophile, we use values given in Daniel et al. (2010)[30] for an enzyme of *Thermus sp.* RT41a, an organism growing best at around 75°C, the highest growth temperature in their dataset. These values are (Extended Data Table 2), $\Delta G_{acat} = 72{,}000$ J mol^{-1} (activation energy of the catabolic reaction), $\Delta H_{eq} = 305{,}000$ J mol^{-1} (equilibrium enthalpy of activated to deactivated enzyme) and $T_{eq} = 363.15$ K (equilibrium temperature of activated to deactivated enzyme). In the Supplementary Results and Discussion, we explore the consequence of varying the values of these parameters. We assume random variation, rather than variation driven by evolutionary adaptation to temperature. Taking evolutionary adaptation into account might result in a larger set of conditions interpreted as habitable and a higher productivity of the modeled population. This might eventually cause significant differences in dihydrogen depletion and methane production in the plume. The consequences of evolutionary adaptation will be addressed in future work, using the approach pioneered in Sauterey et al. (2020)[45].

The specific catabolic rate q_{cat} (per mole of biomolecule carbon C_x, hence in unit $\mathrm{mol}_{eD}\ \mathrm{mol}_{C_x}^{-1}\ \mathrm{s}^{-1}$) is obtained by scaling q_e with the enzymatic fraction τ (dimensionless, or $\mathrm{mol}_{enzyme}\ \mathrm{mol}_{C_x}^{-1}$) of biomass (that is, the number of catabolic enzymes per biomolecule in the cell) so that $q_{cat} = \tau q_e$. The parameter value $\tau \approx 1.73 \times 10^{-5}$ is estimated by fitting the model to data from Taubner et al. (2018)[20] as described later on.

Some of the energy provided by the catabolism goes to the maintenance of the cell and represents a minimal energy for living processes (*e.g.* repairing denatured proteins and damaged DNA). No growth occurs when only this amount of energy is available to the cell. An empirical and general relationship between the cell maintenance energy rate (e_m, kJ d^{-1} $\mathrm{mol}_{C_x}^{-1}$) and temperature was established in Tijhuis et al. (1993)[31]:

$$\begin{aligned} e_m &= 84\, e^{\frac{69{,}400}{R}\left(\frac{1}{T_S}-\frac{1}{T}\right)} \\ q_m &= -\frac{e_m}{\Delta G_{cat}} \end{aligned} \tag{16}$$

where q_m (d^{-1}) is the rate at which the catabolism must run to meet the maintenance needs. When $q_{cat} < q_m$, the cell decays and eventually dies (Fig. 1b). Otherwise, the remaining catabolic energy then goes to fueling the anabolic reaction, which drives the synthesis of cell

building blocks. The anabolic reaction is rarely spontaneous and often requires more energy than catabolism alone may provide. In effect, the cell is able to store energy in the form of ATP until enough is available for anabolism. The rate at which the anabolism runs is much lower than the catabolic reaction rate (think of catabolism as a small gear wheeling the larger gear of anabolism).

More recent observations (Lever et al., 2015[46] and references therein) point out that e_m as modeled in Tijhuis et al. (1993)[31] might be overestimated. Although a variety of underlying processes to explain minimal energy requirements have been proposed[46], a detailed mechanistic model that could explain the available observations is still lacking. Given the persistent uncertainties regarding the minimal energy requirements of cellular life, we chose the conservative approach of using Tijhuis et al.'s model[31] – *i.e.*, the parametrization that disfavors habitability the most.

Kleerebezem & Van Loosdrecht (2010)[12] use a generic chemical composition of biomass that is useful to compute the Gibbs free energy required to synthesize a mol of biomass (ΔG_{ana}). To this energy is added the cost of dissipation that occurs during metabolism. An empirical relationship between characteristics of the metabolism and the amount of dissipated energy (ΔG_{diss}) is[12]:

$$\Delta G_{diss} = 200 + 18(6 - NoC)^{1.8} + exp\left[((-0.2-\gamma)^{2.16})(3.6+0.4NoC)\right] \tag{17}$$

where NoC is the carbon chain length in the carbon source and γ is the oxidation number of carbon in the carbon source. In general $\Delta G_{diss} >> \Delta G_{ana}$. The energy cost of anabolism $\Delta G_{diss} + \Delta G_{ana}$ can therefore be approximated by ΔG_{diss}. The anabolic rate per mol of biomolecule is therefore given by

$$q_{ana} = -\frac{\Delta G_{cat}}{\Delta G_{diss}}(q_{cat} - q_m) \tag{18}$$

For a given metabolism, q_{ana} is strongly dependent on temperature through q_{cat} and q_m, and also depends on the availability of substrates through ΔG_{cat}. The cell accumulates biomolecules at the rate q_{ana} times the number of biomolecules in the cell. We do not represent the number of biomolecules in each cell but rather the mean value for the population. The cells are then set to undergo division at a rate, r, proportional to their internal biomolecule content.

Ecosystem dynamics and steady state

Using quasi-steady-state approximation on internal cell dynamics, the ecosystem dynamics are driven by the dynamics of the cell population and the chemical environment (substrates and products):

$$\begin{array}{rcl} \frac{dN}{dt} & = & N(q_{ana} - d) \\ \frac{\partial C_i}{\partial t} & = & \frac{1}{\varepsilon u_s \rho_o}\left(J_f(C_i^f - C_i^o) + J_c(C_i^o - C_i)\right) + NB^* q_{cat} Y_i^{cat} \end{array} \tag{19}$$

where N is the surface density of the number of individuals in the population per kg of water (thus in $kg^{-1}\ m^{-2}$), d (s^{-1}) is a baseline death rate accounting for density-independent mortality, and B^* (mol_{C_x}) is the steady-state internal cell biomass, or quantity of biomolecules. To our knowledge, there is no estimation of d for methanogens in their natural habitat; we therefore chose to use the value estimated in Connolly & Coffin (1995)[47] for marine plankton, and we explore the consequence of varying its value in Supplementary Results and Discussion.

Solving equation (19) for steady-state, and using equation (13) yields an expression for the steady-state reaction quotient Q^*_{cat}:

$$\begin{aligned} Q^*_{cat} &= exp\left[-\frac{1}{RT}\left(\Delta G_0 + (d + \frac{e_m}{\Delta G_{diss}})\frac{\Delta G_{diss}}{q_{cat}}\right)\right] \\ \log_e Q^*_{cat} &= \log_e K - \frac{1}{RT}\left(d + \frac{e_m}{\Delta G_{diss}}\right)\frac{\Delta G_{diss}}{q_{cat}} \end{aligned} \tag{20}$$

where K (dimensionless) is the equilibrium constant of the reaction. Q^*_{cat} denotes the steady-state value of the reaction quotient, which is the value at which available energy in the cell's environment exactly compensates death (or removal from the environment) and maintenance, and therefore sets a limit value on Q_{cat} for growth, as detailed below.

We assume that if the initial conditions allow for a positive growth rate, the composition will converge toward this non-equilibrium stationary state, as seen in numerical simulations (Supplementary Fig. 1b, c). Initial conditions are favorable to microbial population persistence if Q^*_{cat} is larger than Q_{cat} evaluated at the initial conditions, $Q^0_{cat}|_T$:

$$Q^*_{cat}(T) > Q^0_{cat}|_T \tag{21}$$

which provides a criterion for population viability. In the absence of any other factors, the steady-state concentration of i, C^*_i, is given by

$$C^*_i = C^0_i + Y^{cat}_i (C^0_{e_D} - C^*_{e_D}). \tag{22}$$

The electron donor concentration at steady state, e^*_D, can be solved by using equation (20) and

$$Q^*_{cat} = \frac{1}{C^*_{e_D}} \prod_{i \neq e_D} \left(C^0_i + Y^{cat}_i (C^0_{e_D} - C^*_{e_D})\right)^{Y^{cat}_i} \tag{23}$$

In the case of Enceladus, we apply this bio-ecological model to a generic methanogenic hyperthermophile, with enzyme kinetic properties described hitherto, in a horizontally structured hydrothermal hot spot. The chemical reaction of hydrogenotrophic methanogenesis written per mol of electron donor (H_2) is given by equation (1) and has $\Delta G_{0,S} = -32.6$ kJ mol^{-1}, $\Delta H_{0,S} = -63.2$ kJ mol^{-1} for reactants in gaseous state. To express $\Delta G_{0,cat}$ in the aqueous phase, $\Delta G^{aq}_{0,cat}$, from $\Delta G_{0,cat}$ in the gaseous phase, $\Delta G^{gas}_{0,cat}$, we add a corrective term using enthalpies of dissolution:

$$\Delta G^{aq}_{0,cat} = \Delta G^{gas}_{0,cat} - RT \sum_i Y_i \log_e(H_i) \tag{24}$$

where $\Delta G^{gas}_{0,car}$ is computed using equation (14), $H_i = \frac{C_i}{P_i}$ (mol kg^{-1} Pa^{-1}) are Henry's law coefficients, $H_{\mathrm{H_2}} = 7.8\ 10^{-4}$ and $H_{\mathrm{CH_4}} = 1.4\ 10^{-3}$, taken independent from temperature. The solubility constant of CO_2 is taken as a function of temperature fitted from outputs of the Aspen plus software:

$$\begin{aligned} H_{\mathrm{CO_2}} = \exp\Bigg[&\frac{9345.17}{T} - 167.8108 + 23.3585 \log_e(T) + \\ &(0.023517 - 2.3656\ 10^{-4} T + 4.7036\ 10^{-7} T^2) 35.0 \Bigg] \end{aligned} \tag{25}$$

Regarding anabolism, we assume that the carbon source is also CO_2, hence $NoC = 1$ and $\gamma = 4$ in equation (17) giving ΔG_{diss}.

Estimation of the catabolic enzyme-to-biomass scaling factor

Here, we estimate a credible value for the yet unconstrained enzyme-to-biomass ratio parameter. This ratio is needed in order to derive a specific catabolic rate from enzyme-scale kinetic modeling, as mentioned earlier. To do so, we fit our model to the results of growth experiments in Taubner et al. (2018)[20]. In their study, Taubner et al. (2018)[20] experimented on *M. okinawensis*, *M. marburgensis* and *M. villosus*. We chose to use growth data of *M. villosus*, as *M. villosus* has a higher growth temperature than the other two (*ca.* 80 °C *versus ca.* 65 °C), in accordance with the kinetic parameter values we chose (see the Internal Cell Dynamics section above). Because data in Taubner et al. (2018)[20] measure cell density as optical density, we focus on the initial growth rate, a parameter that appears in the Verhulst logistic model. Our model being analog to a logistic model, we fit parameters n_0, r and K_n (initial population density, growth rate and carrying capacity respectively) of

$$n(t) = \frac{K_n}{1 + (\frac{K_n}{n_0} - 1)e^{-rt}} \tag{26}$$

to the experimental growth curve using least mean square method in the "curve fit" function in the Scipy python package[48]. The estimated value of the growth rate, r, is then used to obtain τ from

$$r \approx q_{ana} = \lambda(\tau q_e - q_m) \tag{27}$$

hence

$$\tau \approx \frac{\frac{r}{\lambda} + q_m}{q_e} \tag{28}$$

where λ, q_e and q_m are obtained from the composition and temperature of the growth medium in the experiment[20]. In Supplementary Fig. 1b we compare a logistic curve using a growth rate generated by the model at experimental temperature (80 ± 1 °C) and chemical conditions ($[H_2] = 1.5$ mmol kg^{-1}, $[CO_2] = 7.7$ mmol kg^{-1}, and negligible $[CH_4] = 10^{-10}$ mol kg^{-1}), with the τ value inferred from the growth data of *M. villosus* in Taubner et al. (2018)[20]. We find $\tau \approx 1.73 \cdot 10^{-5} \text{mol}_{enzyme}\ \text{mol}_{C_x}^{-1}$. We then compared this value to mass-specific resting metabolic rates of unicellular organisms[49]. To do so, we determined the cell mass predicted by our model using Kleerebezem & Van Loosdrecht (2010)[12] which provides an empirical relationship between cell volume V_c (μm^3) and cell structural mass $M_{C_x} = 18 \cdot 10^{-15} \cdot V_c^{0.94}$, in mol_{C_x} (this relationship was also used in Sauterey et al. (2020)[45]). By applying this relation to a typical 1 μm radius cell, we obtain $\log_e(B_0) \approx -6$ (W g$^{-3/4}$) for methanogens in our model at the estimated optimal temperature of ≈ 82 °C in standard chemical conditions (see Fig. 1b). This value falls at the lower end of the empirical range described in Gillooly et al. (2001)[49]. While τ might be an important parameter in determining the temperature growth curve of an organism, other kinetic parameters of catabolism and maintenance described earlier on also contribute and are estimated separately. As a consequence, when using the value obtained from *M. villosus* growth data, we do not aim at modeling this organism *per se*; rather, we model a generic, hypothetical hyperthermophile with temperature apparent properties described in Fig. 1b-d. We explore the consequences of changing the value of parameter τ, along with others in Supplementary Methods, Results and Discussion, as well as in Supplementary Figs. 3-5.

Bayesian inference

Because the model is complex, the likelihood $P(x^0|X)$, $X = A$ or B, cannot be analytically expressed, a problem known as "intractable likelihood". Instead we use an Approximate Bayesian

Computation (ABC) approach[50]. ABC relies on approximating the likelihood by simulating outputs, called pseudo-data, from prior distributions of parameters. The likelihood can then be computed from the distribution of simulation outputs. If θ denotes the parameter vector, x the vector of a pseudo-data point, and x^0 an observation ("true data"), then $P(\theta)$ is specified by the prior distribution, and $P(x^0|\theta)$ is approximated by sampling simulations of a model π linking parameters to the observable, *i.e.* $\pi(\theta) = x$. Then the probability of a parameter set $\hat{\theta}$ given the observation x^0 follows from Bayes theorem:

$$P(\hat{\theta}|x^0) = \frac{P(x^0|\hat{\theta})P(\hat{\theta})}{\int_\theta P(x^0|\theta)P(\theta)d\theta}. \tag{29}$$

Equivalently, instead of parameters, the Bayesian approach can be used to infer so-called classes, corresponding to subsets of the parameter space (Extended Data Fig. 1). See Sisson et al. (2018)[51] for a detailed guide to ABC methods and Csilléry et al. (2010)[50] for examples of applications.

The most basic form of ABC, called "K nearest neighbors" (KNN) ABC, aims at computing the Euclidean distance in the data space to obtain a subset of pseudo-data resembling the observation and the associated subset of the parameter space, so that a posterior distribution of parameters (maximizing the multivariate likelihood of observation) is obtained. Because future characterisation of the habitability or inhabitation of planetary bodies is expected to be a high-dimensional problem, involving a greater number of observables, following Csilléry et al. (2010)[50] we argue that standard ABC, a method suffering the curse of dimensionality, may not be suited. Instead we used a higher-order ABC method, Random Forest ABC (ABC-RF), based on a machine learning method of classification[34]. RF-ABC handles high-dimensional problems much more efficiently than KNN methods[34,50].

A Random Forest is a population of decision trees, that is, oriented graphs (or "flowcharts") in which each node corresponds to a test on a feature (an observable in our context) of the pseudo-data (an example is given in Supplementary Fig. 2a). Depending on the truth value of the test, the pseudo-data are then passed to a branch or another to the next test. For example, in Supplementary Fig. (2a), the first test is $R_1 \leq 2.72$. If this test has a truth value of 1 (resp. 0), the data is passed to the next test : $\Phi_{CH_4} \leq 285,179,380$ mol yr^{-1} (resp. $\Phi_{H_2} \leq 78,416,564$ mol yr^{-1}) and so on. The goal of the decision tree is to split the pseudo-data set in subsets ("leaves") corresponding as faithfully as possible to the classes uninhabitable ($\bar{H}$), abiotic-habitable (H^{ab}) and biotic (B).

The RF is a population of trees that categorizes data points into classes according to the votes of the trees in the decision forest. The pseudo-data (or the training dataset) correspond to a set of features associated with their class (together, they are the reference table) and are used to fit the RF classifier. In other words, the algorithm consists in selecting the decision forest that best classifies the pseudo-data in the training dataset. In order to obtain a population of decision trees that splits the data set well, we evaluate the Gini Impurity, which measures the probability to misclassify a point in a set or subset if classification is done at random. The general expression of the Gini Impurity is:

$$G = \sum_i p(i)(1 - p(i)) \tag{30}$$

where $p(i)$ is the probability of attributing class i to a point when the class of the point is i. The overall Gini Impurity of several subsets is the weighted sum of G in each subset.

A measure of the efficiency of the method is the overall classification error rate (or prior error rate), ζ: the fraction of pseudo-data points that the fitted classifier misclassifies. Conversely, the score $1-\zeta$ measures the success rate of the classifier. One can also measure this score for a specific class and compute how wrong the classifier is when it misclassifies pseudo-data points, for instance "with what class does the classifier confuses *B* the most?". The representation of how the RF classifier confuses different classes is the confusion matrix, of which an example is shown in Extended Data Fig. 2b.

We trained a decision forest on the pseudo-data set of $50,000$ simulations generated from the priors reported in Extended Data Table 1 using the `RandomForestClassifier` function from the python package *ScikitLearn*[52] with parameters reported in Extended Data Table 3. In ABC-RF, the posterior probability is approximated by averaging the class probability of each tree. The class probability of *i* in a tree is the fraction of class *i* in the leaf corresponding to the prediction. A full description of ABC-RF is found in Pudlo et al. (2015)[34].

Building the priors

Input parameters of the ecological model are the composition and temperature of the HF and composition of the ocean in species of interest (H_2, DIC, CH_4), with their distribution defining the prior space θ. In our approach, the temperature of Enceladus' ocean is fixed and thus not drawn from a prior distribution. A well-supported estimate for the temperature of Enceladus' ocean is 275 K, close to thermal equilibrium with ice[3,10]. High ammonia concentrations could lead to a lower freezing point of water, and thus a colder ocean, the consequences of which are discussed in Supplementary Results and Discussion. In this section, we explain the ranges chosen for the concentrations in HF and OW of H_2, DIC and CH_4, and for the temperature of the HF. A summary is given by Extended Data Table 1.

Constraining H_2 in the hydrothermal fluid and in the ocean. Following Waite et al. (2017)[3], we assume that the dihydrogen present in the hydrothermal fluid is produced in Enceladus' core. Choblet et al. (2017)[11] have argued that serpentinization was a plausible source of H_2 in Enceladus' core, and considering other potential sources Waite et al. (2017)[3] concluded that they were likely insufficient to explain observed H_2 flux. We therefore assume that serpentinization is the primary source of dihydrogen. A maximum value for the H_2 concentration in the hydrothermal fluid is the concentration at saturation with the vapor phase produced during serpentinization[32]. This value may be close to 10^{-1} mol kg^{-1} at pressures of the order of 10 MPa, the expected hydrostatic pressure at Enceladus' seafloor. We neglect the hydrogen saturation concentration variation with temperature and thus make the assumption that the concentration of H_2 in the fluid is drawn independently from temperature, according to

$$[H_2]_f \sim \log_e \mathbf{U}(10^{-8}, 10^{-1}) \text{ mol kg}^{-1} \tag{31}$$

where $\mathbf{U}$ denotes the uniform distribution. We note that, depending on the rock and water composition, the maximum dissolved H_2 may be lower than the equilibrium with hydrostatic vapour phase[32,53]. As a consequence, we may be overestimating the production of H_2 from serpentinization.

We assume that the concentration of dihydrogen in the ocean is very low since it is produced in the core, as suggested by Waite et al. (2017)[3] :

$$[H_2]_o \sim \log_e \mathbf{U}(10^{-8}, 10^{-6}) \text{ mol kg}^{-1} \tag{32}$$

Note that assuming low levels of dihydrogen in oceanic waters is equivalent to assuming that most of the dihydrogen escapes or is captured in clathrates instead of recirculating in the core. Choblet et al. (2017)[11] suggest that the whole ocean may be processed in the core at relatively short timescales (25-250 Myr) and that the composition of the ocean may depend on ice melting above hydrothermal hot spots and recrystallisation in other regions.

Constraining CH_4. Methane concentration in the hydrothermal fluid is an important parameter of methanogenesis, as an increased concentration reduces the thermodynamical favorability of the reaction (1). In a first set of simulations, we limit the source of CH_4 to serpentinization chemistry, a source that was suggested in previous studies[13]. The rate of reaction (1) in serpentinization waters is subject to debate but recent experimental work suggests that this reaction conducted abiotically faces significant energy barriers and would be extremely slow across the temperature range considered here[33]. The same study showed that in low-pressure (10 MPa) hydrothermal systems where a H_2 vapor phase can form, the reaction is significantly accelerated. Because the hydrostatic pressure at the bottom of Enceladus ocean might be of that order[20], we assume that CH_4 concentration in the hydrothermal fluid can be as high as reported by McCollom (2016)[33], *i.e.* 0.1 mmol kg^{-1}, so that:

$$[CH_4]_f \sim \log_e \mathbf{U}(10^{-8}, 10^{-4}) \text{ mol kg}^{-1}. \tag{33}$$

Note that we neglect any conversion of H_2 and CO_2 in the mixing layer, where temperatures are lower. In the alternate scenario, we do not limit $[CH_4]_f$ to serpentinization chemistry, and use an arbitrarily higher upper bound so that

$$[CH_4]'_f \sim \log_e \mathbf{U}(10^{-8}, 10^{-1}) \text{ mol kg}^{-1}. \tag{34}$$

In the ocean, and following what was done for H_2, we take

$$[CH_4]_o \sim \log_e \mathbf{U}(10^{-8}, 10^{-6}) \text{ mol kg}^{-1}. \tag{35}$$

Constraining the dissolved inorganic carbon (DIC). On the geological timescale, Earth's oceans are recharged with CO_2 through volcanic outgassing and carbonate dissolution. Glein et al. (2015)[10] suggest that dissolution of carbonates may be Enceladus' main source of ocean dissolved carbon dioxide, whereas magma outgassing seems unlikely given the small and relatively cold core of Enceladus. Dissolved inorganic carbon is the total of dissociated forms of CO_2 in water ($DIC = [CO_2] + [HCO_3^-] + [CO_3^{2-}]$). DIC concentration in Enceladus' ocean was estimated using plumes chemical data from the Cassini mission[10]. The total dissolved inorganic carbon was constrained between 0.1 and 0.004 mol kg^{-1} :

$$[DIC]_o \sim \log_e \mathbf{U}(4 \cdot 10^{-3}, 10^{-1}) \text{ mol kg}^{-1}. \tag{36}$$

Our first order model does not take into account the complex carbonate chemistry in the Enceladus ocean that Glein & Waite (2020)[54] investigate. Instead, we have identified the activity of CO_{2aq} with the total dissolved inorganic carbon. This is a first approximation which should be refined in the future by coupling the current model with that of Glein & Waite (2020)[54]
For our analysis to be conservative, we assume the hydrothermal fluid to be DIC-poor and thus use a much lower range

$$[DIC]_f \sim \log_e \mathbf{U}(4 \cdot 10^{-8}, 10^{-6}) \text{ mol kg}^{-1}. \tag{37}$$

Constraining T_f. In Choblet et al. (2017)[11], temperatures in the core were computed from a range of plausible core porosity (K_p) constrained by the observation of nanometre-sized silica particles and water boiling point (limit of their model). We assume that the water boiling point is not exceeded and that the $\log_e$ of porosity is *a priori* uniformly distributed between 10^{-14} and 10^{-12}. Using the approximation $T \propto \log_e K_p$, we conclude that the hydrothermal fluid temperature is drawn from a uniform distribution :

$$T_f \sim \mathbf{U}(300, 620) \text{ K} \tag{38}$$

For comparison, Earth's hydrothermal systems have been observed venting fluids at 640 K [55]. The temperature of Enceladus ocean is set at a single value of $T_o = 275$ K, close to thermal equilibrium with ice, in agreement with previous work[3,10].

Selection of observables

Choosing reliable and informative observables is a crucial step of the inference framework. Observables must be model outputs that are confidently measured. Because our model captures only a subset of Enceladus geochemical processes that potentially influence the plume composition, we must carefully evaluate whether the model outputs may faithfully translate into measurements that are available from the Cassini mission.

The model takes the composition and temperature of the hydrothermal fluid, and the composition of the ocean, as inputs. Then the model represents the mixing of hydrothermal fluid and ocean water. In this mixing layer, biological methanogenesis might occur, or not. If it does, the local composition of the mixing layer is altered according to the biological model. The mixing layer waters are then mixed together to form an homogeneous oceanic plume water mass. At this point, there is uncertainty in the amount of dilution of these waters with oceanic waters during their ascent across the ocean. Our mixing model predicts an upward advection of 10^4 to 10^5 kg s^{-1} at the hot spot. Using scaling relationships for Europa[27] that have been adapted for Enceladus[11], we find water transports at the top of the column that are three orders of magnitude higher ($\approx 10^8$ kg s^{-1}). This discrepancy between predictions from our mixing model and scaling relations in Goodman et al. (2004)[27] indicates that our approach might greatly underestimate the dilution of ML waters during their ascension towards the ice ceiling, and *a fortiori* neglect the dynamics of ascending waters. To circumvent the challenge of modelling the mixing of the oceanic plume while it travels upward, we select observations that carry information on seafloor processes, independently of dilution in the ascending plumes.

In Waite et al. (2017)[3], the authors assume that the source of dihydrogen in Enceladus compensates for the loss to space from the gas jets at the plume's origin. Following the same approach, we assume that the outward flux is equal to the flux of hydrogen into the ocean (coming out of the mixing layer). Because H_2 is assumed to be produced in the core, we integrate the outward flux from the mixing layer at steady state

$$\Phi_{H_2} = \int_S J_c[H_2]\, dS \text{ (mol yr}^{-1}\text{)} \tag{39}$$

and thus define an observable that should be consistent with model outputs regardless of dilution in the oceanic plume.

Following Bouquet et al. (2015)[23], we assume that a source of CH_4 in the core compensates clathrate dissociation so that the release rate in the ocean is approximately equal to the measured ejection rate

$$\Phi_{CH_4} = \int_S J_c[CH_4]\, dS \ (\text{mol yr}^{-1}) \tag{40}$$

We also use

$$R = \frac{[H_2]}{[CH_4]} \tag{41}$$

as an observable. As the other two observables, R should also be independent of dilution. Using R as an observable, however, requires to further assume that this ratio is conserved when the ocean material travels through the ice. This assumption is discussed in Waite et al. (2017)[3], where it was used to estimate concentrations in H_2, CO_2 and CH_4 in the ocean.

Other observables

Other observables have been considered but were not used in our inference framework. First, the observed flux of CO_2 in the plume, as well as the $H_2 : CO_2$ ratio might contain information on whether hydrogenotrophs might be active or not. We did not use these measurements as observables because the CO_2 source is unknown; as we only assume that CO_2 in the mixing layer comes from the ocean, this makes the observed CO_2 flux and $H_2 : CO_2$ ratio sensitive to dilution of the oceanic plume which, as mentioned above, is a process that is poorly constrained in our model. Furthermore, our knowledge of what might be the CO_2 stock in Enceladus' ocean comes from the Cassini measurements that we aim to use for inference. As a consequence, the observation point would be as informative as the prior knowledge, and thus inference would not be improved.

Second, considering the value in Enceladus's ocean of the Gibbs free-energy associated with reaction (1), Waite et al. (2017)[3] inferred a negative value, and concluded that methanogenesis was feasible in Enceladus ocean. The Gibbs free energy ΔG is computed from the reaction quotient Q and temperature T (equation (12)). The reaction quotient in the oceanic plume might be useful as an observable because its value in the ML is crucial to habitability (equation (21)). In Waite et al. (2017)[3], the concentrations of reactants used to compute Q are inferred by obtaining CO_2 concentration from carbonate equilibrium (using a pH value for Enceladus' ocean of $9-11$) and deriving other concentrations using the ratios $CH_4 : CO_2$ and $H_2 : CO_2$ in the plume. The temperature they use is the putative temperature of Enceladus' ocean (275 K). Here, we did not use the reaction quotient as an observable because of the serious difficulty of computing Q in the plume, and because of its sensitivity to the dilution of the oceanic plume, which is poorly constrained in our model (see above).

Code availability The code of the model presented in the article are available at xxx.gitlab.com.

Data availability Simulated datasets from which the figures were generated are available at xxx.gitlab.com.

References

[45] Sauterey, B., Charnay, B., Affholder, A., Mazevet, S. & Ferrière, R. Co-evolution of primitive methane-cycling ecosystems and early Earth's atmosphere and climate. *Nature Communications* **11**, 1–12 (2020).

[46] Lever, M. A. et al. Life under extreme energy limitation: a synthesis of laboratory-and field-based investigations. *FEMS microbiology reviews* **39**, 688–728 (2015).

[47] Connolly, J. P. & Coffin, R. B. Model of carbon cycling in planktonic food webs. *Journal of Environmental Engineering* **121**, 682–690 (1995).

[48] Virtanen, P. et al. SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python. *arXiv e-prints* , arXiv:1907.10121 (2019).

[49] Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M. & Charnov, E. L. Effects of size and temperature on metabolic rate. *science* **293**, 2248–2251 (2001).

[50] Csilléry, K., Blum, M. G., Gaggiotti, O. E. & François, O. Approximate Bayesian computation (ABC) in practice. *Trends in ecology & evolution* **25**, 410–418 (2010).

[51] Sisson, S. A., Fan, Y. & Beaumont, M. *Handbook of approximate Bayesian computation* (Chapman and Hall/CRC, 2018).

[52] Pedregosa, F. et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

[53] Tutolo, B. M., Seyfried, W. E. & Tosca, N. J. A seawater throttle on H2 production in Precambrian serpentinizing systems. *Proceedings of the National Academy of Sciences* **117**, 14756–14763 (2020).

[54] Glein, C. R. & Waite, J. H. The carbonate geochemistry of Enceladus' ocean. *Geophysical Research Letters* **47**, e2019GL085885 (2020).

[55] Charlou, J., Donval, J., Fouquet, Y., Jean-Baptiste, P. & Holm, N. Geochemistry of high H2 and CH4 vent fluids issuing from ultramafic rocks at the Rainbow hydrothermal field (36 14' N, MAR). *Chemical geology* **191**, 345–359 (2002).