

Formation mechanism of thermally controlled pH gradients

Thomas Matreux ^{1,3}, Bernhard Altaner^{2,3}, Johannes Raith^{2,3}, Dieter Braun ¹, Christof B. Mast ¹✉ & Ulrich Gerland ²✉

Spatial proton gradients create energy in biological systems and are likely a driving force for prebiotic systems. Due to the fast diffusion of protons, they are however difficult to create as steady state, unless driven by other non-equilibria such as thermal gradients. Here, we quantitatively predict the heat-flux driven formation of pH gradients for the case of a simple acid-base reaction system. To this end, we (i) establish a theoretical framework that describes the spatial interplay of chemical reactions with thermal convection, thermophoresis, and electrostatic forces by a separation of timescales, and (ii) report quantitative measurements in a purpose-built microfluidic device. We show experimentally that the slope of such pH gradients undergoes pronounced amplitude changes in a concentration-dependent manner and can even be inverted. The predictions of the theoretical framework fully reflect these features and establish an understanding of how naturally occurring non-equilibrium environmental conditions can drive pH gradients.

¹Systems Biophysics, Ludwig Maximilian University Munich, 80799 Munich, Germany. ²Physics of Complex Biosystems, Technical University of Munich, 85748 Garching, Germany. ³These authors contributed equally: Thomas Matreux, Bernhard Altaner, Johannes Raith. ✉email: christof.mast@physik.uni-muenchen.de; gerland@tum.de

Cells from all kingdoms of life use pH gradients to drive the synthesis of ATP via ATP synthase¹. This is often seen as an indication that pH gradients were also a driving force during the emergence of life. This hypothesis is substantiated by phylogenetic analysis² and geochemical considerations^{3,4}. In non-equilibrium settings, such as heat flows through rock fissures or hydrothermal systems, pH gradients can provide a prebiotic proton-motive force to drive early metabolisms⁵. Prebiotic systems may also benefit from pH gradients to localize pH-selective chemical reactions such as nucleotide adsorption⁶ to certain spatial zones, or to spatially couple the effects of ribozymes that operate at different pH⁷. However, the physicochemical mechanisms that can produce pH gradients in plausible non-equilibrium settings are so far insufficiently understood.

Experimentally, a controlled non-equilibrium setting to study the formation of chemical gradients is created by heat flows across microfluidic chambers^{8,9}. Such experiments mimic ubiquitous natural settings, e.g. in porous rocks near hydrothermal vents¹⁰ (Fig. 1b). A previous study demonstrated that heat flow across narrow (100 μm -scale) pores can induce pH gradients of 2 units along the pore length (mm-scale)¹¹. The gradients were generated by a complex interplay of convection, thermophoresis, electrostatic interactions, and acid-base reactions.

Here, we focus on the physical parameters of this interplay to understand the control of both the amplitude and the direction of the pH gradient (Fig. 1c). We performed an experimental case study (Fig. 1a) of a prebiotically relevant^{12,13} acid-base system consisting of formic acid (HCOOH) and sodium hydroxide (NaOH) diluted in an aqueous solution. The mixture was kept within a microfluidic chamber and subjected to a lateral heat flow. Based on a separation of timescales, we then developed a heuristic theoretical description for this reaction-transport system, which incorporates the electrostatic coupling between the reacting species, as well as their transport due to drift and diffusion processes. This description captures the parameter dependence of the experimentally observed behavior, including a concentration-

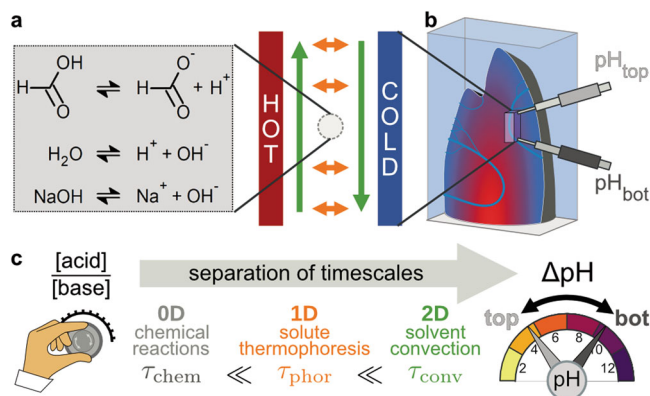


Fig. 1 Heat flows create pH gradients in acid-base mixtures. **a** Dominant dynamic processes in a heat flux chamber: Chemical reactions (gray) remain locally equilibrated ($\tau_{\text{chem}} \lesssim 10^{-3}$ s). Thermophoretic horizontal accumulation ($\tau_{\text{phor}} \sim 10^2$ s, orange arrows) is faster than vertical Rayleigh-Bénard convection ($\tau_{\text{conv}} \sim 10^4$ s, green arrows), given the high aspect ratio (height/width) of the heat flux chamber that mimics, e.g., a rock fissure. Thermophoresis and convection combined drive a selective accumulation of molecules. Their net accumulation is proportional to their effective Soret coefficient $S_H^{\text{eff},0}$, derived in the main text. **b** Heat flows across rock fissures provide a non-equilibrium environment that drives proton gradients. By analyzing local outflows with standard pH-meters, we measure the pH values quasi in situ. **c** By tuning the concentrations of acids and bases, both amplitude and direction of the created pH gradient can be controlled.

dependent steepening and inversion of the pH gradient. The model can be generalized to systems with more reacting species subject to electrostatic coupling, as long as the effective timescales of the local reactions and the global transport are sufficiently separated. The quantitative understanding enables the prediction and the control of thermally induced pH gradients.

Results

Experimental setup and procedure. For this study, we developed a continuous flow microfluidic device with openings that allowed us to probe the local pH at three different outflow positions and using standard pH meters (Fig. 2a; see Methods for a detailed description). This setup avoids complications imposed by a fluorescent pH reporter needed to monitor pH in a closed device¹¹.

The microfluidic structure containing the fluid was cut out of fluorinated ethylene propylene (FEP) foil and subsequently sandwiched between two sapphire windows (chamber dimensions: width $w = 170 \mu\text{m}$, height $h = 50 \mu\text{m}$). This design guaranteed that the liquid sample only has contact to FEP and sapphire, providing an inert environment even for aggressive chemistry. Sapphire was chosen for its transparency (for optical control) and high heat conductivity (to ensure a homogeneous temperature distribution). One of the sapphire windows was

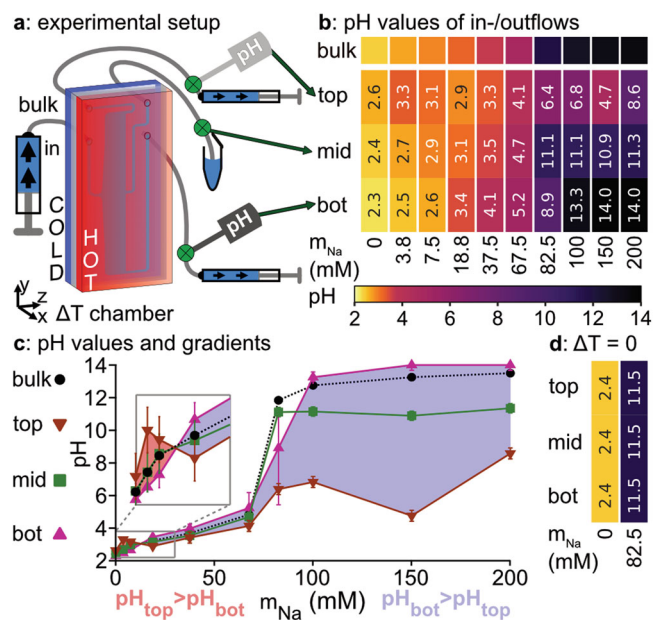


Fig. 2 Experimental setup and results. **a** Experimental setup. The thin thermal reaction chamber is continuously supplied with a premixed solution containing varying concentrations of formic acid and sodium hydroxide. The outflow is split into three parts (top 12%, middle 76% and bottom 12%) and the pH value of each fraction is measured (green circles). For experimental details see Methods. **b** These pH values are depicted for 75 mM formic acid and varying concentrations of NaOH (m_{Na}) for the inlet (bulk) and the different outlets. While the bottom fraction qualitatively follows the bulk, especially the top fraction displays a complex, concentration-dependent behavior both for its value and for the resulting pH gradient (ΔpH)_{vert} between top and bottom of the heat flux cell (measurements for other formic acid concentrations shown in Supplementary Fig. 2). **c** When plotting the pH values of the different fractions (black circle: bulk, brown inverted triangle: top, green box: middle, pink triangle: bottom), a positive gradient ($\text{pH}_{\text{top}} > \text{pH}_{\text{bot}}$, red) for low concentrations of NaOH is observed which is inverted (blue) at higher concentrations of NaOH and increases massively when $m_{\text{Na}} > m_{\text{HCOOH}}$. Error bars are shown as discussed in Methods. **d** In the absence of a temperature gradient, no pH gradient forms ($m_{\text{HCOOH}} = 75 \text{ mM}$).

cooled by an aluminum heat exchanger connected to a cryostat, while the other window was kept at high temperatures by Ohmic heating, resulting in a constant heat flux through the chamber. The device had one inlet at the left side and three outlets at the right side, at different vertical positions ('top', 'middle', 'bottom') where the pH value was measured.

The inlet was used to push a mixture of pure water, acid and base (with initial pH value denoted as pH_0) at a flow rate of 3.28 nl/s into the chamber. Using syringe pumps at each of the outlets, the flux splitting between the three outlets was adjusted to outflow fractions of 12% (top), 76% (middle), and 12% (bottom). These values reflect a trade-off that kept the thermogravitational accumulation within the chamber relatively unperturbed, the total time required for the experiment manageable and the spatial regions from which the top and bottom samples are collected relatively localized (see Methods for a detailed discussion).

Experimental pH gradients. Using this setup, we performed experiments with solutions containing different amounts of formic acid (m_{form}) and NaOH (m_{Na}). The heat flow through the chamber was set by keeping the temperatures at the front and back at 60 ± 2 °C and 21 ± 1 °C, respectively. The pH values of the different fractions obtained with 75 mM formic acid and varying amounts of NaOH are shown in Fig. 2b, c (for other formic acid concentrations see Supplementary Fig. 2).

While the pH in the bottom of the cell qualitatively follows the pH behavior in the equilibrated bulk solution (see Supplementary Fig. 1), the pH in the middle and top fractions displays a markedly different behavior as a function of the NaOH concentration. The pH dependence in the bottom fraction of the cell and the bulk solution is consistent with well-known acid-base characteristics, with a change from acidic to alkaline at equimolar concentrations of NaOH and formic acid. However, the pH behavior in the top fraction is surprisingly complex, leading to a previously not observed¹¹ inversion of the pH difference (ΔpH)_{vert} (see Fig. 3) between the top and bottom fraction at low NaOH concentrations (~ 10 mM). Above equimolarity, the pH difference sharply increases. Despite NaOH being a strong base, even at NaOH concentrations much larger than formic acid concentrations, the pH value of the top fraction is actively pushed into the acidic regime by the heat flux.

Theoretical approach. The complex experimental behavior of the emergent pH gradient arises from the combined effects of diffusion, convection, thermophoresis, electrostatic interactions, and acid-base reactions. The interplay of these effects can be studied by numerically solving the underlying reaction-transport equations, but a full characterization of the system is challenging, due to the large parameter space and the multiple different timescales of the dynamics¹¹.

Due to thermal expansion of the solvent and gravity acting in vertical direction, an approximately rotationally symmetric circular buoyant convection flow is established in the chamber: solutes closer to the hot side will be driven upwards while molecules on the cold side will be transported downwards. Additionally, dissolved molecules are driven along the horizontal temperature gradient by thermophoresis. While the magnitude of the ensuing velocities of both processes are comparable, a small aspect ratio w/h causes the horizontal equilibration to occur on a faster timescale than the vertical accumulation, see Fig. 1a and Methods. Seminal work by P. Debye showed that an expansion in the aspect ratio yields approximate solutions, where concentration gradients along the horizontal dimension shape those along the vertical dimension of the setup¹⁴. Electrostatic interactions were taken into account for electrolytes by Guthrie and coworkers¹⁵. From a theoretical perspective, the task at hand therefore is to

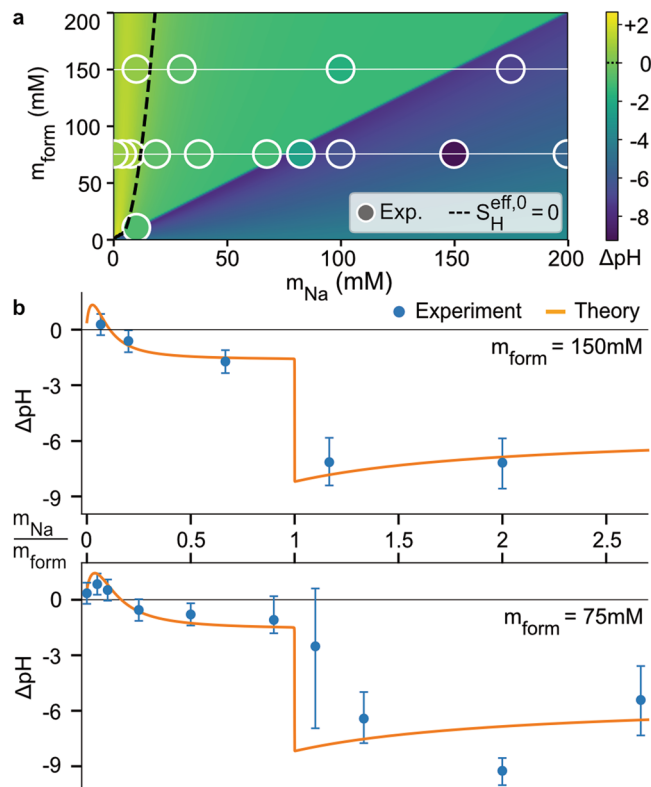
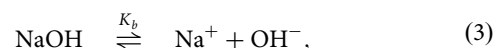
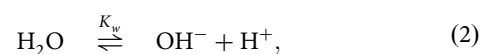


Fig. 3 Comparison of analytically predicted and experimentally measured pH gradients. **a** Heat maps of theoretically calculated pH differences are compared to measurements (circles) for varying concentrations of formic acid (m_{form}) and sodium hydroxide (NaOH, m_{Na}). Our model predicts a complex phase diagram with a positive (yellow), a weakly negative (green) and a strongly negative (blue) pH gradient. The inversion is shown by a black dashed line and corresponds to $S_{\text{H}}^{\text{eff},0} = 0$, see Eq. (9). The experimental results are in very good agreement with the analytical model. **b** Theoretical (orange line) and experimental (blue circles) pH differences for 150 mM and 75 mM formic acid (white lines in **a**), showing again the good agreement between theory and experiment. The error bars include experimental standard deviations from replicate measurements as well as uncertainties accounting for concentration fluctuations from thermal gradient effects in the outflow channels (for more details see Methods).

explicitly integrate chemical reactions to the theoretical description, and to study their effect on the system behavior. Until now, only the case of a chemical reaction of a single species without electrostatic interactions has been considered¹⁶.

In this work, we derive an expression for vertical concentrations gradients, in particular, the proton gradient, as a function of the initial chemical concentrations. In the following, we will rely on a timescale separation between the (zero-dimensional) chemical reaction dynamics, the (one-dimensional, horizontal x -axis) effects of thermophoresis and the (two-dimensional, x - y axis) effects of buoyant convection, see Fig. 1.

Chemical dynamics. Our system is determined by the system of the $\mathcal{R} = 3$ chemical reactions



where we assume that the concentration of H_2O is approximately constant. The equilibrium concentrations c_i^0 of the $\mathcal{N} = 6$ dynamical chemical species X_i are determined by the equilibrium constants for the acid, base, and water, respectively:

$$\begin{aligned} K_a &= \frac{c_{\text{COOH}^-}^0 c_{\text{H}^+}^0}{c_{\text{HCOOH}}^0}, & K_b &= \frac{c_{\text{Na}^+}^0 c_{\text{OH}^-}^0}{c_{\text{NaOH}}^0}, \\ K_w &= c_{\text{OH}^-}^0 c_{\text{H}^+}^0. \end{aligned} \quad (4)$$

In general, closed chemical systems always conserve certain chemical sub-species, so-called moieties¹⁷. We define the moiety concentrations

$$\begin{aligned} m_{\text{form}} &= c_{\text{COOH}^-} + c_{\text{HCOOH}}, & m_{\text{Na}} &= c_{\text{NaOH}} + c_{\text{Na}^+}, \\ m_c &= c_{\text{Na}^+} + c_{\text{H}^+} - c_{\text{COOH}^-} - c_{\text{OH}^-}. \end{aligned} \quad (5)$$

Here, m_{form} and m_{Na} are the total concentration of formate and sodium-containing molecules in the system, which are experimentally controllable parameters. Similarly, m_c is the net charge concentration. Notice that the moieties are not uniquely defined, and any linear combination of a moiety yields another conserved quantity. However, we can always define them in a way that includes the charge moiety and experimentally controllable parameters. Specifying the moiety concentrations m_{Na} and m_{form} and using, due to charge neutrality, $m_c = 0$, together with Eqs. (4) uniquely determines the values of the equilibrium concentrations c_i^0 of all chemical species. Even though the full system is not at equilibrium, given that chemical reactions occur on the fastest timescale $\tau_{\text{chem}} < 10^{-3}$ s, the relations between the homogeneous equilibrium concentrations Eqs. (4), will hold approximately as local relations also in a spatially extended system.

Horizontal currents. Next, we consider the horizontal currents that form due to thermophoresis in an effective one-dimensional system. This is motivated since the timescale of the horizontal dynamics is given by $\tau_{\text{phor}} \gg \tau_{\text{chem}}$ (see Methods). The horizontal flux $J_i(x)$ of molecule species X_i with horizontal concentration profile $c_i(x)$ is given by

$$J_i = -D_i \frac{\partial c_i}{\partial x} - \left(S_i D_i \frac{\partial T}{\partial x} - z_i D_i \frac{FE}{RT} \right) c_i, \quad (6)$$

where D_i , S_i , and z_i are the diffusion constant, Soret coefficient, and valence of X_i , $T(x)$ is a linear temperature profile, F the Faraday constant, and E the electrical field. In a steady state, the divergence of these currents is balanced by a source term given by the chemical reactions. Due to moiety conservation in local chemical reactions, the associated moiety currents are divergence-free. Since in a one-dimensional closed system any divergence-free current must vanish, it follows that the (horizontal) moiety currents I vanish identically. The vanishing electric current, $I_c = J_{\text{Na}^+} + J_{\text{H}^+} - J_{\text{COOH}^-} - J_{\text{OH}^-} = 0$, allows us to solve for the electrical field

$$\frac{FE}{RT} = \frac{\sum_i z_i D_i \left(\frac{\partial c_i}{\partial x} + S_i c_i \frac{\partial T}{\partial x} \right)}{\sum_i z_i^2 c_i D_i}. \quad (7)$$

Inserting this expression into (6) allows us to express the currents $\mathbf{J} = (J_{\text{H}}, J_{\text{OH}}, J_{\text{HCOOH}}, J_{\text{COOH}}, J_{\text{Na}}, J_{\text{NaOH}})$ as linear functions of the horizontal concentration gradients $\partial_x \mathbf{c} = \left(\frac{\partial c_{\text{H}}}{\partial x}, \frac{\partial c_{\text{OH}}}{\partial x}, \frac{\partial c_{\text{HCOOH}}}{\partial x}, \frac{\partial c_{\text{COOH}}}{\partial x}, \frac{\partial c_{\text{Na}}}{\partial x}, \frac{\partial c_{\text{NaOH}}}{\partial x} \right)$ as $\mathbf{J} = \mathbf{D} \partial_x \mathbf{c} + \mathbf{b}$. The matrix \mathbf{D} is a (cross-)diffusion matrix with components

$$\begin{aligned} D_{ij} &= \delta_{ij} D_i + D_j z_j c_j \frac{z_i D_i}{\sum_k z_k^2 c_k D_k} \text{ and } \mathbf{b}_j = -D_j c_j \frac{\partial T}{\partial x} S_j^{\text{ES}}, \text{ where} \\ S_j^{\text{ES}} &= S_j - z_j \frac{\sum_k z_k D_k S_k c_k}{\sum_k z_k^2 c_k D_k} \end{aligned}$$

is the Seebeck coefficient, see also refs. 15,18.

The remaining two horizontal moiety currents $\mathbf{I} = (I_{\text{Na}}, I_{\text{form}}) = \mathbf{L} \mathbf{J}$ are given by a $2 \times \mathcal{N}$ matrix \mathbf{L} as a linear combination of \mathbf{J} . Similarly to the charge current I_c , they are divergence-free and thus vanish, and we can write $\mathbf{I} = 0$ as a linear equation for the horizontal concentration gradients $\partial_x \mathbf{c}$

$$(\mathbf{L} \mathbf{D}) \partial_x \mathbf{c} = \mathbf{L} \mathbf{b}. \quad (8)$$

In order to uniquely solve for the concentration gradients, additional constraints are needed. One of these constraints is given by the condition of approximate local electroneutrality in the bulk, $m_c(x) = 0, \forall x$, which can be systematically derived^{19,20}. Differentiating this condition with respect to the horizontal position yields $\sum_i z_i \partial_x c_i = 0$, which we can write with a $1 \times \mathcal{N}$ matrix \mathbf{Z} with entries z_i as:

$$\mathbf{Z} \partial_x \mathbf{c} = 0. \quad (9)$$

Finally, since chemical reactions are fast as compared to the transport processes, we can assume that their concentrations $c_i(x)$ obey Eqs. (4) as *local equilibrium* relations. Taking the horizontal derivatives yields

$$\mathbf{Q} \partial_x \mathbf{c} = 0, \quad (10)$$

where \mathbf{Q} is the $3 \times \mathcal{N}$ (Jacobian) matrix of the mass-action ratios.

Equations (8), (9) and (10) constitute $\mathcal{N} = 6$ linearly independent equations for the $\mathcal{N} = 6$ concentrations gradients. They can be inverted with a computer algebra system to give the concentrations gradients $\partial_x \mathbf{c}$ as non-linear and convoluted expressions of the concentrations. In particular, we obtain a solution for the proton gradient, which we can formally write as

$$\frac{\partial c_{\text{H}}}{\partial x} = -S_{\text{H}}^{\text{eff}}[\{c_i(x)\}] \frac{\partial T}{\partial x} c_{\text{H}}. \quad (11)$$

with the concentration-dependent effective Soret coefficient for the hydrogen ion $S_{\text{H}}^{\text{eff}}$, whose sign and magnitude determine the horizontal proton gradient (see also Supplementary Note 1 for a complete solution). We consequently expand $c_i(x) = c_i^0 + \gamma_i(x)$ around the (homogeneous) bulk equilibrium concentrations, i.e. $S_{\text{H}}^{\text{eff}} = S_{\text{H}}^{\text{eff},0}(\{c_i^0\}) + \mathcal{O}(\{\gamma_i(x)\})$. For small temperature gradients and small widths of the heat flux chamber, we assume $\gamma_i \ll c_i^0$ and neglect the inhomogeneous part. Since the c_i^0 are uniquely determined by the moieties, i.e. the controlled molar concentrations of sodium hydroxide, m_{Na} , and formic acid, m_{form} , the homogenous effective Soret coefficient $S_{\text{H}}^{\text{eff},0}(m_{\text{Na}}, m_{\text{form}})$ can be calculated from tunable experimental conditions.

The scheme outlined above will work for *arbitrary* chemical reactions networks and is thus general. On first glance, it might be surprising that we obtain exactly \mathcal{N} independent linear equations for the \mathcal{N} gradients of dynamical species. However, this fact can be traced to the algebraic structure of the stoichiometric matrix¹⁷. In particular, situations can arise where not all of Eqs. (4) leading to Eq. (10) are independent. However, the number of independent equations defining the moieties, Eqs. (5), and thus Eqs. (8) are algebraically related, such that we can always find \mathcal{N} linear equations for the gradients of \mathcal{N} dynamical species, regardless of the number \mathcal{R} of reactions involved (see Methods or ref. 17).

Vertical pH gradient. Integrating equation (11) along the horizontal axis yields a horizontal pH difference $(\Delta \text{pH})_{\text{hor}} \approx \frac{\Delta T}{\ln(10)} S_{\text{H}}^{\text{eff},0}$. Given such horizontal gradients, a theory for calculating the resulting

vertical concentration gradient in a Clusius column has been formulated by Debye¹⁴, and adapted later to a situation including electrical fields¹⁵. This approach works, since the corresponding convective timescale τ_{conv} is much larger than the thermophoretic timescale, $\tau_{\text{conv}} \gg \tau_{\text{phor}}$ (Methods). It then turns out that the relation between horizontal and vertical gradients is given by a linear factor $\alpha_{\text{th}} = 358.5$, whose value depends only on material parameters of the solution, the geometric aspect ratio of the chamber, and the applied temperature gradient (see Methods). As such, the effective Soret coefficient $S_{\text{H}}^{\text{eff}}$ for the horizontal dynamics can be used to estimate the vertical pH difference $(\Delta\text{pH})_{\text{vert}}$ measured experimentally,

$$(\Delta\text{pH})_{\text{vert}} \approx \alpha_{\text{th}} \cdot S_{\text{H}}^{\text{eff},0}(m_{\text{Na}}, m_{\text{form}}) \quad (12)$$

Hence, weaker and more plausible temperature gradients in nature can be balanced by equally more plausible larger aspect ratios. As consequence, there is no lower threshold of temperature difference required for the formation of proton gradients, since shallow gradients can be compensated with higher aspect ratio²¹.

Comparison with experiments. The heat map from our theoretical model can roughly be divided into three regions: positive (yellow), weakly negative (green) and strongly negative (blue). As depicted in Fig. 3a, our measurements (shown in circles) fully recapitulate this behavior, both for the amplitude of the pH gradient and its inversion.

To avoid distortion from the flow extraction for pH measurement, we chose a slow flow ($\tau_{\text{flow}} = 13,000$ s). This timescale is the time that a volume element on average remains in the heat flux chamber before it is extracted horizontally by the measurement cross-flow (z -axis in Fig. 2a). The slow extraction can however also slightly change the $m_{\text{Na}}/m_{\text{form}}$ ratio in the effluent channel by thermophoretic accumulation, for which we estimate a systematic error of 2%. This effect has the greatest impact in the vicinity of $m_{\text{Na}} = m_{\text{form}}$, since the local pH then depends sensitively on the concentration ratio. This is reflected in the large error in pH determination at $m_{\text{Na}}/m_{\text{form}} = 1.1$ in Fig. 3b (see Methods).

Discussion

The experimental measurements are in good agreement with our analytical results. The characteristic, concentration-dependent features of an early inversion and massive increase at equimolar concentrations are reproduced without fitting parameters. The used solutes, formic acid and sodium hydroxide, are an example of a prebiotically plausible acid-base reaction system^{22–25}. Possible roles of formic acid in prebiotically relevant synthesis reactions have been investigated^{23,26–28} together with the link to formamide-based reactions^{29,30}.

Our analytical approach provides a more generally applicable theory for the interplay of chemical reactions, electrostatic interactions, and transport processes driven by external gradients. In the present case of thermally driven pH gradients, it allows a deeper understanding, and, compared to the previous computational analysis by Keil et al.¹¹, it permits the treatment of strong acids and bases and much steeper pH gradients. This theoretical advance is complemented by our experimental method to measure local pH values in the heat flux cell using standard pH meters. These can cover a far broader pH spectrum and strong acids/bases compared to fluorescence methods. Together, this enables us to control and predict proton gradients and opens the door to applications involving synthesis reactions and the formation of tunable chemical gradients.

While the local equilibrium assumption is an excellent approximation in the present experimental context, it would be desirable to relax this assumption, in order to be able to apply the

presented theoretical framework also to systems with slower chemical reactions. The behavior of such systems depends explicitly on the Damköhler number, a parameter that measures the relative timescales associated with the transport and reaction kinetics³¹. The inverse Damköhler number may serve as a useful expansion parameter to systematically extend the presented theoretical framework.

An intuitive explanation of the sign change of the pH gradient can be given by using the Soret coefficients listed in Supplementary Table 1 and considering all prevalent electrostatically coupled species. For this, we take into account that ions (H^+ , OH^- , COOH^- , Na^+) show much stronger thermophoresis than uncharged substances, which can thus be neglected in this qualitative discussion. Therefore, at vanishing concentrations of NaOH (m_{Na}), the only relevant ion pair for thermophoresis is COOH^- and H^+ . Its enrichment in the lower part of the chamber increases the proton concentration and correspondingly lowers the pH value. With higher NaOH concentrations, ion pairs such as Na^+ and OH^- become more dominant and by accumulation therefore also more prevalent at the chamber bottom. This results in an increasing local pH value, first compensating the accumulated protons, and finally inverting the pH gradient.

Along this line of reasoning, the respective influence of the contributions of thermophoresis, thermal convection, chemical reactions, and electrostatics can also be understood. While thermophoresis by itself is only a weak effect, it is exponentially amplified by thermal convection¹⁴. Neglecting electrostatic interaction in an ion-bearing system would lead to their accumulation according to their Soret coefficient (see Supplementary Table 1). Consequently, a less complex pH difference profile is achieved with higher pH differences at low concentrations of NaOH due to a larger proton transport that otherwise is counteracted by electrostatic forces (Supplementary Fig. 5c). The temperature dependence of the reaction rate constants has only a minor impact on the formation of the pH gradient and can therefore be neglected in the following considerations (see Supplementary Fig. 5b/c). In contrast, a system devoid of chemical reactions is dominated by thermogravitational accumulation. This leads to a well-defined concentration ratio of the solutes and accordingly to a constant pH difference between the upper and lower part of the chamber irrespective of the absolute NaOH concentration (see Supplementary Fig. 5a). In absence of a thermal gradient, no accumulation and no pH gradient occur (Fig. 2d). Therefore, only the consideration of all these effects together provides a comprehensive understanding of this highly coupled system.

Since heat fluxes are ubiquitous on Early Earth, requirements for proton gradients to occur are significantly less demanding than complex chemical non-equilibria e.g. originating from confluent outflows from reservoirs with different pH³². Cooperative coupling to other gradients such as ion gradients³³ will show interesting effects and help to solve enigmas such as prebiotically required separation of long RNA strands. Recent work by Mariani et al. has evaluated strand separation minimizing hydrolysis by pH oscillations³⁴. This effect was recently used in a system of air-water interfaces to boost the replication of DNA³⁵.

Methods

Preparation of heat flow chambers and microfluidic setup. The FEP-foil (Holscot, Netherlands) with a thickness of 170 μm (width) is cut by an industrial plotter device (CE6000-40 Plus, Graphtec, Germany) and sandwiched between two sapphires (Kyburz, Switzerland) with a thickness of 500 μm (cooled sapphire) and 2000 μm (heated sapphire), respectively. The back sapphire has 4 laser-cut holes with a diameter of 1 mm. The sapphire-FEP-sapphire block is then placed on an aluminum base covered by a heat conducting foil (EYG091203DP, Graphite, 25 μm , 1600 W/mK, Panasonic, Japan) and held in place by a steel frame, mounted by six torque-controlled steel screws for a homogeneous force distribution. A second heat conducting foil (EYGS0811ZLGH, Graphite, 200 μm , 400 W/mK,

Panasonic, Japan) connects the heated sapphire with the electrical heater element, and is also fixed by torque-controlled steel screws. The height of the microfluidic chamber is measured using a confocal micrometer (CL-3000 series with CL-P015, Keyence, Japan) at three positions (top, mid, bottom) to ensure homogeneous width along the chamber. Chambers are pre-flushed using low-viscosity, fluorinated oil (Novec™ 7500 Engineered Fluid, 3M, USA) to check for tightness and push out residual gas inclusions. The assembled chamber is then mounted to a cooled aluminum block. The Ohmic heaters are connected to a 400 W, 24 V power supply that is controlled via Arduino boards with a customized version of the open-source firmware “Repetier”, originally made for 3D printing. All elements were optimized using finite-element simulation to provide a homogeneous temperature profile while minimizing the thickness of the sapphire windows and hence maximizing the thermal gradient.

The in- and outlets of the thermal chamber are connected via tubings to syringes placed on high-precision syringe pumps (Low Pressure Syringe Pump neMESYS 290N with Quadruple Syringe Holder Low-Pressure, cetoni, Germany). For the microfluidic connections, (all from techlab, Germany): connectors (Verbinder zöllig, UP P-702-01), end caps (Tefzel Cap for 1/4-28 Nut, UP P-755), screws (Nut, Delrin, flangeless, VBM 100.823-100.828), ferrules (Ferrule VBM 100.632) and tubings (Tubing Teflon (FEP), KAP 100.969) are used. The flow speeds of the syringe pumps (controlled via neMESYS UserInterface, cetoni, Germany) are chosen such that 12% of the inflow are sucked out at the top and at the bottom outlet, leaving 76% coming out of the middle outlet. The syringes used (all from Göhler-HPLC Syringes, Germany) are: Chemically Resistant Heavy Duty Syringes with PTFE seals: 2606814 and 2606035 (ILS, Germany).

Sample preparation and experimental procedure. The samples are prepared mixing formic acid (Carl Roth GmbH, Germany) and sodium hydroxide (Sigma-Aldrich, USA) with MilliQ-water to reach the given concentration. The pH (for initial pH values as compared to theoretical predictions, see Supplementary Fig. 1) is measured using a Orion™ 9826BN Micro pH Electrode and a 8220BNWP Micro pH Electrode (both Thermo Fisher Scientific, USA), also used for pH measurements after the experiment. For strongly alkaline pH values (>13), measurement is done using pH paper and a larger error was assumed.

The cryostat (TXF-200 R5, Grant, UK) is set to $-30\text{ }^{\circ}\text{C}$ and the heaters to $95\text{ }^{\circ}\text{C}$ which results in a temperature difference from $\sim 20\text{ }^{\circ}\text{C}$ to $60\text{ }^{\circ}\text{C}$ within the microfluidic chamber. Temperatures are measured on the sapphires with a heat imaging camera (ShotPRO Wärmebildkamera, EAN: 0859356006217, Seek Thermal, USA) and a temperature sensor (GTH 1170, Greisinger, Germany and B&B Thermo-Technik 06001301-10, Germany), giving an average cold temperature of $21\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ and an average hot temperature of $60\text{ }^{\circ}\text{C} \pm 2\text{ }^{\circ}\text{C}$. Averaging over the temperature differences throughout all experiments gives a value of $38.9\text{ }^{\circ}\text{C} \pm 1.4\text{ }^{\circ}\text{C}$ (as the higher cold temperatures usually go together with higher warm temperatures). Before starting the experiment, the tubings and thermal chamber are re-flushed with fluorinated oil. The samples are then loaded into the inlet tubing. The experiment is started and left running for 6–7 days. To stop the experiment, the applied temperatures are set to room temperature. The samples are recovered from the tubings and collected for pH measurement.

Experimental uncertainties. The experimental error splits into two main contributions: one standard error from replicate measurements and one to account for concentration fluctuations caused by effects from the thermal gradient in the outflow channels. As described in the main text, our experimental setup was designed to determine the pH value as realistically as possible inside the heat flux cell. For this purpose, the pH values were measured with a standard laboratory pH meter and no pH-dependent dyes were used, which are not designed for use in these extreme temperature conditions and are not compatible with measurements in the entire pH range. However, since measurements using a pH meter are not possible directly inside the cell due to the small dimensions of the microfluidic system, we slowly and continuously extract the cell volume at the top, middle and bottom of the cell respectively to obtain the necessary measurement volume. By doing so and balancing the outflows with respect to each other, the cell volume is split into three fixed volume portions, of which the contents are extracted through the respective outlet. At the same time, we allow the appropriate amount of stock solution to flow through the input channel so that the cell is completely filled with fluid at all times.

The aim of these considerations was to keep the effects of this slow drift flow on the actual experiment as small as possible and thus to obtain an unbiased picture of the pH distribution inside the cell. The corresponding timescale τ_{flow} indicates how long a volume element introduced into the cell by the drift flow remains in the cell on average until extraction. If it is greater than the relaxation time of the accumulation process occurring in the heat flow cell, i.e. if the drift flow is sufficiently small, it can be assumed that the accumulation remains essentially undisturbed. If the drift flux is correspondingly too fast, the accumulation will be disturbed. For the extraction at the lower outlet, the following problem appears: In order to avoid a diffusive loss of accumulated material, we placed the exit hole for the lower output channel not at the bottom of the cell but rather sufficiently high (as shown in the main text, Fig. 2). However, this requires the to-be extracted material to overcome a thermogravitational process, since the vertical part of the exit channel has a convection flow and thermophoresis similar to the main chamber. Thus, if the drift velocity through this channel is too low, the material to

be extracted will remain at the bottom of the heat flow cell and cannot be recovered. The situation here is thus exactly the inverse of the above consideration, so that a trade-off must be found between undisturbed extraction and the lowest possible retention of the material to be extracted in the exit channel.

These considerations made us choose an inflow speed of 3.28 nl/s , resulting in a measurement extraction timescale $\tau_{\text{flow}} \sim 13,000\text{ s}$. In order to avoid a too high retention of the material, we chose an outflow speed of 0.39 nl/s , thus giving a separation into 12% at top and bottom, leaving the remaining 76% flowing out through the middle. The value for top and bottom fractions was still kept relatively low in order to average over as little material as possible and thus to keep as close as possible to the situation present inside the heat flow cell.

To account for this systematic uncertainty, we introduced an additional error in the composition of the bottom fraction. Taking the experimental standard deviation (δ^{exp}), we calculated the point on the graph for both maximal and minimal value and added/subtracted an error $m_{\text{Na}}/m_{\text{form}} = 0.02$ (called $\delta^{\text{distortion}}$). Since the determination of the composition is concentration-dependent, we measured the bottom outflows using ion chromatography and got an average concentration of 3.8 fold the inflow concentration. The errors can be merged by addition since they are independent. The resulting ratio can then be transformed back to the corresponding pH value, resulting in a complete uncertainty, potentially non-symmetrical. The procedure is shown in more details in Supplementary Fig. 3.

We expect this additional distortion to have weak effects on the measurement everywhere except for the equimolar region. The non-linearity near the equivalence line between sodium hydroxide and formic acid makes the measured pH values vulnerable to slight differences in concentration. Around this equimolar point, nearly all molecules are present in their charged state (Na^+ for m_{Na} and HCOO^- for m_{form}). Consequently, the distortion is introduced by different retentions of these ions in the outflow channels and thus dependent on their thermodiffusive properties.

Chemical reaction networks. We closely follow the works of Rao and Esposito¹⁷. Assume we have \mathcal{N} species X_i in the system, subject to \mathcal{R} reactions. The ρ -th reaction is written as

$$\sum_i \nu_{\rho+}^i X_i \xrightleftharpoons[k_{\rho-}]{k_{\rho+}} \sum_i \nu_{\rho-}^i X_i \quad (13)$$

where $\nu_{\rho\pm}^i$ are the stoichiometric coefficients of species X_i in the forward (+) and backward (−) direction of reaction ρ with $k_{\rho\pm}$ the bare forward and backward reaction rates. The dynamical evolution of the concentration c_i of species X_i in a well-mixed (zero-dimensional) system follows

$$\frac{\partial c_i}{\partial t} = \omega_i = \sum_{\rho} \nu_{\rho}^i \left(k_{\rho+} \prod_i c_i^{\nu_{\rho+}^i} - k_{\rho-} \prod_i c_i^{\nu_{\rho-}^i} \right) = \sum_{\rho} \nu_{\rho}^i r_{\rho} \quad (14)$$

with the reaction current r_{ρ} and the total reaction rate of species X_i denoted ω_i . It vanishes if the species are in chemical equilibrium, described by an equilibrium constant K^{ρ}

$$K^{\rho} = \log_{10} \left(\frac{k_{\rho+}}{k_{\rho-}} \right) = \log_{10} \left(\prod_i (c_i^{\text{eq}})^{\nu_{\rho}^i} \right) \quad (15)$$

where c_i^{eq} denote (molar) concentrations in chemical equilibrium. Summarizing the stoichiometric coefficients of all (forward and backward reactions) in a matrix with elements $\nu_{\pm} = \{\nu_{\rho\pm}^i\}$, the stoichiometric matrix reads $\nu := \nu_- - \nu_+$ and has dimensions $\mathcal{N} \times \mathcal{R}$.

Stoichiometric algebra: conservation laws and reaction cycles. Consider an $\mathcal{N} \times \mathcal{R}$ matrix ν . Denote the the dimensions of its right ($\nu \mathbf{v} = 0$) and left ($\mathbf{v}^T \nu = 0$) null-spaces by \mathcal{Z} and \mathcal{C} , respectively. The rank-nullity theorem ensures that

$$\mathcal{Z} = \mathcal{R} - \text{rank}(\nu), \quad \mathcal{C} = \mathcal{N} - \text{rank}(\nu), \quad (16)$$

where $\text{rank}(\nu)$ is the rank of ν . Assume we have a row vector \mathbf{n} of size \mathcal{N} consisting of integer values n_i representing the number of instances of each species X_i present in the (well-mixed) system. In addition, we have a reaction vector \mathbf{q} of length \mathcal{R} with integer entries q_{ρ} . The mapping $\nu \mathbf{q}$ yields the change in the number of species (n_i) after carrying out the ρ -th reaction in forward direction q_{ρ} times. Hence, a non-zero right null-vector of ν can be interpreted as carrying out multiple reactions after which the vector \mathbf{n} remains unchanged. This constitutes a reaction cycle. The number of independent reaction cycles is given by the dimensions of the right nullspace of ν . Each independent reaction cycle $\tilde{\rho}$ however diminishes the number of independent equilibrium constants by linearly relating them, as

$$\sum_{\rho} \tilde{q}_{\rho} K^{\rho} = \sum_{\rho} \tilde{q}_{\rho} \left(\sum_i \nu_{\rho}^i \log_{10} (c_i^{\text{eq}}) \right) = \sum_i \log_{10} (c_i^{\text{eq}}) \sum_{\rho} \nu_{\rho}^i \tilde{q}_{\rho} = 0. \quad (17)$$

Complementary, a mapping $\mathbf{v}^T \mathbf{n}^T$ yields a reaction vector \mathbf{q} which entries q_{ρ} give how often the ρ -th reaction needs to be carried out to change the numbers of species X_i denoted in \mathbf{n} . A non-zero right nullvector $\mathbf{l}^{(j)} = (l_i^{(j)})$ of \mathbf{v}^T hence defines a linear combination of species that remain unchanged by carrying out any reaction. Therefore, these linear combination define conserved quantities, which we

call moeties. The number of these is given by the right nullspace of v^T . Moieity concentrations $m^{(i)} = \sum_i l_i^{(i)} c_i$ are kept constant by the reaction dynamics because $l^{(i)}$ is a right null-eigenvector and hence

$$\frac{\partial}{\partial t} m^{(i)} = \sum_i l_i^{(i)} \frac{\partial c_i}{\partial t} = \sum_i l_i^{(i)} \omega_i = \sum_{i,p} \nu_{i,p}^p l_i^{(i)} r_p = \sum_{i,p} \underbrace{(\nu_{i,p}^p l_i^{(i)})}_{=0} r_p = 0. \quad (18)$$

By the rank-nullity theorem there are always $\mathcal{R} - \mathcal{Z} = \mathcal{N} - \mathcal{C}$ independent equilibrium constants, which give rise to constraints of the form of Eqs. (10). In addition, there are always $\mathcal{C} - 1$ moeties that are linearly independent from the charge moiety and thus give rise to the constraints shown in Eqs. (8). Finally, one can always use differential electroneutrality that gives rise to the single constraint as shown in Eqs. (9). As such, in general we have

$$(\mathcal{N} - \mathcal{C}) + (\mathcal{C} - 1) + 1 = \mathcal{N} \quad (19)$$

independent linear equations for the \mathcal{N} dynamical species. This fact ensures the existence and uniqueness of a solution for general chemical networks.

Chemostated system. In chemostated systems, some of the species are not considered as “dynamic” species but assumed to be set constant. In our system, water is considered a chemostated species and the corresponding equilibrium constant K_w includes this fixed concentration. Algebraically, the above statements still hold. Due to the Rank-Nullity-Theorem each chemostated species either breaks a conservation law or creates a linear dependence between species, reducing the number of independent equilibrium constants by one. Hence, no matter the number of species chemostated, the equations fixing the dynamical species will always match their respective number. For more details see ref. 17.

Thermogravitational accumulation

Thermophoresis. Colloidal particles travel along an external temperature gradient (thermophoresis). In the overdamped limit, this drift is written as

$$J_{DT} = -DS_T c \nabla T, \quad (20)$$

with T a temperature field, D the diffusion coefficient of the colloidal particle and S_T its Soret coefficient, describing the coupling-strength of the particle to the temperature field. The phenomenology of thermophoretic movement can be described thermodynamically³⁶.

2-D description of thermal traps with thermophoresis and convection. A mathematical approach modeling thermal traps was given by Debye^{14,15}. The temporal evolution of the concentration c of a species in a trap is given by the continuity equation

$$\frac{\partial c}{\partial t} = -\nabla J \quad (21)$$

with the flux

$$J = -D \nabla c + (\mathbf{u} - DS_T \nabla T) c. \quad (22)$$

In the bulk of the thermal trap the convective flow field \mathbf{u} mainly consists of its vertical component u_y , with

$$u_y(x) = \frac{\beta g \rho \Delta T}{12 \mu} \left(x - \frac{x^3}{(w/2)^2} \right) \quad (23)$$

with β, ρ, μ the material specific values for the coefficient of thermal expansion, the density and the viscosity, g the gravitational acceleration and ΔT the temperature difference across the short side of the trap with width w . The convective flow is symmetric with respect to the midpoint of the trap. The flow is directed downwards at the cold side and upwards at the hot side of the thermal trap. The average vertical velocity in one half side of the trap can then be calculated to be

$$\bar{u}_y = \frac{1}{b} \int_0^b u_y(x) dx = \frac{\beta g \rho \Delta T}{12 \mu} \int_0^b dx \left(x - \frac{x^3}{b^2} \right) = \frac{\beta g \rho \Delta T}{12 \mu} \frac{1}{16} w^2 \quad (24)$$

The steady state vertical concentration profile is then found to be

$$c(y) = G \exp\left(-\frac{y}{l}\right) \quad \text{with} \quad \frac{1}{l} = \frac{4}{15} \frac{\Delta T S_T \bar{u}_y}{D + \frac{128}{315} \frac{(u_y b)^2}{D}}. \quad (25)$$

A logarithmic concentration difference between bottom and top then writes as

$$\Delta \log_{10}(c) = \frac{h}{l} \log_{10}(e), \quad (26)$$

with h the height of the trap.

Electroneutrality. In order for the bulk electroneutrality assumption to hold to good approximation, electrostatic interactions need to be the strongest forces in the system. In our case they therefore need to be compared with the charge separation processes diffusion and thermophoresis. To estimate the deviation from electroneutrality in one dimension we assume the following situation: a positively charged species is homogeneously distributed around $x = 0$ and a negatively charged

compound is linearly distributed with concentration difference Δc at $x = \pm \Delta x$. The electric field is then—according to Poissons equation—given by

$$-\nabla E = \frac{1}{\epsilon} \rho = \frac{1}{\epsilon} F(c_+ - c_-) \quad (27)$$

Putting in the concentration profile depicted above we can approximate

$$c_- \approx c_+ + \frac{x}{\Delta x} \Delta c \Rightarrow \frac{dE}{dx} \approx \frac{F}{\epsilon} \Delta c \frac{x}{\Delta x} \Rightarrow E \approx \frac{F}{\epsilon} \Delta c \frac{x^2}{\Delta x} \quad (28)$$

the contribution to the fluxes are (disregarding diffusive fluxes)

$$J_{DT,i} = -D_i S_{T,i} \frac{\partial T}{\partial x} c_i \quad \text{and} \quad J_{el,i} = z_i \frac{D_i F}{RT} E c_i \quad (29)$$

with the steady state condition

$$J_{-DT} = J_{-el} \iff D_i S_{T,i} \frac{dT}{dx} c_i = \frac{D_i F}{RT} E c_i \iff S_{T,i} \frac{\partial T}{\partial x} = \frac{F}{RT} \frac{F x^2}{\epsilon \Delta x} \quad (30)$$

regarding the situation at $x = \Delta x$ we get for the deviation from electroneutrality Δc

$$\Delta c = \frac{S_{T,i} RT \epsilon \Delta T}{F^2 \Delta x w} \approx 10^{-10} [\text{mol}/\text{m}^3]. \quad (31)$$

We used $F = 96485.338 \text{ C/mol}$, $R = 8.314 \text{ J}/(\text{mol} \cdot \text{K})$, $\epsilon = \epsilon_r \epsilon_0 = 80 \cdot 8.854 \cdot 10^{-12} \text{ F/m}$ ³⁷. This result shows that by comparing the magnitude of the constants related to the electrostatic flux to those involved in the thermodiffusive flow, the electroneutrality condition is rendered valid. Even though our system is not homogeneous, one can use electroneutrality to calculate the electric field, as explained in more detail by Feldberg²⁰.

Zeroth-order approximation in the 1-D case. Solving the linear Equations (8), (9) and (10) expresses the gradients as nonlinear functions of the concentrations.

$$\frac{\partial c_i}{\partial x} = -S_{T,i}^{\text{eff}}(c_k) \frac{\partial T}{\partial x} c_i. \quad (32)$$

With $S_{T,i}^{\text{eff}}(c_k)$ may in general have some arbitrary non-linear dependence on the steady-state concentrations of (in principle) all species $\{c_i\}$. We define $\gamma = c_i - c_i^0$ the deviation of the steady-state concentration from the initial concentration $c_i^0 = c_i(t = 0)$. The initial concentration is reached, after the constituents have chemically equilibrated. We expand $S_{T,i}^{\text{eff}}(c_i)$ around $\gamma = 0$

$$S_{T,i}^{\text{eff}}(c_i) = S_{T,i}^{\text{eff}}(c_i^0) + \sum_{n=1}^{\infty} \frac{(S_{T,i}^{\text{eff}}(c_i^0))^{(n)}}{n!} \gamma^n = S_{T,i}^{\text{eff},0} + \mathcal{O}[\gamma] \quad (33)$$

For $\gamma \ll 1$ we can make the approximation

$$S_{T,i}^{\text{eff}}(c_i) \approx S_{T,i}^{\text{eff}}(c_i^0) = S_{T,i}^{\text{eff},0}(m_{\text{Na}}, m_{\text{form}}). \quad (34)$$

The steady-state solution is then an exponential (with linear temperature gradient $\frac{\partial T}{\partial x} = \frac{\Delta T}{w}$, w the width, ΔT the temperature difference)

$$c^*(x) = g_0^* \exp\left(-S_{T,i}^{\text{eff},0} \frac{\Delta T}{w} x\right) \quad (35)$$

with the constant g_0^* remains to be fixed (e.g. by global conservation of species). Since we are only interested in pH differences, g_0^* does not play a role for calculating ΔpH .

Homogenous concentrations c_i^0 . The homogenous concentrations to be used to calculate $S_{T,i}^{\text{eff},0}(m_{\text{Na}}, m_{\text{form}})$ are given by the steady state conditions of the well mixed chemical reaction network. Explicitly, the equations for the concentrations are:

$$\begin{aligned} K_a &= \frac{c_{\text{H}^+}^0 c_{\text{COOH}^-}^0}{c_{\text{HCOOH}}^0} \\ K_w &= \frac{c_{\text{H}^+}^0 c_{\text{OH}^-}^0}{c_{\text{H}_2\text{O}}^0} \\ m_{\text{form}} &= c_{\text{HCOOH}}^0 + c_{\text{A}^-}^0 \\ 0 &= c_{\text{H}^+}^0 - c_{\text{COOH}^-}^0 + c_{\text{Na}^+}^0 - c_{\text{OH}^-}^0 \\ K_b &= \frac{c_{\text{Na}^+}^0 c_{\text{OH}^-}^0}{c_{\text{NaOH}}^0} \\ m_{\text{Na}} &= c_{\text{Na}^+}^0 + c_{\text{NaOH}}^0 \end{aligned} \quad (36)$$

with $K_a = 10^{-3.745}$, $K_b = 10^{-0.56}$, $K_w = 10^{-14}$. We can always solve these equations such that we can express the concentrations c_i^0 as functions of the moieties m . Experimentally, the molar moiety concentrations m_{form} (the amount of bare formic acid) and m_{Na} can be controlled.

Separation of timescales. The crucial assumption for the development of a hierarchical theory is the separation of timescales associated to the different dynamic processes. In the following we will give reasoning on why these assumptions are reasonable. For an overview over the timescales involved, see Supplementary Fig. 4.

Reaction timescale τ_{chem} . For the protonation and deprotonation of simple acid-base reactions, it is known that timescales are nearly instantaneous³⁸. As an example, this has been studied for the deprotonation of formic acid where it was confirmed that timescales are $<10^{-3}$ s³⁹.

Thermophoretic timescale τ_{phor} . Since the steady state temperature profile across the thermal trap is linear, the velocity associated with the thermophoretic movement is

$$u_{\text{phor}} = DS_T \frac{\Delta T}{w} \approx 39 \cdot 10^{-6} \text{ m/s}, \quad (37)$$

with $\Delta T = 39$ K and $w = 0.17$ mm. Thus, the time-scale of the phoretic motion in horizontal direction is given as

$$\tau_{\text{phor}} \approx w/u_{\text{phor}}$$

The Soret and diffusion coefficient for the hydrogen ion and the associated phoretic timescales are given in Supplementary Table 1).

Convective timescale τ_{conv} . The average velocity of the convective flow is

$$u_{\text{conv}} = \bar{u}_y = \frac{1}{16} \frac{\beta g \rho \Delta T}{12 \mu} w^2 \approx 5.0 \cdot 10^{-6} \text{ m/s}$$

where we used the values $\beta = 69 \cdot 10^{-6}$ 1/K, $g = 9.81$ m/s², $\rho = 10^3$ kg/m³ and $\mu = 0.80$ mPa · s³⁷. Thus, for a chamber height $h = 50$ mm we obtain a time-scale

$$\tau_{\text{conv}} = h/u_{\text{conv}} \approx 10000 \text{ s}, \quad (38)$$

which is much larger than the typical phoretic time scales.

Measurement (flow) timescale τ_{flow} . The timescale of the measurement process τ_{flow} is defined by the volume of the thermal chamber and the flow rate applied. The volume V is calculated using the height h (50 mm), the width w (170 μ m) and the broadness d (5 mm), giving $V = 42.5$ μ l. Relating to the flow rate used throughout the experiments (3.28 $\frac{\mu\text{l}}{\text{s}}$), we get a timescale

$$\tau_{\text{flow}} = \frac{\text{vol}_{\text{thermalchamber}}}{\text{flow rate}} \approx 13000 \text{ s}, \quad (39)$$

which is somewhat larger than the convective time-scale. Hence, we assume that the concentration dynamics in the convective setting are only slightly disturbed in the experiment.

Interpretation in 2D and 3D

Interpretation in 2D via the Debye factor. While in a purely one-dimensional case, the above expression can directly be used to calculate accumulation profiles of species, in two dimensions, convective transport has to be taken into account. We are able to make qualitative predictions about this situation since the processes happen on different timescales as shown in the section “Separation of timescales”. The initial configuration of the system is homogeneous and chemically equilibrated. As shown above, the two main transport processes, thermophoresis and convection, have similar velocities, however their timescales differ as the length-scale for each process is different by orders of magnitude ($w/h \sim 3.4 \cdot 10^{-4}$). Hence, the concentration field is first equilibrated in the horizontal direction with the exception of the upper and lower end of the trap. Stopping the process and taking a snapshot would reveal a horizontal concentration gradient, while in vertical direction, the species would be largely distributed homogeneously. The convective flow field is zero in the middle of the trap and flows up or downwards on the hot or cold side, respectively. Consequently, species located at either side of the trap are transported upwards or downwards. The horizontal gradient established by thermophoresis within its timescale is an approximation on the magnitude of the expected vertical flow, as it imposes different levels of concentration on either side of the thermal trap. From the slope of the gradient we can therefore estimate what fraction of the total species will be transported either upwards or downwards. This net transport over time accumulates species in the vertical direction until a steady-state of the vertical transport is reached as diffusive and convective currents are balancing each other. Therefore, through the different equilibration timescales of thermophoresis and convection we can make statements about the vertical gradient by looking at the horizontal gradient. The latter in turn is proportional to $S_{\text{H}}^{\text{eff},0} \frac{\Delta T}{w}$. Hence, by calculating $S_{\text{H}}^{\text{eff},0}$ for a given chemical reaction network we are able to make predictions about the outcome of thermal trap experiments in the full dimension.

As introduced in the main text, the effective Soret coefficient relates to the horizontal pH gradient by $(\Delta \text{pH})_{\text{hor}} \approx \frac{\Delta T}{\ln(10)} S_{\text{H}}^{\text{eff},0}$. For the amplitude of the full 2D gradient, we use the scaling $(\Delta \text{pH})_{\text{vert}} \approx \alpha_{\text{th}} \cdot S_{\text{H}}^{\text{eff},0} (m_{\text{Na}}, m_{\text{form}})$. Following the results by Debye¹⁴, we approximate the vertical concentration profile by an exponential function

$$c(y) = G \exp\left(-\frac{y}{l}\right) \quad \text{with} \quad \frac{1}{l} = \frac{4}{15} \frac{\Delta T S \bar{u}_y}{D + \frac{128}{315} \frac{(u_y \bar{u}_y)}{D}}.$$

A logarithmic concentration difference between bottom and top then gives

$\Delta \log_{10}(c) = \frac{h}{l} \log_{10}(e)$ so we can write the factor α_{th}

$$\alpha_{\text{th}} \approx \frac{\Delta T}{\ln(10)} \cdot \frac{4}{15} \frac{h \bar{u}_y}{D + \frac{128}{315} \frac{(u_y \bar{u}_y)}{D}}, \quad \text{where} \quad \bar{u}_y = \frac{\beta g \rho \Delta T}{12 \eta} \frac{1}{16} w^2 \quad (40)$$

With thermal expansion coefficient $\beta = 207 \cdot 10^{-6}$ 1/K, gravitational constant $g = 9.81$ m/s², density of water $\rho = 995.67$ kg/m³ (at 30 °C), viscosity $\eta = 797 \cdot 10^{-6}$ mPa · s, temperature difference $\Delta T = 39$ K, height of trap $h = 0.05$ m, width of trap $w = 170 \cdot 10^{-6}$ m. Influences of formic acid and sodium hydroxide on density and viscosity can be neglected. As diffusion coefficient, we take D_{H} from Supplementary Table 1. We take the values for 30 °C as most material is accumulated towards the colder side. This gives us an α_{th} of 358.5, used in the main text, Fig. 3.

Extrapolation to 3D and flow distortions. In the experimental part however, we measure the pH gradient in a 3D setting, continuously extracting the fluid through 3 outlets which we want to compare to the theoretical predictions. As discussed above, there are multiple uncertainties linked to the experimental setup. In addition and as mentioned before, the measured pH values on the top and bottom outfluxes are averaged over 12 % of the length of the thermal trap. Depending on the specific arrangement of each concentration profile this averaging procedure alters the concentration ratios of each dissolved compound at the outfluxes.

Impact of chemical reactions, electrostatics and temperature dependent reaction rates.

For a better understanding of the individual contributions from chemical reactions, electrostatics and the temperature dependence of reaction rates, we simulated the chemical system in a 1 dimensional finite-element simulation (Comsol 5.4, see Supplementary Data 1). For a detailed study, we had to restrict ourselves to the 1D case since simulations for the 2D or 3D case are not converging when including electrostatic interactions between solutes according to our best knowledge.

For this, we combined the modules “heat transfer in solids”, “general form PDE” and “electrostatics”. In the stationary state and under assumption of insulation, the heat transfer is calculated from

$$\rho \cdot C_p \cdot \mathbf{u} \cdot \nabla T + \nabla \cdot (-\mathbf{k} \cdot \nabla T) = Q \quad (41)$$

with the density ρ , the specific heat capacity C_p , the velocity of the fluid \mathbf{u} , the temperature T , the thermal conductivity $\mathbf{k}_{\text{H}_2\text{O}} = 0.62 \frac{\text{W}}{\text{mK}}$ of water and the heat flux Q . Assuming a very slow fluid velocity, the first term can be neglected.

The result is then coupled to a general PDE and electrostatics. The former gives the following equation:

$$e_a \partial_i^2 c + d_a \partial_i c + \nabla \Gamma = f \quad (42)$$

with the mass coefficients e_a , the concentration vector $\mathbf{c} = [c_{\text{H}}, c_{\text{A}}, \dots]$, the damping coefficients d_a , the flux Γ and the source term f . In our case, the matrix e_a is zero. The flux writes as follows:

$$\Gamma = -D_i \cdot c_i - S_{T,i} \cdot D_i \cdot c_i \cdot \Delta T - \frac{z_i \cdot D_i \cdot F_{\text{const}}}{R_{\text{const}} \cdot T} \cdot \Delta V \cdot c_i \quad (43)$$

with temperature T , electric potential V , gas constant R_{const} and Faraday constant F_{const} . For each species i , the concentration (c_i), the charge (z_i) the diffusion coefficient (D_i), and the Soret coefficient ($S_{T,i}$) are used, for the latter see Supplementary Table 1. The source term includes all reactions, for example in the case of formic acid (AH):

$$f_{\text{AH}} = -k_{\text{on}} \cdot c_{\text{AH}} + k_{\text{off}} \cdot c_{\text{H}} \cdot c_{\text{A}} \quad (44)$$

For the temperature dependence of the dissociation constants of water (K_{W}) and formic acid (K_{A}), we used the values from⁴⁰ for formic acid and⁴¹ for water.

For the electrostatics, the equations write as follows:

$$\nabla \cdot \mathbf{D} = \rho_V \quad \text{and} \quad \mathbf{E} = -\nabla V \quad (45)$$

with the electric displacement $\mathbf{D} = \epsilon_0 \epsilon_r \mathbf{E}$ (including the permittivities), the electric field \mathbf{E} , the relative permittivity ϵ_r , the vacuum permittivity ϵ_0 , the electric potential V and the space charge density ρ_V which writes as follows:

$$\rho_V = F_{\text{const}} \cdot \sum_i c_i \cdot z_i \quad (46)$$

To better understand the individual contributions of electrostatics, temperature dependence etc, we turned them off one by one (Supplementary Fig. 5).

Data availability

All the data that support the plots and the other findings of this study are available from the corresponding authors upon reasonable request. The full details of the finite-element simulation from the Methods section and Supplementary Fig. 5 are documented in the mph file in Supplementary Data 1.

Received: 28 July 2022; Accepted: 30 December 2022;
Published online: 19 January 2023

References

- Lodish, H. F. et al. *Molecular Cell Biology* Vol 4 (WH Freeman, 2006).
- Weiss, M. C. et al. The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1**, 1–8 (2016).
- Martin, W., Baross, J., Kelley, D. & Russell, M. J. Hydrothermal vents and the origin of life. *Nat. Rev. Microbiol.* **6**, 805–814 (2008).
- Lane, N. & Martin, W. F. The origin of membrane bioenergetics. *Cell* **151**, 1406–1416 (2012).
- Lane, N. Proton gradients at the origin of life. *BioEssays* **39**, 1600217 (2017).
- Lawless, J. et al. pH profile of the adsorption of nucleotides onto montmorillonite. *Orig. Life Evol. Biosph.* **15**, 77–88 (1984).
- Le Vay, K., Salibi, E., Song, E. Y. & Mutschler, H. Nucleic acid catalysis under potential prebiotic conditions. *Chem. Asian J.* **15**, 214–230 (2020).
- Budin, I., Bruckner, R. J. & Szostak, J. W. Formation of protocell-like vesicles in a thermal diffusion column. *J. Am. Chem. Soc.* **131**, 9628–9629 (2009).
- Kreysing, M., Keil, L., Lanzmich, S. & Braun, D. Heat flux across an open pore enables the continuous replication and selection of oligonucleotides towards increasing length. *Nat. Chem.* **7**, 203–208 (2015).
- Kelley, D. S. et al. An off-axis hydrothermal vent field near the mid-atlantic ridge at 30° n. *Nature* **412**, 145–149 (2001).
- Keil, L. M., Möller, F. M., Kieß, M., Kudella, P. W. & Mast, C. B. Proton gradients and pH oscillations emerge from heat flow at the microscale. *Nat. Commun.* **8**, 1–9 (2017).
- Becker, S. et al. Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* **366**, 76–82 (2019).
- Kruse, F. M., Teichert, J. S. & Trapp, O. Prebiotic nucleoside synthesis: the selectivity of simplicity. *Chem. Eur. J.* **26**, 14776–14790 (2020).
- Debye, P. Zur theorie des clusiuschen trennungsverfahrens. *Ann. Phys.* **428**, 284–294 (1939).
- Guthrie, G., Wilson, J. N. & Schomaker, V. Theory of the thermal diffusion of electrolytes in a clusius column. *J. Chem. Phys.* **17**, 310–313 (1949).
- Mast, C. B., Schink, S., Gerland, U. & Braun, D. Escalation of polymerization in a thermal gradient. *Proc. Natl Acad. Sci. USA* **110**, 8030–8035 (2013).
- Rao, R. & Esposito, M. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Phys. Rev. X* **6**, 041064 (2016).
- Reichl, M., Herzog, M., Götz, A. & Braun, D. Why charged molecules move across a temperature gradient: The role of electric fields. *Phys. Rev. Lett.* **112**, 198101 (2014).
- MacGillivray, A. D. Nernst-planck equations and the electroneutrality and donnan equilibrium assumptions. *J. Chem. Phys.* **48**, 2903–2907 (1968).
- Feldberg, S. W. On the dilemma of the use of the electroneutrality constraint in electrochemical calculations. *Electrochem. Commun.* **2**, 453–456 (2000).
- Keil, L., Hartmann, M., Lanzmich, S. & Braun, D. Probing of molecular replication and accumulation in shallow heat gradients through numerical simulations. *Phys. Chem. Chem. Phys.* **18**, 20153–20159 (2016).
- Gore, V. Formation of sugars in mixtures of tartaric acid and aldehydes in tropical sunlight. *J. Phys. Chem.* **37**, 745–749 (1933).
- Orgel, L. E. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* **39**, 99–123 (2004).
- Miller, S. L. Production of some organic compounds under possible primitive earth conditions. *J. Am. Chem. Soc.* **77**, 2351–2361 (1955).
- Choughuley, A., Subbaraman, A., Kazi, Z. & Chada, M. A possible prebiotic synthesis of thymine: uracil-formaldehyde-formic acid reaction. *BioSystems* **9**, 73–80 (1977).
- Pine, S. H. & Sanchez, B. L. The formic acid-formaldehyde methylation of amines. *J. Org. Chem.* **36**, 829–832 (1971).
- Cody, G. D. et al. Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate. *Science* **289**, 1337–1339 (2000).
- Becker, S. et al. Origin of life: a high-yielding, strictly regioselective prebiotic purine nucleoside formation pathway. *Science* **352**, 833–836 (2016).
- Saladino, R., Crestini, C., Costanzo, G., Negri, R. & Di Mauro, E. A possible prebiotic synthesis of purine, adenine, cytosine, and 4(3h)-pyrimidinone from formamide: Implications for the origin of life. *Bioorg. Med. Chem.* **9**, 1249–1253 (2001).
- Saladino, R., Crestini, C., Pino, S., Costanzo, G. & Di Mauro, E. Formamide and the origin of life. *Phys. Life Rev.* **9**, 84–104 (2012).
- Göppel, T., Palyulin, V. V. & Gerland, U. The efficiency of driving chemical reactions by a physical non-equilibrium is kinetically controlled. *Phys. Chem. Chem. Phys.* **18**, 20135–20143 (2016).
- Möller, F. M., Krieger, F., Kieß, M., Sojo, V. & Braun, D. Steep pH gradients and directed colloid transport in a microfluidic alkaline hydrothermal pore. *Angew. Chem. Int. Ed.* **56**, 2340–2344 (2017).
- Matreux, T. et al. Heat flows in rock cracks naturally optimize salt compositions for microzymes. *Nat. Chem.* **13**, 1038–1045 (2021).
- Mariani, A., Bonfio, C., Johnson, C. M. & Sutherland, J. D. pH-driven RNA strand separation under prebiotically plausible conditions. *Biochemistry* **57**, 6382–6386 (2018).
- Ianeselli, A. et al. Water cycles in a hadean CO₂ atmosphere drive the evolution of long DNA. *Nat. Phys.* **18**, 1–7 (2022).
- Burelbach, J., Frenkel, D., Pagonabarraga, I. & Eiser, E. A unified description of colloidal thermophoresis. *Eur. Phys. J. E* **41**, 1–12 (2018).
- Lide, D. R. & Baysinger, G. CRC handbook of chemistry and physics: a ready-reference book of chemical and physical data. *Choice Rev. Online* **41**, 41–4368–41–4368 (2004).
- Donten, M. L., VandeVondele, J. & Hamm, P. Speed limits for acid–base chemistry in aqueous solutions. *Chimia* **66**, 182–186 (2012).
- Jolly, G. S., McKenney, D. J., Singleton, D. L., Paraskevopoulos, G. & Bossard, A. R. Rates of hydroxyl radical reactions. part 14. rate constant and mechanism for the reaction of hydroxyl radical with formic acid. *J. Phys. Chem.* **90**, 6557–6562 (1986).
- HwaáKim, M., ShikáKim, C. & WoodáLee, H. et al. Temperature dependence of dissociation constants for formic acid and 2, 6-dinitrophenol in aqueous solutions up to 175° c. *J. Chem. Soc., Faraday Trans.* **92**, 4951–4956 (1996).
- Bandura, A. V. & Lvov, S. N. The ionization constant of water over wide ranges of temperature and density. *J. Phys. Chem. Ref. Data* **35**, 15–30 (2006).

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 364653263—TRR 235 (T.M., J.R., D.B., C.B.M., U.G.) and under Germany's Excellence Strategy—EXC-2094—390783311 (D.B., U.G.). Funding by the Volkswagen Initiative 'Life?—A Fresh Scientific Approach to the Basic Principles of Life' (T.M., C.B.M., D.B.) is gratefully acknowledged. This work was further supported by the Center for Nanoscience Munich (CeNS). We thank A. Schmid and P. Aikkila for experimental support.

Author contributions

T.M., B.A., J.R. and C.B.M. performed the research. T.M., B.A., J.R., D.B., C.B.M. and U.G. designed the research, T.M., B.A., J.R., D.B., C.B.M. and U.G. analyzed the data and wrote the Letter.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42005-023-01126-y>.

Correspondence and requests for materials should be addressed to Christof B. Mast or Ulrich Gerland.

Peer review information *Communications Physics* thanks Ludovic Jullien and the other, anonymous, reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.