PREBIOTIC CHEMISTRY

Replicating towards complexity

Replication of long nucleic acid sequences was required for the evolution of biological complexity during the origin of life; however, short sequences are normally better replicators than long ones. A common physical environment now provides a simple mechanism to reverse this trend and enables long sequences to flourish.

Irene A. Chen

Chemists are notoriously tolerant of long sequences of letters; however, long sequences of nucleotides had a harder time during the origin of life. In addition to being produced in low yields by prebiotic synthesis, long sequences also suffer from an important evolutionary disadvantage: it takes longer to replicate a long sequence than a short one. This intrinsic handicap in Darwinian fitness compounds over time, so, all else being equal, if a pool of nucleic acid sequences containing a range of lengths is left to replicate, short sequences will dominate and long ones will become extinct. The problem would have been particularly acute during the origin of life, when survival of the shortest would prevent the evolution of longer functional RNA sequences. Now, writing in this issue of *Nature Chemistry*, Braun and colleagues demonstrate[1] that a simple physical environment, resembling an asymmetrically heated rock pore, neatly turns the tables, resulting in sharp selection of longer sequences.

The problem of survival of the shortest has been known since 1967. In some of the earliest experiments in molecular evolution, Sol Spiegelman and colleagues used the newly discovered Qβ replicase (an enzyme that copies RNA, making a new strand) to reproduce and evolve a viral RNA genome[2]. Over time, the genome (Spiegelman's 'little monster') ramped up its own replication rate, shrinking almost ten-fold in length and eliminating the ability to encode viruses. Indeed, many known viruses work hard to minimize genome size, such as by overlapping genes in multiple reading frames or translating both strands of the genome into different proteins.

Selection for the shortest replicator would also have important significance during the origin of life. At some point, a primordial chemical system must acquire the ability to copy information, either by purely chemical processes or by spontaneous emergence of ligase or polymerase enzymes. Regardless of the details, any replication process would exhibit a preference for the shortest possible replicators. This would not necessarily prevent the evolution of long sequences, as a strong selection pressure for a function encoded only on long sequences could still favour the replication of these longer sequences. However, in this scenario functional selection must compete with the selection arising from the raw replication rate. Selection for function needs a mechanism, such as compartmentalization, that links the function (for example, enzymatic activity) to the fitness of the template that encodes it. Indeed, functional selection in compartments has been applied to the Qβ replicase, preventing takeover by the shortest replicators[3]. Even so, selection for the shortest still minimizes the genome size given the other selection pressures. Yet expansion in genome length is essential for adding new functions that could increase biological complexity. Survival of the shortest means that genomes will miss opportunities to explore sequence space, putting the brakes on evolutionary innovation.

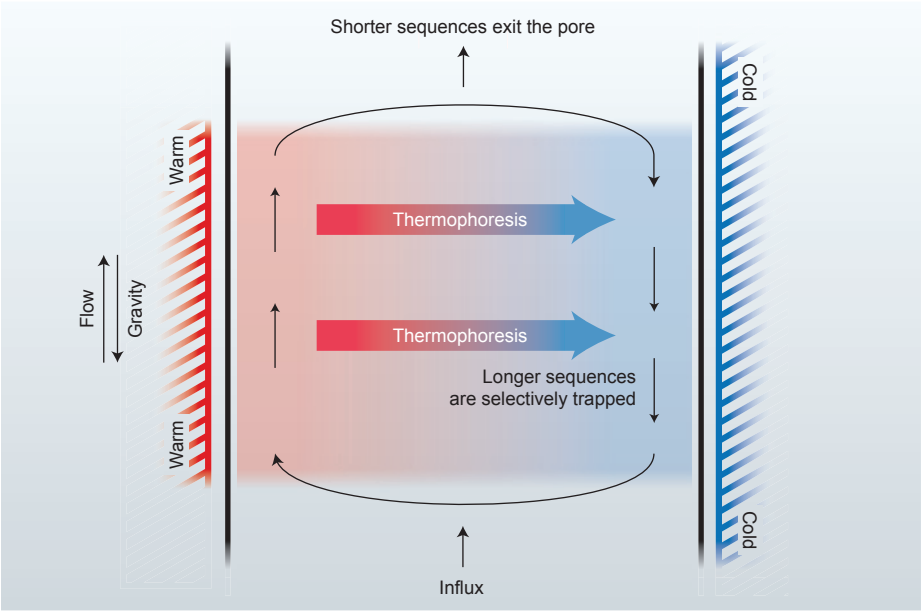

Figure 1 | Survival of the longest. The thermal gradient inside the pore induces the accumulation of longer nucleic acid sequences via a combination of thermophoresis and convection. In contrast, shorter sequences are flushed out of the pore. This combination of thermophoresis and convection also induces thermal cycling, which facilitates replication of the trapped longer sequences. The population of shorter sequences thus 'dies out' while the longer ones continue to accumulate and replicate.

Braun and co-workers show that physical mechanisms can upset this paradigm, leading instead to survival of the longest. Inspired by the impressive thermal and chemical gradients present in hydrothermal systems — which have previously been advocated as a possible location for the origin of life — they applied a horizontal thermal gradient across a vertically positioned glass capillary to simulate a rocky pore. Thermal gradients drive a transport phenomenon known as thermophoresis, in which the solvation energy of a molecule depends on its position in the gradient. The mechanism behind thermophoresis is a subject of active study, but appears to be dominated by the temperature dependence of the energy of the electrical double layer. In addition, in salt solution, thermophoresis of the small ions

themselves establishes an electrical field that causes electrophoretic movement of other charged species, such as DNA[4].

Under the conditions used by Braun and co-workers, DNA (or RNA) drifts towards the cold side of the capillary (Fig. 1). Importantly, the resulting concentration profile depends strongly on charge, such that long DNA sequences exhibit a much steeper profile over the thermal gradient compared with short sequences. To 'feed' the system, the team passed a continuous flow of DNA solution upwards into the capillary, similar to the upward flow through a hydrothermal system. This inflow, combined with convective flow up along the warm side, swept most of the short sequences out of the capillary. In stark contrast, convective flow downwards overpowers the inflow on the cold side, so the long sequences, having steeply accumulated along the cold side, remain trapped and pile up in the bottom, cold corner. This ingenious set-up gave very sharp selection for long sequences, and the threshold length could be tuned by the inflow rate.

To test selection of long sequences with replicators, the authors pitted a long DNA template (75 bp) against a short one (36 bp) in the polymerase chain reaction (PCR). Convection drove thermal cycling, and the inflow brought enzyme and substrates but no additional template. Ordinarily, the short sequence, being a more efficient template, would easily win this competition (as many molecular biologists have discovered during PCR, to their chagrin). However, the simulated hydrothermal pore enables a remarkable reversal of fortunes, with the longer sequence surviving while the shorter sequence dies out. Braun *et al.* thus stage a coup for the traditional underdog: a methodical combination of thermophoresis, convection, flow through an open system, and replication gives strong selection for longer length, providing the raw material needed to evolve new functions on a primitive genome.

Although this microenvironment selects for longer length, there is no guarantee that the added length will be useful. In other words, various long sequences would accumulate alongside each other. Natural selection for function is still needed for a particular nucleotide sequence to become dominant over other similar length sequences. There is also no guarantee that the added length can be faithfully copied. The error rate during copying strictly limits the amount of information that can be maintained on the genome[5]. With recent advances, non-enzymatic processes and ribozymes have been shown to replicate nucleic acids accurately enough to maintain sequences of the length of one or two ribozymes[6–8], but longer genomes are still out of reach for now. The open system reported by Braun *et al.* also requires a constant flow of substrates, so prebiotic nucleotide synthesis must be able to keep up. Such caveats are inevitable when working with a multifaceted puzzle such as the origin of life.

In practical terms, this work also raises the intriguing possibility that thermophoresis with convection could improve PCR results. Indeed, Braun and colleagues previously established microscale thermophoresis as a practical technique for affinity measurements. But the true focus remains experimental models of the origin of life. This elegant work exemplifies the burgeoning set of physical or chemical effects that could favour biological complexity during the origin of life, setting the stage for evolution by natural selection. ❐

Irene A. Chen is in the Department of Chemistry & Biochemistry, University of California, Santa Barbara, California 93106-9510, USA.
e-mail: chen@chem.ucsb.edu

References
1. Kreysing, M., Keil, L., Lanzmich, S. & Braun, D. *Nature Chem.* **7,** 203–208 (2015).
2. Mills, D. R., Peterson, R. L. & Spiegelman, S. *Proc. Natl Acad. Sci. USA* **58,** 217–224 (1967).
3. Ichihashi, N., Usui, K., Kazuta, Y., Sunami, T., Matsuura, T. & Yomo, T. *Nature Commun.* **4,** 2494 (2013).
4. Reichl, M., Herzog, M., Gotz, A. & Braun, D. *Phys. Rev. Lett.* **112,** 198101 (2014).
5. Eigen, M. *Naturwissenschaften* **58,** 465–523 (1971).
6. Rajamani, S. *et al. J. Am. Chem. Soc.* **132,** 5880–5885 (2010).
7. Zhang, S., Blain, J. C., Zielinska, D., Gryaznov, S. M. & Szostak, J. W. *Proc. Natl Acad. Sci. USA* **110,** 17732–17737 (2013).
8. Attwater, J., Wochner, A. & Holliger, P. *Nature Chem.* **5,** 1011–1018 (2013).

INTERFACES

Scientists strike wet gold

The structure of liquid water is intensely studied, but it is not clear what happens to it when a surface is introduced. Now with the aid of X-ray spectroscopy it has been found that water molecules at the interface with a gold electrode have a different structure than in the bulk.

Bernd Winter

The structure of aqueous liquids in the bulk is very much dictated by the hydrogen-bonding network of water[1], however, questions remain as to how this changes when a surface is introduced. So far it has been experimentally difficult to access spectroscopic signals that originate exclusively from the interface without being swamped by signals from the bulk. But understanding such interfaces is very important because many electrochemical processes are governed by both the electronic and geometric structure of the solvent molecules at this interface — in what is known as the electric double layer, or Helmholtz layer (Fig. 1).

The structure of water in the bulk can be characterized by its many hydrogen bonds, with the majority of water molecules fully saturated by four neighbouring water molecules, while a small fraction have unsaturated hydrogen-bond donors — known as dangling hydrogen bonds[2]. Now, using X-ray spectroscopy and first principles calculations, Salmeron and colleagues[3] have found that the population of water molecules with broken hydrogen bonds is substantially larger at the interface with a gold electrode than in bulk liquid water, and in addition the orientation of interfacial water is different from the respective high-vacuum water mono- and multilayer systems. As well as providing insight into the structure of water at an interface, this work poses new and exciting questions to the field of electrochemistry and, on a more technical level, on the coupling mechanism of X-ray excited states to surface electronic structure.

Electronic-structure information from any molecular system is most completely probed by UV/X-ray spectroscopy. In this suite of techniques the absorption of UV/X-ray light leads to the emission of electrons and photons, which ideally