nature chemistry

ARTICLES

PUBLISHED ONLINE: 6 MARCH 2017 | DOI: 10.1038/NCHEM.2739

Simple peptides derived from the ribosomal core potentiate RNA polymerase ribozyme function

Shunsuke Tagami†, James Attwater and Philipp Holliger*

The emergence of functional interactions between nucleic acids and polypeptides was a key transition in the origin of life and remains at the heart of all biology. However, how and why simple non-coded peptides could have become critical for RNA function is unclear. Here, we show that putative ancient peptide segments from the cores of both ribosomal subunits enhance RNA polymerase ribozyme (RPR) function, as do derived homopolymeric peptides comprising lysine or the non-proteinogenic lysine analogues ornithine or, to a lesser extent, diaminobutyric acid, irrespective of chirality or chiral purity. Lysine decapeptides enhance RPR function by promoting holoenzyme assembly through primer–template docking, accelerate RPR evolution, and allow RPR-catalysed RNA synthesis at near physiological (≥1 mM) Mg^{2+} concentrations, enabling templated RNA synthesis within membranous protocells. Our results outline how compositionally simple, mixed-chirality peptides may have augmented the functional potential of early RNAs and promoted the emergence of the first protocells.

Life is widely believed to have descended from a simpler, primordial biology that lacked DNA and proteins, but in which RNA played a central role[1]. A core component of this 'RNA world' would have been an RNA polymerase ribozyme (RPR) to replicate primordial RNA genomes and become encapsulated within membranous compartments to form the first protocells[2]. Modern-day RPR analogues have been developed[3,4], some of which can synthesize ribozymes[4], aptamers[5] or ~200 nucleotide (nt) sequences using favourable RNA templates[6], and even amplify short RNA sequences[5], but their activity is strictly dependent on very high concentrations of available magnesium ions (optimal $[Mg^{2+}] \approx 200$ mM)[7,8]. Such a high $[Mg^{2+}]$ not only greatly accelerates RNA degradation[9]; it also causes amphiphile aggregation, precipitation and membrane destabilization[10,11], suggesting a fundamental incompatibility of RPR activity with the formation of membranous protocells. However, ribozyme function within the RNA world probably emerged not in isolation but in the context of a complex chemical environment comprising a range of other molecular entities potentially including peptides and phospholipids[12]. Functional cooperation between RNA and peptides, in particular, appears to be ancient, and probably reflects deep evolutionary history. For example, in modern biology the complex and manifold interactions between RNAs and (poly)peptides are central to the roles of RNA in gene expression. This includes ribozymes such as the spliceosome[13], ribosome[14] and RNase P (ref. 15), all of which are functionally dependent on association with cognate polypeptides, despite all-RNA catalytic sites. Indeed, previous work has shown that some RNA-binding proteins and derived peptides can promote folding and function, even in simple ribozymes[16,17], and in the case of the RNase P ribozyme core critically reduce the requirement for bivalent magnesium ions[15]. We therefore wondered if RPR function might be similarly enhanced by interaction with peptides and, in particular, if they might reduce the strict Mg^{2+} dependency of RPR function.

Ideally, a critical test of this hypothesis would involve early peptides from the RNA world. As a potential source of ancient peptides we turned to the structure of the ribosome, the most prominent, putative molecular relic from the RNA world. According to accretion models of ribosome evolution[18], the structure of the ribosome comprises a quasi-historical record of different stages of ribosomal evolution. Mg^{2+} ions are enriched as counterions in the innermost core of the large ribosomal subunit reflecting the most ancient stages[19,20], but are progressively replaced during ribosome expansion by the acquisition of unstructured polypeptide 'fingers' with a strikingly biased amino-acid content[21]. It is tempting to speculate that such functional cooperation between RNA and peptides conserved within the ribosomal core structure reflects the deep evolutionary history of the RNA world and its transition to an RNA–peptide (RNP) world[19].

Here we have investigated whether the compositional features of these ribosomal core peptide segments, widely considered to be among the most ancient protein sequences on Earth[22], could confer functions transferable to critical primordial RNA contexts. We identify multiple peptides from both ribosomal subunits that strongly enhance RPR activity. From these peptide sequences we derive a simplified homopolymeric lysine decapeptide (K_{10}), which we find to enhance RPR function at low $[Mg^{2+}]$ regardless of chirality, accelerate RPR evolution, and enable RNA-catalysed templated RNA synthesis within a membranous model protocell.

Results

RPR activation by short ribosomal peptides. To examine whether the unusual compositional and structural features of ribosomal core peptides could confer functional benefits transferable to other RNA contexts, we systematically examined peptide segments from ribosomal proteins of the *Thermus thermophilus* ribosome for their ability to enhance the RNA polymerase activity of the *in vitro*-evolved Z polymerase ribozyme[4]. We focused on extended peptide segments of ribosomal proteins that reach deep into the cores of both subunits, as judged by the structure of the *T. thermophilus* 70S ribosome (PDB: 2XQD, 2XQE)[23], which exhibit a highly biased amino-acid composition (~20% R, K; ~12% G) (Fig. 1a,b). Although many peptides showed no (or very weak) effects, we discovered multiple ribosomal peptides that displayed a

MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge CB2 0QH, UK. †Present address: RIKEN Center for Life Science Technologies, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. *e-mail: ph1@mrc-lmb.cam.ac.uk

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

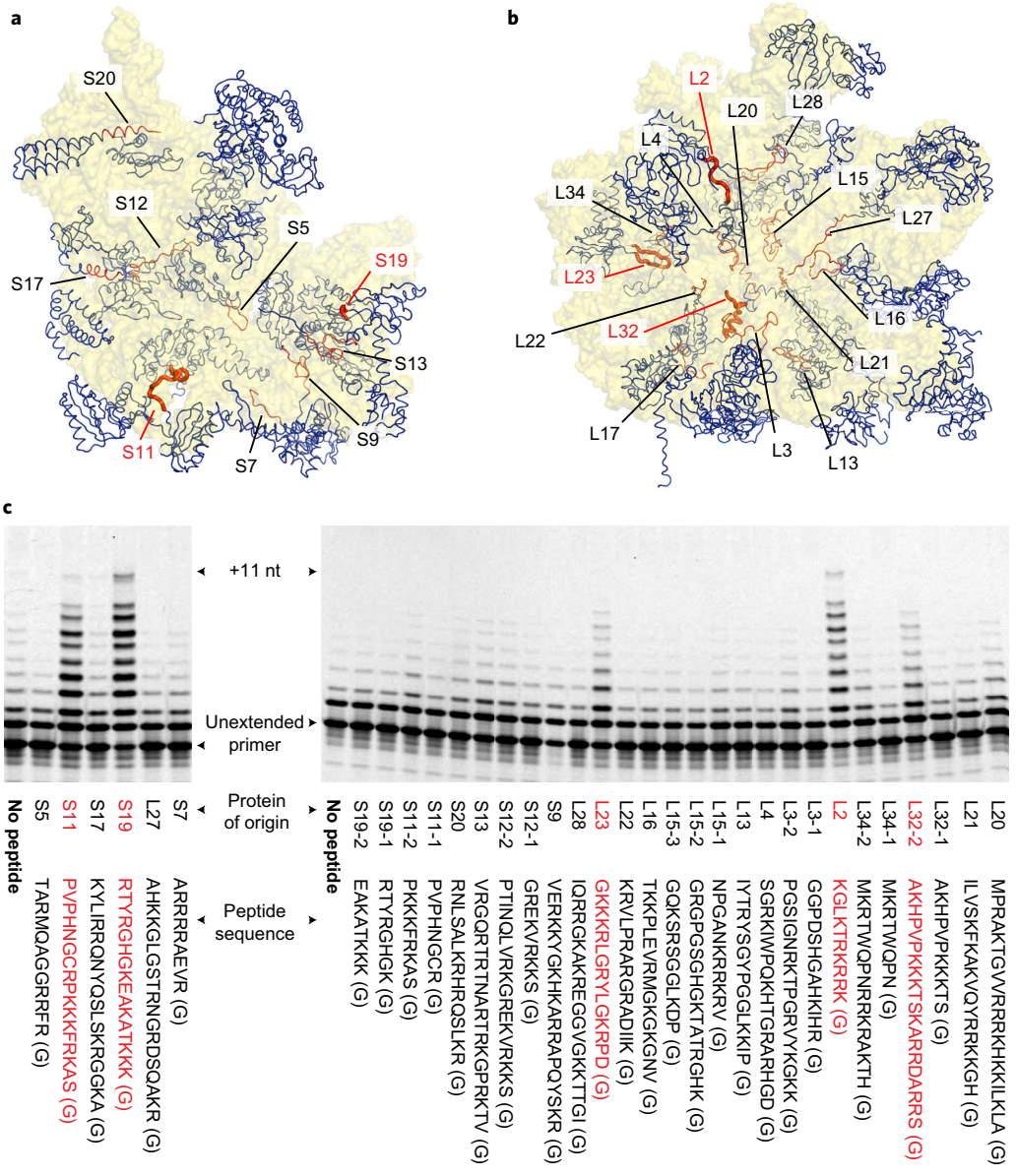


Figure 1 | Ribosomal peptides enhance the activity of an artificial RPR. a,b, Structures of the *T. thermophilus* 30S (**a**) and 50S (**b**) ribosomal subunits, showing proteins (blue) and their peptide extensions (orange) projecting into the centre of each complex. **c**, PAGE analysis of primer extensions on template TI by the RPR Z at a suboptimal [Mg^{2+}] of 22 mM (left, 4 °C for 22 days) or 46 mM (right, 4 °C for 7 days), alone or in the presence of 400 μM of peptides derived from ribosomal proteins (terminating in an extra glycine residue (G) from synthesis). Ribosomal proteins comprising very long unstructured peptide segments give rise to two peptides, for example, L32-1 and L32-2. Peptides enhancing RPR function (red) are highlighted in bold in the ribosomal structures in **a** and **b**.

striking enhancement of RPR activity at suboptimal Mg^{2+} concentrations ([Mg^{2+}] ≪ 200 mM), particularly peptides from ribosomal proteins S11 (PVPHNGCRPKKKFRKAS), S19 (RTYR GHGKEAKATKKK), L2 (KGLKTRKRRK), L23 (GKKKRLGR YLGKRPD) and L32 (AKHPVPKKKTSKARRDARRS) (Fig. 1c and Supplementary Fig. 1).

Next, we dissected the amino-acid contributions of one of the peptides—the S11 peptide segment (S11p: PVPHNG CRPKKKFRKAS(G))—to better understand the critical peptide features responsible for RPR activation. Both N- and C-terminal truncations of S11p retained the ability to enhance RPR activity, provided they contained a central 11-amino-acid segment (NG CRPKKKFRK) (Supplementary Fig. 2) enriched in positively charged amino acids, comprising two arginine (R) and four lysine (K) residues. Overall, the activating peptide segments (Fig. 1c) displayed a significantly higher K content (22 K of 78 amino acids (28%) versus 81 K of 434 amino acids (19%) among all peptides tested, K ≥ 22 $P = 0.027$), but no enrichment of R content (14 R of 78 amino acids (18%) versus 80 R of 434 amino acids (18%)), implicating K as a potential key functional component.

Having identified a range of peptides and established the functional importance of K therein, we sought to establish if the exact ribosomal peptide sequence was of critical functional importance, or if simplified K-rich peptides could recapitulate the activation effect. Although a K_5 pentapeptide provided no functional benefits, both K_{20} and $(KKG)_5$ peptides showed strong RPR activation (Fig. 2a), as did a shorter K_{10} decapeptide (Fig. 2b). Importantly, although able to substitute in part for Mg^{2+} as the counterion, unlike Mg^{2+} the K_{10} peptide appears not to accelerate RNA hydrolysis (Supplementary Fig. 3).

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

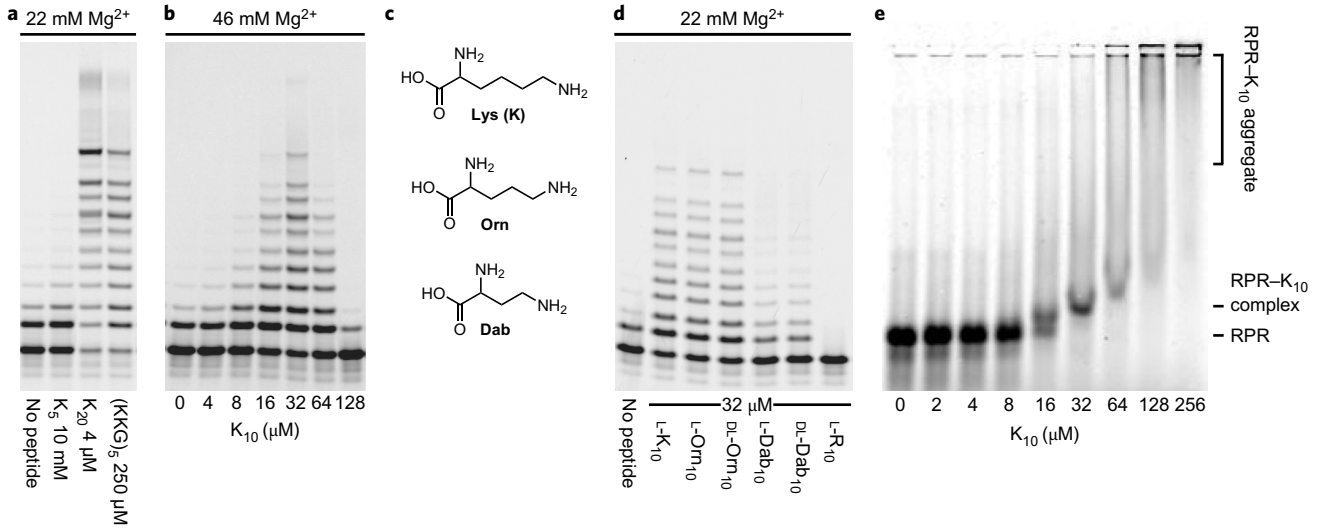


Figure 2 | RPR activation by homopolymeric peptides. PAGE of primer extensions on template TI by the Z RPR with peptides over 3 days at 17 °C at the indicated [Mg^{2+}]. **a**, Varying peptide length strongly influences activation. **b**, Enhancement by K_{10} shows a bell-shaped concentration response with inhibition at higher concentrations. **c**, Structure of the alkyl-amine side-chain amino acids lysine (Lys/K), ornithine (Orn) and diaminobutyric acid (Dab). **d**, Effect on extension enhancement from varying peptide alkyl-amine side-chain length, chirality and chiral purity. **e**, EMSA showing K_{10}–ribozyme interaction using a fluorescently labelled Z ribozyme (Z-Fluor 647) incubated with K_{10} peptide at the indicated concentrations with 22 mM [Mg^{2+}] and resolved on a non-denaturing agarose gel.

We sought to dissect the chemical basis of this RPR activation by varying the peptide side chains and stereochemistry. To our surprise, peptide chirality had little influence upon enhancement, with poly-L-lysine (~K_{700}–K_{1500}) and diastereomeric poly-D-lysine both boosting Z activity to a comparable extent (Supplementary Fig. 4).

The activating effect was also maintained in both chirally pure (L) or racemic (DL) peptides comprising the non-proteinogenic amino acid ornithine (Orn) and to a lesser extent diaminobutyric acid (Dab) (Fig. 2c,d), in which the side chain is shortened and the primary amino group is attached to the δ (Orn) or γ (Dab) carbon, respectively. In contrast, an arginine R_{10} homopeptide only inhibited activity, suggesting that RPR activation is not based on a mere counterion effect mediated by positively charged side chains. Primary amino groups, as present in K, Orn and Dab peptides, appear to be critical, as well as a molecular scaffold of requisite reach (but, surprisingly, not precise stereochemistry). Indeed, an activating effect can be observed even in the absence of a peptide backbone as in polyethyleneimine, but not with simple polyamines such as spermidine and spermine (Supplementary Fig. 4).

Mechanism of RPR activation. We next sought to understand the mechanistic basis for RPR activation by the lysine decapeptide L-K_{10}. K_{10} activation follows a bell-shaped concentration curve with optimal activation between 10 and 50 μM and higher concentrations inhibitory to RPR activity (Fig. 2b). Similar bell-shaped activation curves were observed for other synthetic and ribosomal peptides (Supplementary Fig. 4) with relative activating and inhibitory concentrations dependent on [Mg^{2+}] (Supplementary Fig. 5), suggesting competitive interactions of K_{10} and Mg^{2+} ions with the RPR.

We sought to capture the interaction between the K_{10} peptide and the RPR using native electrophoretic mobility shift assays (EMSA), which indeed show that a discrete upshifted RPR–K_{10} complex of unknown stoichiometry is formed (Fig. 2e) at approximately optimally activating K_{10} concentrations (K_{10}:RPR ratio of ~20:1, although this may not represent the complex composition). This complex is dissociated if Mg^{2+} is present in the gel during electrophoresis, while higher (inhibitory) concentrations of K_{10} (and R_{10}) merely yielded non-specific aggregates of undefined molecular weights (Supplementary Fig. 6). Although this propensity for aggregate formation and inhibition appears to correspond to poly-arginine's greater ability to condense DNA in biological contexts[24], the activity boost by K_{10} and related peptides observed here probably has a different molecular basis.

The EMSA data suggested that RPR activation by K_{10} involved the formation of a peptide–ribozyme complex. Although its role and persistence are unclear, we hypothesized that complex formation might enhance RPR activity by stabilizing some key intermediate in the RPR catalytic cycle. We have previously shown that docking of the primer/template (P/T) duplex with the apo-ribozyme (to form the RPR holoenzyme) is a limiting step in primer extension[7]: RPR constructs loosely tethered or (like Z) untethered to the P/T require cooperative binding of Mg^{2+} ions to low-affinity binding sites (that is, high concentrations of Mg^{2+} ions) for activity, a requirement that can be partially overcome by tight *in cis* tethering of the P/T, which, however, restricts extension length. Examining the relative degree of activation by the K_{10} peptide in such RPR constructs, we found that although K_{10} strongly enhanced extension in untethered and to a lesser degree loosely tethered constructs, it provided only negligible activation when the P/T was tightly tethered (Supplementary Fig. 7). Taken together with the complex formation detected by EMSA (Fig. 2e) and the fact that K_{10} cannot substitute for Mg^{2+} in RPR catalysis (Supplementary Fig. 7), this supports a mechanistic model where K_{10} exerts its activating effect by effecting P/T docking and holoenzyme formation, presumably through bridging between the P/T duplex and the RPR. Examination of the structural context of stimulatory peptide segments, such as L2 and L32, within the ribosome illustrates how they connect distinct RNA segments through interactions involving distal K side chains (Supplementary Fig. 8) providing a potential structural rationalization for holoenzyme formation and the activation effect. Indeed, bridging of RNA segments may be a generic transferable function of such peptides, as K_{10} could also enhance tRNA maturation by the ribozyme component of RNase P, where R_{10} peptides again provided no benefit (Supplementary Fig. 9).

RPR evolution with peptide. To better understand the impact of K_{10} on RPR functional potential, we also investigated *in vitro* RPR

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

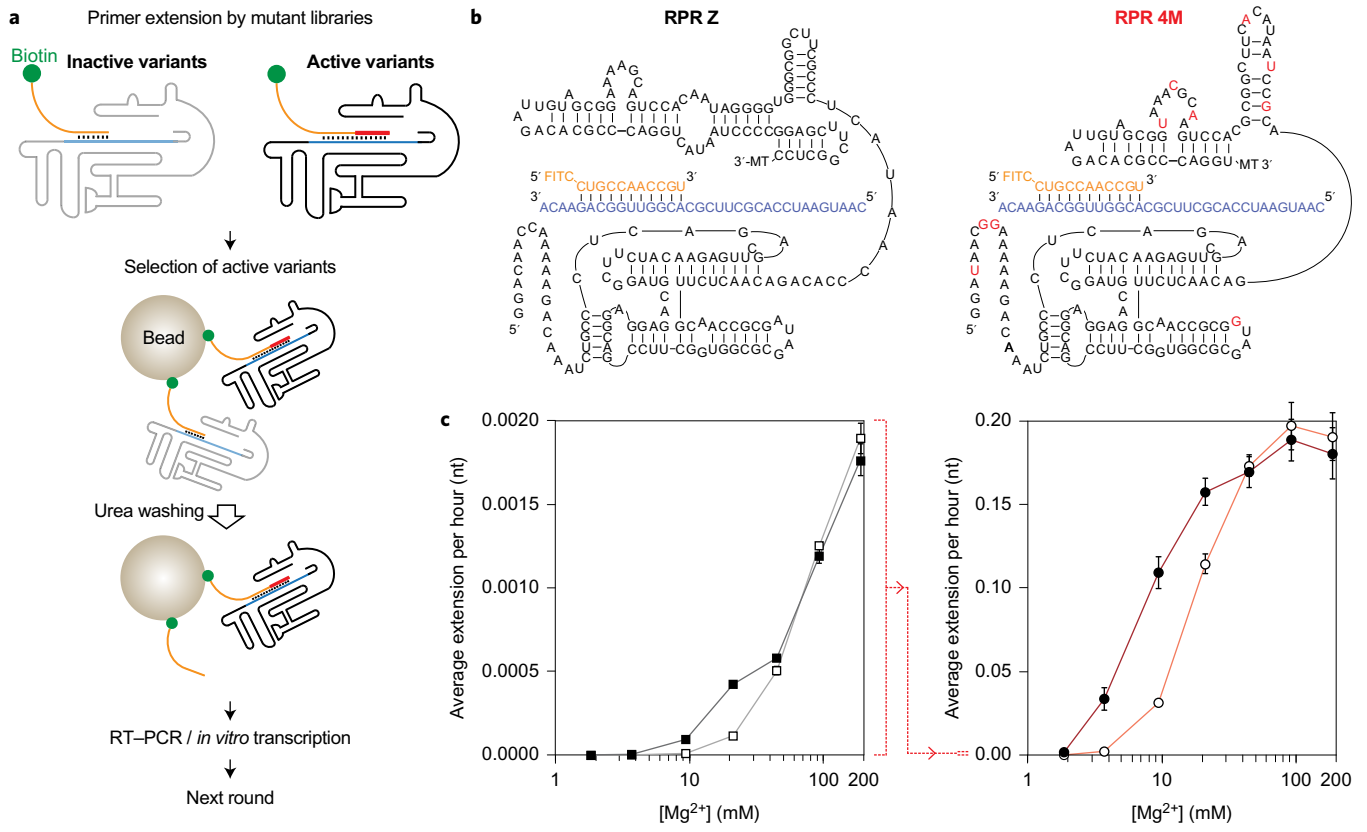


Figure 3 | K_{10} and [Mg^{2+}] dependence of RPR activity by Z and the evolved 4M. a, Outline of the TST selection scheme, showing how *in cis* primer extension by a ribozyme prevents ribozyme loss during a denaturing wash on beads, allowing it to be recovered and amplified. **b**, Putative secondary structures of the Z and 4M ribozymes surrounding the primer (orange)/template (purple) duplex, showing 4M selected mutations in red (3′ MT = tail sequence from selection construct). **c**, Average primer extension per hour on template TI by the Z (left) or 4M (right) RPRs in 5 day (left) or 4 h (right) reactions at varying [Mg^{2+}], with 6 μM K_{10} (filled symbols) or without it (open symbols). Note the change in scale indicated between the two panels (dotted red line) due to the nearly two orders of magnitude faster extension by 4M (error bars represent s.d., $N = 3$).

evolution in the presence and absence of peptide cofactor, starting from a shorter version of the Z RPR (Zd) with a truncated accessory domain[25], which exhibits a similar activity and K_{10} dependence to Z (Supplementary Fig. 10). To this end, we developed a rapid *in vitro* selection system for RPR function, called tethered self-tagging (TST), conceptually similar to the compartmentalized self-tagging method developed for the evolution of proteinaceous polymerases[26] (Fig. 3a and Supplementary Fig. 11). TST test selections at low [Mg^{2+}] (25 mM) revealed a striking impact of K_{10} peptide on RPR evolution, with selections in the presence of peptide progressing almost an order of magnitude faster, requiring less total incubation time for all seven rounds of selection than the first round in the absence of peptide (Supplementary Fig. 12). Such peptide-mediated enhancements could also sensitize initially weak activities to recovery using *in vitro* evolution.

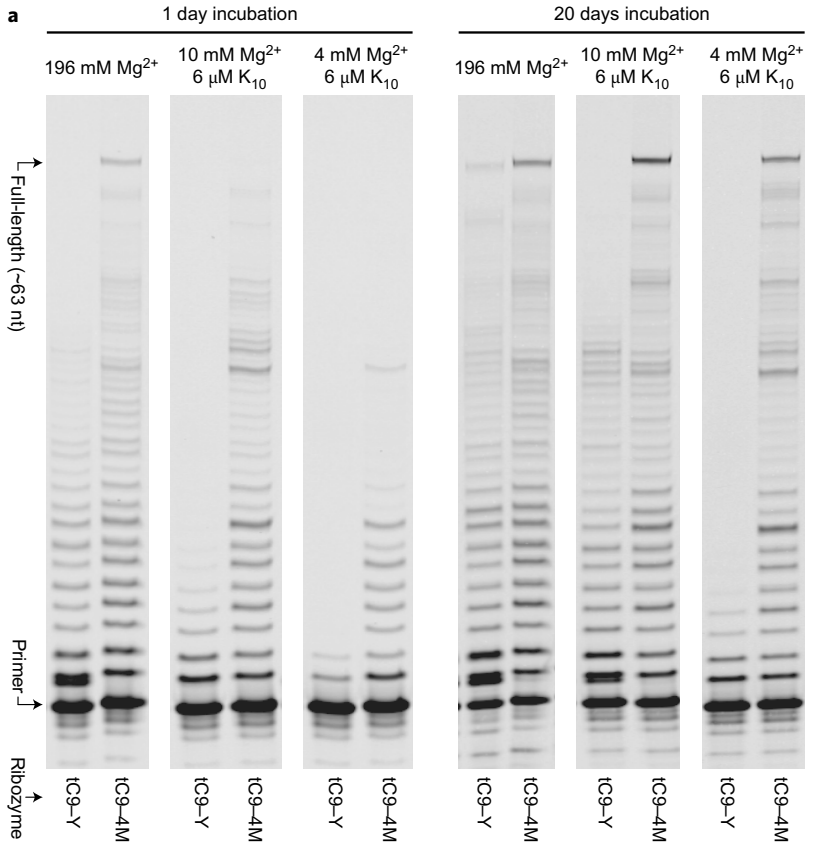
Encouraged by these efficiency gains, we carried out TST selections in the presence of K_{10} peptide under progressively lower [Mg^{2+}] conditions, yielding RPR variants increasingly capable of RNA synthesis under such challenging conditions (Supplementary Fig. 13). Selection at much lower [Mg^{2+}] (2 mM) yielded the RPR variant 4M (Fig. 3b), which exhibited a greatly reduced Mg^{2+} dependence, displaying clear activity with boosts from K_{10} peptide at [Mg^{2+}] between 2 and 30 mM. Furthermore, 4M outperformed Z by approximately two orders of magnitude over all Mg^{2+} concentrations, both with and without K_{10} peptide (Fig. 3c and Supplementary Fig. 14). 4M mutations (including a number of insertions) cluster near the 5′ end and in the accessory domain within a purine-rich loop (thought to be positioned near the active site[25]) and within the linker connecting the core to the accessory domain, potentially reshaping this linker region into a new hairpin domain (Fig. 3b).

Next, we appended a previously described 5′ sequence (5′-$GUCAUUGA_8$ (tC9)[6]) that boosts long RNA synthesis, yielding tC9–4M, which could perform long RNA synthesis down to 2 mM Mg^{2+} in the presence of K_{10} or at a near-physiological [Mg^{2+}] of 1 mM in the presence of molecular crowding agents such as PEG (Supplementary Fig. 15). tC9–4M greatly outperformed a previous top RPR, the fast and accurate tC9–Y (ref. 6), under all conditions, synthesizing (at high [Mg^{2+}], 196 mM) more full-length (~63 nt) RNA products (0.8%) in a single day than tC9–Y did in 20 days (0.4%) (Fig. 4a and Supplementary Fig. 16).

Over long incubations at low [Mg^{2+}] (4 and 10 mM), tC9–4M with K_{10} peptide yielded 2.3 and 4.6% full-length products (the latter a higher yield than the 2.9% it made at optimal high [Mg^{2+}] (196 mM), presumably due to reduced RPR and template degradation), whereas tC9–Y was barely active under these conditions. tC9–4M could also synthesize RNAs longer than itself (177 nt) at low [Mg^{2+}] (10 mM) with K_{10}, but provided much better yields (2.4% of ~195 nt) at optimal high [Mg^{2+}] (196 mM) (versus the 0.37% observed previously for tC9–Y under optimal conditions[6]; Supplementary Fig. 17).

Finally, deep sequencing of total extension products showed that tC9–4M exhibits an essentially identical overall fidelity compared to tC9–Y (>97.5% per position)[6] and that neither low [Mg^{2+}] nor K_{10} peptide significantly alter RPR fidelity (Fig. 4b and Supplementary Fig. 18).

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.



b tC9–Y, 196 mM Mg^{2+} — 8.9 × 10^6 positions

Incorporated base \ Encoded base	A	C	G	U	Total insertion
A	**93.29**	0.23	0.13	0.42	
C	0.47	**98.98**	0.13	0.30	0.065
G	2.69	0.08	**99.52**	0.55	Overall
U	0.13	0.63	0.09	**98.54**	fidelity
-	3.42	0.08	0.12	0.19	97.48

tC9–Y, 10 mM Mg^{2+}, 10 μM K_{10} — 2.9 × 10^6 positions

Incorporated base \ Encoded base	A	C	G	U	Total insertion
A	**94.37**	0.52	0.31	0.30	
C	0.87	**98.58**	0.33	0.37	0.151
G	2.88	0.25	**99.00**	0.44	Overall
U	0.29	0.54	0.23	**98.76**	fidelity
-	1.59	0.12	0.13	0.13	97.51

tC9–4M, 10 mM Mg^{2+}, 10 μM K_{10} — 6.4 × 10^6 positions

Incorporated base \ Encoded base	A	C	G	U	Total insertion
A	**94.86**	0.16	0.10	0.64	
C	0.39	**98.45**	0.09	0.23	0.035
G	4.00	0.08	**99.65**	0.42	Overall
U	0.06	1.26	0.07	**98.58**	fidelity
-	0.68	0.05	0.10	0.13	97.83

tC9–4M, 4 mM Mg^{2+}, 10 μM K_{10} — 9.0 × 10^6 positions

Incorporated base \ Encoded base	A	C	G	U	Total insertion
A	**95.50**	0.18	0.11	0.53	
C	0.38	**98.49**	0.13	0.25	0.034
G	3.57	0.10	**99.58**	0.41	Overall
U	0.08	1.18	0.08	**98.67**	fidelity
-	0.48	0.06	0.10	0.14	98.01

Figure 4 | Robust RNA synthesis during long incubations. a, PAGE of primer extensions by RPRs tC9-Y and tC9-4M tethered to the long repeat template I-6; the K_{10} peptide allows tC9-4M extensions at low [Mg^{2+}] to catch up or exceed those at high [Mg^{2+}] after long incubations. **b**, Relative incorporation of correct and incorrect nucleotides and deletions for each nucleobase, derived from similar extensions on these templates in Supplementary Fig. 18. The overall fidelity, calculated as a geometric mean of the fidelity at each nucleobase minus the rate of insertions, was broadly similar across all ribozymes and conditions. The total positions sequenced for each ribozyme condition are indicated at the upper right of each table.

Peptide-mediated RNA synthesis in protocells. Encapsulation of RNA synthesis and replication within compartments is widely considered an important transition in early evolution as well as a key stepping stone in generating a simple synthetic cell[2,27], but high [Mg^{2+}] can disrupt membrane integrity and cause precipitation of some membrane-forming amphiphiles. Simple ribozymes with low [Mg^{2+}] requirements such as the hammerhead endonuclease ribozyme have been shown to operate at a reduced rate in membranous protocells formed from fatty acids (indicating ~1 mM free Mg^{2+})[11,28], but the exceptionally high [Mg^{2+}] requirement for RPR activity had previously rendered enzymatic RNA synthesis fundamentally incompatible with such membranous compartments.

To begin to explore if the reduced Mg^{2+} requirements of tC9–4M in combination with K_{10} peptide could allow enzymatic RNA synthesis within model membranous compartments, we generated stable vesicles formed from the modern membrane lipid 1-palmitoyl-2-oleoylphosphatidylcholine (POPC), which could encapsulate RNA even in the presence of 10 mM Mg^{2+} with and without K_{10} peptide (Fig. 5a). Only in the presence of K_{10} peptide did we observe robust templated full-length RNA synthesis by tC9–4M within POPC compartments (Fig. 5b and Supplementary Fig. 19 (variant protocol)). When only the primer/template and K_{10} were encapsulated and tC9–4M was added later to these preformed vesicles, no RNA synthesis was observed, confirming that the POPC vesicles were intact and able to compartmentalize macromolecules without leakage while still supporting RNA synthesis exclusively within the vesicular lumen.

Discussion

Prebiotic chemistry provides increasingly clear indications that early RNA replication arose not in isolation but in a chemically diverse environment, which very probably included building blocks for simple peptides and lipids[12]. Here, we have explored to what extent short peptides can interact with and augment the function of RPRs and can facilitate primordial RNA replication and the formation of model protocells.

Our results show that unstructured peptide segments from several ribosomal proteins, thought to be among the most ancient evolved coded protein sequences in biology[22], are able to potently enhance the activity of an evolutionarily unrelated RPR. It has been suggested that the replacement of Mg^{2+} by peptides in the outer and less ancient layers of the ribosome may reflect a transition from inorganic metal ions to peptides as counterions during ribosomal evolution[19]. Our data indeed indicate how these peptides might have promoted such a transition in proto-ribosomal function and how similar peptides might have helped early RPRs (and other ribozymes) achieve a similar shift from a high to a low [Mg^{2+}] regime. Such a transition is crucial for the emergence of self-replication and cellular evolution, as high Mg^{2+} concentrations, while helpful in promoting RNA folding and catalysis, are destructive both to RNA stability and to the integrity of prebiotically plausible membranes, placing conflicting demands on plausible solute compositions for early RNA replication. Although membranes can be protected to some extent by chelating agents such as citrate, which has been shown to sequester Mg^{2+} in a form still able to promote non-enzymatic RNA synthesis from

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

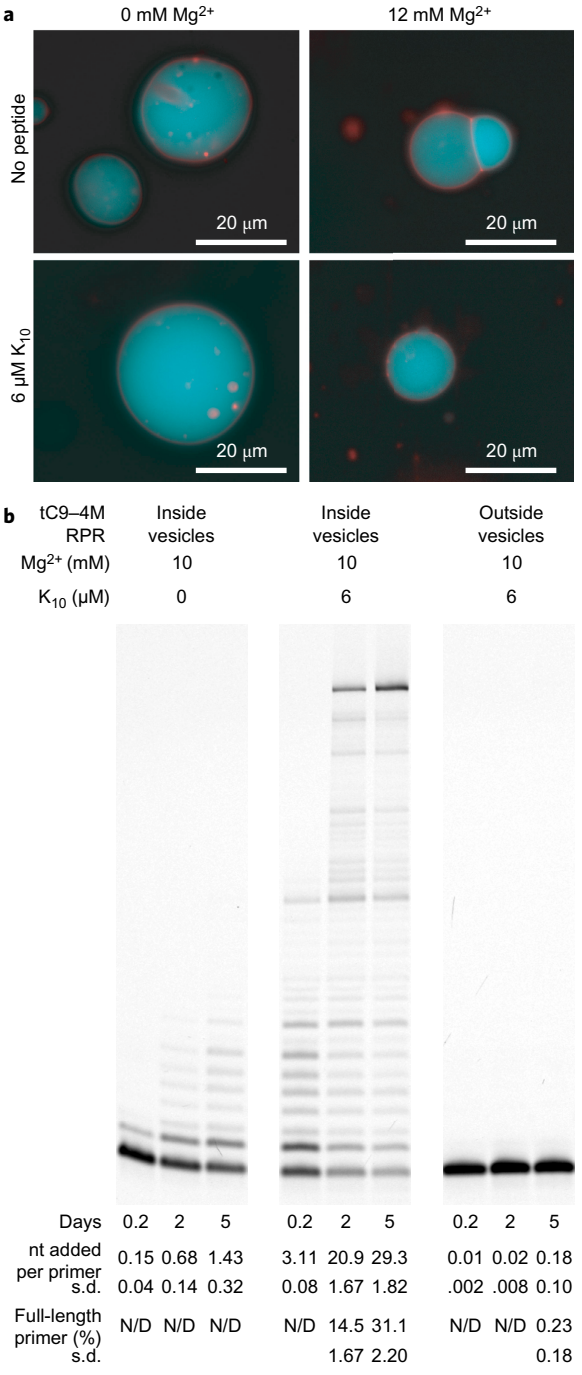


Figure 5 | Ribozyme-catalysed RNA synthesis in phospholipid protocells. **a**, Fluorescence micrograph of POPC vesicles containing RNA and NTPs, which can form across a range of conditions both with and without Mg^{2+} and K_{10} peptide; the membrane is stained with Rhodamine G (red) and encapsulated RNA is labelled with fluorescein (blue). **b**, PAGE of extensions at 17 °C on the long repeat template I-6 encapsulated within POPC vesicles, by tC9–4M added externally or encapsulated within the vesicles. Quantification ($N = 4$; N/D = not detected) of extension per primer and full-length extension are displayed beneath. K_{10} allows efficient full-length RNA synthesis by the ribozyme within POPC vesicles.

phosphorimidazolides[29], citrate-chelated Mg^{2+} is inaccessible for RPR catalysis (Supplementary Fig. 20).

From the salient sequence features of activating ribosomal peptides we derived a maximally simple homopolymeric lysine decapeptide K_{10} that could relieve the dependency of RPRs on high [Mg^{2+}] and potentiate RPR function, stability and evolution, enabling templated RNA synthesis down to near-physiological (≥1 mM) Mg^{2+} concentrations. In analogy to the function of peptides within the ribosomal core and the ability of other peptides to promote simple ribozyme assembly[17], the K_{10} peptide does not substitute for Mg^{2+} in catalysis, but instead appears to act as a molecular bridge and counterion promoting RPR holoenzyme formation, that is, the docking of the polyanionic RNA primer–template duplex to the polyanionic RPR apoenzyme. The ability of peptides to fulfil this highly Mg^{2+}-dependent role may be a critical element in realizing ribozyme-catalysed RNA synthesis in primordial protocells based on fatty acid or related amphiphiles, in addition to the more Mg^{2+}-tolerant model POPC phospholipid membranes examined here.

A long-standing question has been whether and how early peptides could have conferred a beneficial heritable phenotype in the absence of encoded synthesis. The simple peptides described herein confer a critical phenotypic and adaptive advantage irrespective of chirality (or chiral purity), which is dependent largely on length and composition rather than sequence. Furthermore, other roles are emerging for other compositionally biased peptides, for example, in promoting RNA template accessibility or membrane localization[30,31]. Unlike peptides with a defined sequence, such compositionally simple peptides would not depend on encoded synthesis, but could plausibly emerge from prebiotic sources via a biased chemical (or non-specific enzymatic) peptide synthesis pathway. As proposed by Cech, these peptides could become heritable in the form of peptidyl-transferase ribozymes with single amino-acid specificity[32] akin to the modern-day D-Ala-D-Ala ligase enzymes involved in bacterial cell wall synthesis. Such dedicated (homo)-peptide synthetase ribozymes would provide a conceptually simple linkage of peptide-encoded phenotypes to RNA-encoded genotypes in the absence of translation and offer a plausible evolutionary rationale for the emergence of a peptidyl-transferase ribozyme prior to its recruitment into a nascent proto-ribosomal system of encoded protein synthesis[33,34].

Although lysine is not usually considered to be among the most common amino acids emerging from prebiotic chemistry[12], we find that decapeptides formed from the non-proteinogenic Lys analogues Orn and Dab (proposed as components of an early genetic code[35]) are functional substitutes and provide similar RPR activation. It has been postulated that the earliest amino acids incorporated into a primeval genetic code must have included the amino acids with the highest 'catalytic potential'[36]. Amino acids bearing side-chain primary amino groups (such as the above) certainly hold unique functional potential and might have benefited early protocellular entities in ways other than enhancing RPR function and RNA replication. For example, K-rich peptides have been shown to be capable of forming primitive catalysts with decarboxylase and retro-aldolase activities[37,38], potentially providing simple metabolic functions. Furthermore, they can promote liquid–liquid demixing and coacervate formation at higher concentrations[39], potentially a key mechanism for prebiotic substrate concentration and protocellular (and subcellular) molecular organization[40].

In summary, our results suggest that the advent of RNA self-replication could have been promoted by simple mixed-chirality peptides that reduced RNA dependence on inorganic counterions, accelerated RNA evolution, and enabled the partition of RNA replication within membranous compartments, a prerequisite for Darwinian evolution. The ancient molecular symbiosis between nucleic acids and peptides, a molecular memory of which remains captured within the structure of the ribosome, may thus have been a key factor in the emergence of the first protocells.

Methods

Primer extension assays. Standard primer extension reactions by the RPRs were performed with 250 nM each of ribozyme, template and 5′-FITC-labelled primer in

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

extension buffer containing 50 mM Tris·HCl buffer (pH 8.3), 0–200 mM $MgCl_2$, 0.5 mM of each ribonucleotide triphosphate (NTP) and at 17 °C (unless indicated otherwise). Each NTP chelates one to two Mg^{2+} ions (with K_D values at 37 °C of 30 μM ($[NTP^{4-}][Mg^{2+}]/[MgNTP^{2-}]$) and 25 mM ($[MgNTP^{2-}][Mg^{2+}]/[Mg_2NTP]$))[41] and the predicted unchelated accessible Mg^{2+} concentrations (free $[Mg^{2+}]$) are displayed in the figures and legends unless otherwise specified. Initially, RNAs (Z^{LT}, template TI (5′-CAAUGAAUCCACGCUUCGCACGGUUGGCAGAACA-3′) and primer FITC11 (5′- FITC-CUGCCAACCGU-3′) unless indicated otherwise) were annealed in H_2O (50 °C for 5 min, 17 °C for 10 min). They were then mixed with the extension buffer, including peptides (oligomer concentrations indicated in each figure or figure legend). In some conditions where extension by the ribozyme was particularly limited (Supplementary Figs 2, 4, 10 and 15), 8% PEG-6000 was added to the extension buffer to accelerate the primer extension by molecular crowding. After the incubation time indicated in the figure legends, the reactions were stopped by adding three volumes of stop buffer (75 mM EDTA, 8 M urea and 10× (RNA) or 100× (DNA) excess of competing oligonucleotide able to hybridize to the template sequence), heated to 94 °C for 4 min and resolved by urea–PAGE gels (20% polyacrylamide, 8 M urea). The gels were analysed using a Typhoon Trio Scanner (GE Healthcare). Quantification of extension products was performed as previously described[6], with the exception of those in Fig. 5b, where bands beyond +11 were grouped into 11 nt 'bins' weighted by their average length. All nucleic acid sequences used are described in the three Supplementary Tables.

Formation and observation of POPC giant vesicles. Giant vesicles were prepared by the droplet transfer method[42]. First, an Eppendorf tube (2 ml) was filled with 0.5 ml of the 'lower solution' containing 50 mM Tris-HCl pH 8.3, 500 mM glucose and 2 μM Rhodamine 6G (Sigma-Aldrich). The interfacial phase (0.5 ml mineral oil (Sigma-Aldrich) containing 0.5 mM 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC, Anatrace)) was then added on the lower solution. The Eppendorf tube was left at room temperature for a few minutes to let POPC make a single layer between the outer solution and interfacial phase. Finally, a water/oil emulsion was prepared by pipetting the 'inner phase' (50 μl, 4 μM RNA primer FITC11, 50 mM Tris HCl pH 8.3, 0–12 mM $MgCl_2$, 500 mM sucrose, ±6 μM K_{10} peptide) in 1.0 ml of 0.5 mM POPC in mineral oil and added gently on the interfacial phase. The tube was immediately centrifuged at 300*g* for 10 min at room temperature. After centrifugation, mineral oil and debris between the mineral oil and aqueous phase were removed, and the aqueous phase was centrifuged again (2,000*g*, 5 min, room temperature). The supernatant was then removed and POPC vesicles were washed with 200 μl of 50 mM Tris HCl pH 8.3, 500 mM glucose, and vesicles were collected by centrifugation (2,000*g*, 2 min, room temperature). Finally, vesicles were suspended in 50 μl of the 'outer phase' (50 mM Tris HCl pH 8.3, 0–12 mM $MgCl_2$, 500 mM sucrose, ±6 μM K_{10} peptide) and observed under the fluorescent microscope.

RNA primer extension in POPC giant vesicles. tC9–4M^{MT} (12.5 pmol) and the RNA primer/template (FITC11/I-6) were annealed in water separately. The primer/template was then mixed with reaction buffer (50 mM Tris HCl pH 8.3, 12 mM $MgCl_2$, 0.5 mM each NTP, 500 mM sucrose, ±6 μM K_{10} peptide) with/without tC9–4M^{MT} to make 50 μl of the inner phase. Vesicles were then formed as described in the previous section (without adding Rhodamine 6G to the lower solution). The vesicles were washed three times with 200 μl of 50 mM Tris HCl pH 8.3, 500 mM glucose and suspended in 50 μl of the outer phase (50 mM Tris HCl pH 8.3, 12 mM $MgCl_2$, 0.5 mM each NTP, 500 mM sucrose, ±6 μM K_{10} peptide) with/without tC9–4M^{MT}. The vesicles were incubated at 17 °C for the periods indicated in Fig. 5b and then washed three times with 200 μl of 50 mM Tris HCl pH 8.3, 500 mM glucose, and collected by centrifugation. The vesicles were mixed with three volumes of the stop buffer (95% formamide, 7.5 mM EDTA and 10× excess of competing RNA able to hybridize to the template sequence), heated to 94 °C for 4 min, and resolved by urea–PAGE gels (10% polyacrylamide, 8 M urea). The gels were analysed using a Typhoon Trio Scanner (GE Healthcare). A variant vesicle preparation method (Supplementary Materials and Methods) using paraffin oil and higher POPC concentrations also yielded in-vesicle extension (Supplementary Fig. 19)[43].

Data availability. All relevant data such as ribozyme and peptide sequences, primer and template sequences and detailed experimental conditions are included within the manuscript, specifically in the Supplementary Tables and Supplementary Materials and Methods.

Received 13 July 2016; accepted 20 January 2017; published online 6 March 2017

References

1. Gesteland, R., Cech, T. & Atkins, J. (eds) *The RNA World* 3rd edn (Cold Spring Harbor Laboratory, 2006).
2. Szostak, J. W., Bartel, D. P. & Luisi, P. L. Synthesizing life. *Nature* **409,** 387–390 (2001).
3. Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E. & Bartel, D. P. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* **292,** 1319–1325 (2001).
4. Wochner, A., Attwater, J., Coulson, A. & Holliger, P. Ribozyme-catalyzed transcription of an active ribozyme. *Science* **332,** 209–212 (2011)
5. Horning, D. P. & Joyce, G. F. Amplification of RNA by an RNA polymerase ribozyme. *Proc. Natl Acad. Sci. USA* **113,** 9786–9791 (2016)
6. Attwater, J., Wochner, A. & Holliger, P. In-ice evolution of RNA polymerase ribozyme activity. *Nat. Chem.* **5,** 1011–1018 (2013).
7. Attwater, J. *et al.* Chemical fidelity of an RNA polymerase ribozyme. *Chem. Sci.* **4,** 2804–2814 (2013).
8. Muller, U. F. & Bartel, D. P. Improved polymerase ribozyme efficiency on hydrophobic assemblies. *RNA* **14,** 552–562 (2008).
9. Attwater, J., Wochner, A., Pinheiro, V. B., Coulson, A. & Holliger, P. Ice as a protocellular medium for RNA replication. *Nat. Commun.* **1,** 1–8 (2010).
10. Monnard, P. A., Apel, C. L., Kanavarioti, A. & Deamer, D. W. Influence of ionic inorganic solutes on self-assembly and polymerization processes related to early forms of life: implications for a prebiotic aqueous medium. *Astrobiology* **2,** 139–152 (2002).
11. Chen, I. A., Salehi-Ashtiani, K. & Szostak, J. W. RNA catalysis in model protocell vesicles. *J. Am. Chem. Soc.* **127,** 13213–13219 (2005).
12. Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D. & Sutherland, J. D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat. Chem.* **7,** 301–307 (2015).
13. Valadkhan, S., Mohammadi, A., Jaladat, Y. & Geisler, S. Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc. Natl Acad. Sci. USA* **106,** 11901–11906 (2009).
14. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289,** 920–930 (2000).
15. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35,** 849–857 (1983).
16. Coetzee, T., Herschlag, D. & Belfort, M. *Escherichia coli* proteins, including ribosomal protein S12, facilitate *in vitro* splicing of phage T4 introns by acting as RNA chaperones. *Genes Dev.* **8,** 1575–1588 (1994).
17. Herschlag, D., Khosla, M., Tsuchihashi, Z. & Karpel, R. L. An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis. *EMBO J.* **13,** 2913–2924 (1994).
18. Bokov, K. & Steinberg, S. V. A hierarchical model for evolution of 23S ribosomal RNA. *Nature* **457,** 977–980 (2009).
19. Hsiao, C., Mohan, S., Kalahar, B. K. & Williams, L. D. Peeling the onion: ribosomes are ancient molecular fossils. *Mol. Biol. Evol.* **26,** 2415–2425 (2009).
20. Klein, D. J., Moore, P. B. & Steitz, T. A. The contribution of metal ions to the structural stability of the large ribosomal subunit. *RNA* **10,** 1366–1379 (2004).
21. Smith, T. F., Lee, J. C., Gutell, R. R. & Hartman, H. The origin and evolution of the ribosome. *Biol. Direct.* **3,** 16 (2008).
22. Alva, V., Soding, J. & Lupas, A. N. A vocabulary of ancient peptides at the origin of folded proteins. *eLife* **4,** e09410 (2015).
23. Voorhees, R. M., Schmeing, T. M., Kelley, A. C. & Ramakrishnan, V. The mechanism for activation of GTP hydrolysis on the ribosome. *Science* **330,** 835–838 (2010).
24. DeRouchey, J., Hoover, B. & Rau, D. C. A comparison of DNA compaction by arginine and lysine peptides: a physical basis for arginine rich protarnines. *Biochemistry* **52,** 3000–3009 (2013).
25. Wang, Q. S., Cheng, L. K. & Unrau, P. J. Characterization of the B6.61 polymerase ribozyme accessory domain. *RNA* **17,** 469–477 (2011).
26. Pinheiro, V. B. *et al.* Synthetic genetic polymers capable of heredity and evolution. *Science* **336,** 341–344 (2012).
27. Attwater, J. & Holliger, P. A synthetic approach to abiogenesis. *Nat. Methods* **11,** 495–498 (2014).
28. Engelhart, A. E., Adamala, K. P. & Szostak, J. W. A simple physical mechanism enables homeostasis in primitive cells. *Nat. Chem.* **8,** 448–453 (2016).
29. Adamala, K. & Szostak, J. W. Nonenzymatic template-directed RNA synthesis inside model protocells. *Science* **342,** 1098–1100 (2013).
30. Kamat, N. P., Tobe, S., Hill, I. T. & Szostak, J. W. Electrostatic localization of RNA to protocell membranes by cationic hydrophobic peptides. *Angew. Chem. Int. Ed.* **54,** 11735–11739 (2015).
31. Jia, T. Z., Fahrenbach, A. C., Kamat, N. P., Adamala, K. P. & Szostak, J. W. Oligoarginine peptides slow strand annealing and assist non-enzymatic RNA replication. *Nat. Chem.* **8,** 915–921 (2016).
32. Cech, T. R. Evolution of biological catalysis: ribozyme to RNP enzyme. *Cold Spring Harb. Symp. Quant. Biol.* **74,** 11–16 (2009).
33. Krupkin, M. *et al.* A vestige of a prebiotic bonding machine is functioning within the contemporary ribosome. *Philos. Trans. R. Soc. Lond. B* **366,** 2972–2978 (2011).
34. Petrov, A. S. *et al.* History of the ribosome and the origin of translation. *Proc. Natl Acad. Sci. USA* **112,** 15396–15401 (2015).
35. Hartman, H. & Smith, T. F. The evolution of the ribosome and the genetic code. *Life (Basel)* **4,** 227–249 (2014).
36. Szathmary, E. The origin of the genetic code: amino acids as cofactors in an RNA world. *Trends Genet.* **15,** 223–229 (1999).

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

37. Johnsson, K., Allemann, R. K., Widmer, H. & Benner, S. A. Synthesis, structure and activity of artificial, rationally designed catalytic polypeptides. *Nature* **365,** 530–532 (1993).
38. Muller, M. M., Windsor, M. A., Pomerantz, W. C., Gellman, S. H. & Hilvert, D. A rationally designed aldolase foldamer. *Angew. Chem. Int. Ed.* **48,** 922–925 (2009).
39. Koga, S., Williams, D. S., Perriman, A. W. & Mann, S. Peptide–nucleotide microdroplets as a step towards a membrane-free protocell model. *Nat. Chem.* **3,** 720–724 (2011).
40. Mann, S. Systems of creation: the emergence of life from nonliving matter. *Acc. Chem. Res.* **45,** 2131–2141 (2012).
41. Thomen, P. *et al.* T7 RNA polymerase studied by force measurements varying cofactor concentration. *Biophys. J.* **95,** 2423–2433 (2008).
42. Miele, Y., Bánsági, T. Jr, Taylor, A. F., Stano, P. & Rossi, F. in *Advances in Artificial Life, Evolutionary Computation and Systems Chemistry* Vol. 587 (eds Rossi, F. *et al.*) Ch. 18, 197–208 (Springer, 2016).
43. Fujii, S. *et al.* Liposome display for *in vitro* selection and evolution of membrane proteins. *Nat. Protoc.* **9,** 1578–1591 (2014).

Acknowledgements

The authors thank S. James for assistance with fidelity data processing. This work was supported by postdoctoral fellowships from JSPS (Japanese Society for the Promotion of Science) and HFSP (Human Frontiers Science Program) (S.T.) and by the Medical Research Council (J.A., P.H.; programme no. MC_U105178804).

Author contributions

S.T. and P.H. conceived and designed the experiments. S.T. performed all experiments together with J.A. for selection design, fidelity measurement and peptide assays. All authors discussed the results and jointly wrote and commented on the manuscript.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.H.

Competing financial interests

The authors declare no competing financial interests.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.