A prebiotically plausible scenario of an RNA–peptide world

The RNA world concept is one of the most fundamental pillars of the origin of life theory. It predicts that life evolved from increasingly complex self-replicating RNA molecules. The question of how this RNA world then advanced to the next stage, in which proteins became the catalysts of life and RNA reduced its function predominantly to information storage, is one of the most mysterious chicken-and-egg conundrums in evolution. Here we show that non-canonical RNA bases, which are found today in transfer and ribosomal RNAs, and which are considered to be relics of the RNA world, are able to establish peptide synthesis directly on RNA. The discovered chemistry creates complex peptide-decorated RNA chimeric molecules, which suggests the early emergence of an RNA–peptide world from which ribosomal peptide synthesis may have emerged. The ability to grow peptides on RNA with the help of non-canonical vestige nucleosides offers the possibility of an early co-evolution of covalently connected RNAs and peptides, which then could have dissociated at a higher level of sophistication to create the dualistic nucleic acid–protein world that is the hallmark of all life on Earth.

A central commonality of all cellular life is the translational process, in which ribosomal RNA (rRNA) catalyses peptide formation with the help of transfer RNAs (tRNA), which function as amino acid carrying adapter molecules. Comparative genomics suggests that ribosomal translation is one of the oldest evolutionary processes, which dates back to the hypothetical RNA world. The questions of how and when RNA learned to instruct peptide synthesis is one of the grand unsolved challenges in prebiotic evolutionary research.

The immense complexity of ribosomal translation demands a stepwise evolutionary process. From the perspective of the RNA world, at some point RNA must have gained the ability to instruct and catalyse the synthesis of, initially, just small peptides. This initiated the transition from a pure RNA world into an RNA–peptide world. In this RNA–peptide world, both molecular species could have co-evolved to gain increasing ‘translation’ and ‘replication’ efficiency.

To gain insight into the initial processes that may have enabled the emergence of an RNA–peptide world, we analysed the chemical properties of non-canonical nucleosides, which can be traced back to the last universal common ancestor and, as such, are considered to be ‘living molecular fossils’ of an early RNA world.

This approach, which can be called ‘palaeochemistry’, enabled us to learn about the chemical possibilities that existed in the RNA world and, therefore, sets the chemical framework for the emergence of life. In contrast to earlier investigations of the origin of translation, we naturally occurring non-canonical vestige nucleosides and conditions compatible with aqueous wet–dry cycles were used to study the chemical possibilities that existed in the RNA world.

**Peptide synthesis on RNA**

In modern tRNAs, the amino acids that give peptides are linked to the CCA 3′ terminus via a labile ester group. Some tRNAs, however, contain additional amino acids in the form of amino acid-modified nucleosides, for example, g6A (ref. 3), t6A (ref. 4) and m2tA (ref. 5), which are found directly next to the anticodon loop at position 37. Other non-canonical vestige nucleosides are present in the wobble position 34 as nm5U and mnm5U (refs. 6–8).

A kinetic analysis shows that the nature of the amino acid affects the rate constant (kapp) of the coupling reaction (Fig. 2d). For example, G (in 1a) couples to 2c with an apparent rate constant (kapp) of 0.1 h⁻¹. For the amino acids L (in 1d), I (in 1e) and M (in 1h) a fourfold higher rate constant (≈0.4 h⁻¹) was determined. The highest rate was measured for E (in 1g) with kapp > 1 h⁻¹. These differences establish a pronounced amino acid selectivity in the coupling reaction, probably as a result of distinct pre-organizations. We next reduced the length of the RNA donor strand to five, and finally to three, nucleotides (Supplementary Information).

We detected coupling even with a trimer...
RNA donor strand, although it required duplex-enforcing high salt and low temperature conditions (1 M NaCl and 0 °C). The interaction of three nucleotides on the donor with the corresponding triplet on the acceptor seems to be the lower limit for productive coupling. Interestingly, this is the size of the codon–anticodon interaction in contemporary translation.

We next investigated coupling of the nitrile derivative of (m⁶)aa⁶A (aa, amino acid), which is present at position 37 in certain tRNAs. General oligonucleotides are simplified and only terminal nucleobases are drawn.

We investigated whether longer peptides can also be generated by mRNA 5′-acceptor strands as starting materials (Supplementary Information). The synthesized acceptor strands were hybridized to the donor strand 1a. After carboxylic acid activation, rapid formation of elongated hairpin-type intermediates with yields between 40% and 60% was observed (Fig. 3b). We found that the coupling yields did not drop substantially with increasing peptide length, suggesting that other factors, such as the RNA hybridization kinetics, are rate limiting. In all cases, the subsequent urea cleavage (pH 4, 90 °C) affords dipeptide- to hexapeptide-decorated RNAs in 10–15% yield. These modest yields are the result of substantial RNA degradation, driven by the pH and temperature conditions that were used. The decomposition of RNA, however, can be overcome by using 2′-Ome nucleotides (see Stepwise growth of peptides on RNA), which are also vestiges of the early RNA world. During urea cleavage we detected competing formation of hydantoin side products, depending on the pH and temperature (Fig. 3a). Under mildly acidic conditions (pH 6, 90 °C), exclusive formation of the hydantoin product, cyclic-5c, was observed. Reducing the temperature and a shift to higher acidity (pH 4, 60 °C) led to the preferential formation of the peptide product, 5c (approximately 7:1 5c:cyclic-5c ratio).

**Fragment coupling on RNA**

We investigated whether longer peptides can also be generated by fragment coupling chemistry with RNA donor strands containing an already longer peptide (m⁶-peptide⁶A). This is essential because an RNA–peptide world, with initially low chemical efficiency, might have been limited to the synthesis of smaller peptides. We found that the required adenosine nucleosides, containing a whole peptide attached to the 5′-position, are available if the peptides that are produced by RNA degradation of the RNA–peptide chimeras, for example, can react with nitrosated N⁶-methylurea adenosine (Fig. 4a). When we treated N⁶-methylurea adenosine with NaN₃ (5% H₃PO₄) and added the solution to triglycine (pH 9.5), we obtained the peptide-coupled adenosine nucleoside 4,1,3⁶A.
Fig. 2 | Peptide synthesis on RNA with terminal (m)mm5U and m5asA nucleotides. a, Reaction scheme for 1a (5′-m6g6A-RNA-3′) and 2a (3′-mm5U-RNA-5′) with coupling (1) and cleavage (2). b, HPLC chromatograms of the crude mixture, obtained after coupling of 1a with 2a using DMTMM-CI (see reaction condition b) and cleavage of 3a (100 mM DMTM buffer pH 6, 100 mM NaCl, 90 °C, 6 h). HPLC peaks of RNAs are coloured: donor in blue; acceptor in red; hairpin-type intermediate in purple; and cleaved donor strand in pale blue. The insets show MALDI-TOF data (negative mode) of the isolated products 3a and 5a. Calcd., calculated. c, Coupling results obtained with different activators for 1a and 1i with 2a–2c. d, Coupling reactions with different donors 1a–1i and acceptors 2a–2c, and apparent rate constants (k_app) of selected coupling reactions with 2c. All coupling reactions were carried out using a concentration of 50 μM for 1a–1i and 50 μM for 2a–2c (100 mM NaCl, 25 °C, 50 mM DMTMM buffer pH 6, 24 h). 750 mM DMTMM-CI (100 mM DMTM buffer pH 6, 24 h), 50 mM MeNC (50 mM DCI buffer pH 6, 5 days). e50 mM DTT (100 mM borate buffer pH 9, 24 h). The two yields with 1j (aa, D) describe the reaction of the aspartic acid α-COOH and of the side chain COOH. An assignment was not performed. RT, room temperature; ND, not determined.

**Effect of base pairing**

To investigate the importance of sequence complementarity, we added two RNA donor strands of different lengths (7-mer: 5′-m6g6A-RNA-3′ and 11-mer: 5′-mm5U-RNA-3′) to an acceptor strand with a mm5U at the 3′ end (11-mer: 2c) (Fig. 5b, left). On the basis of the melting temperatures of the two possible duplexes (approximately 30 °C for the 7-mer-11-mer and 59 °C for the 11-mer-11-mer, see Supplementary Information), only formation of the VV-dipeptide RNA conjugate, derived from the thermodynamically more stable duplex, was observed. Finally, we mixed two RNA donor strands of identical length (7-mer). The first contained a 5′-mm5U and the second a 5′-mm5U, together with two mismatches. We added this mixture to an RNA acceptor strand (11-mer: 2c) with a 3′-mm5U nucleotide (Fig. 5b, right). In this experiment, exclusive formation of the LV-dipeptide was found, generated from the fully complementary strands and thus the more stable duplex. Collectively, these results suggest that full complementarity is needed for efficient peptide synthesis.

**Stepwise growth of peptides on RNA**

We finally investigated whether one-pot stepwise growth of a peptide on RNA is possible (Fig. 5c). To increase the stability of the RNA towards phosphodiester hydrolysis, as needed for this experiment, we used the RNA acceptor strand 2g, in which the contemporary canonical bases were replaced by the non-canonical 2′-Ome nucleotides: A_m, C_m, G_m and U_m. The strand 2g was equipped with an additional 3′-mm5U

---

**Multiple peptide growth on RNA**

We next investigated whether peptide growth is possible at different RNA positions simultaneously. To this end, we examined the simultaneous binding of different donor strands to one or two acceptor strands. We hybridized two donor strands (7-mer: 5′-m6g6A-RNA-3′ and 10-mer: 5′-mm5U-RNA-3′) to a single RNA acceptor strand (21-mer) with a central mm5U and a 3′ terminal mm5U (Fig. 5a, left). On activation of the carboxylic acids, a GG-dipeptide was synthesized in the centre of the RNA, whereas a valine amino acid was attached to the 3′ end of the acceptor strand. In a different experiment, we hybridized an RNA donor strand (22-mer), containing both a 3′-m6g6A and a 5′-mm5U, to two different acceptor RNAs, containing a central mm5U (21-mer) and a 3′ terminal mm5U (11-mer) (Fig. 5a, right). On activation, we observed formation of a central GV- and a terminal YY-dipeptide.
structures, such as ggg 6A.

Fig. 4 | Capture of peptides by nitrosated

For the experiment we used the same amount of donor strand for all coupling steps and performed filtration steps to remove remaining activator. After two couplings, two urea cleavages and two filtrations, we observed, by high-performance liquid chromatography (HPLC) analysis, the presence of the product 3'-gmm 5U-RNA-5' 7g (Fig. 5c, left). The circumvented material consuming isolation steps

Fig. 3 | Growth of longer peptide structures on RNA. a. Scheme for the reaction of 1a (5'-m6g6A-RNA-3') with 2c (3'-gmm 5U-RNA-5') including coupling (1) and cleavage (2). b. Coupling reactions between 1a and RNA–peptide acceptor strands using EDC/Sulfo-NHS (see reaction condition a in Fig. 2) and cleavage reactions of the coupled compounds (100 mM acetate buffer pH 4, 100 mM NaCl, 90 °C, 6 h). MALDI-TOF data (negative mode) of the isolated products are given.

Fig. 4 | Capture of peptides by nitrosated N¼-methyl urea adenosine for fragment condensation. a. Prebiotically plausible formation of peptide l-A structures, such as ggg ggg 6A. b. Coupling reactions between RNA–peptide conjugates using EDC/Sulfo-NHS (see reaction condition a in Fig. 2) and cleavage reactions of the coupled compounds (see reaction conditions in Fig. 3). HPLC chromatograms show the crude mixtures of the coupling reactions. The RNA signals are coloured: donor in blue; acceptor in red; and hairpin–type intermediate in purple. MALDI-TOF data (negative mode) are shown for the isolated products, together with the 5'-m6A-RNA-3' strand 4 and the hydantoin side product (cyclic form) in the case indicated.
The plausible formation of catalytically competent and self-replicating RNA structures without the aid of proteins is one of the major challenges for the model of the RNA world. It is difficult to imagine how an RNA world with complex RNA molecules could have emerged without the help of proteins and it is hard to envision how such an RNA world transitions into the modern dualistic RNA and protein world, in which RNA predominantly encodes information whereas proteins are the key catalysts of life.

We found that non-canonical vestigial nucleosides, which are key components of contemporary RNAs, are able to equip RNA with the ability to self-decorate with peptides. This creates chimeric structures, in which both chemical entities can co-evolve in a covalently connected form, generating gradually more and more sophisticated and complex RNA–peptide structures. Although, in this study, we observe peptide coupling on RNA in good yields, the efficiency will certainly improve if we allow optimization of the structures and sequences of the RNA–peptides by chemical evolution. The simultaneous presence of the chemical functionalities of RNA and amino acids certainly increases the chance of generating catalytically competent structures. The stabilization of RNA by incorporation of 2′-OMe nucleotides significantly improved the urea cleavage yield.

Interestingly, in the coupling step we observed large differences in the rate constants, which suggests that our system has the potential to preferentially generate certain peptides. We also found that peptides can simultaneously grow at multiple sites on RNA on the basis of rules determined by sequence complementarity, which is the indispensable requirement for efficient peptide growth.

(Extended Data Fig. 2) enabled us to obtain the product in an overall yield of about 18%. A final, third coupling reaction with the 5′-m6gggA donor strand 1g furnished the FGG-hairpin intermediate 8g in approximately 10% overall yield (Fig. 5c, right).

We next studied fragment condensation with the 5′-m6gggA-RNA-3′ donor strand and the complementary 3′-agggmmU-RNA-5′ acceptor strand, consisting only of 2′-OMe nucleotides. Here, coupling with approximately 50% and urea cleavage with approximately 85% generated the product 3′-agggagmmU-RNA-5′, together with some of the hydantoin side product (Supplementary Information). Together these data show that, with the help of 2′-OMe nucleotides, peptides can grow on RNA in a stepwise fashion and via fragment condensation to generate higher complexity.

**Discussion**

The plausible formation of catalytically competent and self-replicating RNA structures without the aid of proteins is one of the major challenges for the model of the RNA world. It is difficult to imagine how an RNA world with complex RNA molecules could have emerged without the help of proteins and it is hard to envision how such an RNA world...
All these data together support the idea that non-canonical vestigial nucleosides in RNA have the potential to create peptide self-decorating RNAs and hence an RNA–peptide world. The formed RNA–peptide chimeras are comparatively stable, and so it is conceivable that some of these structures learned, at some point, to activate amino acids by adenylation and to transfer them onto the ribose OH groups to capture the reactivity in structures that were large and hydrophobic enough to exclude water. This would then have been the transition from the non-canonical nucleoside-based RNA–peptide world to the ribosome-centred translational process that is a hallmark of all life on Earth today.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-022-04676-3.
Methods

General method for the peptide coupling reactions
The RNA donor and acceptor strands (1:1 ratio, 5 nmol of each strand) were annealed with NaCl (5 μl from a 1 M aqueous solution) by heating at 95 °C for 4 min, followed by cooling down slowly to room temperature. After that, MES buffer pH 6 (25 μl from a 400 mM aqueous solution) and NaCl (5 μl from a 1 M aqueous solution) were added to the oligonucleotide solution. Finally, carboxylic acid or nitrile activator/s (10 μl of each component from a 500 mM aqueous solution) and water (100 μl of total reaction volume) were added to the solution mixture. The peptide coupling reaction was incubated at 25 °C for 24 h. The crude reaction mixtures were analysed by HPLC and MALDI-TOF mass spectrometry.

General method for the urea cleavage reactions
The hairpin-type intermediate (0.5 nmol) was diluted with MES buffer pH 6 or acetate buffer pH 4 (12.5 μl from a 400 mM aqueous solution), NaCl (5 μl from a 1 M aqueous solution) and water (50 μl of total reaction volume). The urea cleavage reaction was incubated at 60–90 °C at different time intervals. The crude reaction mixtures were analysed by HPLC and MALDI-TOF mass spectrometry.

Data availability
The data that support the findings of this study are available within the paper and its Supplementary Information.

Acknowledgements
We thank the Deutsche Forschungsgemeinschaft for supporting this research through the DFG grants: CA275/11-3 (ID: 326039064), CRC1309 (ID: 325871075, A4), CRC1032 (ID: 201269156, A5) and CRC1361 (ID: 393547639, P2). We thank the Volkswagen Foundation for funding this research (grant Ev0Rib). This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 741912 (EFRI) and under the Marie Skłodowska-Curie grant agreement no. 881381 (Nature-ETN). L.E. thanks the Alexander von Humboldt Foundation for a postdoctoral fellowship (ESFP1214218 HFST-P).

Author contributions
F.M., L.E., F.X. and E.W. synthesized the modified phosphoramidites and RNA strands and performed the peptide coupling and urea cleavage experiments. M.N. synthesized RNA donor strands and performed preliminary experiments. T.A. refined and developed mechanistic concepts and performed initial proof-of-principle studies. C.-Y.C. and A.P. synthesized modified phosphoramidites. T.C. conceived the project and directed the research. All authors contributed to the analysis of the results and writing of the manuscript.

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/s41586-022-04676-3.
Correspondence and requests for materials should be addressed to Thomas Carell.
Peer review information Nature thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.
Reprints and permissions information is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | Analytical data of the growth of longer peptides on RNA. a, HPLC chromatograms show the crude mixtures of the coupling reactions (100 mM MES buffer pH 6, 100 mM NaCl, 50 mM EDC/Sulfo-NHS, 25 °C, 24 h) between 5'-m6g6A-RNA-3' and RNA-peptide acceptor strands. b, MALDI-TOF mass spectra (negative mode) are shown for the isolated products obtained after the cleavage reactions (100 mM acetate buffer pH 4, 100 mM NaCl, 90 °C, 6 h) of the coupled compounds. In the HPLCs, the RNA strands are coloured: donor in blue; acceptor in red and hairpin-type intermediate in purple.
Extended Data Fig. 2 | RNA-peptide synthesis cycles using a 2′-OMe acceptor strand.  

**a**, Two RNA-peptide synthesis cycles in which the product of each step was separated and added into the next reaction (coupling conditions: 100 mM MES buffer pH 6, 100 mM NaCl, 50 mM DMTMM•Cl, 25 °C, 24 h; cleavage conditions: 100 mM acetate buffer pH 4, 100 mM NaCl, 90 °C, 24 h).

**b**, HPLC chromatograms show the crude mixtures of the coupling and cleavage reactions. In the HPLCs, peaks of RNA strands are coloured as in the reaction scheme. The product 3′-ggmmnm5U-RNA-5′ was obtained in ≈ 6% overall yield.