



ARTICLE

DOI: 10.1038/s41467-017-00463-1

OPEN

# Emergence of native peptide sequences in prebiotic replication networks

Jayanta Nanda <sup>1,2</sup>, Boris Rubinov<sup>1</sup>, Denis Ivnitski<sup>1</sup>, Rakesh Mukherjee<sup>1</sup>, Elina Shtelman<sup>1</sup>, Yair Motro<sup>1</sup>, Yifat Miller<sup>1</sup>, Nathaniel Wagner<sup>1</sup>, Rivka Cohen-Luria<sup>1</sup> & Gonen Ashkenasy <sup>1</sup>

Biopolymer syntheses in living cells are perfected by an elaborate error correction machinery, which was not applicable during polymerization on early Earth. Scientists are consequently striving to identify mechanisms by which functional polymers were selected and further amplified from complex prebiotic mixtures. Here we show the instrumental role of non-enzymatic replication in the enrichment of certain product(s). To this end, we analyzed a complex web of reactions in  $\beta$ -sheet peptide networks, focusing on the formation of specific intermediate compounds and template-assisted replication. Remarkably, we find that the formation of several products in a mixture is not critically harmful, since efficient and selective template-assisted reactions serve as a backbone correction mechanism, namely, for keeping the concentration of the peptide containing the native backbone equal to, or even higher than, the concentrations of the other products. We suggest that these findings may shed light on molecular evolution processes that led to current biology.

<sup>1</sup>Department of Chemistry, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel. <sup>2</sup>Present address: Department of Chemistry, Indian Institute of Engineering Science and Technology, Shibpur, Howrah 711103, India. Jayanta Nanda and Boris Rubinov contributed equally to this work. Correspondence and requests for materials should be addressed to G.A. (email: [gonenash@bgu.ac.il](mailto:gonenash@bgu.ac.il))

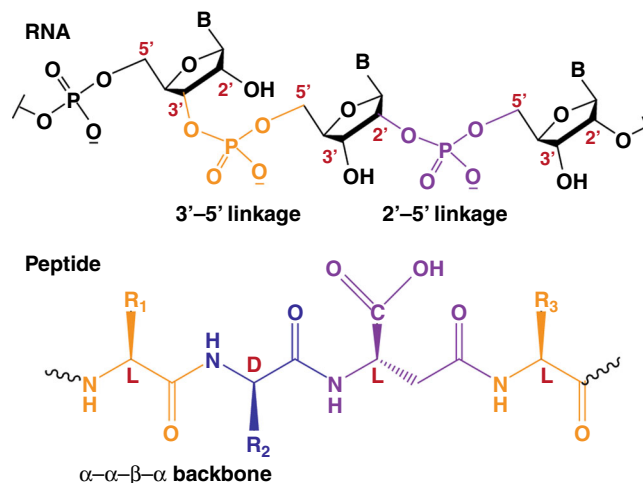
Cells use a surprisingly small number of molecules, primarily nucleic acids, proteins, fatty acids, and sugars, to perform an immense repertoire of functions. Even more remarkable is the phenomenon that only small subsets of these molecular families, containing distinctive backbone connectivity and chiral arrangements, have evolved to be operative. For example, today's DNA and RNA molecules contain the D form of (deoxy)ribose attached solely at the 3'-5' positions of successive backbone units, and proteins are made up almost entirely of L type  $\alpha$ -amino acids, connected through their  $\alpha$ -carboxy and  $\alpha$ -amino termini. The evolution of these unique polymer sequences—in preference to others—from a huge prebiotic configurational space remains a mystery<sup>1–3</sup>, since in the absence of highly evolved genetic machinery or enzyme catalysis most polymer condensation reactions result in mixtures containing a large number of products, typically isomers of the “desired” molecule, each produced in low yield.

Many experiments directed at spontaneous polymerization of biomolecules have indeed yielded mixtures containing RNA molecules with heterogeneous backbones<sup>4–6</sup>, or peptides with mixed chirality<sup>7–10</sup>, in some cases elongated at their side chains (Fig. 1). Elucidating the mechanisms by which functional prebiotic polymers were selected and further amplified may offer ways to resolve this “prebiotic chemist's nightmare”<sup>3, 11, 12</sup>. Achieving this aim may explain the early molecular evolution processes that led to biology as we know it today.

In the context of the emergence of an RNA world, researchers have suggested that RNA was preceded by transient informational oligomers that were compatible with pre-biological pathways, which, in turn, could support the transition to RNA itself<sup>1, 13</sup>. Based on RNA replication experiments, it has also been proposed that the difficulty presented by chemical heterogeneity to the understanding of early evolution has been exaggerated, since some chemical heterogeneity is actually tolerable in polymerization and replication processes<sup>5, 14</sup>. With respect to chirality, several studies have shown that even minute differences in abundance (i.e., enantiomeric excess) can be amplified when various self-assembly processes, including crystallization, occur<sup>8, 15–20</sup>. In addition, it has been postulated that the delivery of enantio-rich amino acids from outer space was the origin for homo-chirality<sup>21, 22</sup>. Active polymerization-racemization processes<sup>18</sup>, and differential breakdown of incorrect product(s) followed by channeling resources into the correct product<sup>23–25</sup>, comprise additional putative pathways for the formation of chiral peptide oligomers. Furthermore, regio-enantioselective in the steps of chain elongation by growing  $\beta$ -sheet peptides yielded isotactic sequences<sup>8, 26, 27</sup>, albeit often forming racemic aggregates. Interestingly, a recent study (with RNA) has shown that both cross-catalytic and anti-cross catalytic of one handedness or the other are tolerable<sup>28</sup>.

Herein, we put forward another selection mechanism relevant to early chemical evolution, since we probe the instrumental significance of non-enzymatic replication for the amplification and fixation of certain products within heterogeneous prebiotic mixtures. In particular, we demonstrate that  $\beta$ -sheet-driven peptide replication can be used for significant enrichment of a native condensation product in preference to some of its regio-isomers and stereo-isomers and hence for correction of the evolutionary drift towards non-functional heterogeneous mixtures.

The involvement of peptides in early evolution has been advocated several times<sup>29–33</sup> and has also been suggested as a complementary or alternative route to the RNA world, leading to primitive life-like functions<sup>34–39</sup>. Recent studies have indeed revealed multiple assembly pathways, yielding amphiphilic peptide fibrils or nanotubes and, consequently, the emergence of



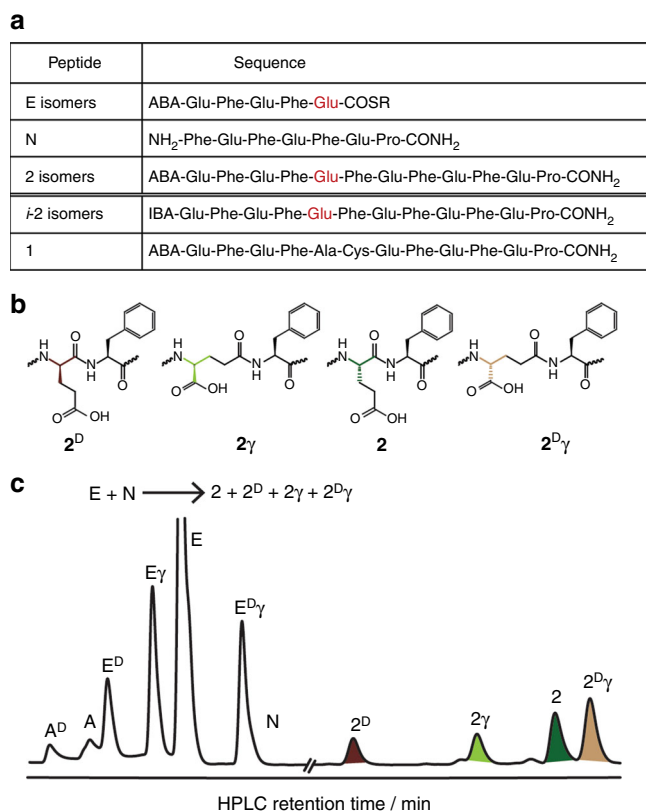
**Fig. 1** illustration of RNA and peptide molecules with heterogeneous backbones

new functions, such as catalysis, binding of nucleic acids and charge transfer<sup>40–47</sup>. Of particular relevance to studying the origin of life is the search for short oligopeptides, spontaneously formed and evolved as auto-catalysts<sup>48–50</sup>. Accordingly, we have previously studied the replication of a simple amphiphilic peptide (**1**, Fig. 2a) that, when forming  $\beta$ -sheet architectures, can efficiently accelerate its own formation from a solution containing the appropriate precursors<sup>48, 51, 52</sup>. The current study takes a different direction, in that here we applied original design and experimental conditions for inducing spontaneous evolution of a mixture containing isomeric compounds, each presenting a different phenotype following self-assembly. We thoroughly analyzed the complex web of reactions in this network via experiments and simulation with the aim to elucidate the role of the initial conditions (seeded precursors), the stability of the intermediate compounds yielding different products, and template-assisted replication. Remarkably, we find that when the replicating molecule is also the condensation kinetic product, modest replication exerted, but when the network reaction is initiated with materials leading primarily to other products with a non-homogenous backbone, efficient correction takes place by replication and enrichment of the native peptide at the expense of other network components. We conclude this paper with a brief discussion of some implications of our findings, highlighting plausible relevant pathways in early molecular evolution.

## Results

### Spontaneous evolution of a peptide network via glutamic acid ligation.

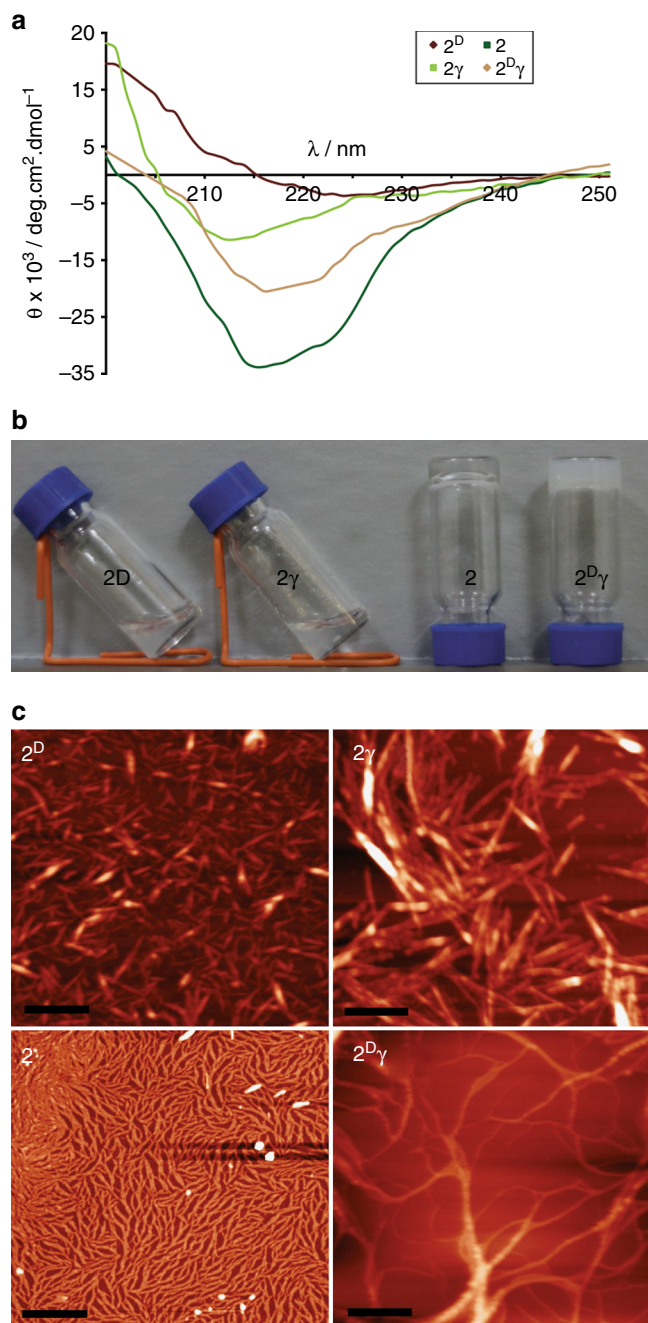
This study focuses on the unique non-enzymatic replication that takes place in a small network made up of peptide **2** and its isomers (Fig. 2). The sequence of this peptide contains only repeating Glu-Phe residue dyads and capping groups. It is, accordingly, simpler than peptide **1** and potentially more prebiotically relevant. The main differences between the replication of **2** and that observed earlier for **1** originate from the mechanism of their precursors' ligation. The replication of **1** was driven by the efficient and selective “native chemical ligation” reaction<sup>53</sup>, propelled by an alanine thioester electrophile and a dedicated Cys residue at the nucleophile N terminus. The ligation of **2**, in contrast, is achieved by strong activation of the electrophile **E** as a glutamic acid thioester, while the free Phe amine of nucleophile **N** remains non-activated (Fig. 2). As a result, the formation of **2** takes place together with the appearance of side products, structural isomers of **2** (Fig. 2),



**Fig. 2** Peptides utilized for the current study. **a** Peptide names and their sequences. Sequences with a glutamic acid residue marked in red represent the four isomeric species discussed in the text and shown in **b**. ABA 4-acetamidobenzoate, IBA 4-iso-butylamide benzoate, and SR 4-mercaptophenylacetic acid (MPAA). **b** Absolute configuration and backbone connectivity of the ligation site Glu residue in **2** and its isomers **2<sup>D</sup>**, **2<sup>γ</sup>**, and **2<sup>Dγ</sup>**. **c** Representative HPLC chromatogram (detected at 270 nm) obtained during a ligation reaction initiated by 200 μM **E** and 300 μM **N**. Note that schemes detailing the ligation reaction steps are given in Fig. 5 and Supplementary Fig. 13. The chromatogram allows identification of all the species forming the reaction network, including the products **2**, **2<sup>D</sup>**, **2<sup>γ</sup>**, and **2<sup>Dγ</sup>**, the isomeric electrophile molecules (**E<sup>D</sup>**, **E<sup>γ</sup>**, and **E<sup>Dγ</sup>**), and the intermediate anhydride species (**A** and **A<sup>D</sup>**, Supplementary Fig. 1a). The electrophile and product isomers were all analyzed by LCMS, and their elution times were compared to those of the respective originally synthesized compounds (Supplementary Fig. 1b). **A** and **A<sup>D</sup>** peptides were identified by LCMS. Note that peptide **N** could not be well detected, since it lacks ABA labeling, and that two electrophile hydrolysis products, eluted at early times, are not shown (see SI Supplementary Methods)

as was also observed previously in similar Asp-thioester- or Glu-thioester-based ligation reactions<sup>54, 55</sup>. The three isomers differ from **2** in terms of epimerization at the ligation site Glu residue (**2<sup>D</sup>** product) and/or in terms of non-canonical connectivity to the next Phe residue via the Glu γ-carbonyl (**2<sup>Dγ</sup>** and **2<sup>γ</sup>**, respectively). Remarkably, we observed significant production of these isomers, which implies that the epimerization and isomerization processes are fast in comparison with the strategic ligation.

The substantially different elution times in the high-performance liquid chromatography (HPLC) product region (Fig. 2c) mean that the hydrophobicity of peptide **2** and its three isomers was markedly different. Remarkably, further analysis using a set of common β-sheet characterization tools (Fig. 3) shows that even small structural changes next to the ligation site lead to formation of isomers (“prebiotic mutants”)



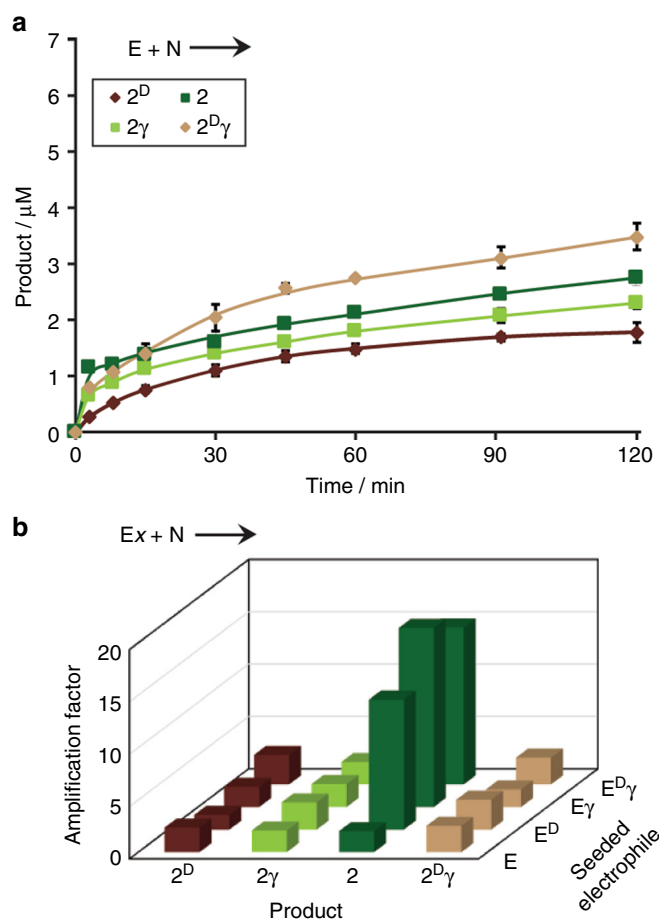
**Fig. 3** Structural characterization of the four isomeric peptides **2**, **2<sup>D</sup>**, **2<sup>γ</sup>**, and **2<sup>Dγ</sup>** formed spontaneously during the ligation reaction. **a** CD spectra of 50 ± 2 μM peptides at pH 7. **b** Gelation analysis for 5.6 mM aqueous solutions at pH 7. **c** AFM images of the self-assembly products formed in 50 ± 2 μM solutions. Samples were drop cast on a mica surface after 10 min of sonication, followed by 30 min of equilibration in MOPS buffer, pH 7. All scale bars are 1 μm

exhibiting different phenotypes. Inspection of the 216-nm minima in the CD spectra of the isomers (Fig. 3a) indicates efficient assembly into β-sheets for **2**, and to a lesser extent for **2<sup>Dγ</sup>**, while a practically unstructured arrangement was detected for **2<sup>D</sup>** and **2<sup>γ</sup>**. A qualitative gelation study yielded similar results in that, at a designated concentration, peptide **2** formed a transparent gel phase, whereas **2<sup>Dγ</sup>** gelled into an opaque phase, and **2<sup>D</sup>** and **2<sup>γ</sup>** remained soluble (Fig. 3b). A further indication of this assembly trend was evident in atomic force microscopy

(AFM) images (Fig. 3c) showing that peptide **2** formed denser fibers, and peptide  $2^{\text{D}\gamma}$  formed longer fibers, vs.  $2^{\text{D}}$  and  $2\gamma$ . Since peptide **2** possesses an amphiphilic sequence and a native homogenous backbone, its strain-free assembly into  $\beta$ -sheets is intuitively clear. From time-dependent cryo-transmission electron microscopy (TEM) measurements (Supplementary Figs. 2 and 3), we find that the dynamic assembly of **2** proceeds via a pathway similar to that observed earlier for peptide **1**<sup>51</sup>, namely, first the formation of elongated fibril structures and then, at longer times ( $t \geq 60$  min), rearrangement into hollow tubes. Our recent molecular dynamic simulations revealed that well-defined fibrils of **2** are formed by peptide bilayers in which the monomers are oriented anti-parallel to each other, both within the same layer and in opposite layers<sup>52</sup>. A logical explanation for the variability in assembly propensity of the other isomers may be obtained from simple modeling of their backbone orientations (Supplementary Fig. 4). Such an analysis did indeed show that native bond angles within the singly isomerized isomers  $2^{\text{D}}$  and  $2\gamma$  would force the incorporation of the hydrophilic Glu residue side chains into the fibril hydrophobic core, with overall destabilization of the fibril structure. In contrast, the backbone of  $2^{\text{D}\gamma}$ , isomerized twice with respect to **2**, can tolerate the assembly of stable fibers, since their hydrophobic core is occupied only by Phe residues, with the mutated Glu residue presenting its  $\alpha$ -carboxylate toward the water phase.

**Mechanisms of the network reaction and dynamic isomerization.** We then followed the kinetics of product formation during ligation between peptides **E** and **N** with the aims to expose the mechanistic features that lead to the formation of the four-isomer network and, furthermore, to explain the observed product distribution (Fig. 4a). At the early stage of the reaction (up to 15 min), the direct ligation product, **2**, formed faster than the other products. However, as the reaction progressed,  $2^{\text{D}\gamma}$  formed efficiently and, overall, no strong preference for any one of the products was observed. The HPLC chromatograms (as in Fig. 2c) reveal that during the reaction the thioester electrophile **E** is readily converted into five reactive species, namely, three thioester isomers ( $\text{E}^{\text{D}}$ ,  $\text{E}\gamma$ , and  $\text{E}^{\text{D}\gamma}$ ) and two glutaric anhydride stereo-isomers (**A** and  $\text{A}^{\text{D}}$ ; Supplementary Figs. 1a and 5e). Direct ligation of the electrophiles with **N** leads to the formation of the four products described above. We note that the two anhydrides **A** and  $\text{A}^{\text{D}}$  constantly convert from one to the other. Consequently, these species play a central role in spontaneous mutation, as they form and epimerize very fast and can also ligate with **N**, via their  $\alpha$ -carbonyl or  $\gamma$ -carbonyl carbons, to form all four products, or alternatively, react back with MPAA to form the four thioester isomers (Fig. 5).

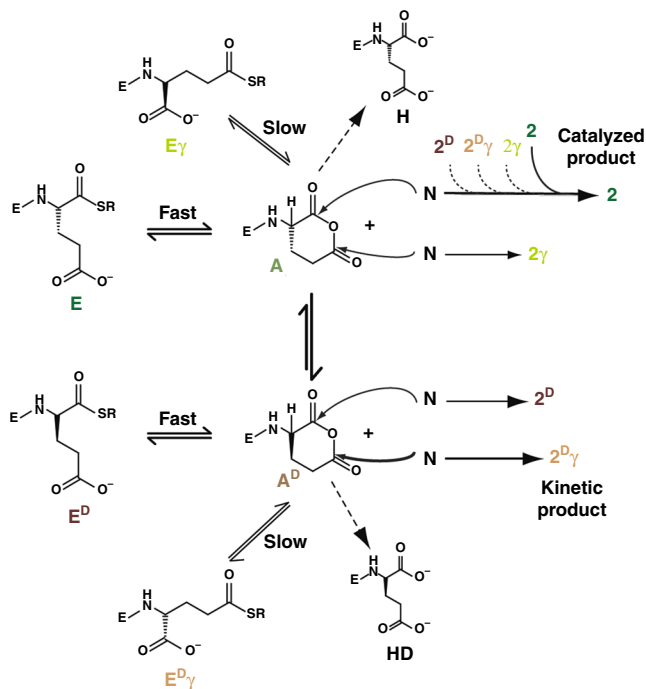
Further insight into the network reaction mechanisms was obtained by following the time-dependent growth of the four isomeric products in reactions initiated with **N** and each of the four (pre-synthesized) electrophiles separately (Supplementary Fig. 5). The results show more efficient (total) product formation for reactions initiated with the  $\alpha$ -carboxy thioesters **E** or  $\text{E}^{\text{D}}$ , irrespective of their end terminal Glu chirality, than for the  $\gamma$ -carboxy thioesters  $\text{E}\gamma$  or  $\text{E}^{\text{D}\gamma}$ . This observation is supported by the observed easier formation of the anhydrides (**A** and  $\text{A}^{\text{D}}$ ) from the  $\alpha$ -carboxy thioesters than from the  $\gamma$ -carboxy thioesters (Supplementary Figs. 5e and 6). In reactions initiated with  $\text{E}^{\text{D}}$  or  $\text{E}\gamma$ , similarly to the above-discussed reaction initiated with **E**, the kinetic product from direct attack on the seeded thioester formed rapidly in the early stages, but its relative abundance diminished at longer times when  $2^{\text{D}\gamma}$  had become the most abundant product in the mixture. The putative reaction



**Fig. 4** Product formation in template-free reactions. **a** Time-dependent formation of the four isomeric products in a reaction between  $200 \mu\text{M}$  **E** and  $300 \mu\text{M}$  **N**. **b** Amplification factors observed in the formation of the isomeric peptides in reactions initiated with **N** and each of the four electrophiles, **Ex**, separately. The amplification factors were calculated from ratios of product formation at 120 min and 15 min for reactions initiated with the  $\alpha$ -carboxy thioesters (**E** or  $\text{E}^{\text{D}}$ ; Supplementary Fig. 5), and the ratios at 240 and 90 min for reactions initiated with the  $\gamma$ -carboxy thioesters ( $\text{E}\gamma$  and  $\text{E}^{\text{D}\gamma}$ ; Supplementary Fig. 5). All reactions were carried out in MOPS buffer at pH 7. Product concentrations were calculated from the HPLC chromatogram by using an ABA-labeled peptide (ABA-Ala-NH<sub>2</sub>) sample of known concentration as an internal standard

pathways leading to efficient formation of  $2^{\text{D}\gamma}$  (Fig. 5 and Supplementary Fig. 5) were reinforced by quantitative computational analysis of the intermediate anhydride structures (Supplementary Fig. 7), which revealed that  $\text{A}^{\text{D}}$  is more stable than **A** and that the attack on the  $\gamma$ -carbonyl of  $\text{A}^{\text{D}}$  is less sterically hindered than the attack on its  $\alpha$ -carbonyl. Most importantly, the formation of **2** was very slow in the early stages of all reactions initiated with the non-native electrophiles ( $\text{E}^{\text{D}}$ ,  $\text{E}\gamma$ , and  $\text{E}^{\text{D}\gamma}$ ), but it was significantly enhanced in later stages, with **2** reaching a final abundance similar to that of the other products (Supplementary Fig. 5). Figure 4b emphasizes that **2** was the only compound in the mixture amplified in a non-linear fashion, providing strong evidence for a error correction mechanism, via autocatalytic or cross-catalytic template-assisted replication, as further elaborated below.

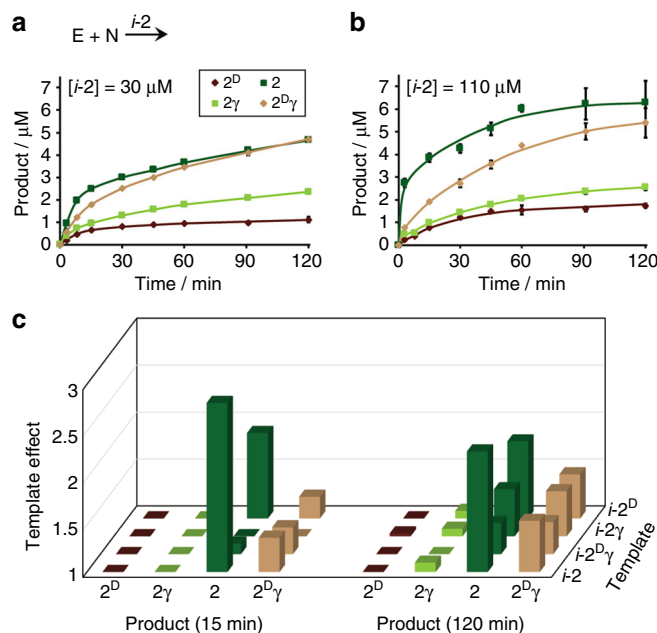
**Template-directed self-replication and cross-catalytic backbone correction.** The high-sequence similarity between peptide **2** and the replicator peptide **1** as well as the tendency of both molecules



**Fig. 5** General mechanism of a network reaction initiated by **N** and any one of the electrophiles **Ex**. The scheme illustrates the central role of the glutaric anhydride intermediates **A** and **A<sup>D</sup>** in the dynamic exchange between the starting material and the product isomers, as well as the kinetically favored pathways leading to product **2<sup>D</sup> $\gamma$**  and the catalyzed pathways leading to product **2** containing the native backbone

to form assemblies of similar  $\beta$ -sheet architectures strongly suggest that **2** (and/or its isomers) can also replicate. Indeed, equilibration of **2** in the presence of excess amounts of **E** and **N** peptides stabilized short-lived  $\beta$ -plate and fibril structures (cryo-TEM in Supplementary Fig. 3), potentially facilitating self-replication activity not hampered by transition to inactive nanotubes<sup>51</sup>. During a network reaction initiated by an **E** type isomer and **N**, such replication can take place either auto-catalytically, by a product formed in situ enhancing the formation of an identical molecule, or cross-catalytically with one isomer product templating the formation of another one.

The template-assisted effects in reactions initiated by **E** and **N** were investigated by seeding in the external templates *i-2*, *i-2<sup>D</sup>*, *i-2 $\gamma$* , or *i-2<sup>D</sup> $\gamma$*  (Fig. 2a) very close analogs of **2**, **2<sup>D</sup>**, **2 $\gamma$** , or **2<sup>D</sup> $\gamma$** , respectively, assembled into identical fibril structures (Supplementary Fig. 8). Use of *i-2* as a template enhanced the initial production rate of **2** by up to threefold as compared with the template-free reaction (Fig. 6). This modest increase in rate correlated monotonically with increasing the amount of seeded template, and at high template concentrations **2** became the most abundant product in the network (Fig. 6b). A smaller increase in the production rate (~50%) was observed for **2<sup>D</sup> $\gamma$** , and the growth of the other two products (**2<sup>D</sup>** and **2 $\gamma$** ) remained unchanged or even slightly reduced, probably due to fast and guided consumption of the common intermediates, **A** and **A<sup>D</sup>**, by **N**. The selective enhancement of product **2** formation by *i-2* may be explained by the perfect matching between the template fibrils as catalysts and the product. Seeding experiments initiated with *i-2<sup>D</sup> $\gamma$* , *i-2 $\gamma$* , or *i-2<sup>D</sup>* induced an increase in rate of production of **2** (up to twofold in case of seeding with *i-2<sup>D</sup>*), and lower or equal rate enhancement of the respective autocatalytic product (Fig. 6c). This observation clearly advocates for that



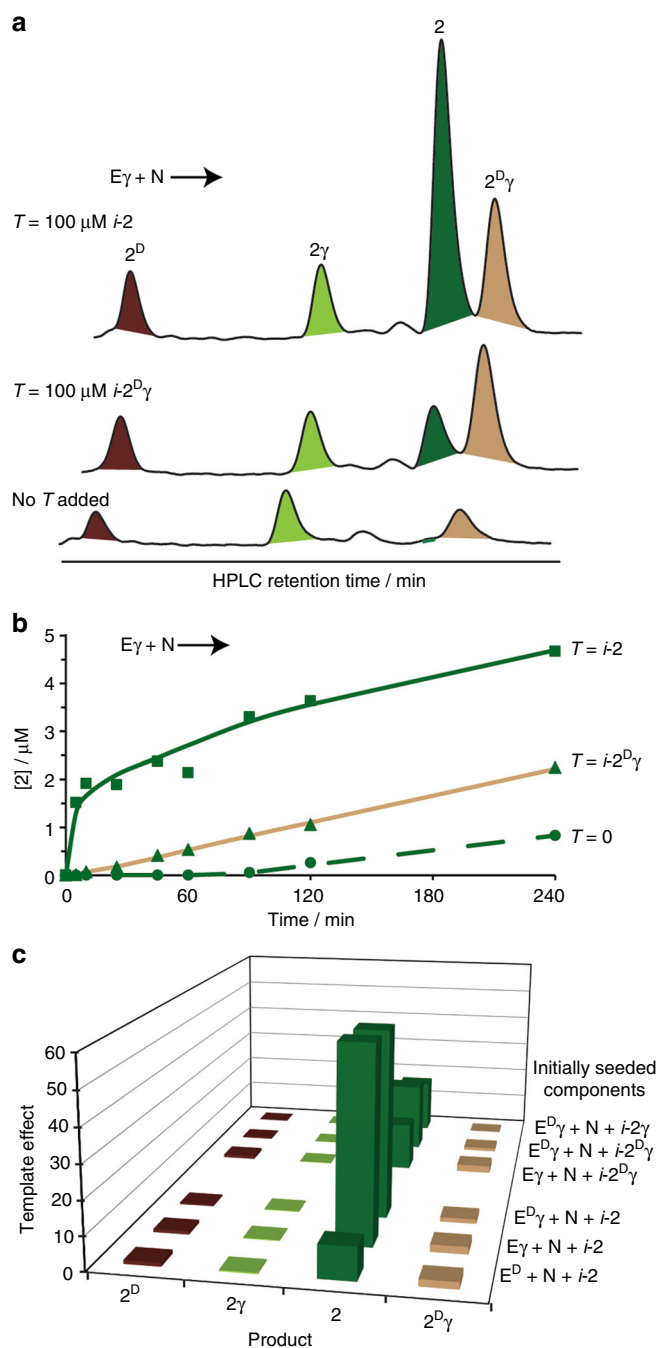
**Fig. 6** Template-assisted product formation in reactions between **E** and **N**. **a**, **b** Time-dependent product formation in reactions seeded with different *i-2* concentrations. **c** Bar diagram showing the template effect on the four products in the early (15 min) and late (120 min) stages of reactions seeded with  $100 \pm 10 \mu\text{M}$  of each one of the respective templates, separately. Template effect at a particular time is the ratio of product formation in a template-assisted reaction (this figure **b** and Supplementary Fig. 9) to that in the template-free reaction (Fig. 4)

primitive correction mechanism drives the network progress, since each of the non-native isomers (**2<sup>D</sup> $\gamma$** , **2 $\gamma$** , and **2<sup>D</sup>**) is apparently less effective as an autocatalyst than as a cross-catalyst producing the native sequence **2**.

The above-discussed template-free reactions initiated by **N** and each of the non-native thioesters, **E<sup>D</sup>**, **E $\gamma$** , and **E<sup>D</sup> $\gamma$**  (Fig. 4b), exhibited non-linear formation of peptide **2**, strongly suggesting that efficient template-assisted processes facilitate its growth in such mixtures. This hypothesis was directly probed by performing a comprehensive set of experiments, in which the network reaction was initiated with **N** and one of the non-native thioesters in the presence of one of the four product isomer analogs (*i-2*, *i-2<sup>D</sup>*, *i-2 $\gamma$* , or *i-2<sup>D</sup> $\gamma$* ; Fig. 7 and Supplementary Figs. 10–12). Marked rate enhancement (up to 60-fold) in the formation of **2** was observed in reactions seeded with *i-2* vs. the template-free reactions. The efficient “quasi” self-replication of the native peptide **2** diverted the original reaction pathway, leading to the formation of **2** as a major product in the mixture (Fig. 7a). Remarkably, the formation of **2** was selectively enhanced, while the formation rates of the three other isomers were similar, or only slightly higher, vs. those observed in the template-free reactions (up to threefold initial rate enhancement observed for **2<sup>D</sup> $\gamma$** ). Interestingly, seeding in any of the other templates (*i-2<sup>D</sup>*, *i-2 $\gamma$* , or *i-2<sup>D</sup> $\gamma$* ; Fig. 7) also resulted in rate enhancement of the production of **2**, albeit with lower efficiency, revealing a second latent mechanism for correction via cross-catalysis.

#### Simulation analysis of the replication network topology.

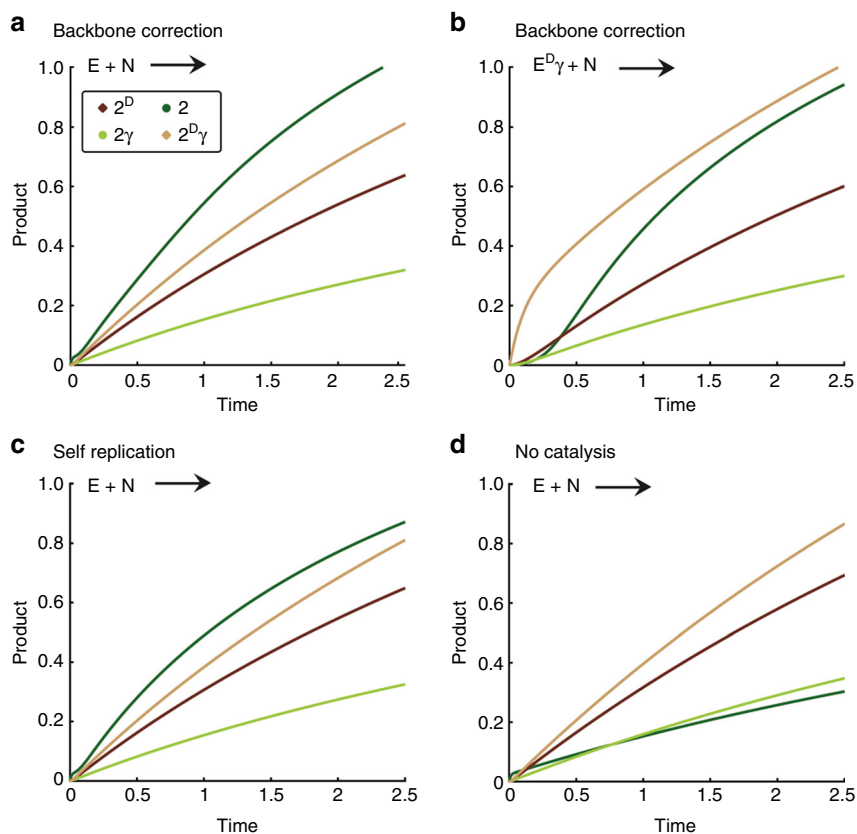
The various product distribution patterns observed in the template-free and template-seeded network reactions allow us to propose an overall network topology facilitating the synthesis of



**Fig. 7** Template-assisted product formation in reactions initiated with **N** and each of the non-native electrophile thioesters ( $E^D$ ,  $E\gamma$ , and  $E^D\gamma$ ) separately. **a** HPLC chromatograms comparing product formation at 90 min in template-free and template-seeded reactions initiated with  $E\gamma$  and **N** (seeding with  $i-2$  or  $i-2^{D\gamma}$ ). **b** Time-dependent formation of the native product **2** in template-free and template-seeded reactions initiated with  $E\gamma$  and **N**. **c** Bar diagram representing the template effects in reactions initiated with **N** and each of the non-native electrophile thioesters, separately, and seeded each time with a different template ( $100 \pm 10 \mu\text{M}$ ). Template effects were calculated from the respective product formation data at 15 min for reactions initiated with  $E^D$ , and at 90 min for reactions initiated with  $E\gamma$  or  $E^D\gamma$  in template-free (Supplementary Fig. 5) and template-assisted (**b** of this figure and Supplementary Figs. 10–12) reactions

four isomeric peptides and, in particular, the amplification of product **2** possessing the native backbone structure (Fig. 5). While the most efficient template-free reaction pathways lead to formation of the  $2^{D\gamma}$  product, template-assisted reactions—both self-replication and cross-catalysis—enhance the formation of **2**. Remarkably, as typically seen for network-level correction processes, the cross catalytic reactions, in which the non-native products ( $2^D$ ,  $2\gamma$ , or  $2^{D\gamma}$ ) catalyze peptide **2**, were found to be more efficient than the autocatalytic reactions leading to their own formation. When the reaction is initiated with the non-native isomer electrophiles, or when the anhydride analogs of the latter are formed in situ, these efficient and selective template-assisted reactions play a role in keeping the concentration of **2** equal to, or even higher than, the concentrations of the other products. Note that a similar scenario, in which condensation reactions of pre-equilibrating starting materials and intermediates are channeled through self-replication (only) to form one major product, has been described using organic an abiotic system<sup>56</sup>.

Based on the kinetic model in Fig. 5, we ran a simulation in Matlab, using the set of equations shown in Supplementary Fig. 13. This network analysis makes use of and expands on the model we previously applied in studying the fibril-dependent self-replication of peptide **1**<sup>51</sup>. Accordingly, multiple fibril growth and assembly steps (Eqs. 1–3), and template free (Eqs. 7) and template assisted (Eqs. 8) ligation reactions, account here for the formation of the four template/product isomers, including their cross reactivity when one template facilitates the replication of another one. In addition, since the new system proceeds via formation of the glutaric anhydride analogs (**A** and **A**<sup>D</sup>, Fig. 5), the relevant steps producing, isomerizing, and consuming these anhydrides were introduced as well (Eqs. 4–6). The initial concentrations applied for the simulation directly match the concentrations we used in the experiments, while the assigned kinetic constants were grouped together to simplify the analysis when possible, as explained in the SI (Supplementary Methods). The results of this simple modeling (Fig. 8 and Supplementary Fig. 14) can remarkably reproduce many of the network kinetic features we observed in the experiments. Fig. 8a, b, and Supplementary Fig. 14a, describe the network growth profile when initiated with **N** and either one of the electrophiles  $E\gamma$ , under the conditions we defined as “backbone correction”, emphasizing that both self-replication of **2** and cross catalytic production of **2** by the other isomers take place. As was found in the above-described experiments (Fig. 4), the native product **2** formed relatively fast right from the beginning when the reaction started with **E** and **N** (Fig. 8a), while it was formed very slowly in the early stages of reactions that started with **N** and one of the other isomer electrophiles ( $E^D$ ,  $E\gamma$ , and  $E^D\gamma$ ) and only later on sped up due to efficient replication processes (e.g., Fig. 8b). The fast production of  $2^{D\gamma}$ , due mainly to template-free processes, was also observed in these cases. The simulation can also nicely account for the observed network growth profiles under the same correction reaction conditions when seeded with the various products as templates (Supplementary Fig. 14, panels c–f). Most importantly, we have also utilized the simulation to screen for the network growth kinetics had they been governed by a different set of interactions, e.g., with much more dominant self-replication of **2** (Fig. 8c), or with no template assisted processes taking place at all (Fig. 8d). We find that very efficient self-replication of **2** may also lead to its formation as the most abundant compound, but the observed preference over production of the other isomers is lower. When no template-assisted processes take place, the kinetic product  $2^{D\gamma}$  is the most abundant while **2** is formed very slowly.



**Fig. 8** Simulation results based on the mechanism shown in Fig. 5 and the reaction model given in Supplementary Fig. 13. Experimental initial concentrations were assigned for **E** and **N** (250  $\mu$ M each), and reaction constants were applied as specified in the SI. **a, b** Backbone correction growth mechanism observed when the reactions were initialized with **E** (a) or **E<sup>D</sup> $\gamma$**  (b). **c, d** Alternate growth behavior observed when the network kinetics were governed by different sets of interactions, e.g., with more dominant self-replication **c**, or with no templated assisted processes **d**

## Discussion

The common practice of sequence-specific oligomer synthesis in organic chemistry makes use of very selective coupling reactions and the incorporation of protecting groups into reactive side chains. DNA polymerization in biology is, however, controlled by a different mechanism, namely post replication error-correction, proceeding by removal of mis-incorporated units and their replacement for the correct ones. Ill-defined protein sequences are also doomed by the cell machinery to post-synthesis digestion, enriching the population of functional proteins. Both the synthetic and biological strategies were obviously not applicable to polymer synthesis on early Earth. Scientists were thus looking for other universal mechanisms that could support sequence specific oligomer synthesis, and were accordingly attracted by the idea of post synthesis correction mechanisms. Nevertheless, in the absence of efficient catalysis by nucleases or peptidases, the reversible processes needed for breaking the nucleic acid or peptide backbone, and for recycling their fragments as building blocks in correct synthesis, are very slow. For that matter, dynamic covalent chemistry (DCC) was developed recently, enabling continuous equilibration of building blocks and products, thus leading to selection of the most fit compound(s) in the mixture, depending on the applied environmental conditions, templating molecules and sometimes even function<sup>57, 58</sup>. However, as of now, DCC cannot support reversible chemistry of native peptide bonds and is consequently only partially useful for studying prebiotic protein evolution. We have consequently searched for another systems chemistry approach<sup>59</sup>, whereby the enrichment of one

“master sequence” is a result of a collective network behavior manifested by multiple template-assisted processes. Similar to previously reported results with coiled-coil peptide networks<sup>38</sup>, the synthesis of a native (or wild-type) peptide in our system is significantly enhanced in a small isomer population. Our results hint at an alternative backbone correction mechanism that could have developed within simple chemical networks through the emergence of shorter  $\beta$ -sheet polymers with efficient and selective catalytic behavior. In such a scenario, the formation of several products in a mixture is not critically harmful, since this is accompanied by a slow progression of the production of one molecule at the expense of the others, leading finally to the formation of a dominant functional polymer backbone. In the future, we wish to study the relevance of our approach to replication systems facilitating the condensation of more than two peptide units, and ultimately even the sequence specific addition of amino acids. Naturally, the success in developing such system will depend on further improvement in selectivity toward the desired sequence. Finally, we note that our results are in line with the old notion that the emergence of a primitive replicator was crucial and potentially sufficient for the origin of life further down the road<sup>60</sup>.

**Data availability.** The data that support the findings of this study are available from the corresponding author upon request.

Received: 25 December 2016 Accepted: 30 June 2017

Published online: 05 September 2017

## References

1. Eschenmoser, A. Chemical etiology of nucleic acid structure. *Science* **284**, 2118–2124 (1999).
2. Orgel, L. E. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* **39**, 99–123 (2004).
3. Szostak, J. W. An optimal degree of physical and chemical heterogeneity for the origin of life? *Philos. Trans. R. Soc., B* **366**, 2894–2901 (2011).
4. Monnard, P.-A. & Ziock, H. Eutectic phase in water-ice: a self-assembled environment conducive to metal-catalyzed non-enzymatic RNA polymerization. *Chem. Biodiversity* **5**, 1521–1539 (2008).
5. Engelhart, A. E., Powner, M. W. & Szostak, J. W. Functional RNAs exhibit tolerance for non-heritable 2'-5' versus 3'-5' backbone heterogeneity. *Nat. Chem* **5**, 390–394 (2013).
6. Monnard, P.-A., Kanavarioti, A. & Deamer, D. W. Eutectic phase polymerization of activated ribonucleotide mixtures yields quasi-equimolar incorporation of purine and pyrimidine nucleobases. *J. Am. Chem. Soc.* **125**, 13734–13740 (2003).
7. Leman, L., Orgel, L. & Ghadiri, M. R. Carbonyl sulfide-mediated prebiotic formation of peptides. *Science* **306**, 283–286 (2004).
8. Weissbuch, I., Illos, R. A., Bolbach, G. & Lahav, M. Racemic  $\beta$ -Sheets as templates of relevance to the origin of homochirality of peptides: lessons from crystal chemistry. *Acc. Chem. Res.* **42**, 1128–1140 (2009).
9. Danger, G. et al. 5(4H)-oxazolones as intermediates in the carbodiimide- and cyanamide-promoted peptide activations in aqueous solution. *Angew. Chem., Int. Ed.* **52**, 611–614 (2013).
10. Forsythe, J. G. et al. Ester-mediated amide bond formation driven by Wet-Dry Cycles: a possible path to polypeptides on the prebiotic Earth. *Angew. Chem., Int. Ed.* **54**, 9871–9875 (2015).
11. Pressman, A., Blanco, C. & Chen, I. A. The RNA world as a model system to study the origin of life. *Curr. Biol.* **25**, R953–R963 (2015).
12. Sutherland, J. D. The origin of life - out of the blue. *Angew. Chem., Int. Ed.* **55**, 104–121 (2016).
13. Krishnamurthy, R. On the emergence of RNA. *Isr. J. Chem.* **55**, 837–850 (2015).
14. Trevino, S. G., Zhang, N., Elenko, M. P., Luptak, A. & Szostak, J. W. Evolution of functional nucleic acids in the presence of nonheritable backbone heterogeneity. *Proc. Natl Acad. Sci. USA* **108**, 13492–13497 (2011).
15. Soai, K., Shibata, T., Morioka, H. & Choji, K. Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. *Nature* **378**, 767–768 (1995).
16. Saghatelian, A., Yokobayashi, Y., Soltani, K. & Ghadiri, M. R. A chiroselective peptide replicator. *Nature* **409**, 797–801 (2001).
17. Blackmond, D. Asymmetric autocatalysis and its implications for the origin of homochirality. *Proc. Natl Acad. Sci. USA* **101**, 5732–5736 (2004).
18. Plasson, R., Bersini, H. & Commeyras, A. Recycling frank: spontaneous emergence of homochirality in noncatalytic systems. *Proc. Natl Acad. Sci. USA* **101**, 16733–16738 (2004).
19. McBride, J. M. & Tully, J. C. Did life grind to a start? *Nature* **452**, 161–162 (2008).
20. Soai, K., Kawasaki, T. & Matsumoto, A. Asymmetric autocatalysis of pyrimidyl alcohol and its application to the study on the origin of homochirality. *Acc. Chem. Res.* **47**, 3643–3654 (2014).
21. Pizzarello, S. & Weber, A. L. Prebiotic amino acids as asymmetric catalysts. *Science* **303**, 1151 (2004).
22. Glavin, D. P. & Dworkin, J. P. Enrichment of the amino acid L-isovaline by aqueous alteration on CI and CM meteorite parent bodies. *Proc. Natl Acad. Sci. USA* **106**, 5487–5492 (2009).
23. Blackmond, D. G. & Matar, O. K. Re-Examination of reversibility in reaction models for the spontaneous emergence of homochirality. *J. Phys. Chem. B.* **112**, 5098–5104 (2008).
24. Walker, S. I., Grover, M. A. & Hud, N. V. Universal sequence replication, reversible polymerization and early functional biopolymers: a model for the initiation of prebiotic sequence evolution. *PLoS ONE* **7**, e34166 (2012).
25. Vaidya, N., Walker, S. I. & Lehman, N. Recycling of informational units leads to selection of replicators in a prebiotic soup. *Chem. Biol.* **20**, 241–252 (2013).
26. Blocher, M., Hitz, T. & Luisi, P. L. Stereoselectivity in the oligomerization of racemic tryptophan N-Carboxyanhydride (NCA-Trp) as determined by isotope labeling and mass spectrometry. *Helv. Chim. Acta* **84**, 842–848 (2001).
27. Wagner, N., Rubinov, B. & Ashkenasy, G.  $\beta$  sheet induced chirogenesis in polymerization of oligopeptides. *Chemphyschem.* **12**, 2771–2780 (2011).
28. Szczepanski, J. T. & Joyce, G. F. A cross-chiral RNA polymerase ribozyme. *Nature* **515**, 440–442 (2014).
29. Brack, A. & Orgel, L. E.  $\beta$  structures of alternating polypeptides and their possible prebiotic significance. *Nature* **256**, 383–387 (1975).
30. Maury, C. P. J. Self-propagating  $\beta$ -sheet polypeptide structures as prebiotic informational molecular entities: the amyloid world. *Orig. Life Evol. Biosph.* **39**, 141–150 (2009).
31. Danger, G., Plasson, R. & Pascal, R. Pathways for the formation and evolution of peptides in prebiotic environments. *Chem. Soc. Rev.* **41**, 5416–5429 (2012).
32. Childers, W. S., Ni, R., Mehta, A. K. & Lynn, D. G. Peptide membranes in chemical evolution. *Curr. Opin. Chem. Biol.* **13**, 652–659 (2009).
33. Carny, O. & Gazit, E. A model for the role of short self-assembled peptides in the very early stages of the origin of life. *FASEB J.* **19**, 1051–1055 (2005).
34. Ura, Y., Beierle, J. M., Leman, L. J., Orgel, L. E. & Ghadiri, M. R. Self-assembling sequence-adaptive peptide nucleic acids. *Science* **325**, 73–77 (2009).
35. Turk, R. M., Illangasekare, M. & Yarus, M. Catalyzed and spontaneous reactions on ribozyme ribose. *J. Am. Chem. Soc.* **133**, 6044–6050 (2011).
36. Goodwin, J. T., Mehta, A. K. & Lynn, D. G. Digital and analog chemical evolution. *Acc. Chem. Res.* **45**, 2189–2199 (2012).
37. 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).
38. Severin, K., Lee, D. H., Martinez, J. A., Vieth, M. & Ghadiri, M. R. Dynamic error correction in autocatalytic peptide networks. *Angew. Chem. Int. Ed.* **37**, 126–128 (1998).
39. Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K. & Ghadiri, M. R. A self-replicating peptide. *Nature* **382**, 525–528 (1996).
40. Tovar, J. D. Supramolecular construction of optoelectronic biomaterials. *Acc. Chem. Res.* **46**, 1527–1537 (2013).
41. Rufo, C. M. et al. Short peptides self-assemble to produce catalytic amyloids. *Nat. Chem.* **6**, 303–309 (2014).
42. Zhang, C. et al. Self-assembled peptide nanofibers designed as biological enzymes for catalyzing ester hydrolysis. *ACS Nano* **8**, 11715–11723 (2014).
43. Amit, M. et al. Hybrid proton and electron transport in peptide fibrils. *Adv. Funct. Mater.* **24**, 5873–5880 (2014).
44. Tena-Solsona, M. et al. Emergent catalytic behavior of self-assembled low molecular weight peptide-based aggregates and hydrogels. *Chem. Eur. J.* **22**, 6687–6694 (2016).
45. Ivnitski, D. et al. The strong influence of structure polymorphism on the conductivity of peptide fibrils. *Angew. Chem. Int. Ed.* **55**, 9988–9992 (2016).
46. Singh, N., Kumar, M., Miravet, J. F., Ulijn, R. V. & Escuder, B. Peptide-based molecular hydrogels as supramolecular protein mimics. *Chem. Eur. J.* **23**, 981–993 (2017).
47. Guler, M. O. & Stupp, S. I. A self-assembled nanofiber catalyst for ester hydrolysis. *J. Am. Chem. Soc.* **129**, 12082–12083 (2007).
48. Rubinov, B., Wagner, N., Rapaport, H. & Ashkenasy, G. Self-replicating amphiphilic  $\beta$ -sheet peptides. *Angew. Chem. Int. Ed.* **48**, 6683–6686 (2009).
49. Carnall, J. M. A. et al. Mechanosensitive self-replication driven by self-organization. *Science* **327**, 1502–1506 (2010).
50. Sadownik, J. W., Mattia, E., Nowak, P. & Otto, S. Diversification of self-replicating molecules. *Nat. Chem.* **8**, 264–269 (2016).
51. Rubinov, B. et al. Transient fibril structures facilitating non-enzymatic self-replication. *ACS Nano* **6**, 7893–7901 (2012).
52. Raz, Y. et al. Effects of mutations in de novo designed synthetic amphiphilic  $\beta$ -sheet peptides on self-assembly of fibrils. *Chem. Commun.* **49**, 6561–6563 (2013).
53. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. H. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994).
54. Villain, M., Gaertner, H. & Botti, P. Native chemical ligation with aspartic and glutamic acids as C-terminal residues: scope and limitations. *Eur. J. Org. Chem.*, 3267–3272 (2003).
55. Dang, B., Kubota, T., Mandal, K., Bezanilla, F. & Kent, S. B. H. Native chemical ligation at Asx-Cys, Glx-Cys: chemical synthesis and high-resolution X-ray structure of ShK toxin by racemic protein crystallography. *J. Am. Chem. Soc.* **135**, 11911–11919 (2013).
56. Sadownik, J. W. & Philp, D. A simple synthetic replicator amplifies itself from a dynamic reagent pool. *Angew. Chem. Int. Ed.* **47**, 9965–9970 (2008).
57. Lehn, J.-M. Perspectives in chemistry—steps towards complex matter. *Angew. Chem., Int. Ed.* **52**, 2836–2850 (2013).
58. Li, J., Nowak, P. & Otto, S. Dynamic combinatorial libraries: from exploring molecular recognition to systems chemistry. *J. Am. Chem. Soc.* **135**, 9222–9239 (2013).
59. Ashkenasy, G., Hermans, T. M., Otto, S. & Taylor, A. F. Systems chemistry. *Chem. Soc. Rev.* **46**, 2543–2554 (2017).
60. Orgel, L. E. Molecular replication. *Nature* **358**, 203–209 (1992).

## Acknowledgements

We acknowledge support from the COST actions CM1304 and TD1308, a PBC post-doctoral fellowship (J.N.), and BGU Kreitman School fellowships for post-doctoral (B.R.) and PhD (D.I.) research. J.N. also acknowledges DST-INSPIRE faculty award (DST/INSPIRE/04/2016/001969).

## Author contributions

J.N., B.R., E.S., and R.M. conceived, designed, and performed the experiments. D.I. performed the AFM experiments. Y. Motro and Y. Miller performed the dynamic simulation analysis, and N.W. the kinetic analysis. G.A. supervised the overall project.



J.N., R.C.-L., and G.A. co-wrote the paper. All the authors discussed the results and commented on the manuscript.

### Additional information

**Supplementary Information** accompanies this paper at doi:10.1038/s41467-017-00463-1.

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**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/>.

© The Author(s) 2017