

Prebiotically Plausible Autocatalytic Peptide Amyloids

Saroj K. Rout,^[a] David Rhyner,^[a] Roland Riek,^{*,[a]} and Jason Greenwald^{*,[a]}

Abstract: The prebiotic emergence of molecules capable both of self-replication and of storing information was a defining event at the dawn of life. Still, no plausible prebiotic self-replication of biologically relevant molecules has been demonstrated. Building upon the known templating nature of amyloids, we present two systems in which the products of a peptide-bond-forming reaction act as self-replicators to

enhance the yield and stereoselectivity of their formation. This first report of an amino acid condensation that can undergo autocatalysis further supports the potential role of amyloids in prebiotic molecular evolution as an environment-responsive and information-coding system capable of self-replication.

Introduction

The accumulation of amyloid aggregates is a characteristic pathology of numerous diseases,^[1] however, amyloids are also known to carry out a wide range of native biological functions.^[2] This has led to several hypotheses on the possible role of amyloids in prebiotic molecular evolution.^[3–6] In addition to its prebiotic plausibility^[4] and catalytic potential^[7–10] the unique and repetitive structure of the cross- β -sheet-rich amyloid fold^[11,12] is of particular interest to studies of prebiotic replication as it provides the basis for the well-studied conformational templating ability of amyloids.^[12,13] In fact the chemical replication of amyloids has even been demonstrated for a few peptide systems.^[14–16] In each of these reports, two shorter non-amyloidogenic peptides ligate via an electrophilic thioester to produce an amyloidogenic parent peptide that as an amyloid can then template further ligations. Similarly, β -structured oligomers provide the structural basis for several reports of self-replicating molecules. For example, β -structured peptides that undergo a reversible aldehyde/amine condensation can evolve to a specific chain length^[17] and certain ring sizes of thiol-linked β -structured peptide macrocycles can self-replicate while competing for a common peptide feedstock.^[18] Interestingly, the latter system has also been found to catalyze cleavage of Fmoc-Gly, the alkene product of which promotes the thiol oxidation required for self-replication, thereby creating a protometabolic cycle.^[19] More recently peptide-nucleobase

conjugates have also been shown to spontaneously form self-replicating macromolecules which also depends on β -structuring of the templating product.^[20] Each of these examples involves the condensation of preformed peptides, however in an effort to demonstrate templated monomer addition to a peptide, we recently found that amyloids could act as templates to direct the sequential addition of single amino acids to a growing peptide.^[21,22] In particular, this amyloid templating functions in a complementary fashion analogous to the RNA/DNA transcription/replication and can occur in a prebiotically plausible condensation reaction. This sequence-selective, stereoselective and regioselective templating inspired us to study a prebiotically plausible form of amyloid self-replication, in which a parent amyloid templates its own production through the sequential addition of amino acids. This basic concept (Figure 1A) is that the product of an amino acid condensation with a peptide will yield an amyloidogenic peptide that can selectively template further condensations.

Results and Discussion

We have previously observed that the peptide FRFRFR-NH₂ (further referred to as (FR)₄) is soluble at neutral pH yet remains aggregation prone at slightly basic pH or in the presence of a complementary acidic peptide like (FE)₄.^[22] Yet even without a complement peptide, it was expected that both pH and NaCl concentration could be modulated in order to induce amyloid aggregation through a reduction and shielding of the electrostatic charges on the peptide. We hypothesized that if such conditions were also amenable to peptide bond formation, then (FR)₄ could potentially act as a template for its own synthesis through sequential amino acid additions. To test this, we examined reactions between carbonyldiimidazole (CDI)-activated D-L-Phe and RFRFRFR-NH₂ (R(FR)₃). If the L-Phe (F) condensation product (FR)₄ can form a templating amyloid but the D-Phe (f) product fR(FR)₃ cannot, then the reaction rate and stereoselectivity should change as the reaction proceeds. Therefore, we monitored the time-dependent formation of the (FR)₄ and fR(FR)₃ products at pH 6.5 and 7.1 and from 0 to 3 M NaCl.

[a] S. K. Rout, D. Rhyner, R. Riek, J. Greenwald
Laboratory of Physical Chemistry
Swiss Federal Institute of Technology, ETH Hönggerberg
Vladimir-Prelog-Weg 2, 8093 Zürich (Switzerland)
E-mail: roland.riek@phys.chem.ethz.ch
jason.greenwald@phys.chem.ethz.ch

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/chem.202103841>

© 2021 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

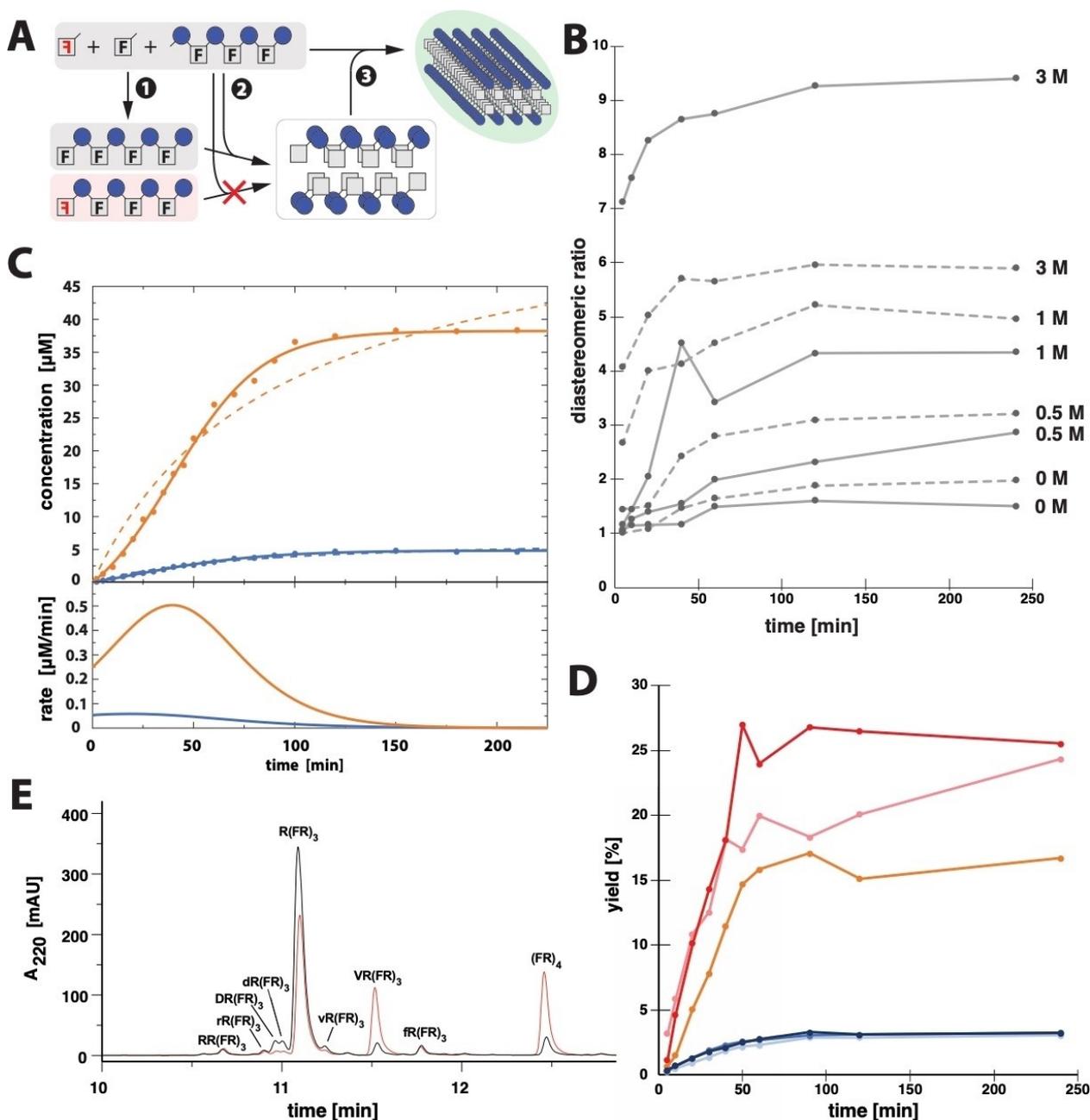


Figure 1. Autocatalysis with the (FR)₄ peptide amyloid. A) A schematic depicting the mechanism of amyloid-templated autocatalytic condensation of activated amino acids. The black F and the mirror-image red F represent the amino acid enantiomers in the reaction. B) The evolution of the product *dr* during the reaction of 200 μM activated D,L-Phe with 100 μM R(FR)₃ at pH 6.5 (—) and 7.1 (<L_>) at four different NaCl concentrations. C) Kinetics of product formation at 1 M NaCl, pH 6.5 for (FR)₄ (orange) and its diastereomer fR(FR)₃ (blue; independent measurement from B). The upper plot shows the concentration of the diastereomeric products as a function of time with fits to an autocatalytic reaction mechanism (—) and second-order reaction mechanism (<L_>). The lower plot shows the derivative of the autocatalytic fits in the upper plot, depicting the reaction rate as a function of time. D) Kinetics of (FR)₄-seeded condensation reactions with 200 μM D,L-Phe and 100 μM R(FR)₃ in 1 M NaCl at pH 6.5. The yield of the D product fR(FR)₃ is plotted in shades of blue, and the L-product (FR)₄ yield, after subtraction of the seed concentration, is plotted in orange for the no-seed control, pink for 5 μM (FR)₄, and red for 20 μM (FR)₄. E) Reversed-phase HPLC chromatogram showing the condensation products of four racemic amino acids (100 μM each of Asp, Arg, Val, and Phe) with 100 μM R(FR)₃ at pH 6.5. The black trace is for no NaCl, and the red trace is at 3 M NaCl with the peaks labeled by the single-letter code with lower-case letters for the D-configured residues. The elution time of each product was identified in individual reactions with each enantiomer of each amino acid.

We found that the yield and rate of formation of (FR)₄ but not fR(FR)₃ is strongly enhanced as a function of NaCl concentration and that this enhancement is generally greater at pH 7.1 versus 6.5 (Figure S1 in the Supporting Information). Of particular

interest to our hypothesis is that stereoselectivity increases as the reaction proceeds (Figure 1B), hinting at an underlying autocatalytic process. The kinetics of the reaction were measured in more detail at 1 M NaCl (Figure 1C), revealing a

clear sigmoidal form that can be described by a simple two-step mechanism (see the Supporting Information). The reaction rate derived from the fit for the L-Phe condensation product formation reaches a maximum about 40 min after the start of the reaction (Figure 1C), a feature also observed in the other tested reaction conditions (Figure S1). By comparison, the D-Phe condensation product reaction kinetics are more typical of a second-order reaction.

The highest stereoselectivity with a final diastereomeric ratio (*dr*) of 9.4 was observed for 3 M NaCl, while the largest change in the *dr* throughout the reaction (from 1.2 to 4.3) was observed for 1 M NaCl at pH 6.5. This suggests that at 3 M NaCl the aggregation kinetics are faster than the nontemplated reaction, or in other words the aggregation occurs so quickly that as soon as we can detect product, the reaction is already being templated and therefore highly stereoselective. However in 1 M NaCl the aggregation of product is slow enough to allow for a nonselective start but fast enough to take part in the reaction before the substrate is depleted. This is consistent with the finding that at 3 M NaCl, (FR)₄ aggregates within seconds, whereas R(FR)₃ remains soluble for hours (Figure S2). Also, in addition to the β -structure observed in the reaction products at 3 M NaCl, the TEM images of these products contained fibrillar aggregates (Figure S3). Also of note is that the strong *dr* enhancement was almost absent if the reactions were not agitated (Figure S4). As mechanical agitation is known to affect the rate of amyloid formation^[23,24] as well as the outcome of other oligomerization-based self-replicating systems,^[18] the observed dependence on agitation is also consistent with an underlying amyloid-based templating mechanism. In addition to agitation, the surface to volume ratio (SVR) of the reaction was also found to be important. The results from reactions in a range of volumes and vessel dimensions indicate that the volume needs to be high enough to allow for efficient agitation, but low enough to maintain a high SVR (Figure S5). A high SVR increases the air–water interface which has been shown to be important for the amyloid aggregation of A β , IAPP and α -synuclein.^[25–27] Also, while the autocatalytic nature of the reaction is highly reproducible, the early kinetics are somewhat variable, consistent with the stochastic nature of a nucleated aggregation process. For example, for three independent reactions from three different days, the kinetics and yield are similar but the standard deviation of the data is significant (Figure S6). Further evidence that β -aggregation is required for the enhanced yield and stereochemistry of the reaction is that condensation reactions with the shorter peptide R(FR)₂ lack the autocatalytic signatures, displaying a constant and low stereoselectivity at 3 M NaCl (Figure S7). That the homogeneous reactions with R(FR)₂ are actually less efficient at high salt concentration hints at an inhibitory effect of NaCl on CDI-activated amino acid condensations, an effect which we also observed in other homogeneous reactions (Figure S8). Considering the negative impact of NaCl on these reactions, the increased yield of (FR)₄ at 3 M NaCl is even more remarkable.

The role of the product in the reaction was further probed by the addition of sub-stoichiometric (FR)₄ to the reaction at 1 M NaCl, pH 6.5. We observed that 0.05 or 0.2 equivalents of

(FR)₄ in a reaction with 100 μ M R(FR)₃ increased the yield and *dr* of the product (Figure 1D). This is particularly noteworthy as (FR)₄ could in principle compete with R(FR)₃ for activated DL-Phe but instead appears to cooperate. Thus, it appears that (FR)₄ can template its own synthesis and that the templating is dependent on the pH and the salt concentration.

Unexpected was the evolution in selectivity at 0 M NaCl (increase in *dr* from \sim 1 to \sim 2) because at this condition (FR)₄ is soluble and with a non-beta secondary structure similar to R(FR)₃ (Figure S9). Even though the *dr* and its change are small, the effect is reproducible. In an effort to find a possible interaction between R(FR)₃ and (FR)₄ that could explain this result, we measured the diffusion of these peptides by DOSY NMR in the absence of NaCl but found that they both appear to be non-interacting in solution (Figure S10). If there had existed a templating-competent oligomeric species that was in fast exchange with monomer, then this would have led to a decrease in the observed diffusion of the peptide compared to a purely monomeric species. Alternatively, it could be an interface phenomenon by which the ordering of (FR)₄ at the air–water interface leads to a β -structure that can template stereoselective condensation with L-Phe. To test this hypothesis, we performed a reaction with a smaller surface to volume ratio and found that the *dr* decreased, consistent with an interface phenomenon (Figure S11). A similar air–water interface-induced stereoselectivity has been observed for other peptides with sequences of alternating hydrophobicity.^[28] Thus, there is good reason to conclude that the small selectivity observed in the absence of NaCl is not amyloid-based templating in the bulk solution, although it could be the result of interface-stabilized amyloids.

The self-templating of amino acids with R(FR)₃ is also sequence-selective in that for a racemic mixture of four activated amino acids (Asp, Arg, Val, Phe) only L-Val and L-Phe show a NaCl-dependent increase in yield and stereoselectivity (Figures 1E and S12). Interestingly, the yield and stereoselectivity of L-Val condensation with R(FR)₃ is enhanced by the presence of DL-Phe (Figures S12 and S13). The results of these two independent experiments (4-amino-acid mixture or 2-amino-acid mixture) indicate that (FR)₄ is a better template than VR(FR)₃ and that the former is able to cross-template L-Val condensations. Sequence selectivity was further investigated by using a peptide (FR)₃ that is two residues shorter than the (FR)₄ template. In a reaction of activated DL-Phe and DL-Arg with (FR)₃, the products contained significantly more (FR)₄ in 3 M NaCl compared to the 0 M control (Figure S14). In this case the initial template is formed by the sequential nontemplated addition of L-Arg and then L-Phe, thereby forming the self-replicating peptide (FR)₄, one of 16 possible double addition products. This experiment also revealed that FF(FR)₃ is also able to self-template in a competition with (FR)₄.

To explore the general feasibility of self-replicating peptide amyloids, we examined an analogous system composed of amino acids that are potentially more prebiotically relevant.^[29,30] Here the substrate peptide (OV)₄ consists of valine and ornithine with the sequence OVOVOV-NH₂ and is also soluble at neutral pH with a net positive charge. Furthermore, reactions

with $(OV)_4$ have the added complexity that the four ornithine sidechains and the N terminus of the peptide present five reactive sites for amino acid condensations (Figure 2A). The potential of $V(OV)_4$ to form an amyloid template for valine condensation with $(OV)_4$ was confirmed in a screen of these peptides' NaCl-dependent solubilities (Figure S15). Interestingly, the presence of the $V(OV)_4$ reduced the solubility of $(OV)_4$ indicating the formation of a co-aggregate. However, in the reactions of activated D-L-Val with $(OV)_4$ in 4 M NaCl it appeared as though there was no templating: both the *dr* and regioselectivity remained stable as the reaction proceeded (Figure S16). Suspecting that this result could be due to slow aggregation kinetics of $V(OV)_4$, we measured its aggregation by CD spectroscopy and found that it indeed takes several hours to aggregate in 4 M NaCl pH 6.3, even at a concentration of 100 μM peptide (Figure S17). Because the reactions with CDI-activated amino acids are essentially complete in 6 h at 20 °C and the concentration of the $V(OV)_4$ product never exceeded 15 μM , its aggregation under these conditions appears to be too slow for it to play a role in the reaction between Val and $(OV)_4$. In an effort to uncouple the reaction and aggregation kinetics, we carried out the reaction in steps with aliquots of activated amino acid added every 12 h. This method produced a clear increase in the regioselectivity and stereoselectivity as a function of the extent of reaction (Figure 2B, C). That is, a single 1 mM addition of activated D-L-Val produced less $V(OV)_4$ and more side chain and D-Val condensation products than the reaction carried out with five 200 μM additions (Figure S18). As

with the $(FR)_4$ replicator, agitation of the reaction is required for the autocatalytic stereo- and regioselective amino acid additions (Figure S19). TEM images of the reaction products revealed fibrillar aggregates (Figure S20) and the involvement of the product in the reaction and the requirement that it aggregates in order to act as a template was further demonstrated by $V(OV)_4$ "seeded" reactions (Figure 2D). In these reactions performed at 3 and 4 M NaCl and with and without preincubation of $V(OV)_4$ (to allow aggregation to occur), only the preincubated reaction at 4 M NaCl produced more $V(OV)_4$ product than was consumed by its reaction with D-L-Val. This verifies that the salt-dependent aggregation of $V(OV)_4$ is critical for the regioselectivity and stereoselectivity of the reaction that produces more $V(OV)_4$. It is important to keep in mind that in this "seeded" $V(OV)_4$ reaction, there are ten unique amines and two activated molecules (L- and D-Val) that together can produce 20 unique single condensation products in addition to numerous multiple condensation products. Under such circumstances, any increase in the amount of the N-terminal L-Val condensation product is noteworthy.

While investigating the NaCl-dependence of the D-L-Val + $(OV)_4$ condensation reaction, we discovered another type of NaCl-induced templating. At sufficiently high NaCl concentration (>4.5 M), the reaction CDI-activated D-L-Val with $(OV)_4$ becomes significantly more regioselective and stereoselective even in a single step reaction and without "seeding" by $V(OV)_4$ (Figure S21). That this selectivity enhancement is also dependent on a pre-incubation of $(OV)_4$ in NaCl indicates that an

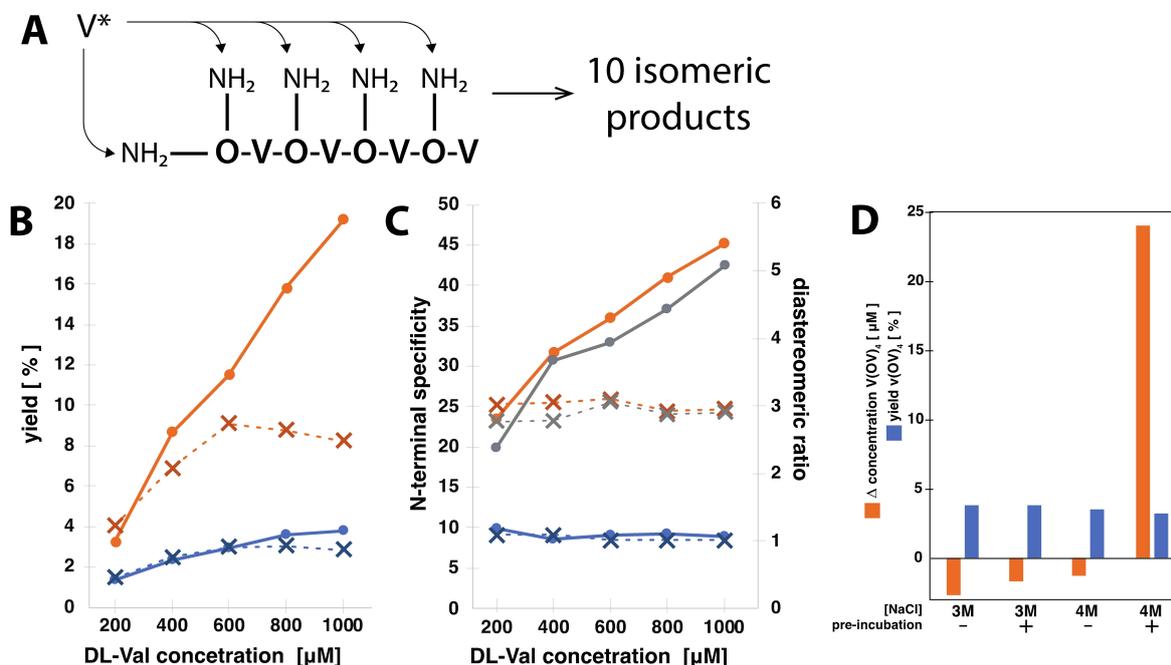


Figure 2. Autocatalysis with the $V(OV)_4$ peptide amyloid. A) A schematic depicting the five reactive amines in the $(OV)_4$ substrate peptide. B) The evolution of the N-terminal L-Val (orange) and D-Val (blue) condensation product yields upon sequential addition (circles) or single bolus addition (crosses) of activated D-L-Val. Samples were collected and stabilized in 6 M guanidine-HCl 12 h after addition of activated D-L-Val. C) Evolution of the N-terminal specificity colored as in B and of the *dr* of the N-terminal addition products (gray). The N-terminal specificity is the amount of N-terminal product as a percent of all single condensation products. D) The effect of NaCl and preincubation on $V(OV)_4$ -seeded reactions. The plot shows the change in the concentration of $V(OV)_4$ in reactions seeded with $V(OV)_4$ as well as the final concentration of the D enantiomer condensation product $v(OV)_4$.

aggregate form of (OV)₄ can on its own also act as a template for the production of V(OV)₄.

From the origin of life context, the self-replication of molecules appears to be a necessary event which must have taken place on the prebiotic earth to increase their abundance and to encode and store information. As the amyloid templating mechanism is compatible with the chemistry of COS activation,^[21] we conclude that the self-replication presented here also presents a plausibly prebiotic scenario for peptide formation. Taken together with our earlier reports on amyloid-templating,^[22] the presented results on the autocatalytic replication of peptide amyloids highlight an interesting feature of amyloid-templating: that the preferred product, whether it be a self-peptide or a complement-peptide, can be dictated by the environmental conditions of NaCl concentration and pH.

Acknowledgements

We would like to thank the ETHZ ScopeM microscopy center for help with the EM images and Dr. Harindranath Kadavath for help with the DOSY NMR. This work was supported by an ETH grant. Open access funding provided by Eidgenössische Technische Hochschule Zurich.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

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

Keywords: aggregation · autocatalysis · amyloids · origin of life · self-replication

[1] F. Chiti, C. M. Dobson, *Annu. Rev. Biochem.* **2006**, *75*, 333–366.

[2] D. Otzen, R. Riek, *Cold Spring Harbor Perspect. Biol.* **2019**, *11*, a033860.

- [3] O. Carny, E. Gazit, *FASEB J.* **2005**, *19*, 1051–1055.
[4] J. Greenwald, W. Kwiatkowski, R. Riek, *J. Mol. Biol.* **2018**, *430*, 3735–3750.
[5] C. P. J. Maury, *Cell. Mol. Life Sci.* **2018**, *75*, 1499–1507.
[6] L. E. Orgel, *PLoS Biol.* **2008**, *6*, e18.
[7] M. P. Friedmann, V. Torbeev, V. Zelenay, A. Sobol, J. Greenwald, R. Riek, *PLoS One* **2015**, *10*, e0143948.
[8] Z. Lengyel, C. M. Rufo, Y. S. Moroz, O. V. Makhlynets, I. V. Korendovych, *ACS Catal.* **2018**, *8*, 59–62.
[9] O. V. Makhlynets, P. M. Gosavi, I. V. Korendovych, *Angew. Chem. Int. Ed.* **2016**, *55*, 9017–9020; *Angew. Chem.* **2016**, *128*, 9163–9166.
[10] C. M. Rufo, Y. S. Moroz, O. V. Moroz, J. Stöhr, T. A. Smith, X. Hu, W. F. DeGrado, I. V. Korendovych, *Nat. Chem.* **2014**, *6*, 303–309.
[11] W. T. Astbury, S. Dickinson, K. Bailey, *Biochem. J.* **1935**, *29*, 2351–2360.
[12] R. Riek, D. S. Eisenberg, *Nature* **2016**, *539*, 227–235.
[13] D. M. Fowler, A. V. Koulouf, C. Alory-Jost, M. S. Marks, W. E. Balch, J. W. Kelly, *PLoS Biol.* **2005**, *4*, e6.
[14] J. Nanda, B. Rubinov, D. Ivnicki, R. Mukherjee, E. Shtelman, Y. Motro, Y. Miller, N. Wagner, R. Cohen-Luria, G. Ashkenasy, *Nat. Commun.* **2017**, *8*, 434.
[15] B. Rubinov, N. Wagner, H. Rapaport, G. Ashkenasy, *Angew. Chem. Int. Ed.* **2009**, *48*, 6683–6686; *Angew. Chem.* **2009**, *121*, 6811–6814.
[16] Y. Takahashi, H. Mihara, *Bioorg. Med. Chem.* **2004**, *12*, 693–699.
[17] C. Chen, J. Tan, M.-C. Hsieh, T. Pan, J. T. Goodwin, A. K. Mehta, M. A. Grover, D. G. Lynn, *Nat. Chem.* **2017**, *9*, 799–804.
[18] J. M. A. Carnall, C. A. Waudby, A. M. Belenguer, M. C. A. Stuart, J. J.-P. Peyralans, S. Otto, *Science* **2010**, *327*, 1502–1506.
[19] J. Ottelé, A. S. Hussain, C. Mayer, S. Otto, *Nat. Catal.* **2020**, *3*, 547–553.
[20] A. K. Bandela, N. Wagner, H. Sadihov, S. Morales-Reina, A. Chotera-Ouda, K. Basu, R. Cohen-Luria, A. de la Escosura, G. Ashkenasy, *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2015285118.
[21] R. Bomba, S. K. Rout, M. Bütikofer, W. Kwiatkowski, R. Riek, J. Greenwald, *Orig. Life Evol. Biospheres* **2019**, *49*, 213–224.
[22] S. K. Rout, M. P. Friedmann, R. Riek, J. Greenwald, *Nat. Commun.* **2018**, *9*, 234.
[23] K. M. Batzli, B. J. Love, *Mater. Sci. Eng. C* **2015**, *48*, 359–364.
[24] F. Grigolato, C. Colombo, R. Ferrari, L. Rezabkova, P. Arosio, *ACS Nano* **2017**, *11*, 11358–11367.
[25] S. Campioni, G. Carret, S. Jordens, L. Nicoud, R. Mezzenga, R. Riek, *J. Am. Chem. Soc.* **2014**, *136*, 2866–2875.
[26] L. Jean, C. F. Lee, D. J. Vaux, *Biophys. J.* **2012**, *102*, 1154–1162.
[27] A. Morinaga, K. Hasegawa, R. Nomura, T. Ookoshi, D. Ozawa, Y. Goto, M. Yamada, H. Naiki, *Biochim. Biophys. Acta BBA-Proteins Proteomics* **2010**, *1804*, 986–995.
[28] H. Zepik, E. Shavit, M. Tang, T. R. Jensen, K. Kjaer, G. Bolbach, L. Leiserowitz, I. Weissbuch, M. Lahav, *Science* **2002**, *295*, 1266–1269.
[29] P. G. Higgs, R. E. Pudritz, *Astrobiology* **2009**, *9*, 483–490.
[30] L. M. Longo, D. Despotović, O. Weil-Ktorza, M. J. Walker, J. Jabłońska, Y. Fridmann-Sirkis, G. Varani, N. Metanis, D. S. Tawfik, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 15731–15739.

Manuscript received: October 25, 2021

Accepted manuscript online: November 23, 2021

Version of record online: December 20, 2021