Systems chemistry of aminoacyl phosphates: Spontaneous and selective peptide oligomerisation in water driven by phase changes

Kun Dai,[1‡] Mahesh D. Pol,[1‡] Lenard Saile,[1] Arti Sharma,[1] Bin Liu,[2] Ralf Thomann,[1,3] Johanna L. Trefs,[1,4] Danye Qiu,[5] Sandra Moser,[5] Stefan Wiesler,[5] Bizan N. Balzer,[1,3,4] Thorsten Hugel,[1,3,4] Henning J. Jessen[1,5] and Charalampos G. Pappas[1*]

[1] DFG Cluster of Excellence Cluster of Excellence *liv*MatS @FIT – Freiburg Center for Interactive Materials and Bioinspired Technologies, University of Freiburg, Georges-Köhler-Allee 105, 79110, Freiburg, Germany. [2]College of Materials and Chemical Engineering, China Three Gorges University, Yichang, Hubei 443002, China. [3]Freiburg Materials Research Center (FMF), University of Freiburg, Stefan-Meier-Strasse 21, 79104, Freiburg, Germany. [4]Institute of Physical Chemistry, University of Freiburg, Albertstrasse 21, 79104, Freiburg, Germany. [5]Institute of Organic Chemistry, University of Freiburg, Albertstrasse 21, 79104, Freiburg, Germany.

‡These authors contributed equally to this work. *e-mail: charalampos.pappas@livmats.uni-freiburg.de

Abstract: The pathways and mechanisms that allow spontaneous and selective peptide elongation in aqueous abiotic systems remain unclear. Herein we work to uncover those pathways by following the systems chemistry of aminoacyl phosphates, the synthetic counterpart of aminoacyl adenylates. Thus, oligomers up to dodecamers were synthesized in one flask and on the minute time scale, where consecutive additions activated autonomous phase changes. The resulting phases arose from the high polar and reactive nature of the aminoacyl phosphates, which can be dissolved in water in concentrations of up to 300 mM. Upon elongation, short oligomers are prevalent in solution, while in the aggregated phase, longer oligomers are selected based on their aggregation propensity. We furthermore demonstrated that the solution phases can be isolated and act as a new environment for continuous elongation, by adding various phosphates. These findings suggest that the systems chemistry of aminoacyl phosphates can activate a selection mechanism for peptide formation, by merging aqueous synthesis and self-assembly.

Introduction: In the realm of biology, amino acids are activated using Adenosine Triphosphate (ATP), where elongation, sequence control and termination mechanisms are achieved through complex biomolecular machinery.[1] After protein synthesis, additional processes involving multiple and dynamic phase changes[2] give rise to an interplay between folding, aggregation and binding to other compound classes.[3] Inspired by the dynamics found in biological systems involving multicomponent interactions, systems chemistry[4-8] focusses on mechanisms and pathways that enable coupling of chemical reactions with self-assembly using simple building blocks.[9, 10] The convergence of covalent synthesis and physical assembly often involve chemical reaction networks[11] and the use of activated precursors to access thermodynamically stable,[12] kinetically trapped structures[13] and assemblies that function only by constant supply of energy.[14-18] Regardless of the pathway followed to induce self-assembly, phase changes (liquid/liquid or liquid/solid) give rise to dynamic reconfiguration of different species, enabling self-selection of assemblies,[19] template-driven synthesis[20] and activation of anabolic pathways.[21] The presence of a stimulus or a chemical messenger alters the distribution in different phases, assisting selectivity[22] and autocatalysis.[23] In order to achieve and further impact phase transitions within synthetic systems, the chemical information is often installed within short peptide sequences, which represent simplified versions of known biological structures,[24] involving the fabrication of helices and β-sheets. Such structural motifs emerge from minimal amino acid repeats, where aromatic and aliphatic dyads have been associated with gelation,[25-27] while tripeptides incorporating proline residues adopt helical conformations.[28] Thus, the chemical information and the length of the repeat have a significant impact on the self-assembly, even by exchanging single amino acids in short peptide sequences. The impact of sequence is further evident in longer repeats of hydrophilic and charged residues, which

nowadays gain significant attention towards phase separation,[29, 30] making peptides[31-34] ideal candidates for aqueous nanotechnology.[35, 36] However, abiotic pathways that couple autonomous oligomerisation of single amino acids[37] with self-assembling pathways remain elusive.

Outside of biology, chemical reactions yielding peptide bonds have gained significant interest. [38] Some peptide oligomerisation strategies use amino acids, which are appended at the C-terminus with various activated ester moieties.[39] Aminolytic reactions then achieve sequence elongation in organic media but often encounter challenges associated with the formation of stable species, such as cyclic dipeptides (diketopiperazines).[40] The development of native chemical ligation[41, 42] and solid phase peptide synthesis (SPPS)[43] methodologies has undoubtedly revolutionized peptide coupling, where a few challenges remain, including hydrophobic sequences, organic solvents and extensive use of starting materials. Non-activated approaches have been applied towards peptide coupling using gases,[44] mechanical forces[45] and wet-dry cycles,[46, 47] however these strategies may require extreme experimental conditions. The challenges associated with the solubility of amino acid precursors in water and the unfavorable thermodynamics of peptide condensation have been a roadblock to exploring the effect of sequence, pathway and self-assembly in spontaneous and selective peptide oligomerisation.

In this work, we use aminoacyl phosphates - soluble activated amino acid precursors, which contain enough chemical information to trigger phase changes during autonomous peptide oligomerisation. This strategy leads to the formation of a pool of oligomers through aqueous phase peptide synthesis, in which short oligomers are selected in solution, while long and more hydrophobic residues are protected in the aggregated phase. We control the pathway of reactive intermediates, leading to continuous elongation, by furthermore reducing the formation of cyclic dipeptides - a common issue of protecting-group-free synthesis (Fig. 1a, b). By continuously adding various phosphate monomers, oligomers of different length and composition were formed, capable of forming highly ordered supramolecular structures (Fig. 1c). We showed that the chemical nature and the length of the amino acid repeat is directly related to the hydrophobicity of the monomers, which in turn affects selectivity and the distribution in different phases, hinting towards a selection mechanism in aqueous peptide oligomerisation (Fig. 1d).

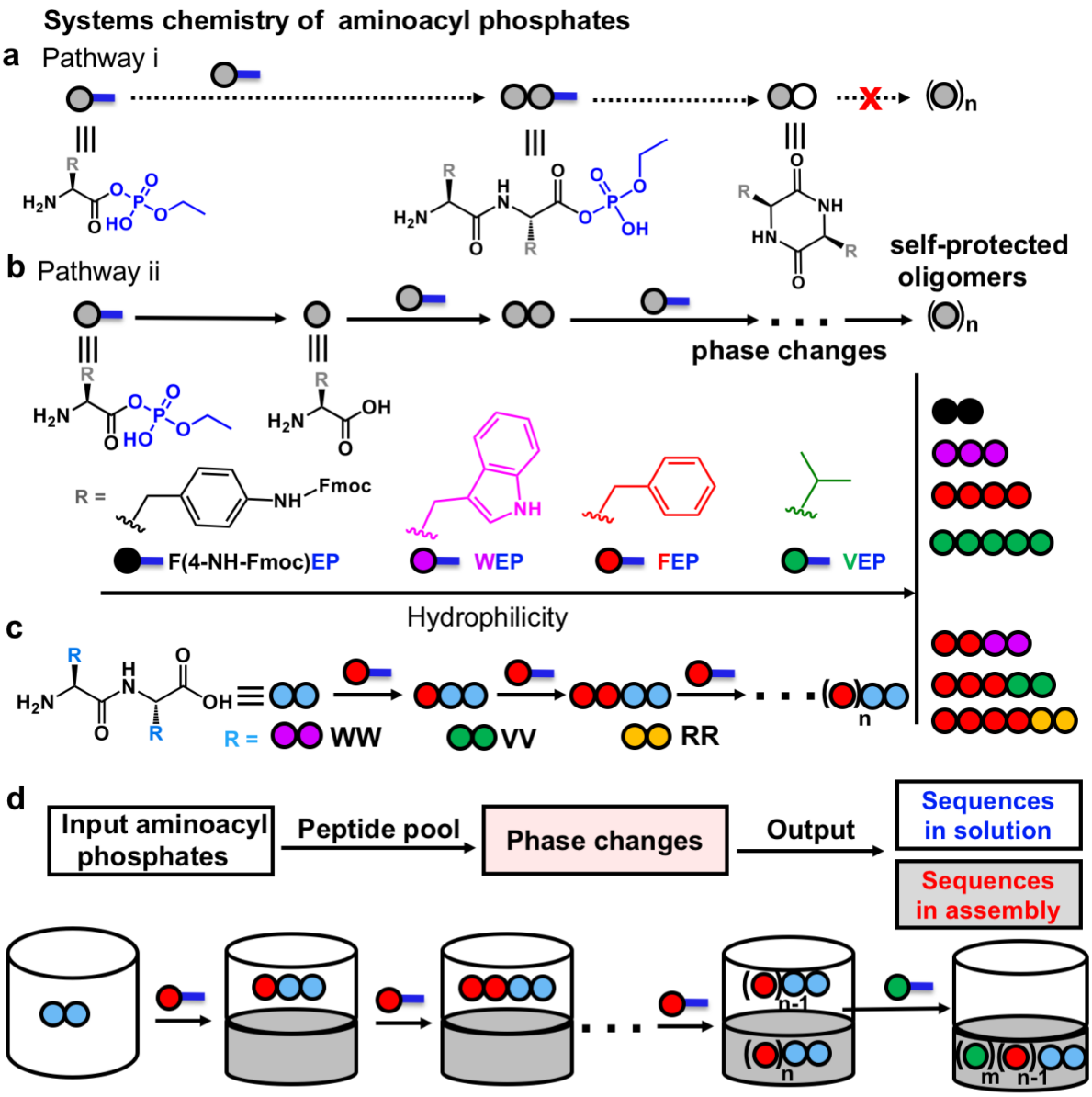

Fig. 1: Systems chemistry of aminoacyl phosphates: Pathway-dependent oligomerisation through **a** reactive phosphate intermediates, resulting in the formation of cyclic dipeptides, **b** consecutive addition of activated monomers giving rise to sequence elongation. Chemical structure of aminoacyl phosphates involved in self-assembly, where length-specific oligomers are formed based on the hydrophilicity of the monomers used. **c** Hetero-coupling utilizing mixtures of dipeptides and aminoacyl phosphates. **d** Schematic representation of continuous sequence elongation using various phosphate monomers. At early stages, sequentially adding activated amino acids enables the formation of hydrophilic sequences in solution (white boxes). At later stages and through consecutive additions, hydrophobic sequences are protected in the assembled phase (dark grey boxes), as a result of autonomous phase changes.

Results

Pathway-dependent reactivity: Water-soluble precursors featuring phosphate moieties have been used to construct peptide bonds.[48-50] Specifically, mixed aminoacyl-phosphate anhydrides have been reported from Kluger[51] to yield peptide bonds with protected amino acids residues. However, these

systems are rarely coupled to self-assembly.[52] Throughout this study, we used mixed aminoacyl-phosphate anhydrides, where the phosphate esters act as solubility tags, producing overall a highly polar and reactive amino acid derivative. The charge in the phosphates esters enables us to reach concentrations of up to 300 mM in an aqueous environment, which is challenging to achieve with other type of activated moieties.[53] The high water solubility of the phosphate precursors allows us in turn to explore autonomous phase transitions from the self-assembly of the oligomers formed. During elongation, the solubility tag is removed, giving rise to the construction of nonpolar and unreactive peptides. We use the single letter code for amino acids throughout this work. We started with phenylalanine ethyl phosphate (**FEP**), as an activated monomer and investigated the degree of oligomerisation in 0.6 M borate buffer pH 9.1. We used basic conditions to enhance solubility and assist the formation of oligomers, considering the pKa of the primary amine of **FEP**.[51] By dissolving 80 mM of **FEP** and by performing direct injection experiments using mass spectrometry, we observed the spontaneous formation of oligomers of up to 12 amino acid residues (Fig. 2a), accompanied with a significant amount of cyclic dipeptide ($\mathbf{F_{cy}}$). Time-dependent Ultra-Performance Liquid Chromatography-Mass Spectrometry (UPLC-MS) analysis enabled us to track the transient formation of oligomers carrying the ethyl phosphate moiety ($\mathbf{F_nEP}$) and the presence of cyclic dipeptide. The latter is mainly formed through intramolecular cyclisation of the dipeptide phosphate ester (**FFEP**) (Fig. 2b and Supplementary Fig. 1). These findings are in agreement with previous observations, in which dipeptide esters have been used in spontaneous cyclisations.[54] In order to assist oligomerisation and decrease cyclisation, we introduced competition from another pathway, by reducing the initial concentration to 10 mM and by sequentially adding **FEP** (8 times total). This pathway involves initially faster hydrolysis to single amino acid (**F**), as the pH is not drastically decreased compared to the one-pot addition. Therefore, the hydrolysis product (**F**) is a stronger nucleophile than **FEP** itself, minimizing the formation of **FFEP** and consequently the construction of $\mathbf{F_{cy}}$. In this environment, elongation mainly occurs through nucleophilic attack from **F** and **F**-containing oligomers to **FEP** (Fig. 2d). Analysis of the libraries using UPLC-MS revealed that at early stages of the process, the dimer ($\mathbf{F_2}$) was the dominant product (50%), where the system macroscopically behaved as transparent solution. Upon consecutive additions, **F** and $\mathbf{F_2}$ decrease, accompanied with an increase of trimer ($\mathbf{F_3}$) and tetramer ($\mathbf{F_4}$). After making 4 additions, $\mathbf{F_3}$ reached up to 32% conversion, while $\mathbf{F_4}$ kept increasing until 40%. The formation of cyclic dipeptide ($\mathbf{F_{cy}}$) was significantly reduced (10%) following the pathway of consecutive additions (Fig. 2c).

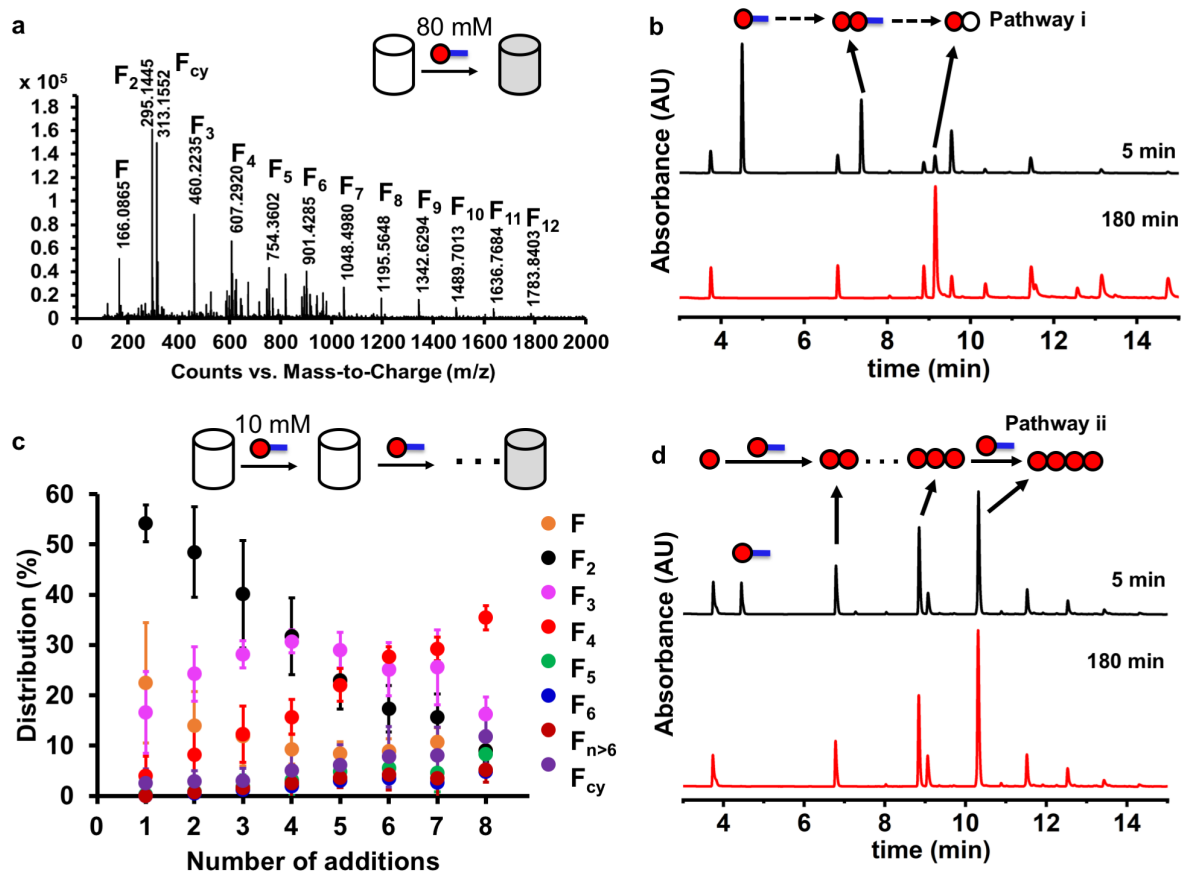

Fig. 2: Pathway-dependent oligomerisation. a Mass spectrum of 80 mM **FEP** in 0.6 M borate buffer pH 9.1. **b** UPLC chromatograms (absorption at 214 nm) of one-pot addition of 80 mM **FEP** at 5 and 180 min, highlighting the formation of dipeptide-phosphate intermediate (**FFEP**) resulting in the construction of cyclic dipeptide (F_{cy}). **c** Distribution of oligomers throughout **FEP** additions (8 times). Error bars represent the standard deviation of three experiments conducted at different batches of **FEP**. **d** UPLC chromatograms (absorption at 214 nm) on the 8th addition of **FEP** (80 mM total) at 5 and 180 min, minimizing the formation of **FFEP**.

Length-selective oligomerisation: Constructing chemically activated peptide libraries from a single component and the pathway of reactive intermediates prompted us to investigate the effect of hydrophobicity on the oligomerisation process. Thus, we varied the amino acid side chains in the structure of aminoacyl phosphates using aromatic (tryptophan) and aliphatic (valine) residues. Non-natural amino acids carrying bulky aromatic groups (fluorenylmethoxycarbonyl) have been utilized to assist aromatic stacking interactions, which have previously been shown to enhance the assembly of short peptide amphiphiles.[55] The synthesis of all aminoacyl phosphates is presented in the Supplementary Information (Supplementary Information- synthesis section). We used the retention time in the UPLC chromatograms for the different phosphate monomers as an indication of their hydrophobicity, with the **F(4-NH-Fmoc)EP** being the most hydrophobic as expected (Supplementary Information- characterization of **XEP** by UPLC-MS). Then similarly to **F** libraries, we performed 8 additions for **WEP** and **VEP**, while for **F(4-NH-Fmoc)EP**, we stopped at the 4th addition, considering the macroscopically observed strong gel formation after initial additions. Analysis of libraries revealed a significant amount of dimer for the **F(4-NH-Fmoc)-**containing building block, comprising 50% of the final product distribution. As the hydrophobicity decreased in the case of **WEP**, initial additions resulted in a reduction of **W** and W_2, leading to increase of trimer (W_3), reaching finally maximum abundance to

38%. For **VEP**, a significant amount of pentamer (40% of V_5) was obtained (Fig. 3a and Supplementary Fig. 2). UPLC-MS analysis of all libraries is presented in the SI (Supplementary Figs. 3-10). These results indicate a direct correlation between the chemical information in the structures of aminoacyl phosphates with the degree of oligomerisation and with the termination step. As the hydrophobicity decreased from **F(4-NH-Fmoc)EP**, to **WEP, FEP** and **VEP**, the final dominant oligomer increased from dimer ($F(4\text{-NH-Fmoc})_2$), to trimer (W_3), tetramer (F_4) and pentamer (V_5) respectively. Moreover, the formation of cyclic dipeptide was enhanced for sequences carrying smaller amino acid side chains, where in the case of the **V**-containing libraries it reached 25%. Importantly, the amount of V_{cy} could be reduced to 10%, by building libraries at a higher pH (9.7), demonstrating that environmental conditions can be applied to tune library composition and increase the conversion to larger oligomers (Fig. 3b, Supplementary Fig. 11).

Self-protected oligomers: The fact that different oligomers formed in combination with macroscopically observed transitions motivated us to investigate the impact of supramolecular interactions upon consecutively adding activated precursors. Macroscopic changes were confirmed using turbidity measurements. Monitoring the absorbance at 600 nm, we observed enhanced turbidity as the length of the oligomers increased for the different libraries (Supplementary Fig. 12). These changes coincided with a more chiral environment of the structures formed after oligomerisation. Circular Dichroism (CD) spectra for the **V** libraries revealed a continuous enhancement of the signal at 218 and 197 nm, which are characteristic bands for the formation of β-sheet structures.[23] A more chiral ordering for the arrangement of the aromatics was noticed for the **F** and **W** libraries, evidenced from the appearance of red-shifted structures to 230 and 240 nm respectively. The supramolecular organization of the chromophores for the **F(4-NH-Fmoc)**-containing hydrogel is in line with the enhancement of the CD signal at 295 and 307 nm (Supplementary Fig. 13). Notably, the mechanical properties of the samples were altered upon oligomerisation. Rheological experiments revealed the formation of weak gels for, **F**, **V** and the **F(4-NH-Fmoc)**-containing libraries while for the **W** oligomers, no gelation was noticed (Fig. 3d). Supramolecular reconfiguration of assemblies was monitored using Atomic Force Microscopy (AFM) and Transmission Electron Microscopy (TEM). After the first addition of **FEP**, analysis of the libraries using AFM (Supplementary Fig. 14) revealed the presence of small fibers. With the hydrophobicity of the amino acid side chain increased, there was a transition to bulkier and spherical structures, which was confirmed and further investigated using TEM. More specifically, upon consecutive **FEP** additions, fibers increased both in length and density, assembling finally into an entangled network. For **V** libraries a transition from ill-defined aggregates to rod-like assemblies was noticed, while for **W** oligomers spherical aggregates were visualized that did not reconfigure throughout the different additions (Fig. 3c and Supplementary Fig. 15). This agrees with previous reports on the self-assembly of short peptides containing tryptophan residues.[12] For the **F(4-NH-Fmoc)**-containing hydrogel, a dense fibrillar network was visualized. These observations highlight that during the sequential additions, different supramolecular structures are formed and reconfigured, as the oligomers differ in length and composition. Thus, these libraries can be used to investigate the selection of oligomers and consequently supramolecular motifs from activated single amino acids.

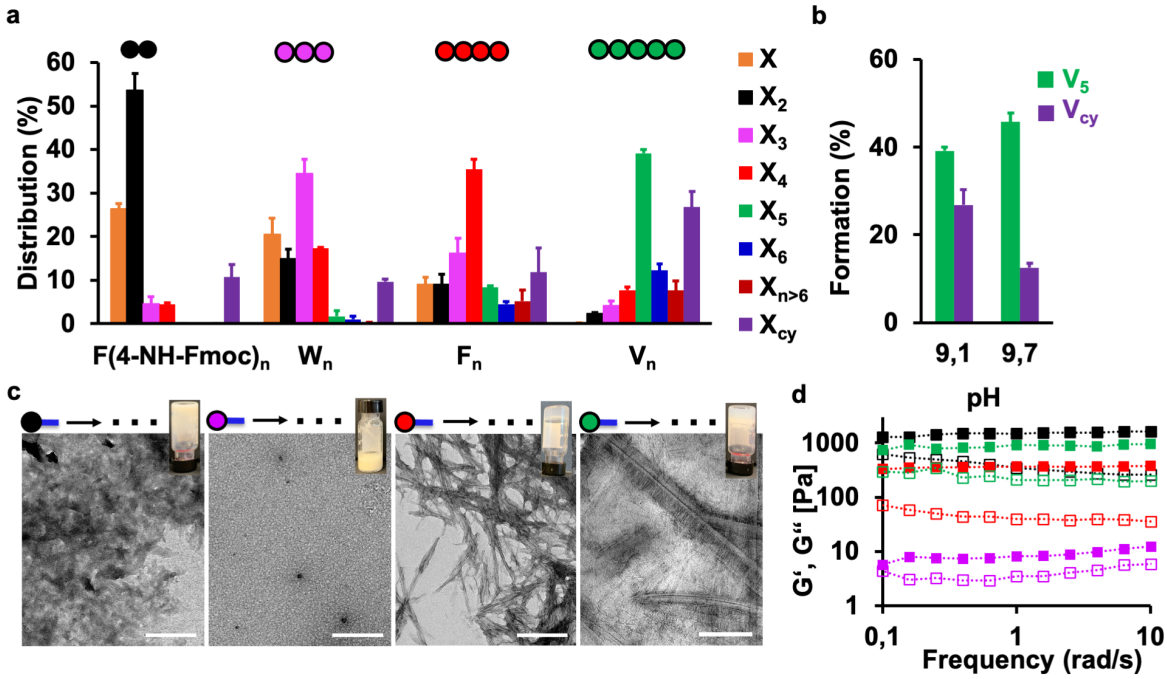

Fig. 3: Effect of amino acid side chain on oligomerisation and self-assembly. a Distribution of oligomers after making 4 additions of 10 mM **F(4-NH-Fmoc)EP** and after 8 additions of 10 mM **WEP**, **FEP** and **VEP** in 0.6 M borate buffer pH 9.1. **b** Comparison of the formation of V_5 and V_{cy} after several additions of 10 mM **VEP** in 0.6 M borate buffer at different pH (8 times from pH 9.1 and 10 times from pH 9.7). Error bars represent the standard deviation of three experiments conducted at different batches of precursors. **c** TEM images of libraries, scale bar 500 nm. Sample photos represent the macroscopic behaviour of libraries after various additions. **d.** Storage and loss modulus of the libraries after 8 additions of monomer. The mechanical properties of the samples for the **F(4-NH-Fmoc)**-libraries was monitored after 4 additions of **F(4-NH-Fmoc)EP**. The solid squares represent the storage modulus (G'), while open squares represent the loss modulus (G'').

Having established the effect of supramolecular interactions on the assembly of the peptide oligomers, we next identified the exact library composition in solution and in the aggregated state for the peptide libraries. We performed centrifugation experiments after 8 additions of **FEP** and analyzed the sample using UPLC (Figs. 4a, c). The solution contained oligomers up to trimers, whereas the aggregated phase is dominated by F_4. The majority of F_4 is present in the assembled phase (98%). Similarly, in analogous experiments for the **W** and **V** libraries, we found that 95% of W_3 and 99% of V_5 were also in the assembled phase (Figs. 4b, d, Supplementary Fig. 16). We moreover carried out the centrifugation experiments after 4 additions of **FEP** and **VEP**, where F_4 is not the dominant oligomer and V_5 competes with other **V**-containing oligomers. Similar to the outcome after 8 additions of monomer, **F**-F_3 and **V**-V_4 are present in solution, while F_4 and V_5 are protected in the assembled phase (Supplementary Fig. 17). These findings highlight the self-assembly propensity of the more hydrophobic oligomers at the initial additions. In order to further support the importance of hydrophobic interactions, we altered the solvent environment in the case of the **F** library after 4 additions of **FEP**. We split the libraries in two parts, by adding 20% of acetonitrile (ACN) into one of them and we monitored the behavior of F_4. UPLC analysis suggested that 80% of F_4 is transferred to the solution phase (Fig. 4e). Upon consecutive **FEP** additions, the presence of co-solvent increases F_5 (more than two-fold) compared to amounts found in the aqueous phase, suggesting that part of the assembly is disrupted and its components redissolved, enabling further elongation (Fig. 4f). Attempts to further increase the amount of organic solvent and

consequently the formation of larger oligomers were hampered by the low solubility of borate buffer in ACN.

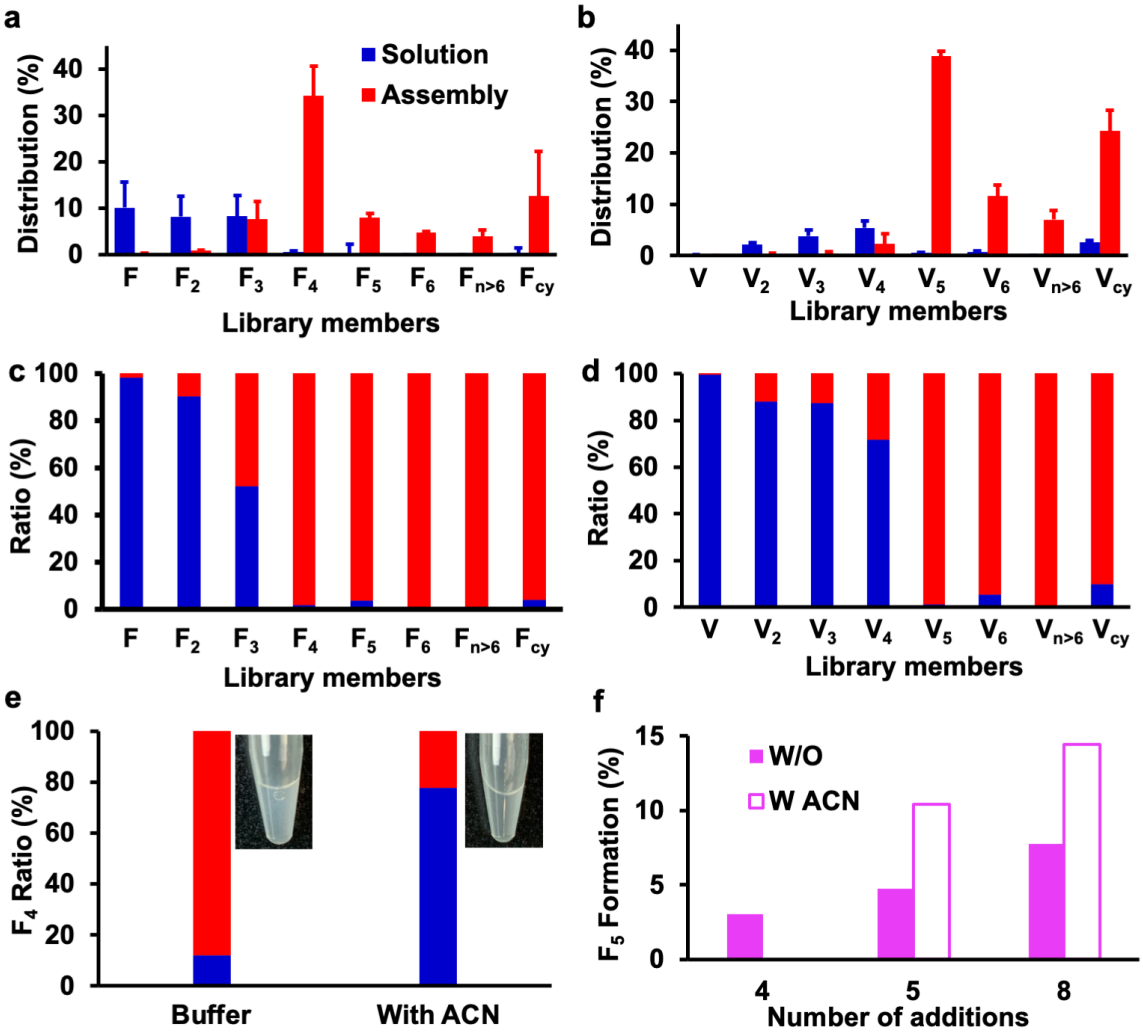

Fig. 4: Self-protected oligomers. a, b Distribution and c, d Ratio of oligomers in solution and in assembled phase after 8 additions of FEP and VEP in 0.6 M borate buffer pH 9.1 .The concentration used in every addition was 10 mM. Centrifugation was applied after the 8th addition. Error bars represent the standard deviation of three experiments conducted at different batches of precursors. e Ratio of F_4 in solution and in assembled phase in the presence and absence of 20% ACN after the 4th FEP addition. Inset photos represent the macroscopic behavior of libraries. f Formation of F_5 in the presence and absence of 20% ACN. Libraries were split in two parts after the 4th addition, where consecutive additions of 10 mM FEP were introduced in two different solvent environments. The conversion to F_5 is represented with solid bars for the libraries without ACN and with open bars for libraries containing 20% ACN.

Phase changes in libraries containing various phosphate monomers: In our final step, we shifted our focus to diversifying the sequences of the peptide oligomers, by performing sequential additions of aminoacyl phosphates into the presence of soluble dipeptides. This approach enabled us to apply the concept of selective oligomerisation driven by self-assembly into longer sequences featuring hydrophobic and hydrophilic residues (Fig. 5a). Thus, we built libraries of 10 mM FEP in the presence of 20 mM of dyads to favour hetero-coupling. The dipeptide sequences consisted of aromatic (W_2), aliphatic (V_2), and charged (R_2) amino acid residues. Similar to the libraries formed by a single

component, the degree of oligomerisation and the termination step were directly related to the hydrophobicity of the structures used in the mixtures. Thus, after performing 8 additions of **FEP**, the dominant oligomers in the final F_nX_2 sequences were F_2W_2, F_3V_2 and F_4R_2 for W_2, V_2 and R_2 mixtures respectively (Fig. 5b, Supplementary Figs. 18-23). Notably, while homo-coupling of **F** lead to mainly tetramer and fibrillar assemblies, we were able to incorporate **F** via hetero-coupling into oligomers which had different lengths and compositions, while retaining the same supramolecular structure, as evidenced using TEM (SI Fig. 24).

We then tried to apply the strategy of spontaneous phase changes in systems featuring various aminoacyl phosphates. We used the dyads of **V** and **R** in mixtures that initially contained **FEP**. In the first addition, 20 mM of V_2 was coupled with 10 mM **FEP**, giving rise to 55% conversion of the FV_2 tripeptide. In order to increase the length of the oligomers, we added **FEP** until we observed phase changes (solution to solid), which in turn could have resulted in the selective formation of a hydrophobic (assembled phase) or a hydrophilic sequence (solution phase). More specifically, after 5 additions of **FEP**, the library was dominated by F_2V_2 and F_3V_2 (Supplementary Figs. 25, 26). Subsequently, centrifugation experiments and an analysis of the solution phase using UPLC revealed that the shorter peptide sequence (F_2V_2) remained in solution (42%), while the more hydrophobic oligomer (F_3V_2) was protected in the assembled phase (40%). We have isolated the solution phase in this step and we added 10 mM **VEP**. Adding **VEP** gave rise to the formation of new sequences featuring valine residues also at the N-terminus ($V_mF_nV_2$). After centrifugation, unreacted F_2V_2, and other soluble oligomers remained in solution, while VF_2V_2 could be obtained from the assembled phase with only one addition of **VEP** (Fig. 5c, Supplementary Figs. 27, 28). The distribution of the assembled phase demonstrated that VF_2V_2 (40%) was the dominant species. By replacing V_2 with R_2, 50% of the tripeptide FR_2 was obtained from the first addition of **FEP**. Since F_4R_2 was the dominant sequence (Fig. 5b), less hydrophobic peptides (FR_2, F_2R_2 and F_3R_2) were selected in solution. Based on this principle and after adding **FEP** 3 times, F_2R_2 was present in solution, while after adding 6 times, the solution and the assembled phase was enriched in F_3R_2 and F_4R_2 respectively (Supplementary Figs. 29-31). By isolating the solution phase (after 6 times of **FEP** addition) and by adding **VEP**, new sequences of $V_mF_nR_2$ were obtained, where the assembled phase was dominated by 40% $V_2F_3R_2$ (Supplementary Figs. 32, 33).

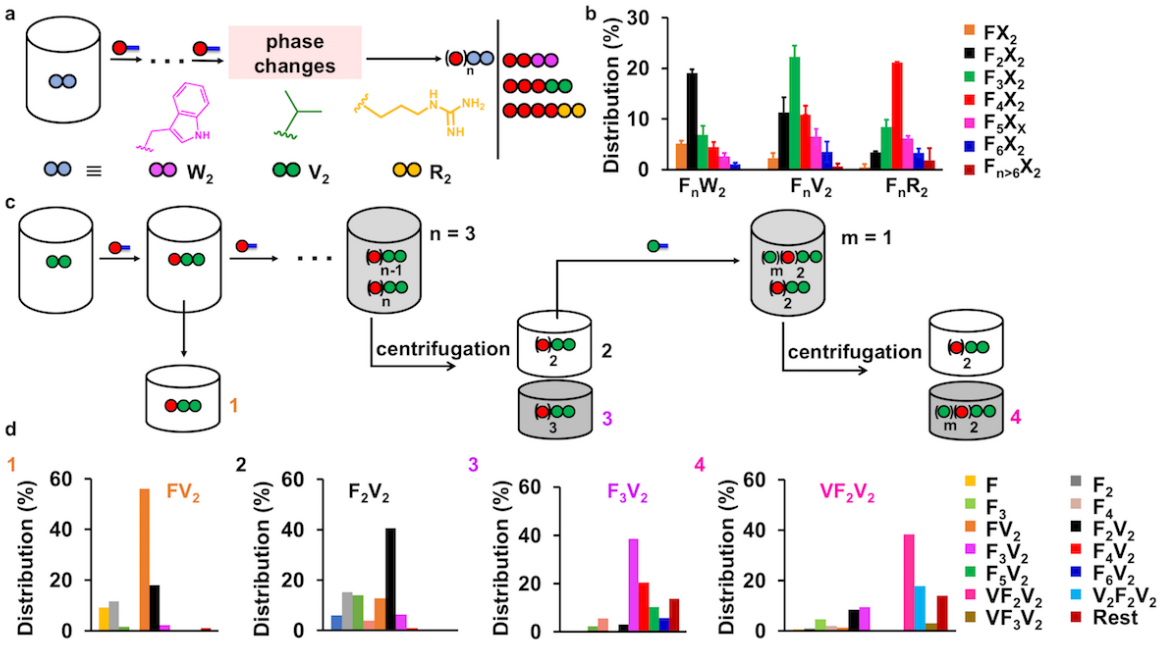

Fig. 5: Phase changes in multicomponent libraries. a Chemical structures of dipeptides coupled to aminoacyl phosphates. **b** Distribution of oligomers after adding 10 mM of **FEP** 8 times in the presence of 20 mM W_2, V_2 and R_2 in 0.6 M borate buffer pH 9.1. Error bars represent the standard deviation of three experiments conducted at different batches of **FEP**. **c** Schematic representation of continuous aqueous synthesis of libraries containing **FEP** and V_2 leading initially to the formation of F_nV_2 oligomers. Adding **VEP** results in the construction of $V_mF_nV_2$. The UPLC analyses highlight the distribution in solution and in the assembled phase at different stages of **FEP** and **VEP** additions. White boxes represent the solution phase and dark grey boxes the assembled phase. **d**. Oligomer distribution for the V_2-containing libraries. From left to right after the 1st **FEP** addition (1), after the 5th addition in solution (2), after the 5th addition in the assembled phase (3) and after adding **VEP** in the solution phase of (2), demonstrating the composition in the assembled phase (4). Centrifugation was applied at the 5th **FEP** addition and after adding **VEP** at the final step.

These findings suggest that our approach can be applied towards selectively oligomerising peptides with different lengths and compositions in different phases. In general, we observed that when the assembled phase is dominated by oligomers having **n** number of amino acid residues, the solution is enriched in **n-1** oligomers (Fig. 6). Thus, shorter homo-oligomers are present in solution (W_1 and W_2, F_1-F_3, V_1-V_4), while longer and more hydrophobic amino acid repeats are protected in the aggregated phase (W_3, F_4, V_5). The same rule applies for the multicomponent libraries, where hydrophobic peptides (F_3V_2, F_4R_2, VF_2V_2 and $V_2F_3R_2$) can be exported from the assembled phase. This outcome depends on the number of additions we performed, the chemical structure of the aminoacyl phosphates and the relative aggregation of the peptides formed. Compared to the V_2-containing libraries, the more hydrophilic precursors and the oligomers which they produced (R_2-containing libraries), allowed coupling and oligomerisation to longer peptides, where self-assembly requires more additions to be achieved. The resulting oligomers are capable of self- and co-assembling into well-defined architectures. These observations are in agreement with the selectivity and with the degree of oligomerisation obtained from activated amino acids using amyloid-like sequences.[20, 56]

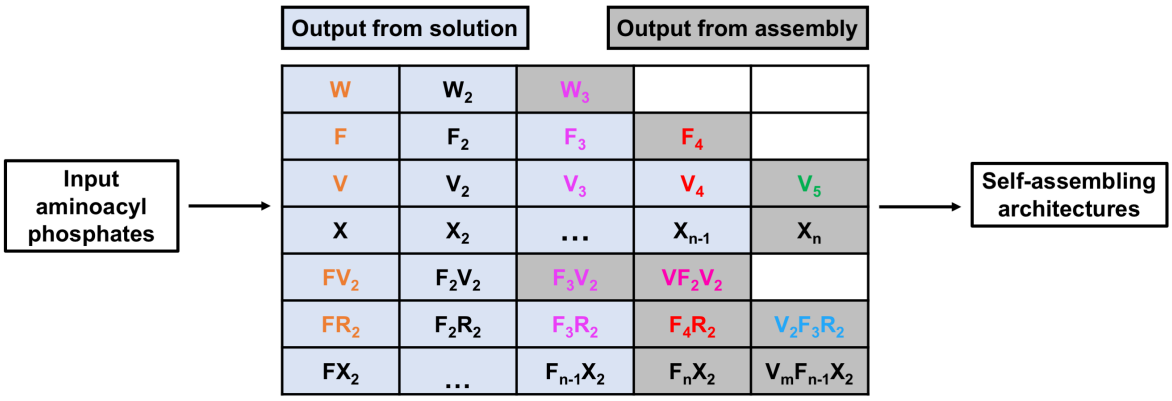

Fig. 6: Input and output of spontaneous peptide oligomerisation. Output of oligomers resulting from autonomous oligomerisation using aminoacyl phosphates as input. The white boxes indicate oligomers that are selected in solution, while the dark grey boxes represent sequences which are protected in the assembled phases, capable of adopting well-defined structures.

Discussion: We have reported the use of aminoacyl phosphates as activated monomers, where the phosphate esters offer high water solubility and trigger the spontaneous formation of peptide oligomers. In these chemically activated peptide libraries, the highly polar and reactive phosphate precursors are reconfigured to unreactive and less polar oligomers. We demonstrated that the

construction of reactive intermediates can be switched off following kinetic pathways, canceling cyclic products that block elongation. In these pathways, the chemical information of the amino acid side chains dictates the degree of the oligomerisation, where trimers, tetramers, pentamers and hexamers were formed and were capable of adopting well-defined architectures. Consecutive additions triggered phase changes, enabling the selective formation of hydrophobic oligomers in the assembled phase, while shorter and polar oligomers were present in solution. The solution phases can be isolated and utilized towards continuous coupling of targeted sequences that have different lengths, compositions and self-assembling behaviours. The approach demonstrated herein could be extended towards the one-pot synthesis and assembly of chimeric oligomers, featuring peptides and nucleobases, non-natural amino acids and other reactive nucleophiles. The peptide libraries and the resulting phases could be further impacted using templates and surfactants, which can in turn alter the composition, create new phases and induce selectivity. The sequential additions can be potentially automated, through the fabrication of a microreactor, which would encapsulate automatically injected amino acids inside droplets containing the phosphate precursors, leading to programmable peptide oligomerisation coupled to material screening. While aminoacyl phosphates have received significant attention in prebiotic chemistry and in research on the origins of life[57], our work uses them in peptide systems chemistry with potential applications in aqueous phase synthesis.

Methods:

Library preparation. Aminoacyl phosphates (**XEP**) were dissolved in 0.6 M borate buffer at pH 9.1. Consecutive additions of **XEP** were carried out after the reaction finished (monomer was fully consumed), where the vortexed and sonicated sample as a whole was transferred to a new vial containing the aminoacyl phosphate. The concentration used for **XEP** in all libraries was 10 mM. For the libraries containing dipeptides (X_2), aminoacyl phosphates (**XEP**, 10 mM) were dissolved in 0.6 M borate buffer pH 9.1 containing W_2, V_2 or R_2. The concentration used for the dipeptides was 20 mM. Additions were performed in the same manner to the systems without dipeptides.

Centrifugation. For the libraries yielding homo-oligomers, centrifugation was applied at the end of the additions (after 8 times of **FEP**, **VEP** and **WEP**) in order to have a better separation of the solution and the assembled phase (Fig. 4a-d). In case of **FEP** and **VEP**, centrifugation was also applied after 4 additions as an additional control experiment to extract the composition in the two phases. For the libraries containing dipeptides, centrifugation was introduced at different stages to also maximize the selectivity of different oligomers, as shown in Fig: 5c and Supplementary Fig. 29. Centrifugation experiments were performed using the VWR Micro Star microcentrifuge 12 for 5 min at 13000 RPM.

UPLC analysis. UPLC analyses were performed on a Waters Acquity UPLC H-Class Bio system, equipped with a photodiode array detector at a detection wavelength of 214 nm. Samples were injected on an Acquity UPLC CSH-C18 (150×2.1 mm) column, using ULC-MS grade water (eluent A) and ULC-MS grade acetonitrile (eluent B), which contained 0.1% trifluoroacetic acid as the modifier. A flow rate of 0.3 ml min^{-1} and a column temperature of 35 °C were applied. For the non-centrifuged samples vortex (30 seconds) and sonication (30 seconds) were performed prior to UPLC injection to ensure a homogeneous phase. Vortex and sonication (for 30 seconds) was furthermore applied after dilution, prior to UPLC injection. Samples were prepared by taking 10 µl from the reaction vial and diluting (100-500 times) into H_2O with 0.1% trifluoroacetic acid.

Oligomer distribution in libraries. With distribution (%), we refer to the peak area (%) of single amino acids and oligomers relative to the absorbance of the monomeric unit. For the centrifugation experiments, the total component distribution before and after centrifugation (for the solution phase)

was analyzed using UPLC. The distribution in the assembled phase is attained by the component peak area before centrifugation, deducting the corresponding area in solution phase. For example, if the peak area of F_3V_2 before centrifugation equals to **T** and the peak area of F_3V_2 in solution phase equals to **S**, then the peak area of F_3V_2 in assembled phase could be calculated from **A** = **T-S** .

UPLC-MS analysis. Ultra-Performance Liquid Chromatography-Mass Spectrometry (UPLC-MS) experiments were performed on an Agilent 6546 LC/Q-TOT equipped with an infinity 1290 II in the LC section. We used the same UPLC column, as this is described in the UPLC analysis section above. The Q-TOF was equipped with a dual AJS ESI source. The experiments were conducted at a VCap voltage of 4000 V, a sheath gas temperature of 300°C and a fragmentor voltage of 120 V. An internal reference was used.

NMR. The ^{1}H NMR spectra were recorded on a Bruker 300 MHz spectrometer and Bruker Avance Neo 400 MHz with broadband cryoprobe Prodigy. The ^{31}P NMR spectra were recorded on 162 MHz and Bruker 122 MHz spectrometers using ^{1}H-broad band decoupling in the indicated deuterated solvent. Chemical shifts were reports as delta values from standard peaks.

Rheology. Rheological measurements were carried out with an Anton Paar MCR 302 rheometer at 20 °C, using a 25 mm cone-plate geometry (CP25-1, Anton Paar) and a measuring gap of 0.047 mm. Samples were prepared as previously mentioned and placed on the bottom plate. The mechanical properties of the samples for the **F(4-NH-Fmoc)-**libraries were monitored after adding monomer 4 times, while for **FEP**, **WEP** and **VEP** after 8 times. Frequency sweep tests were performed at a strain amplitude of γ = 1.0% over a frequency range of 0.1–10 rad s^{-1}. Data was recorded at 25 °C.

Circular Dichroism (CD). CD spectra were recorded using a J-810 circular dichroism spectrometer from Jasco at 20 °C, established by a PTC-423S Peltier controller. 100 times dilution of the peptide libraries was introduced and the samples were placed into a quartz cuvette with 1 mm path length. Each spectrum was obtained by scanning wavelength from 300 nm to 190 nm at a scanning rate of 100 nm/min. Three successive wavelength scans were taken to average for each sample. The spectral region below 195 nm and over 330 were omitted in the analyses.

Turbidity. Turbidity measurements were carried out at 29 °C on a Microplate Spectrophotometer (Tecan). Measurements were performed in a 96-well plate. Absorbance was measured at 600 nm.

Transmission Electron Microscopy (TEM). Imaging was performed after the addition of the activated monomers, where no oligomers carrying ethyl phosphate moiety (F_nEP) were observed using UPLC (typically one day). Libraries were vortexed and sonicated prior to sample preparation for imaging. Small drops of the solution, the gel or the suspension were applied to a carbon-coated Cu grid for 30 seconds incubation, followed by two drops of water wash and one drop 5 μl of 2% (w/v) uranyl acetate solution for 30 seconds staining. Excess solution was removed by blotting the grid with a piece of filter paper and left to air dry. Imaging was performed using a FEI Talos 120C at 120 kV operating voltage. Images were taken using a Ceta 16 megapixel camera.

Atomic Force Microscopy (AFM). The samples were analyzed using AFM-based imaging after the first addition of active monomers, which were also used for TEM imaging. Measurements were carried out on a HOPG (HOPG-ZYA, 0.4° mosaic spread, Mikromasch) surface, which was cleaved and immersed into the sample solution for 60 sec. Then two successive washing steps were performed for 2 sec with ultrapure water (18.2 MΩ cm, Purelab Chorus 1, Elga LabWater). The surface was dried and stored in a desiccator. HOPG with the dried structures on it were imaged in air using the Cypher ES (Asylum Research, an Oxford Instruments Company) with a sample heating/cooling stage at a constant temperature of 25 °C, operating in intermittent contact mode (AC mode) using blueDrive

(photothermal excitation) with Sense 70 (NuNano) cantilevers. Image analysis and processing was performed using the free SPM analysis software Gwyddion, version 2.61.[58] The images were levelled by a mean-plane subtraction, rows were aligned to median value and two or three degree leveling was performed. For some images few horizontal scares were corrected and the minimum data value was set to zero.

Author Contributions

C.G.P. supervised the overall project. K.D. and M.D.P. contributed equally to this work. K.D. conceived and designed the study. K.D. build the libraries and analyzed them by CD, rheology and UV. M.D.P. synthesized all the aminoacyl phosphates. K.D., M.D.P. and L.S. performed the UPLC experiments. K.D. and R.T. conducted the TEM experiments. M.D.P., A.S., B.L., S.M. and S.W. carried out the purity tests by UPLC and NMR. S. W. provided helpful suggestions on the synthesis of aminoacyl phosphates. K.D. and D.Q. performed the UPLC-MS characterization. J.L.T. performed the AFM experiments and analyzed them together with B.N.B. and T.H. H.J.J. and T.H. provided valuable suggestions throughout the manuscript. C.G.P. and K.D. co-wrote the manuscript with the support of all authors.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2193/1-390951807 via “Living, Adaptive and Energy-Autonomous Materials Systems” (livMatS). We acknowledge further support by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the project PA 3783/2-1). We thank Prof. Michael Müller and Daniela Bjarnesen for the assistance with the CD measurements and Prof. Bastian Rapp and David Böcherer for the assistance with rheology. We thank Dr. Claudia Jessen-Trefzer for the assistance with the UV measurements and Christoph Warth for analytical support.

Data Availability

All data generated or analyzed during this study are included in this article and its supplementary information files. Source data are provided with this paper.

Competing interests

The authors declare no competing interests.

References

1. Francklyn C. S., Mullen P. Progress and challenges in aminoacyl-tRNA synthetase-based therapeutics. *J. Biol. Chem.* **294**, 5365-5385 (2019).
2. Tan J., Zhang L., Hsieh M.-C., Goodwin J. T., Grover M. A., Lynn D. G. Chemical control of peptide material phase transitions. *Chem. Sci.* **12**, 3025-3031 (2021).
3. Aumiller W. M., Keating C. D. Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. *Nat. Chem.* **8**, 129-137 (2016).
4. Bai Y., Chotera A., Taran O., Liang C., Ashkenasy G., Lynn D. G. Achieving biopolymer synergy in systems chemistry. *Chem. Soc. Rev.* **47**, 5444-5456 (2018).
5. Ashkenasy G., Hermans T. M., Otto S., Taylor A. F. Systems chemistry. *Chem. Soc. Rev.* **46**, 2543-2554 (2017).
6. Chatterjee A., Reja A., Pal S., Das D. Systems chemistry of peptide-assemblies for biochemical transformations. *Chem. Soc. Rev.* **51**, 3047-3070 (2022).

7. Grzybowski B. A., Huck W. T. The nanotechnology of life-inspired systems. *Nat. Nanotechnol.* **11**, 585-592 (2016).
8. Sheehan F.*, et al.* Peptide-Based Supramolecular Systems Chemistry. *Chem. Rev.* **121**, 13869-13914 (2021).
9. Vantomme G., Meijer E. W. The construction of supramolecular systems. *Science* **363**, 1396-1397 (2019).
10. Walther A. Viewpoint: From Responsive to Adaptive and Interactive Materials and Materials Systems: A Roadmap. *Adv. Mater.* **32**, 1905111 (2020).
11. Klemm B., Lewis R. W., Piergentili I., Eelkema R. Temporally programmed polymer – solvent interactions using a chemical reaction network. *Nat. Commun.* **13**, 6242 (2022).
12. Pappas C. G.*, et al.* Dynamic peptide libraries for the discovery of supramolecular nanomaterials. *Nat. Nanotechnol.* **11**, 960-967 (2016).
13. Kemper B., Zengerling L., Spitzer D., Otter R., Bauer T., Besenius P. Kinetically Controlled Stepwise Self-Assembly of AuI-Metallopeptides in Water. *J. Am. Chem. Soc.* **140**, 534-537 (2018).
14. Sharko A., Livitz D., De Piccoli S., Bishop K. J. M., Hermans T. M. Insights into Chemically Fueled Supramolecular Polymers. *Chem. Rev.* **122**, 11759-11777 (2022).
15. Klajn R. Dissipative Self-Assembly: Fueling with Chemicals versus Light: Fueling with Chemicals versus Light. *Chem* **7**, 23-37 (2021).
16. Ragazzon G., Prins L. J. Energy consumption in chemical fuel-driven self-assembly. *Nat. Nanotechnol.***13**, 882-889 (2018).
17. van Rossum S. A. P., Tena-Solsona M., van Esch J. H., Eelkema R., Boekhoven J. Dissipative out-of-equilibrium assembly of man-made supramolecular materials. *Chem. Soc. Rev.* **46**, 5519-5535 (2017).
18. Heuser T., Weyandt E., Walther A. Biocatalytic Feedback-Driven Temporal Programming of Self-Regulating Peptide Hydrogels. *Angew. Chem. Int. Ed.* **54**, 13258-13262 (2015).
19. Tena-Solsona M., Wanzke C., Riess B., Bausch A. R., Boekhoven J. Self-selection of dissipative assemblies driven by primitive chemical reaction networks. *Nat. Commun.* **9**, 2044 (2018).
20. Rout S. K., Friedmann M. P., Riek R., Greenwald J. A prebiotic template-directed peptide synthesis based on amyloids. *Nat. Commun.* **9**, 234 (2018).
21. Omosun T. O.*, et al.* Catalytic diversity in self-propagating peptide assemblies. *Nat. Chem.* **9**, 805-809 (2017).
22. Hafezi N., Lehn J.-M. Adaptation of Dynamic Covalent Systems of Imine Constituents to Medium Change by Component Redistribution under Reversible Phase Separation. *J. Am. Chem. Soc.* **134**, 12861-12868 (2012).
23. Chen C.*, et al.* Design of multi-phase dynamic chemical networks. *Nat. Chem.* **9**, 799-804 (2017).
24. Wei G.*, et al.* Self-assembling peptide and protein amyloids: from structure to tailored function in nanotechnology. *Chem. Soc. Rev.* **46**, 4661-4708 (2017).
25. Adler-Abramovich L., Gazit E. The physical properties of supramolecular peptide assemblies: from building block association to technological applications. *Chem. Soc. Rev.* **43**, 6881-6893 (2014).
26. de Groot N. S., Parella T., Aviles F. X., Vendrell J., Ventura S. Ile-Phe Dipeptide Self-Assembly: Clues to Amyloid Formation. *Biophys. J.* **92**, 1732-1741 (2007).
27. Garcia A. M.*, et al.* Nanoscale Assembly of Functional Peptides with Divergent Programming Elements. *ACS Nano* **15**, 3015-3025 (2021).
28. Bera S., Mondal S., Xue B., Shimon L. J. W., Cao Y., Gazit E. Rigid helical-like assemblies from a self-aggregating tripeptide. *Nat. Mater.* **18**, 503-509 (2019).
29. Abbas M., Lipiński W. P., Wang J., Spruijt E. Peptide-based coacervates as biomimetic protocells. *Chem. Soc. Rev.* **50**, 3690-3705 (2021).
30. Baruch Leshem A.*, et al.* Biomolecular condensates formed by designer minimalistic peptides. *Nat. Commun.* **14**, 421 (2023).
31. Hendricks M. P., Sato K., Palmer L. C., Stupp S. I. Supramolecular Assembly of Peptide Amphiphiles. *Acc. Chem. Res.* **50**, 2440-2448 (2017).
32. Zhang S. Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.* **21**, 1171-1178 (2003).
33. Ghadiri M. R., Granja J. R., Milligan R. A., McRee D. E., Khazanovich N. Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **366**, 324-327 (1993).
34. Hamley I. W. Self-assembly of amphiphilic peptides. *Soft Matter* **7**, 4122-4138 (2011).
35. Aida T., Meijer E. W., Stupp S. I. Functional supramolecular polymers. *Science* **335**, 813-817 (2012).
36. Levin A., Hakala T. A., Schnaider L., Bernardes G. J. L., Gazit E., Knowles T. P. J. Biomimetic peptide self-assembly for functional materials. *Nat. Rev. Chem.* **4**, 615-634 (2020).
37. Chavali S., Singh A. K., Santhanam B., Babu M. M. Amino acid homorepeats in proteins. *Nat. Rev. Chem.* **4**, 420-434 (2020).
38. Pattabiraman V. R., Bode J. W. Rethinking amide bond synthesis. *Nature* **480**, 471-479 (2011).

39. Johnson B. J. Synthesis, structure, and biological properties of sequential polypeptides. *J. Pharm. Sci*. **63**, 313-327 (1974).
40. Canavelli P., Islam S., Powner M. W. Peptide ligation by chemoselective aminonitrile coupling in water. *Nature* **571**, 546-549 (2019).
41. Ruff Y., Garavini V., Giuseppone N. Reversible Native Chemical Ligation: A Facile Access to Dynamic Covalent Peptides. *J. Am. Chem. Soc.* **136**, 6333-6339 (2014).
42. Miao X., Paikar A., Lerner B., Diskin-Posner Y., Shmul G., Semenov S. N. Kinetic Selection in the Out-of-Equilibrium Autocatalytic Reaction Networks that Produce Macrocyclic Peptides. *Angew. Chem. Int. Ed.* **60**, 20366-20375 (2021).
43. Merrifield R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149-2154 (1963).
44. Sauer F., Haas M., Sydow C., Siegle A. F., Lauer C. A., Trapp O. From amino acid mixtures to peptides in liquid sulphur dioxide on early Earth. *Nat. Commun.* **12**, 7182 (2021).
45. Stolar T., Grubešić S., Cindro N., Meštrović E., Užarević K., Hernández J. G. Mechanochemical Prebiotic Peptide Bond Formation. *Angew. Chem. Int. Ed.* **60**, 12727-12731 (2021).
46. Rodriguez-Garcia M*., et al.* Formation of oligopeptides in high yield under simple programmable conditions. *Nat. Commun.* **6**, 8385 (2015).
47. Campbell T. D., Febrian R., McCarthy J. T., Kleinschmidt H. E., Forsythe J. G., Bracher P. J. Prebiotic condensation through wet–dry cycling regulated by deliquescence. *Nat. Commun.* **10**, 4508 (2019).
48. Doppleb O*., et al.* Determining the Diastereoselectivity of the Formation of Dipeptidonucleotides by NMR Spectroscopy. *Chem. Eur. J.* **27**, 13544-13551 (2021).
49. Biron J.-P., Pascal R. Amino Acid N-Carboxyanhydrides: Activated Peptide Monomers Behaving as Phosphate-Activating Agents in Aqueous Solution. *J. Am. Chem. Soc.* **126**, 9198-9199 (2004).
50. Roberts S. J., Liu Z., Sutherland J. D. Potentially Prebiotic Synthesis of Aminoacyl-RNA via a Bridging Phosphoramidate-Ester Intermediate. *J. Am. Chem. Soc.* **144**, 4254-4259 (2022).
51. Dhiman R. S., Opinska L. G., Kluger R. Biomimetic peptide bond formation in water with aminoacyl phosphate esters. *Org. Biomol. Chem.* **9**, 5645-5647 (2011).
52. Gibard C., Bhowmik S., Karki M., Kim E.-K., Krishnamurthy R. Phosphorylation, oligomerization and self-assembly in water under potential prebiotic conditions. *Nat. Chem.* **10**, 212-217 (2018).
53. van der Helm M. P., Wang C.-L., Fan B., Macchione M., Mendes E., Eelkema R. Organocatalytic Control over a Fuel-Driven Transient-Esterification Network. *Angew. Chem. Int. Ed.* **59**, 20604-20611 (2020).
54. Pappas C. G*., et al.* Spontaneous Aminolytic Cyclization and Self-Assembly of Dipeptide Methyl Esters in Water. *ChemSystemsChem* **2**, e2000013 (2020).
55. Morris K. L*., et al.* Chemically programmed self-sorting of gelator networks. *Nat. Commun.* **4**, 1480 (2013).
56. Kwiatkowski W., Bomba R., Afanasyev P., Boehringer D., Riek R., Greenwald J. Prebiotic Peptide Synthesis and Spontaneous Amyloid Formation Inside a Proto-Cellular Compartment. *Angew. Chem. Int. Ed.* **60**, 5561-5568 (2021).
57. Muchowska K. B., Varma S. J., Moran J. Nonenzymatic Metabolic Reactions and Life's Origins. *Chem. Rev.* **120**, 7708-7744 (2020).
58. Nečas D., & Klapetek, P. Gwyddion: an open-source software for SPM data analysis. *Cent. Eur. J. Phys.* **10**, 181-188 (2012).