ChemSystemsChem

Review
doi.org/10.1002/syst.202200040

Chemistry Europe
European Chemical Societies Publishing

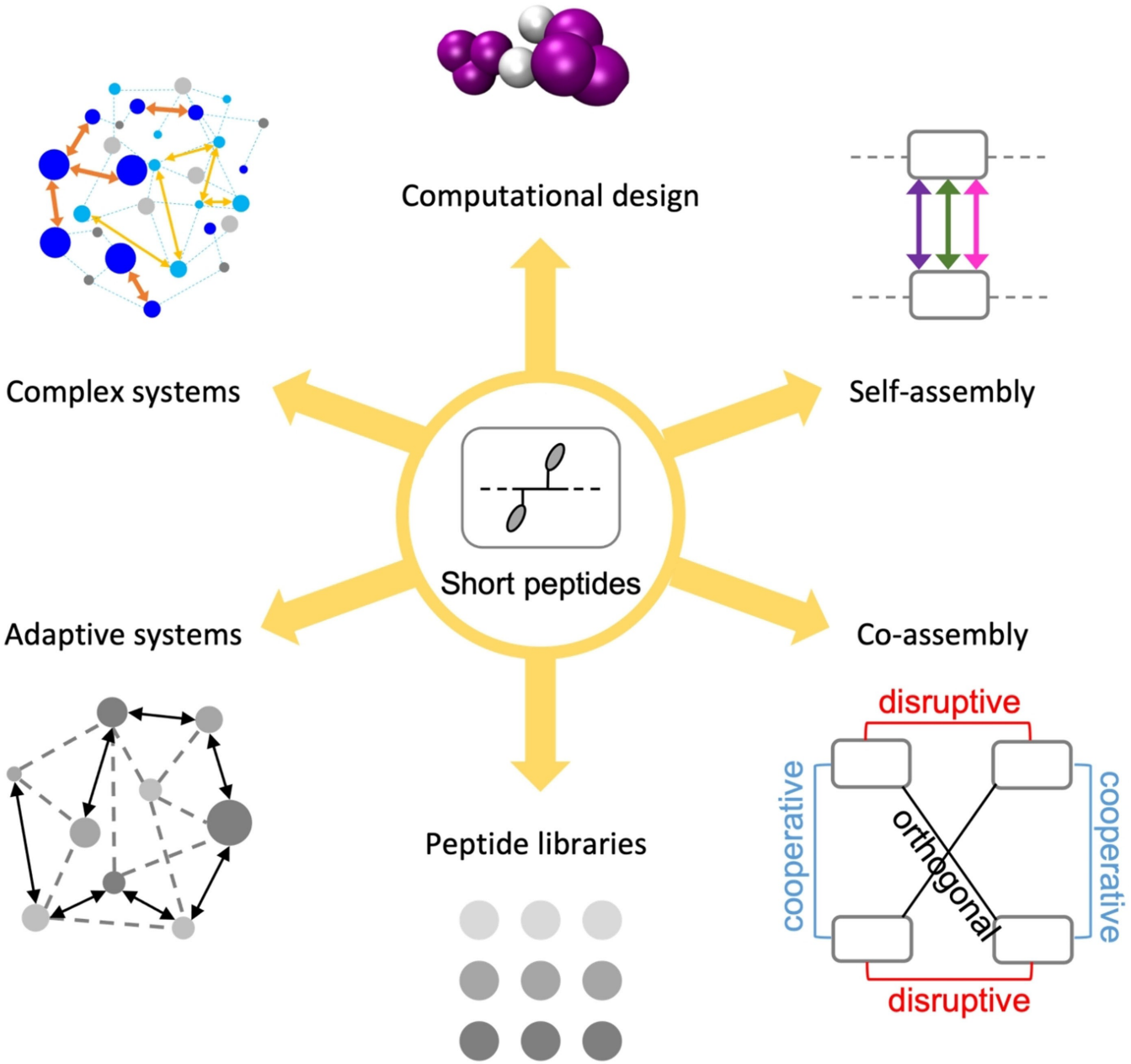
www.chemsystemschem.org

VIP *Very Important Paper*

Special Collection

Designed Complex Peptide-Based Adaptive Systems: A Bottom-Up Approach

Salma Kassem*[a] and Rein V. Ulijn*[a, b, c]



Systems chemistry represents a shift from studying molecules in isolation to studying the collective behavior of mixtures or ensembles of reacting and interacting molecules. This research direction provides new ways to build and understand collective chemical properties and emergent functions that are inaccessible using conventional, reductionist, chemistry approaches. This field exemplifies fundamental connections between chemistry and biology, and has the potential to give rise to completely new ways for chemists to approach and resolve wide-ranging issues related to all aspects of living systems, that themselves are complex systems. Short peptides, while simple in design, carry rich and versatile information encoded in sequence specific side-chain interactions, that makes it possible to access a high level of complex, information rich systems with emergent properties from Life's simplified building blocks. This review focuses on recent developments in the synthesis of complex adaptive systems based on peptides, where multiple components are designed to cooperatively interact, react and collectively adapt to their context.

1. Introduction

Complex systems are collections of interacting components, that are inter-connected, with unique properties and functionalities that are dictated not only by the nature of the components, but also by their collective dynamics.[1] In particular, adaptive complex systems adapt to their environments and respond to changes by re-organizing and amplifying certain species, to survive and sustain their structure after change-inducing events.[2] Defined as such, it becomes clear that complex adaptive systems are ubiquitous and fundamental to the functioning of cells, organs, organisms, populations and the planet. One could argue that considering the natural world with a complex systems view is critical to understanding and managing all aspects of human and planetary life, including avoiding unintended or unexpected consequences of interventions. At the most fundamental level, this understanding is one of mixtures of reacting and interacting molecules. Chemistry, the central science, is therefore key to enhancing understanding in this field and especially the emerging field of Systems Chemistry,[3] as it provides opportunities to synthesize complexity systematically and analyze how components' interactions and reactions collectively give rise to new structures, properties, and functionalities.

The evolving properties of systems are not easily rationalized by individual analysis of their components because even with a few interacting elements, the number of variables quickly becomes large, as these interactions are typically linked and inter-dependent.[4] The attributes of dynamic complex systems are therefore, by definition, not well grasped by intuitive and traditional chemistry approaches that focus on the final outcome or product rather than the understanding and analysis of the evolving collective behavior of mixtures. By taking inspiration from biology, and with the advances in the field of supramolecular chemistry, chemists have for decades started to embrace these aspects, for instance, by focusing on non-equilibrium systems,[5] and studying multi-component[6] and heterogenous mixtures[7] instead of focusing on pure solids, liquids or solutions. Despite this progress, chemistry as a field still has far to go in embracing a complex and adaptive system view, where mixtures of components evolve through learning, and are able to adapt to their environment by constantly re-organizing and modifying their composition and encoded information in response to feedback.[8]

From a systems' point of view, self-assembly can be considered as an emergent property of an ensemble of interacting monomers. Short peptides, while being simplified versions of Nature's building blocks, are minimalistic yet modular and versatile motifs that have been used in self-assembly for decades.[9] The 20 canonical amino acids cover a comprehensive range of the chemical space at their side chains through which tunable non-covalent interactions encode sequence dependent information. Many of the reported examples provide deep theoretical and experimental understandings in the combination and balancing of inter- and intra-molecular interactions and their effects on self-assembly[10–18] and materials properties.[19–25]

In order to design a complex adaptive system, we must first start by having a fundamental understanding of the simple components that make up the system. It is easy to see how lessons from peptide interactions and self-assembly could be valuable for the design and development of multi-component networks with dynamic and flexible connections. It is exciting to consider the development of designed systems with evolving peptide sequence compositions where new information (and consequently, functionality) is spontaneously created in response to feedback and adaptation, and truly capitalize on the richness and complexity that is accessible through the amino-acid side-chain interaction space (Figure 1). In this review, we present how adhering to Feynman's concept of "*what I cannot create I do not understand*" and "*bottom-up*" approach[26] provides a promising guideline to designing peptide-based complex adaptive systems. We first review some important papers that explore the peptide sequence space and the effects of balancing combinatorial and competitive interactions on the self-assembly and co-assembly of short peptides. We then

[a] *Dr. S. Kassem, Prof. Dr. R. V. Ulijn*
Nanoscience Initiative at Advanced Science Research Center (ASRC)
The Graduate Center of the City University of New York
85 St. Nicholas Terrace, New York, NY 10031 (USA)
E-mail: skassem@gc.cuny.edu
rulijn@gc.cuny.edu
[b] *Prof. Dr. R. V. Ulijn*
Ph.D. Programs in Biochemistry and Chemistry
The Graduate Center of the City University of New York
365 Fifth Avenue, New York, NY 10016 USA
[c] *Prof. Dr. R. V. Ulijn*
Department of Chemistry Hunter College
City University of New York
695 Park Avenue, New York, NY 10065 (USA)

Special Collection *An invited contribution to a Special Collection on Systems Chemistry in the USA.*

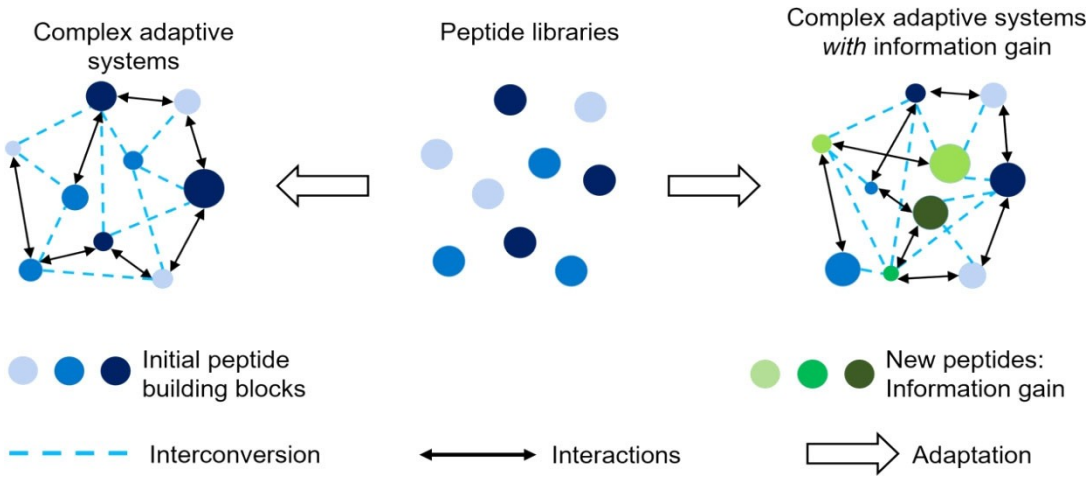


Figure 1. Peptide systems. Complex adaptive systems are composed of peptide libraries (middle), where components interact with each other and are interconvertible (left and right). Left: interconversion re-organizes pre-existing peptide building blocks that interact with each other but no new sequences are formed. Right: systems where new peptides are synthesized from the starting building blocks in a dynamic and reversible way can access new sequence-encoded information. In both cases, the distribution of components, which is represented by different size circles to reflect differential abundance, can adapt to changes in conditions and environment through learning and amplification of the most stable structures.

present different approaches for designing combinatorial peptide libraries and highlight examples that embrace the use of collections of short peptides to build, from the bottom-up, adaptive systems and complex networks where the encoded information is enriched. Finally, we provide our perspective on future directions.

2. Computational Approaches for Screening Peptide Spaces

The amino-acid interaction space is wide and versatile with multiple possible interactions (electrostatic, π-π, van der Waals) and sites (side chains, termini, amide back-bone), making peptides composed of just a few amino-acids (from the 20 canonical amino-acids) an ideal platform for the systematic study of combinatorial dynamic interactions (Figure 2). Despite the plethora of possibilities and combinations, computational approaches can streamline the design process by providing design principles/guidelines as long as parameters and descriptors geared for the desired properties are set. The following examples showcase computational approaches in search of side-chain interaction dictated properties.

In a long-standing collaboration with Tell Tuttle we computationally assessed the self-assembly propensities of all 400 dipeptides[10] and later all 8,000 tripeptides[11] using a coarse-grained molecular dynamic (MD) simulation (Martini).[27] While in this coarse-graining approach, where every 3–4 atoms are replaced with a 'bead' to represent collective interactions, atomistic detail and subtleties around non-covalent interactions are clearly lost, it has a major advantage of speeding up these simulations. Aggregation propensity scores were calculated for all 8,000 tripeptides and allowed us to extract some simple rules

Salma Kassem is a postdoctoral research associate in Prof. Rein Ulijn's group at the Advanced Science Research Center (ASRC) at the Graduate Center of the City University of New York (CUNY). She received her BSc and MSc in Chemistry from The University of Geneva and obtained a PhD from The University of Manchester working on artificial molecular machines. Her work is focused on developing complex and adaptive peptide systems as well as the design of novel peptide motifs for self-assembly and supramolecular materials design.

Rein Ulijn's lab is trying to figure out how the molecular building blocks, design and discovery concepts of living systems can be simplified and repurposed to produce materials, systems, and molecular technologies with adaptive functionalities that cannot be achieved using existing chemical design approaches. He is founding Director of the Nanoscience Initiative at the Advanced Science Research Center at CUNY, New York. Prof. Ulijn has held several personal fellowships and has won a number of awards, including the Vannevar Bush Faculty Fellowship, the RSC Norman Heatley Medal, Royal Society Merit Award, and was elected as a Fellow of the Royal Society of Edinburgh. He is the Einstein Professor of Chemistry of Hunter College at CUNY.

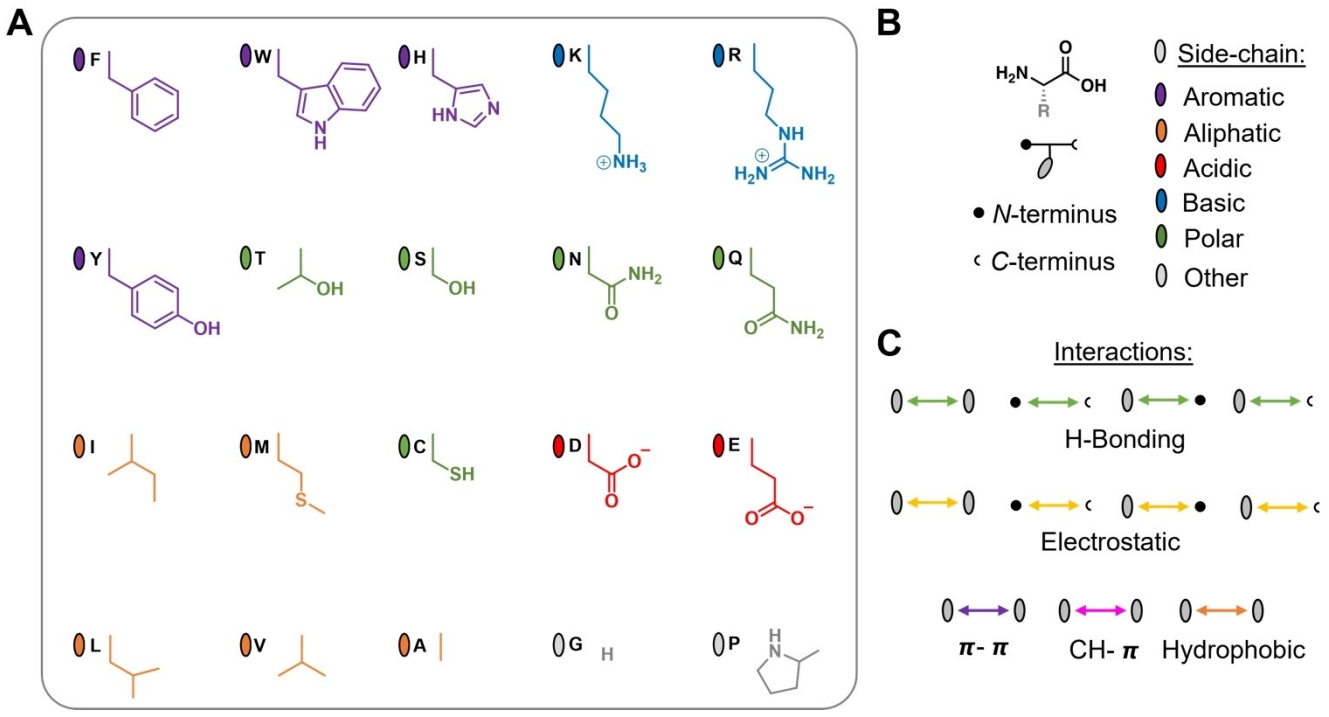


Figure 2. Amino-acid side chains and interactions. A. Chemical structure of the 20 canonical amino-acid side chains. **B.** Color-coded classification of the amino-acids according to their properties. **C.** Color-coded representation of the amino-acid interaction space at different possible interaction sites. Double headed arrows represent an interaction between two moieties, color-coded depending on the nature of interaction.

for tripeptide self-assembly, such that combining two (paired) aromatics and coupling charges from side chains and termini favor self-assembly. This understanding of sequence-encoded properties led to the discovery of a number of short peptides that give rise to hydrogel formation, such as KYF, and others that could be used as emulsifiers, including FFD.[28] Recently, Tang *et al.* also used a Martini coarse-grained molecular dynamic simulation model to predict the liquid-liquid phase separation (LLPS) behavior of the dipeptide sequence space.[12] In addition to the aggregation capabilities of dipeptides, they introduced a new fluidity parameter to calculate LLPS scores of all 400 dipeptides and were able to classify phase-separating dipeptides into three classes depending on their composition: (i) dipeptides containing Y and a hydrophobic residue (L/I/V/C); (ii) dipeptides composed of W and a polar residue (Q/N) or slightly hydrophobic residue (A/G); and (iii) dipeptides with a large dipole moment (KE/KD/RE/RD).

As can be expected, predicting self-assembly of longer peptides or even co-assembly of more complex (dynamic) mixtures will require exhaustive enumerations of simulations and seems almost impractical from a time/cost perspective. Therefore, there has been a growing effort in implementing machine learning algorithms in the discovery of functional peptides, specifically active learning methods.[29] Active, or sequential, learning is a systematic data-driven screening approach where all data collected to date is used to inform the "next-best" measurement or candidates to screen. In this way, an iterative and specific order in which to screen peptide sequences, or potential candidates, can be established, while only performing MD simulations on a fraction of the sequence space.

Van Teijlingen *et al.* developed a two-step active learning process to identify self-assembling water-soluble peptides with the highest aggregation propensities in the tetra-, penta- and hexa-peptide sequence space.[30] The two-step method relies on, first, a low-resolution prescreen which searches the entirety of the sequence space and selects the peptides with the highest AP scores based on only 10 parameters. The long list of potential candidates then undergoes a second-step of high-resolution screening which considers a more extensive list of parameters and uses active learning to improve both low and high-resolution models in search of the sequences with the highest self-assembling propensities. They were also able to focus the search on water-soluble aggregating peptide and differentiate between insoluble peptides, highly aggregating ones, which typically contain F/Y/W, and peptides that form hydrogels, by using restricted data sets with a defined log P limit, showing that their active learning model can be tuned to search for peptides (and beyond) with specified properties.

Experimental findings also need to be considered to encompass the adaptability and collective behavior of systems and correct for approximations and simplifications inherent to the high-throughput computational processes. Shmilovich *et al.* used a data-driven hybrid computational/experimental active learning approach to design π-conjugated peptides that self-assemble into pseudo-1D nanoaggregates.[31] They had previously reported an active learning approach to study the self-assembly of π-conjugated peptides where an electronically

active core (1,4-distyrylbenzene) is flanked by DXXX- and -XXXD tetrapeptide wings and were able to identify sequences with the most promising stacking efficiency in the aggregates while only running simulations on 2.3% of the search space.[32] In their hybrid approach, they integrated experimental feedback into their data-driven search to construct a sequence-property model as an active learning component, informing their next selection of the candidates for computational and experimental screening. Such hybrid and systematic approaches offer a much deeper and exhaustive exploration of the sequence space.

One of the main challenges for systems chemists is how to design dynamic combinatorial peptide libraries for supramolecular systems, starting from a minimalistic pool of diverse components, and analyze the arising diversity and complexity. Whereas there are many strategies for creating peptide libraries for fitness screening,[33] including phage display,[34] dynamic peptide libraries differ from conventional combinatorial high-throughput libraries in that the starting components are designed, and the focus is on the collective behavior rather than finding the best performing candidate. Kalafatovic's group developed an algorithm-supported approach to design peptide libraries that are diversity-focused rather than size-oriented.[35] Their methodology was developed with the anticipated complex task of analysis in mind, which for peptide identification and sequence deconvolution is typically liquid chromatography coupled mass spectrometry (LCMS). They used an evolutionary computing approach based on a multi-objective genetic algorithm to rationally design a chemically diverse peptide library that maintains chemical diversity while minimizing mass redundancy for simplified analysis. The composition of the library, peptide length and variable positions, are user-defined and can be extended to comprise skeletal, structural and stereo-chemical modifications. Their method showed that genetic algorithms can optimize the library search with the highest number of peptides based on multiple objectives such as chemical and mass diversity. Their system could be extended to other search objectives and their combinations as they've later shown by introducing a sequence-based property evaluation step, such as polarity, hydrophobicity and/or H-bonding.[36] Their work opens up the possibilities of user-defined, rationally designed peptide libraries by exploring the sequence space in an informed and targeted way.

These systematic learning models allow for a faster and more efficient screening of large sequence spaces to discover molecules with desired functions and properties. As they generally look at static systems after equilibration, although, in some cases, certain information about the dynamics of the system can be extracted from the trajectories of the simulations, they can provide the starting point when designing complex dynamic systems. An inherent challenge in this context however, is the lack of quality training sets, that are critical to the success of machine learning in other contexts, *e.g.* the impressive predictive protein folding from sequence by Alphafold, which used the entire protein data bank as a training set.[37]

3. Self-Assembly of Short Peptides

3.1. Balance of interactions in self-assembling short-peptides

The field of peptide self-assembly has been strongly influenced by protein secondary structures, with a focus on back-bone interactions in beta sheet,[38] alpha helix[39] or collagen like[40] folds. For adaptive systems, there may be advantages in focusing more on side chain interactions instead of backbone folds, inspired less by protein folding, but by protein-protein interactions and dynamic features, such as induced fit binding. In this context, it is interesting to consider how to design side-chain interactions to encode specific properties and how the balance of interactions moderates the systems' behavior on different size scales (Figure 3).

Because of the modularity of their design and synthesis, simple changes in peptide sequence can be made to appropriately position amino acids for certain side chain interactions. There have been a number of examples where a supramolecular assembly motif, the so-called F/F zipper,[41] which promotes self-assembly through highly organized π-π stacking, was engineered to influence assembly. Marchesan's group demonstrated how a simple factor, such as amino acid chirality, can dictate the self-assembly propensity and morphology of FF containing tripeptides.[13] They showed that the inversion of chirality of the *N*-terminus amino acids of non-assembling homochiral tripeptides VFF, FFV and LFF leads to the formation of hydrogels by reducing steric hinderance and allowing the formation of extended F/F zippers (Figure 3a). Later, they demonstrated that inversion of chirality at different positions of tripeptide sequences can have very contrasting effects on nanostructure and morphology.[14] They were able to use this strategy to resolve competing self-assembly instructions in tripeptides and control intermolecular interactions and gelation.[15] By investigating the differences in morphologies and nanostructures of the eight stereoisomers of PFF, which is the most aggregating tripeptide according to our 2015 study[11] and was found to assemble in unusual helical motifs,[24] they showed that amino acid chirality is the decisive factor in determining the domination between competing instructions from β-breaker P and β-forming FF. Inversion in P chirality led to the formation of a gel while monodisperse spherical nanoparticles were formed when the chirality of the central F was inverted and non-gelling nanotapes were obtained with the inversion of the *C*-terminus F. The contrasting differences in nanostructures were elucidated by all-atom MD simulations which showed that the inversion of P chirality introduces a kink in the peptide's backbone allowing it to form parallel stacks with segregated hydrophilic (P) and hydrophobic (F) faces. While the inversion of chirality of the central and terminal F also resulted in the formation of stacks, these stereoconfigurations did not cause a net segregation between the hydrophilic and hydrophobic faces, leading only to short-range self-organization patterns. The predominance of the self-assembling instruction, is therefore different for each stereoisomer and affects the peptide's ability to form short, medium and long-range assemblies.

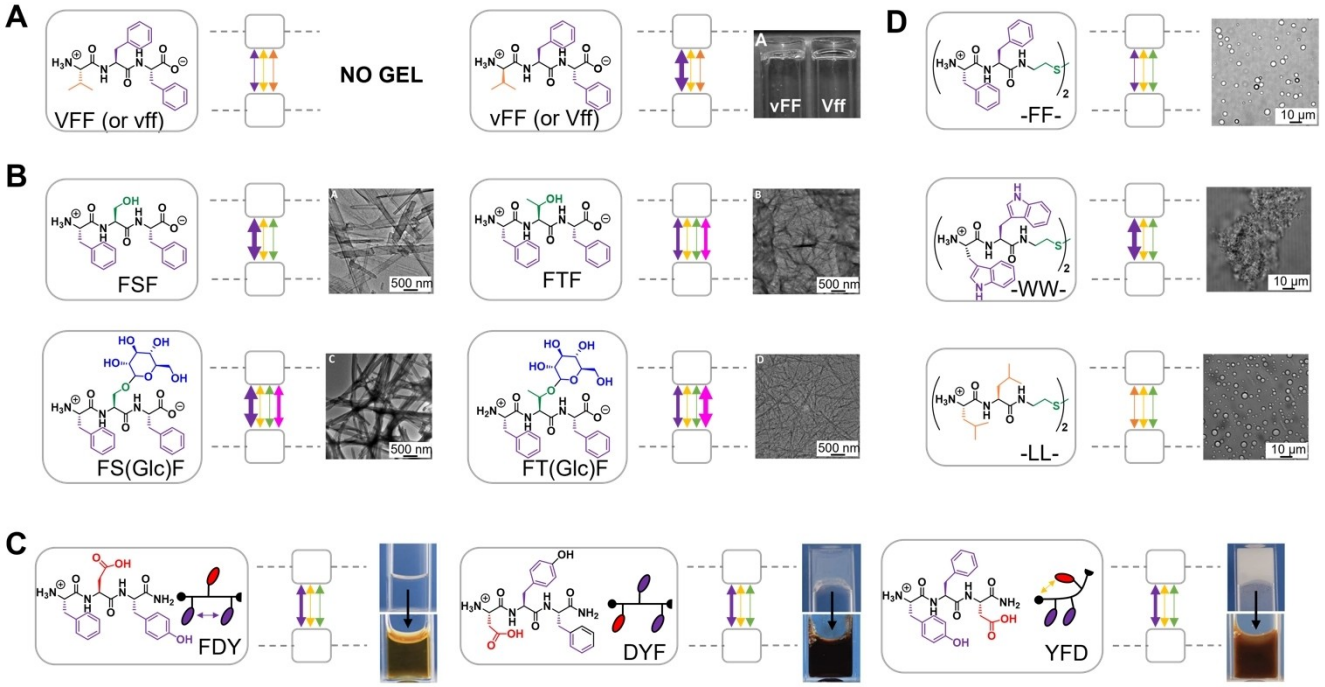


Figure 3. Balancing side-chain interactions in short-peptide self-assembly. Chemical structures, simplified inter-molecular interactions representation and microscopic/ macroscopic characterization of a selection of self-assembling short-peptides where balance of side-chain interactions dictates material properties. Interactions are represented by double headed arrows, color-coded according to figure 1, between components (grey boxes). A) Chirality dictates gelation in tripeptides containing FF. Inversion of chirality at the *N*-terminus determines the extent of π-π stacking and aromatic interactions in V/F/F tripeptides. Images of self-supporting hydrogels of vFF and Vff. Reproduced with permission.[14] Copyright 2015, the Royal Society of Chemistry. B) CH-π interactions compete with π-π stacking and affect fiber thickness with the introduction of T or *O*-glycosylation as shown in the TEM images. Reproduced with permission.[16] Copyright 2021, American Chemical Society. C. Intra-molecular interactions dictate peptide conformations which are retained in the supramolecular assemblies and determine the order/disorder balance of the assembly, which in turn controls the degree of polymerization upon enzymatic oxidation of Y. Macroscopic images from unoxidized peptide materials (top) and after 24 h of oxidation (bottom). Reproduced with permission.[25] Copyright 2017, The American Association for the Advancement of Science. D. Balance of aggregating and polar interactions in self-coacervating peptides. -FF- and -LL- peptides form coacervates while the aggregating interactions overpower the self-coacervation propensity in -WW- peptide. Microscopic images of solutions. Reproduced with permission.[18] Copyright 2021, Springer Nature.

Rappaport's group showed that the side chains in the FF motif can be separated by a hydrophilic amino acid (FXF) yet still promote self-assembly through aromatic side chain stacking.[42] Tripeptides FTF, FCF, FEF and FKF all showed evidence of F/F zippers that macroscopically extended to the formation of hydrogels. To accommodate for optimal π-stacking, the peptides adopt a conformation where the central amino acids remain exposed to water and is amenable to alternative types of interactions and further functionalization. In collaboration with Tuttle, Pires and Pashkuleva, we recently used the FXF platform as a minimalistic model to expand the amino-acid interaction space by studying the effect of *O*-glycosylation on self-assembly.[16] Glycosylation is known to cause changes in protein aggregation and bioactivity but the underlying mechanisms and interactions are poorly understood.[43] Minimalistic models such as these glycotripeptides provide valuable insights on the rebalance of interactions at play. FSF and FTF were chosen to allow covalent attachment of a carbohydrate moiety (glucose) to the central amino acid *via* *O*- linkage. While the tripeptides and their glycosylated analogs were all expected to self-assemble through the formation of F/F zippers, the expansion of the interaction space resulted in sequence specific changes in morphology and mechanical properties of the resulting materials (Figure 3b). CH-π interactions between the T methyl group and the F aromatic group were found, both computationally and experimentally, to disrupt the F/F zipper-driven aggregation. This resulted in changes in self-assembly morphology from nanotapes in the case of FSF to nanofibrils for FTF. The effect of diversification of supramolecular interactions on self-assembly was even more pronounced with the glycosylation of the peptides. MD simulation showed that the carbohydrate moieties are present both at the solvent exposed surface as well as the core, meaning that the conformational space of the glycosylated peptides is expanded compared to the non-glycosylated peptides to accommodate for additional CH-π stabilizing interactions. Beyond serving as a minimalistic model for understanding the rebalancing of interactions in glycoproteins, this work also highlights how the expansion of the peptide space can introduce further handles and controls on supramolecular interactions in a network context.

We've demonstrated that the self-assembled morphology of sequence isomers of *C*-terminal amidated tripeptides composed of D, Y and F, dramatically changes depending on sequence.[25] Aggregates, gels and extended crystalline fibers were formed from solutions of peptides with paired aromatic amino acids

(DFY, DYF, YFD, FYD) whereas those where aromatic residues are separated by an aspartic acid, remained clear solutions (FDY, YDF). Crystallographic data as well as FT-IR and molecular dynamic simulations showed that the positioning of D plays an important role in determining the relative orientation of the two aromatic units (*syn* for FDY and *anti* for DYF) as well as the extent and contribution of intra-molecular salt-bridge formation (D at the C-terminus in YFD) (Figure 3c). The sequence specific balance of intra-molecular π-π (side chains) and electrostatic (side-chain and termini salt-bridges) interactions dictates the preferred conformation of each peptide which is translated in the extant of crystal packing and inter-molecular π-stacking in the self-assembled structures. The supramolecular order of the assembled peptides in turns dictates the degree of polymerization that occurs upon the enzymatic oxidation of Y. This resulted in materials with strikingly different properties, such as UV absorbance, storage capacity, color, morphology and nanostructure, obtained from short-peptides composed of the same amino acids and showcases how sequence specific combinations of interactions and their balance, can have cascading effects on different size scales.

The examples above show how the information encoded in the sequence of short-peptides balances inter-molecular/component interactions from side-chain, termini and backbone to produce materials and systems with sequence-dictated properties. The designs rely on strong stable interactions that result in ordered nanostructures. Until the major advances in the field of intrinsically disordered proteins (IDPs),[44] disordered structures had historically been overlooked and regarded as unfunctional as they typically arise from weaker, therefore more dynamic, side-chain interactions. Nature takes advantage of these attributes to promote intracellular compartmentalization through LLPS and the formation of coacervates.[45] The discovery and characterization of IDPs have inspired efforts to develop minimalistic versions and establish design rules for short-peptide LLPS.

In ground breaking-work, Martin *et al.* proposed a design based on the sticker-spacer model,[46] where hydrophobic interactions (the stickers), that typically lead to ordered and solid structures, are counter-balanced with hydrophilic components (the spacers) to produce disordered liquid condensates instead.[17] This model was later used by Abbas *et al.* to design a short peptide synthon that undergoes self-coacervation, composed of two hydrophobic dipeptide stickers connected through a flexible hydrophilic spacer.[18] In a peptide made of two FF units linked by a disulfide bond through an appended cystamine moiety, they showed that π-stacking interactions that typically lead to the formation of F/F zippers can be counter-balanced by the flexibility and polarity of the cystamine moiety to produce micrometer-sized liquid droplets that are stable in a range of environmental conditions. The hydrophilic spacers are able to solubilize the otherwise aggregating FF units and maintain them in a hydrated liquid state. They expanded their design to demonstrate that the approach can be generalized to dimers of aromatic and aliphatic dipeptides (LL, LF and FL) and different polar spacers while showing that a mis-match in balance can occur with strongly aggregating sequences (WF, FW and WW) and/or poorly solubilizing apolar spacers (Figure 3d).

These fundamental understandings, experimentally and computationally, are crucial in laying the groundwork for designing and building complexity in multi-component peptide systems and highlight how subtle changes in composition and positioning of amino acids can be leveraged when designing complex systems.

3.2. Multicomponent assemblies

Over the past few decades, the tendency in self-assembly of functional nanostructures is to incorporate additional features into molecular building blocks to achieve functionality, so resulting in more complex building blocks. For example, recent examples include impressive systems with targeting, imaging, catalysts, enzyme-responsive and self-assembling molecules incorporated into each building block.[47] A fundamental aspect of systems thinking is that the property of a mixture is not a mere sum of its components. With this mind set, a logical alternative approach is to instead explore multi-component co-assembly to investigate when mixtures of molecules gain sufficient complexity to display emergent functions. Strategies for peptide-peptide co-assembly have been previously reviewed,[48] here we highlight a few that harness complimentary interactions to access new functionalities and self-assembly modes, and provide some insights into design rules for expanding on single components' interaction space.

By studying the co-assembly of 4 peptide amphiphiles (Pyrene-YL (Py-YL), Py-S, fluorenylmethoxycarbonyl-YL (Fmoc-YL) and Fmoc-S) in different two-component combinations, we were able to extract key features of peptide co-assembly modes, depending on similarities and differences in their interaction space (Figure 4A).[49] Surfactants Py-S and Fmoc-S mainly assemble through aromatic stacking while hydrogen bonding interactions in the YL peptides reinforce the π-stacking structures with β-sheet-type arrangements and result in the formation of gels. By mixing the peptides in different combinations, three distinct co-assembly modes were observed: (I) segregated and orthogonal assemblies in structurally different peptides (Py-YL/Fmoc-S and Fmoc-YL/Py-S); (II) cooperative co-assembly for structurally similar peptides (Py-YL/Fmoc-YL and Py-S/Fmoc-S); and (III) compromised and disrupted self-assembly for peptides that have similar π-stacking abilities, but differential H-bonding interactions (Py-YL/Py-S and Fmoc-YL/Fmoc-S). These rules allowed us to design a system where a new functional hydrogel is obtained from co-assembly of a functional (GHK) and structural (FFD) tripeptides (Figure 4B).[50] GHK is an extracellular matrix derived peptide that has been considered as an active ingredient in cosmetic products, and was chosen for its copper-binding abilities. FFD was chosen for its ability to self-assemble and form defined structures as well as bearing a complimentary charge to GHK. The peptides do not form self-sustainable hydrogels at neutral pH neither individually nor when combined in equimolar ratio. When Cu ions are introduced, the FFD/GHK mixture spontaneously forms

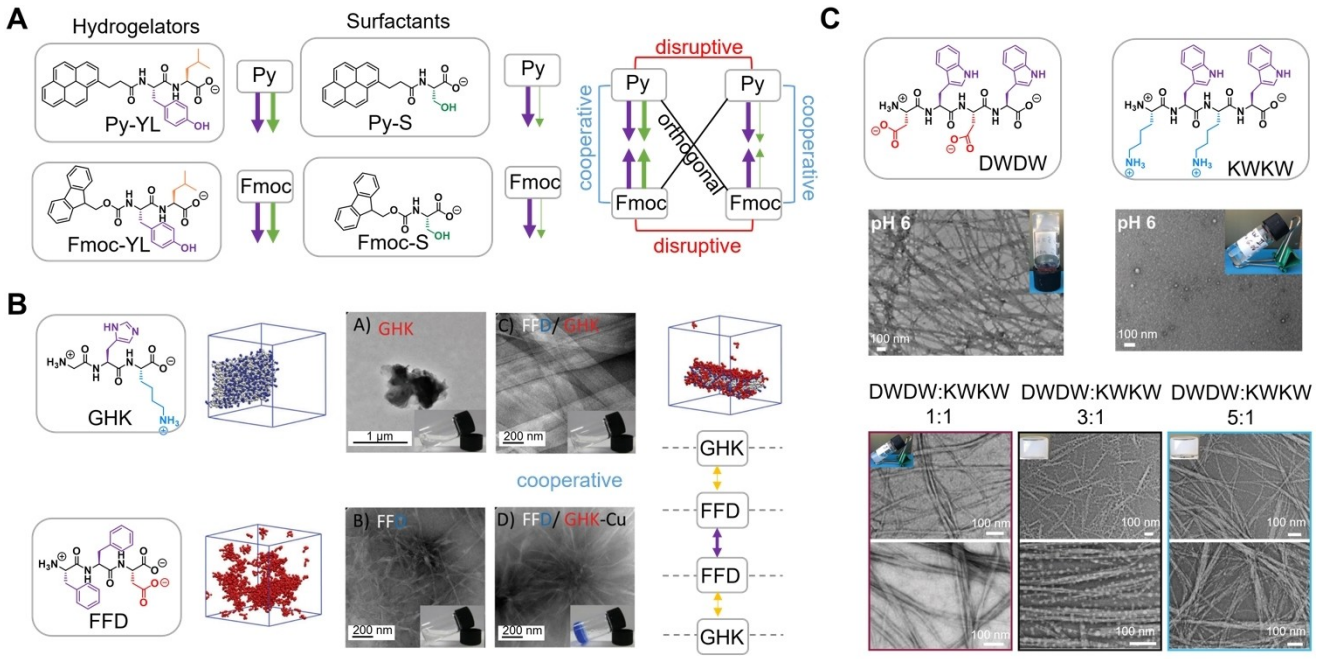


Figure 4. Multi-components co-assembly of short-peptides. Chemical structures, simplified inter-molecular interactions representation and microscopic/macroscopic characterization of a selection of co-assembling short-peptide where the nature of side-chain interactions dictates the mode of co-assembly. Interaction are represented by double headed arrows, color-coded according to figure 1, between components (grey boxes). A) Differential co-assembly modes of hydrogelators Py-YL and Fmoc-YL and surfactants Py–S and Fmoc-S depend on similarity/differences in chemical structure and cooperative self-assembling interactions. Adapted from ref. 49. B) Cooperative co-assembly of tripeptides GHK and FFD in the absence and presence of Cu ions. TEM and macroscopic images showing morphological changes upon co-assembly. Final frame (~9.6 μs) of computational time course of the self-assembly simulations of the tripeptides in isolation and co-assembled. Reproduced with permission.[50] Copyright 2017, the Royal Society of Chemistry. C) Co-assembly of charge complimentary tetrapeptides DWDW and KWKW. TEM and macroscopic images showing the nanostructures of each peptide individually and co-assembles structures at different molar ratios resulting in orthogonal and cooperative assemblies. Reproduced with permission.[51] Copyright 2019, the Royal Society of Chemistry.

a hydrogel. Coarse-grained MD simulations revealed that FFD forms a bilayer structure, with a hydrophobic core (FF) and negatively charged hydrophilic (D) surface. Complimentary side-chain electrostatic interactions result in a cooperative co-assembly structure with GHK where GHK peptides are organized on the surface of FFD. It is worth noting that this process is sequence-specific, as no hydrogel formation was observed with DFF or FDF tripeptides.

Complimentary electrostatic interactions do not always result in cooperative co-assemblies. This depends on the strength and balance of competing interactions. Sahoo *et al.* showed that when mixed, tetrapeptides DWDW and KWKW self-sort and co-assemble orthogonally, each retaining their individual fibrillar and spherical nanostructures, respectively (Figure 4C).[51] DWDW forms a hydrogel, sustained by π-π stacking and β-sheet-like hydrogen bonding whereas KWKW forms less structured aggregates with no spectral signatures from β-sheets. Self-supporting hydrogels were obtained when the peptides were mixed at pH 6 in ratios of higher DWDW composition (3 : 1, 5 : 1). At a 1 : 1 ratio, a viscous solution was obtained instead. TEM micrographs showed that, in the gelling mixtures, the surface of DWDW filaments is decorated with spherical KWKW nanostructures indicating cooperative co-assembly. When the peptides are mixed in an equimolar ratio, the extent of surface association decreases and segregation of the two morphologies increases (orthogonal co-assembly). Extensive spectral and mechanical characterizations revealed that the co-assembly results in electrostatic surface association which reduces the charge repulsion between DWDW fibers and promotes nanofiber bundling. This results in an enhancement in stiffness and mechanical properties which can be tuned by varying the ratio and association of the two peptides.

Besides complementarity of interactions, crystalline inter-layer distancing was proposed by Gazit's group as an additional co-assembly mode determining factor.[52] Despite being chemically diverse, all 20 coded amino acids form layered assemblies stabilized by hydrogen bonds between α-amine and α-carboxyl groups of individual amino-acids. Depending on the interlayer distances, amino acids can be classified in three categories which correlate with their β-sheet forming propensities and are influenced by hydrophobicity, steric bulkiness, and folding. They studied the co-assembly of amino-acids from the same crystal packing subgroup (F/I and F/M) as well as from different subgroups (F/G and F/A) and found that comparable interlayer separation results in interactive co-assemblies yielding different structural architecture while the mismatch in interlayer distances from amino acids in different subgroups results in orthogonal self-sorted co-assemblies where each amino acid retains its properties. Individually, F forms elongated micron-long fibers, I and M large hexagonal crystals, G one-dimensional

long crystal and A flake-like aggregates. When mixed in equimolar ratios, the F/I mixture formed spherical structures not observed from individual amino-acid, indicating that the formation of a new architecture. Similarly, new flake-like aggregates were observed as a result of inter-component interactions in F/M mixtures. These results contrast the self-sorting and segregation of morphologies observed in F/G and F/A mixtures, where both parent morphologies were still obtained in the mixed systems indicating orthogonal co-assembly. This study offered an additional factor to consider when designing multi-component systems, based on not only chemical interactions, but also supramolecular ordering and nano-structure.

The systems and design rules described above, consider the co-assembly of two-components at a static, thermodynamically equilibrated stage. Examples where adaptive and dynamic morphological transitions occur in multi-component systems are still rare. Inostroza-Brito *et al.* described a dynamic co-assembly process of a positively charged peptide amphiphile ($C_{15}H_{31}$CONH-VVVAAAKKK-CONH$_2$) and a negatively charged Elastin-like polypeptide, consisting of simple repeats of penta-peptides (VPGXG, X = V/E).[53] The interactions between the two components, which can be assimilated to short-peptide interactions, are governed by surface electrostatic interactions as well as hydrophobic interactions between their aliphatic cores. When mixed together, initial electrostatic complexation causes a re-configuration and opening of the protein structure and immediately creates a dynamic hierarchical membrane. The non-equilibrium structure can be maintained for long periods of time and is responsive to perturbations and instabilities. Hydrophobic interactions then drive self-assembly at the interface which is displaced by the formation and growth of tubes resulting in tubular morphogenesis of the assembly. Here, the co-assembly process guides the protein to access different hierarchies and architectures producing novel complex and functional materials.

The balance between order and disorder does not have to be mutually exclusive. In biology, many multi-component condensates are not homogeneous and have internal solid or solid-like architectures that are key to their structural integrity and functioning.[45] We recently investigated how to interface and balance order and disorder in designed peptide-based condensates, based on R_9·ATP complex coacervation.[54] We designed a hybrid peptide, LVFFAR$_9$, that has an aggregating component, LVFFA, covalently linked to a coacervation component, R_9. By controlling the amount of hybrid peptide in the system, we were able to tune the balance between fully disordered systems, containing 5% or less of LVFFAR$_9$, and fully aggregated systems with 50% or more LVFFAR$_9$. In between these two extremes, we observed fiber nucleation and growth exclusively inside the coacervates and proved that fibers and droplets co-exist and have mutual effects on each other's properties. Often times, heterogeneous condensates are not stable and undergo morphological liquid-to-solid transitions, a process associated with neurodegenerative diseases. Remarkably, in our system, coacervates remained fluid through fiber growth and maturation, yet the process endowed them with enhanced mechanical properties such as increase in stiffness. The strategy of covalently linking the two order and disorder inducing motifs stabilizes the interface between coacervate and fiber for extended time periods. This approach can be extended to the dynamic covalent linking of amino-acids or peptides with different functionalities, which can provide a tunable chemical environment around the fibres, and enhance their stability and performance e.g. in catalytic reactions.

Multi-component examples are still conservative in terms of their complexity, typically not exceeding two components. There exists a significant untapped opportunity to study much more complex mixtures, of 10s or 100s of components, that would represent an information-rich interaction space, likely to be highly sensitive and responsive to perturbations. However it is noted that analysis of such systems poses major challenges, and would likely be driven by a significant computational component.

4. Complex Peptide Systems

A fundamental principle of complex adaptive systems is the interconversion of components and collective adaptability and self-replication. In complex adaptive systems found in the natural world, such as the brain, an additional aspect is that the system itself changes the nature of the components. For example, the strengthening of neuronal connections involves physical changes in the physiology in the relevant area of the brain. This concept can be studied using Chemistry, by enabling the formation of new molecules (new information) from the starting mixture and generating dynamic libraries where new sequences that were not part of the starting mixture can emerge and, ultimately, self-assemble to form new structures (new *stored* information). These newly created molecules display new functionalities and information that were not there to start with and therefore represent true molecular adaptation. The major developments in dynamic covalent chemistry have paved the way for the development of dynamic combinatorial peptide libraries where components combine to give rise to a new network of interacting and interconverting species.[55]

4.1. Combinatorial peptide libraries

Many strategies for the generation of combinatorial short-peptide libraries have been developed (Figure 5). The idea was first published by Sanders and co-workers in 2005 where they use hydrazone exchange to generate a library of macrocycles and supramolecular structures starting from dipeptide derivatives.[56] To facilitate hydrazone exchange, they modified the dipeptide PF to display a hydrazide moiety at the *C*-terminus and a protected aldehyde group at the *N*-terminus. Under acidic conditions, deprotection of the aldehyde and hydrazone condensation reactions take place and exchange between the monomer and longer oligomers generating a thermodynamically controlled distribution of macrocycles of

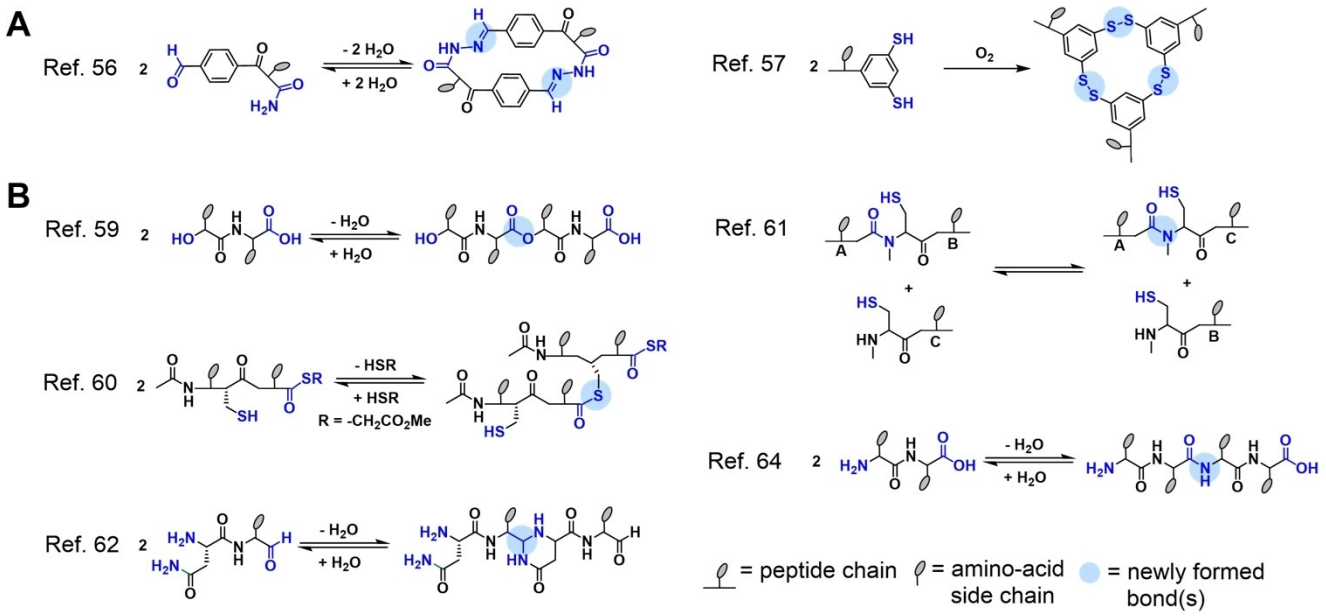


Figure 5. Strategies for combinatorial peptide libraries. A) Approaches relying on dynamic covalent chemistry. B) Approaches relying on dynamic amide bonds and substitutes.

different sizes with up to 6 newly formed dynamic covalent bonds.

Otto's group has pioneered the use of disulfide exchange to generate combinatorial libraries.[57] Peptides are modified at the *N*-terminus with an aromatic di-thiol moiety to facilitate disulfide formation under oxidative conditions, followed by reversible disulfide exchange, thus producing a mixture of macrocycles of varying sizes (up to 23-mers).[58]

The idea of generating diverse combinatorial peptide libraries based on changes in primary sequence, so by using amide bond exchange in unmodified peptides based on natural amino-acids, would provide compatibility of information generated with bio-relevant applications.

The intrinsic stability of the covalent peptide bond and the high kinetic barrier to amide bond formation has driven many groups to find dynamic substitutes that are still biologically relevant (Figure 5B). For instance, Hud's and Williams' groups have jointly used dynamic ester bond formation to facilitate the oligomerization of depsipeptides from dimers of amino acid-sand α-hydroxy amino acids using dehydration-hydration cycles.[59] Although the process does not occur in physiological conditions, they were able to produceoligomers with up to 8 heterodimer units with prolonged heating(dehydration phase) and show that the process is reversible by rehydration and hydrolysis. Ghosh *et al.* proposed the use of thioester exchange to generate thiodepsipeptide libraries from cysteine-containing tetrapeptides bearing a thioester at the *C*-terminus.[60] Thiol-thioester exchange occurs at neutral pH and generates dimeric macrocycles as the major product. Ruff *et al.* reported a dynamic Native Chemical Ligation (NCL) approach that uses *N*-(methyl) cysteine residues incorporated in the peptide sequences to facilitate amino-acid exchange between two peptides.[61] The exchange occurs at the *N*-(methyl) cysteine site through internal N→S acyl shift followed by trans-thioesterification and are efficient under physiological conditions for a variety of peptides of different lengths and compositions. Lynn and co-workers used reversible imine condensation to oligomerize peptides bearing a N at the *N*-terminus and an aldehyde modification at the *C*-terminus.[62] Imine condensation-cyclisation of the N-side chain onto the aldehyde terminus produces a mixture of linear and cyclic polymers, linked together with a tetrahydro-4-pyrimidinone unit.

The common goal of these approaches is finding an amide bond substitute that is reversible under physiological conditions, dynamic and requires minimum synthetic modifications. Our group demonstrated the ability for the amide bond itself to serve as a dynamic covalent bond-forming handle through the use of proteases that catalyze the formation and hydrolysis of peptide bonds.[63] In aqueous conditions, the thermodynamic equilibrium is largely in favor of the hydrolysis reaction making the exchange of peptide sequences a very inefficient process. We've shown however that driving forces, such as self-assembly[64] and side chain interactions,[65] can be designed to drive the equilibrium towards preferential formation of amide bonds and synthesis of products, making it possible to generate combinatorial peptides from unmodified amino-acids. To aid enzyme activation and selectivity, dipeptides bearing a hydrophobic group on the *N*-terminus are used as the site for peptide-bond formation and further trans-amidation reactions lead to the generation of a library of longer oligomers of varying sizes and compositions.

4.2. Adaptive peptide libraries

Combinatorial peptide libraries operate under thermodynamic control to generate divergent networks of interconverting and interacting molecules. Given their dynamic attribute, the thermodynamic distribution of products can be influenced by interactions between them, but also by changes in environmental conditions or by the inclusion of a guest molecule that preferentially binds only certain members of the dynamically exchanging network.

In pioneering work, Sanders and co-workers showed that the addition of acetylcholine to their hydrazone-exchange mediated library of macrocycles, templates the formation of a new [2]-catenane, which was not observed in its absence.[56] The addition of the guest molecule not only leads to the formation of a new species, it also shifts the product distribution as the [2]-catenane becomes the major species at the expense of the other library components.

In the systems developed by the Otto group, they leverage concepts from peptide self-assembly to drive the stacking and aggregation of their dynamic macrocycles. Peptide motifs decorate the macrocycle rings formed through dynamic disulfide bond exchange. By using β-sheet-forming sequences such as GLKLK,[57] peptide-peptide interactions drive the self-assembly of macrocycles and provide the thermodynamic drive that shifts the product distribution towards macrocycle sizes with optimum stacking stabilization, resulting in the formation of fibers. In this manner, the ring-size distribution becomes dictated by the extent of stabilization from non-covalent peptide-peptide interactions, while the peptide components remain compositionally unchanged. Given that all species are interconvertible, the self-assembling macrocycles become the predominant species and promote their own formation at the expense of other non-stacking species showing that peptide self-assembling interactions dictate product selectivity and distribution. They've demonstrated that their systems are highly dynamic and adaptive to several factors including mechanical cues[57] and spontaneously diversify depending on composition,[66] seeding and templating.[67]

Considerable progress has been made in establishing and expanding the systems chemistry toolbox to be able to encompass and analyze the complexity, properties and dynamics of emergent libraries. A number of reports from the Otto and Sanders groups provide a strong foundation to aid the understanding the relationship between library component amplification through guest-binding and the thermodynamic and numerical analysis of the binding process.[68] They developed a model to simulate and fit library product distribution to binding affinities in libraries containing up to more than 30 components without the need for individual separation of components or alteration of experimental conditions. Their methodology showed good correlation between experimental and simulated results in different systems including aggregation driven[69] and guest binding dynamic combinatorial libraries.[70] These detailed and extensive studies of the evolutionary dynamics in self-replicating systems reveals fundamental aspects pertaining to selection principles and self-replication that are considered relevant to the emergence of (synthetic) life using a bottom-up approach[6] and using systems chemistry for the discovery of novel dynamic sensors and receptors.

Only a handful of examples, described below, generate combinatorial libraries where sequence-encoded information and sequence-dictated functionality evolve with the chemical evolution of the system through covalent bond breakage/formation and non-covalent side-chain interactions. Such system are truly able to learn from the components' collective behavior and adapt their peptide composition to the evolving/changing conditions by providing feedback on the most stable structures to amplify.

In the work reported by Chen *et al.*, they describe the evolution of a dynamic chemical network that emerges through imine condensation driven polymerization of the dipeptide NF-CHO, and the changes in self-assembly and morphology that accompany the changes in chemical composition of the system (Figure 6a).[62] At each stage of the polymerization process, the distribution of components and peptide length determines the systems' morphology and drives the equilibrium towards the most stable structures. After an initial equilibration stage, analysis by TEM reveals the formation of nanoparticles corresponding to a distribution of linear and cyclic dimers, along with traces of monomers and linear and cyclic trimers. As the system evolves further, fibrous structures can be seen to nucleate inside the particles and eventually, the growth of fibers dominates the system. At this stage, the linear trimer is the predominant and almost exclusive species observed. Each morphological stage is characteristic of a specific distribution of the chemical network and evolves with the growth of peptide oligomers. The system can also adapt to the introduction of templating seeds, by amplifying the formation of sequences that propagate the seeding morphology.

We've shown that with our enzyme-mediated combinatorial peptide ligation approach, we are able to navigate energy landscapes of systems where open-ended chemical evolution, through amide bond formation from unprotected amino-acids, allows for a much broader accessibility of peptides varying in lengths and compositions.[64,65] Depending on the desired property, the "best" performing sequences prevail once the system reaches equilibrium, allowing us to discover new functional peptides and identify cooperative behaviors starting from simple short building blocks. This approach led to the selection of ionic complementary octapeptide mixtures, that self-assemble into antiparallel β-sheet rich fibers and lead to the *in situ* formation of gels, synthesized through a cascade of enzyme-mediated hydrolysis/condensation reactions starting from a tetrapeptide (FEFK).[64a]

We later extended this approach to explore a wider peptide sequence space, using homo- and heterodipeptides with aromatic, aliphatic, polar and charged side-chains in search of self-assembling sequences.[64b] Rich combinatorial peptide libraries were obtained from single and multi-component systems where product distribution was biased towards the most stable self-assembling structures (Figure 6b). The self-assembly process, which is mediated by hydrogen bonding and hydrophobic interactions, and the products' free energies, which determine

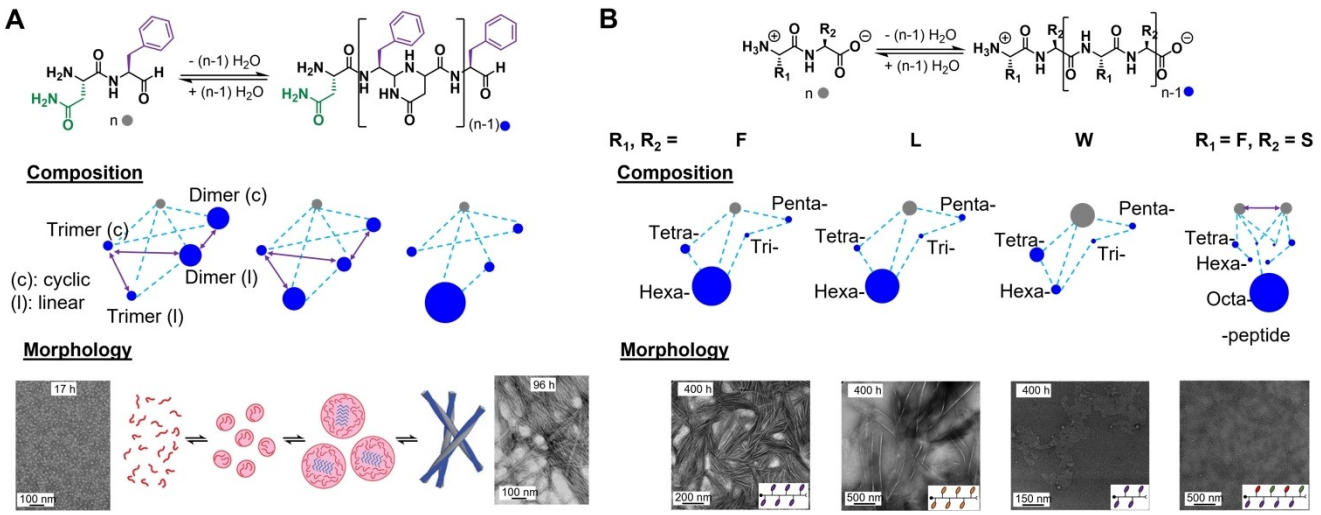


Figure 6. Examples of sequence-selection and amplification in complex peptide systems. Schemes of reversible condensation and polymerization through imine condensation A) and enzyme-mediated peptide reversible ligation B) of complex adaptive systems. Representative mixture compositions: starting materials (grey circles) and products (blue circles) are interconvertible (cyan dashed line) through reversible dynamic condensation. Purple arrows indicate aromatic (stacking) interactions between components. The size of the circle is representative of the component's relative abundance. Morphology: TEM micrographs of mixtures showing distinct morphologies for each mixture. A. The system undergoes morphological transitions from aggregates to fibrous network with the chemical evolution and lengthening of the polymers. Reproduced with permission.[62] Copyright 2017, Nature Publishing Group. B. Sequence-specific selection of self-assembling oligomers from enzyme-mediated dynamic peptide libraries starting from single components and a two-component mixture. Reproduced with permission.[64b] Copyright 2016, Nature Publishing Group.4.3. Compositionally enriched dynamic peptide libraries

their distribution, are dictated by the relative contributions of these interactions. We were able to show that this balance can be altered by changes in the experimental conditions, such as the introduction of organic co-solvents and Hofmeister series ions, leading to significant changes in product selection and distribution in response to environmental triggers. We also demonstrated the efficiency of the process in self-selecting and amplifying the synthesis of the strongest self-assembling sequences, even as long as octapeptides. Starting from an equimolar ratio of dipeptides FD and FS, octapeptide FDFSFDFS was found to be the exclusively formed and amplified sequence, outcompeting all other possible combinations. The system transitioned from a homogeneous solution to a self-supporting hydrogel composed of an entangled fiber network of the octapeptide.

These observations led us to believe that adaptive combinatorial peptide libraries represent a new avenue and a way to connect designed peptide libraries and evolving chemical networks. We've recently expanded on this concept by investigating and analyzing the collective behavior of emergent dynamic networks from complex peptide libraries starting from a mixture of 15 dipeptides (Figure 7).[65] We designed a highly diverse chemical space, and looked at the evolution of patterns of interactions and information gain as the system adapts to the introduction of ATP as a guest molecule. Each dipeptide was composed of a structural (G, A, V) and functional residue, the design and choice of which were informed by mining the binding sites of ATP binding proteins in the protein data bank and selecting the over-represented ATP-interacting residues (D, H, K, R, S). Under enzymatic exchange conditions, a library that includes 225 tetrapeptides is generated and analyzed by LCMS deconvolution. It becomes clear that with as many interacting species, new analysis and representation methods are needed to identify collective patterns of interaction and adaptation. In the absence of strong structural information (G and A dipeptides), the system relies on stabilization from intermolecular interactions with ATP to shift the equilibrium towards condensation products, in contrast to self-assembly as previously discussed. Since the equilibrium between condensation and hydrolysis of a given sequence is dictated by its strength of interaction with ATP, non-interacting tetrapeptides were expected to hydrolyze back to the starting dipeptides while interacting tetrapeptides were amplified. We found that, in presence of ATP, cationic tetrapeptides were amplified while neutral and anionic ones were downregulated, reflecting the ability of the side-chain residues to accommodate or oppose electrostatic ATP-binding interactions, and, as a result, inducing the redistribution of the collective sequence composition. In a way, this can be assimilated to learning mechanisms, where systems evolve through feedbacks from productive and degenerate information. A unique feature of this system, is that the balancing of supra-molecular interactions causes a re-adaptation in the chemical composition of the peptide sequence space. When a competing interaction is encoded in the starting sequence, the system behaves differently. At 37 °C, the presence of ATP has little to no effect on libraries emerging from V containing dipeptides as the contribution of hydrophobic and aggregating interactions influences the tetrapeptide distribution to a greater extent. At higher temperature, electrostatic ATP-binding interactions become more dominant and the system's distribution is re-configured to amplify ATP-binding sequences. This work provides a significant step in the chemical

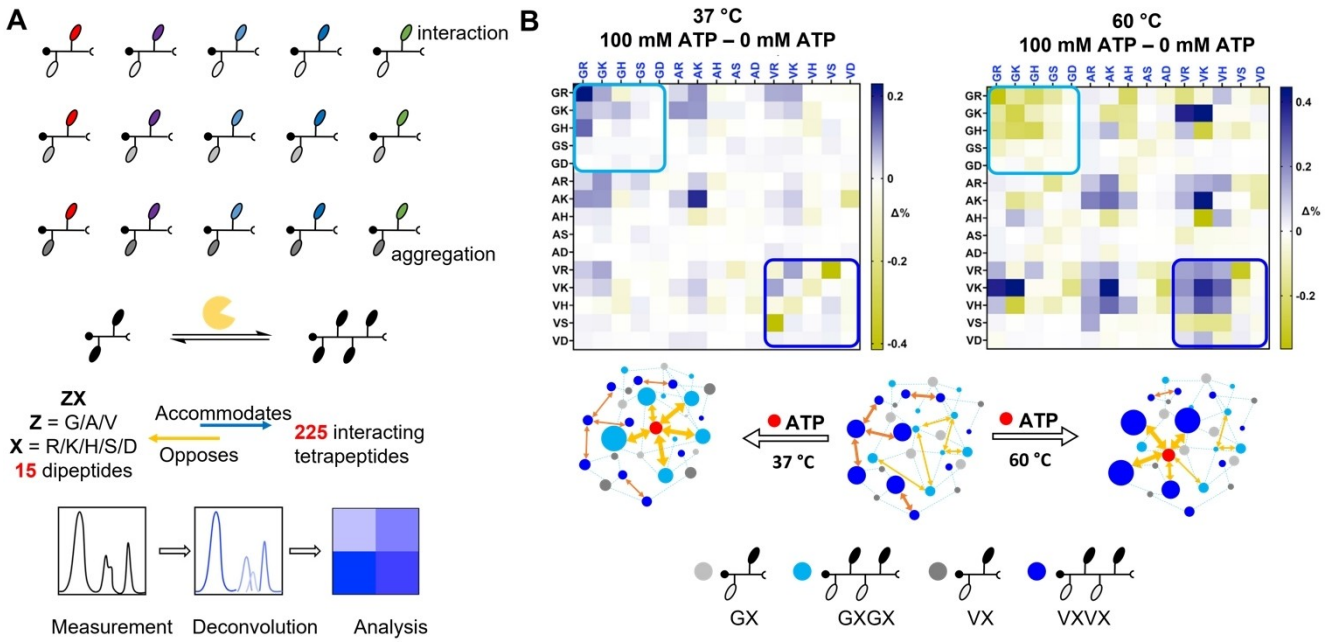


Figure 7. Collective adaption in a complex peptide system. A) Designed chemical space of 15 interacting dipeptides, each composed of a functional (interaction) and structural (aggregation) amino-acid. Enzyme-mediated peptide ligation produces a mixture of 225 interacting tetrapeptides. To analyze and represent the collective behavior of the system, a heat map of the deconvoluted peptide mass-spectrometry (MS) areas is extracted. B) Heatmaps with direct comparison of product distribution reveal patterns of cooperative behavior. Representative composition adaptation of GXGX and VXVX libraries. Starting materials (grey circles) and products (blue circles) are interconvertible (cyan dashed line) through reversible peptide bond formation. Governing inter-component interactions are represented with double headed arrows, color-coded according to figure 1. Size of circle and width of arrow indicate dominant product/interactions. Two distinct distributions are obtained after the introduction of a stressor (temperature or ATP (red circle)). The system adapts to these stressors by a product re-distribution and amplification of stabilizing structures. Reproduced with permission.[65] Copyright 2022, Elsevier Inc.

evolution of peptide-based systems chemistry, where sequence-encoded information evolves and adapts to the environment. We believe that this proof-of-concept, draws parallel to biological systems in showcasing the immense and, at time underrated, potential of dynamic side-chain interactions in providing positive and negative feedback through adaptive sequence-encoded information in the evolution of chemical networks.

5. What's Next

Living systems teach us that a few dozen chemically simple building blocks, when organized covalently and non-covalently in space and over time, can provide tremendous functionality that goes far beyond what can be achieved using existing design approaches used for materials and devices. Biological systems are however extremely complex and we are currently far from able to design materials that approximate their complexity. While building blocks are individually simple and highly conserved, they are connected through huge numbers of covalent and non-covalent interactions that all influence each other. We discuss that one approach to start to address the challenge of biomolecular design of multicomponent, flexible and beyond-biology systems, is to start by looking at simpler molecules that share some of the same features of proteins, and are made of the same amino acid building blocks, but only contain a small number of amino acids. The combinatorial and adaptive approaches highlighted here investigate how networks of these interacting and inter-converting species behave and use feedback mechanisms to evolve and select for the best fit and how the systems re-adapt and reconfigure in response to stressors. Many draw parallels between Systems Chemistry and Origin of Life research. But while these systems operate under thermodynamic equilibrium, it is known that Life is an out-of-equilibrium process. With the promises held by artificial out-of-equilibrium[71] and chemically fueled systems,[72] it is not impossible to imagine out-of-equilibrium dynamic peptide libraries where transient sequence-encoded information is used to relieve the effect of a stressor, to sense the presence or absence of a ligand or signal a disregulation in chemical balance. Systems chemistry also promises radically new materials, sensors and devices that incorporate features, such as functional reconfiguration, adaptation through learning in changing conditions, and even molecular memory, that are normally only associated with living matter and are unimaginable using conventional chemistry approaches. While chemistry as a field still has far to go in embracing a complex system view when analyzing and synthesizing new functional materials, we strongly believe that the opportunities and possibilities that peptide-based complex adaptive systems have to offer are immense.

Acknowledgements

R.V.U. acknowledges funding from the Office of Naval Research for the Vannevar Bush Faculty Fellowship (grant N00014-21-1-2967). We thank the Air Force Office of Scientific Research (AFoSR) for funding R.V.U. and S.K. (grant FA9550-21-1-0091).

Conflict of Interest

The authors declare no conflict of interest.

[1] a) M. Mitchell Waldrop, *Complexity: the emerging science at the edge of order and chaos*, Simon & Shuster, New York, USA **1992**; b) M. A. Fuentes, *Entropy* **2014**, *16*, 4489–4496.
[2] J. Jost, *Theory Biosci.* **2004**, *123*, 69–88.
[3] a) G. Ashkenasy, T. M. Hermans, S. Otto, A. F. Taylor, *Chem. Soc. Rev.* **2017**, *46*, 2543–2554; b) O. Š. Miljanić, *Chem* **2017**, *2*, 502–524; c) R. F. Ludlowa, S. Otto, *Chem. Soc. Rev.* **2008**, *37*, 101–108.
[4] a) A. Jain, A. Calò, D. Barceló, M. Kumar, *Anal. Bioanal. Chem.* **2022**, *414*, 5105–5119; b) M. A. Grover, M.-C. Hsieh, D. G. Lynn, *Life* **2019**, *9*, 55–66.
[5] J. Boekhoven, W. E. Hendriksen, G. J. M. Koper, R. Eelkema, J. H. Van Esch, *Science* **2015**, *349*, 1075–1079.
[6] S. Otto, *Acc. Chem. Res.* **2022**, *55*, 145–155.
[7] F. Sheehan, D. Sementa, A. Jain, M. Kumar, M. Tayarani-Najjaran, D. Kroiss, R. V. Ulijn, *Chem. Rev.* **2021**, *121*, 13869–13914.
[8] G. M. Whitesides, R. F. Ismagilov, *Science* **1999**, *284*, 89–92.
[9] a) X. Hu, M. Liao, H. Gong, L. Zhang, H. Cox, T. A. Waigh, J. R. Lu, *Curr. Opin. Colloid Interface Sci.* **2020**, *45*, 1–13; b) A. Levin, T. A. Hakala, L. Schnaider, G. J. L. Bernardes, E. Gazit, T. P. J. Knowles, *Nat. Chem. Rev.* **2020**, *4*, 615–634; c) N. J. Sinha, M. G. Langenstein, D. J. Pochan, C. J. Kloxin, J. G. Saven, *Chem. Rev.* **2021**, *121*, 13915–13935.
[10] P. W. J. M. Frederix, R. V. Ulijn, N. T. Hunt, T. Tuttle, *J. Phys. Chem.* **2011**, *2*, 2380–2384.
[11] P. W. J. M. Frederix, G. Scott, Y. Abul-Haija, D. Kalafatovic, C. G. Pappas, N. Javid, N. T. Hunt, R. V. Ulijn, T. Tuttle, *Nat. Chem.* **2015**, *7*, 30–37.
[12] Y. Tang, S. Bera, Y. Yao, J. Zeng, Z. Lao, X. Dong, E. Gazit, G. Wei, *Cell Reports Physical Science* **2021**, *2*, 100579–100596.
[13] S. Marchesan, C. D. Easton, F. Kushkaki, L. Waddington, P. G. Hartley, *Chem. Commun.* **2012**, *48*, 2195–2197.
[14] S. Marchesan, K. E. Styan, C. D. Easton, L. Waddington, A. V. Vargiuc, *J. Mater. Chem. B* **2015**, *3*, 8123–8132.
[15] A. M. Garcia, M. Melchionna, O. Bellotto, S. Kralj, S. Semeraro, E. Parisi, D. Iglesias, P. D'Andrea, R. De Zorzi, A. V. Vargiu, S. Marchesan, *ACS Nano* **2021**, *15*, 3015–3025.
[16] A. Brito, D. Dave, A. Lampel, V. I. B. Castro, D. Kroiss, R. L. Reis, T. Tuttle, R. V. Ulijn, R. A. Pires, I. Pashkuleva, *J. Am. Chem. Soc.* **2021**, *143*, 19703–19710.
[17] E. W. Martin, A. S. Holehouse, I. Peran, M. Farag, J. J. Incicco, A. Bremer, C. R. Grace, A. Soranno, R. V. Pappu, T. Mitta, *Science* **2020**, *367*, 694–699.
[18] M. Abbas, W. P. Lipiński, K. K. Nakashima, W. T. S. Huck, E. Spruijt, *Nat. Chem.* **2021**, *13*, 1046–1054.
[19] a) A. J. Scott, A. Niitsu, H. T. Kratochvil, E. J. M. Lang, J. T. Sengel, W. M. Dawson, K. R. Mahendran, M. Mravic, A. R. Thomson, R. L. Brady, L. Liu, A. J. Mulholland, H. Bayley, W. F. DeGrado, M. I. Wallace, D. N. Woolfson, *Nat. Chem.* **2021**, *13*, 643–650; b) A. R. Thomoson, C. W. Wood, A. J. Burton, G. J. Bartlett, R. B. Sessions, R. L. Brady, D. N. Woolfson, *Science* **2014**, *346*, 485–488.
[20] N. H. Joh, T. Wang, M. P. Bhate, R. Acharya, Y. Wu, M. Grabe, M. Hong, G. Grigoryan, W. F. DeGrado, *Science* **2014**, *346*, 1520–1524.
[21] a) Y. Suzuki, G. Cardone, D. Restrepo, P. D. Zavattieri, T. S. Baker, F. A. Tezcan, *Nature* **2016**, *533*, 369–373; b) E. Golub, R. H. Subramanian, J. Esselborn, R. G. Alberstein, J. B. Bailey, J. A. Chiong, X. Yan, T. Booth, T. Baker, F. A. Tezcan, *Nature* **2020**, *578*, 172–176.
[22] P. S. Huang, S. Boyken, D. Baker, *Nature* **2016**, *537*, 320–327.
[23] R. Freeman, M. Han, Z. Álvarez, J. A. Lewis, J. R. Wester, N. Stephanopoulos, M. T. McClendon, C. Lynsky, J. M. Godbe, H. Sangji, E. Luijten, S. I. Stupp, *Science* **2018**, *362*, 808–813.
[24] S. Bera, S. Mondal, B. Xue, L. J. W. Shimon, Y. Cao, E. Gazit, *Nat. Mater.* **2019**, *18*, 503–509.
[25] A. Lampel, S. A. McPhee, H. A. Park, G. G. Scott, S. Humagain, D. R. Hekstra, B. Yoo, P. W. J. M. Frederix, T.-D. Li, R. R. Abzalimov, S. G. Greenbaum, T. Tuttle, C. Hu, C. J. Bettinger, R. V. Ulijn, *Science* **2017**, *356*, 1064–1068.
[26] R. P. Feynman, *Eng. Sci.* **1960**, *23*, 22–36.
[27] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, S.-J. Marrink, *J. Chem. Theory Comput.* **2008**, *4*, 819–834.
[28] G. G. Scott, P. J. McKnight, T. Tuttle, R. V. Ulijn, *Adv. Mater.* **2016**, *28*, 1381–1386.
[29] G. Schneider, W. Schrödl, G. Wallukat, J. Müller, E. Nissen, W. Rönspeck, P. Wrede, R. Kunze, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12179–12184.
[30] A. van Teijlingen, T. Tuttle, *J. Chem. Theory Comput.* **2021**, *17*, 3221–3232.
[31] K. Shmilovich, S. S. Panda, A. Stouffer, J. D. Tovar, A. L. Ferguson, *Digital Discovery* **2022**, *1*, 448–462.
[32] K. Shmilovich, R. A. Mansbach, H. Sidky, O. E. Dunne, S. S. Panda, J. D. Tovar, A. L. Ferguson, *J. Phys. Chem. B* **2020**, *124*, 3873–3891.
[33] K. Bozovičar, T. Bratkovič, *Int. J. Mol. Sci.* **2020**, *21*, 215–246.
[34] J. Pande, M. M. Szewczyk, A. K. Grover, *Biotech. Adv.* **2010**, *28*, 849–858.
[35] D. Kalafatovic, G. Mauša, T. Todorovski, E. Giralt, *J. Cheminf.* **2019**, *11*, 25–40.
[36] D. Kalafatovic, G. Mauša, D. R. Maslov, E. Giralt, *Molecules* **2020**, *25*, 3316–3331.
[37] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583–589.
[38] X. Zhao, S. Zhanc, *Chem. Soc. Rev.* **2006**, *35*, 1105–1110.
[39] D. N. Woolfson, *Subcell. Biochem.* **2017**, *82*, 35–61.
[40] L. E. O'Leary, J. A. Fallas, E. L. Bakota, M. K. Kang, J. D. Hartgerink, *Nat. Chem.* **2011**, *3*, 821–828.
[41] M. Reches, E. Gazit, *Science* **2003**, *300*, 625–627.
[42] Z. Shlomo, T. P. Vinod, Raz Jelinek, H. Rapaport, *Chem. Commun.* **2015**, *51*, 3154–3157.
[43] D. Shental-Bechor, Y. Levy, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 8256–8261.
[44] R. van der Lee, M. Buljan, B. Lang, R. J. Weatheritt, G. W. Daughdrill, A. K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D. T. Jones, P. M. Kim, R. W. Kriwacki, C. J. Oldfield, R. V. Pappu, P. Tompa, V. N. Uversky, P. E. Wright, M. M. Babu, *Chem. Rev.* **2014**, *114*, 6589–631.
[45] C. M. Fare, A. Villani, L. E. Drake, J. Shorter, *Open Biol.* **2021**, *11*, 210137–210156.
[46] T. S. Harmon, A. S. Holehouse, M. K. Rosen, R. V. Pappu, *eLife* **2017**, *6*, e30294.
[47] B. J. Kim, D. Yang, B. Xu, *Trends Chem.* **2020**, *2*, 71–83.
[48] a) P. Makam, E. Gazit, *Chem. Soc. Rev.* **2018**, *47*, 3406–3420; b) B. O. Okesola, A. Mata, *Chem. Soc. Rev.* **2018**, *47*, 3721–3736; c) D. M. Raymond, B. L. Nilsson, *Chem. Soc. Rev.* **2018**, *47*, 3659–3720; d) E. R. Draper, D. J. Adams, *Chem. Soc. Rev.* **2018**, *47*, 3395–3405.
[49] S. Fleming, S. Debnath, P. W. J. M. Frederix, N. T. Hunt, R. V. Ulijn, *Biomacromolecules* **2014**, *15*, 1171–1184.
[50] Y. M. Abul-Haija, G. G. Scott, J. K. Sahoo, T. Tuttle, R. V. Ulijn, *Chem. Commun.* **2017**, *53*, 9562–9565.
[51] J. K. Sahoo, M. A. VandenBerg, E. E. R. Bello, C. D. Nazaretha, M. J. Webber, *Nanoscale* **2019**, *11*, 16534–16543.
[52] S. Bera, S. Mondal, Y. Tang, G. Jacoby, E. Arad, T. Guterman, R. Jelinek, R. Beck, G. Wei, E. Gazit, *ACS Nano* **2019**, *13*, 1703–1712.
[53] K. E. Inostroza-Brito, E. Collin, O. Siton-Mendelson, K. H. Smith, A. Monge-Marcet, D. S. Ferreira, R. P. Rodríguez, M. Alonso, J. C. Rodríguez-Cabello, R. L. Reis, F. Sagués, L. Botto, R. Bitton, H. S. Azevedo, A. Mata, *Nat. Chem.* **2015**, *7*, 897–904.
[54] A. Jain, S. Kassem, R. S. Fisher, B. Wang, T.-D. Li, T. Wang, Y. He, S. Elbaum-Garfinkle, R. V. Ulijn, *J. Am. Chem. Soc.* **2022**, *144*, 15002–15007.
[55] J. Li, P. Nowak, S. Otto, *J. Am. Chem. Soc.* **2013**, *135*, 9222–9239.
[56] R. T. Lam, A. Belenguer, S. L. Roberts, C. Naumann, T. Jarrosson, S. Otto, J. K. M. Sanders, *Science* **2005**, *308*, 667–669.
[57] J. M. Carnall, C. A. Waudby, A. M. Belenguer, M. C. Stuart, J. J. Peyralans, S. Otto, *Science* **2010**, *327*, 1502–1506.
[58] C. G. Pappas, P. K. Mandal, B. Liu, B. Kauffmann, X. Miao, D. Komáromy, W. Hoffmann, C. Manz, R. Chang, K. Liu, K. Pagel, I. Huc, S. Otto, *Nat. Chem.* **2020**, *12*, 1180–1186.

[59] Martin C, M. Frenkel-Pinter, K. H. Smith, V. F. Rivera-Santana, A. B. Sargon, K. C. Jacobson, A. Guzman-Martinez, L. D. Williams, L. J. Leman, C. L. Liotta, M. A. Grover, N. V. Hud, *JACS Au* **2022**, *2*, 1395–1404.
[60] S. Ghosh, L. A. Ingerman, A. G. Frye, S. J. Lee, M. R. Gagne, M. L. Waters, *Org. Lett.* **2010**, *12*, 1860–1863.
[61] Y. Ruff, V. Garavini, N. Giuseppone, *J. Am. Chem. Soc.* **2014**, *136*, 6333–6339.
[62] C. Chen, J. Tan, M.-C. Hsieh, T. Pan, J. T. Goodwin, A. K. Mehta, M. A. Grover, D. G. Lynn, *Nature Chem* **2017**, *9*, 799–804.
[63] A. K. Das, A. R. Hirsth, R. V. Ulijn, *Faraday Discuss.* **2009**, *143*, 293–303.
[64] a) J.-B. Guilbaud, E. Vey, S. Boothroyd, A. M. Smith, R. V. Ulijn, A. Saiani, A. F. Miller, *Langmuir* **2010**, *26*, 11297–11303; b) C. G. Pappas, R. Shafi, I. Sasselli, H. Siccardi, T. Wang, V. Narang, R. Abzalimov, N. Wijerathne, R. V. Ulijn, *Nat. Nanotechnol.* **2016**, *11*, 960–967.
[65] A. Jain, S. A. McPhee, T. Wang, M. Narayanan Nair, D. Kroiss, T. Z. Jia, R. V. Ulijn, *Chem* **2022**, *8*, 1894–1905.
[66] a) M. Altay, Y. Altay, S. Otto, *Angew. Chem. Int. Ed.* **2018**, *57*, 10564–10568; *Angew. Chem.* **2018**, *130*, 10724–10728; b) B. Bartolec, M. Altay, S. Otto, *Chem. Commun.* **2018**, *54*, 13096–13098.
[67] a) C. G. Pappas, B. Liu, I. Marić, J. Ottelé, A. Kiani, M. L. van der Klok, P. R. Onck, S. Otto, *J. Am. Chem. Soc.* **2021**, *143*, 7388–7393; b) Y. Altay, M. Tezcan, S. Otto, *J. Am. Chem. Soc.* **2017**, *139*, 13612–13615; c) M. Malakoutikhah, J. J.-P. Peyralans, M. Colomb-Delsuc, H. Fanlo-Virgós, M. C. A. Stuart, S. Otto, *J. Am. Chem. Soc.* **2013**, *135*, 18406–18417.
[68] P. T. Corbett, S. Otto, J. K. M. Sanders, *Chem. Eur. J.* **2004**, *10*, 3139–3143.
[69] R. A. R. Hunt, R. F. Ludlow, S. Otto, *Org. Lett.* **2009**, *11*, 5110–5113.
[70] a) R. Ludlow, J. Liu, H. Li, S. Roberts, J. Sanders, S. Otto, *Angew. Chem. Int. Ed.* **2007**, *46*, 5762–5764; *Angew. Chem.* **2007**, *119*, 5864–5866; b) P. T. Corbett, S. Otto, J. K. M. Sanders, *Chem. Eur. J.* **2008**, *14*, 2153–2166.
[71] a) J. H. van Esch, R. Klajn, S. Otto, *Chem. Soc. Rev.* **2017**, *46*, 5474–5475; b) S. A. P. van Rossum, M. Tena-Solsona, J. H. van Esch, R. Eelkema, J. Boekhoven, *Chem. Soc. Rev.* **2017**, *46*, 5519–5535.
[72] a) A. Sharko, D. Livitz, S. De Piccoli, K. J. M. Bishop, T. M. Hermans, *Chem. Rev.* **2022**, *122*, 11759–11777; b) K. Das, L. Gabrielli, L. J. Prins, *Angew. Chem. Int. Ed.* **2021**, *60*, 20120–20143; c) K. Gentile, A. Somasundar, A. Bhide, A. Sen, *Chem.* **2020**, *6*, 2174–2185.

Manuscript received: October 2, 2022
Accepted manuscript online: November 3, 2022
Version of record online: November 24, 2022