

## Review article



# Context dependence in assembly code for supramolecular peptide materials and systems

Kübra Kaygisiz<sup>1</sup>, Deborah Sementa<sup>1</sup>, Vignesh Athiyarath<sup>1</sup>, Xi Chen<sup>1,2,3,4</sup> & Rein V. Ulijn<sup>1,3,5,6</sup>✉

## Abstract

Living systems provide the most sophisticated materials known. These materials are created from a few dozen building blocks that are driven to self-organize by covalent and non-covalent interactions. Biology's building blocks can be repurposed for the design of synthetic materials that life has not explored. In this Review, we examine the bottom-up design, discovery and evolution of self-assembling peptides by considering the entire supramolecular interaction space available to their constituent amino acids. Our approach focuses on sequence context, or how peptide sequence and environmental conditions collectively influence peptide self-assembly outcomes. We discuss examples of peptides that assemble through multimodal backbone, side chain and water interactions. We conclude that a more systematic (comparing sequences side-by-side), integrated (pairing computation and experiment) and holistic (considering peptide, solvent and environment) approach is required to better understand and fully exploit amino acids as a universal assembly code. This goal is particularly timely, because laboratory automation and artificial intelligence now have the potential to accelerate discoveries in these highly modular and complex materials, beyond the limited sequence space that biology uses.

## Sections

Introduction

Interaction modes,  
conformational flexibility  
and sterics of side chains

Designing sequence  
context-dependent order  
and disorder

Environment-dependent  
assembly

In situ editing of  
peptide sequence

Outlook

<sup>1</sup>Nanoscience Initiative at Advanced Science Research Center of the Graduate Center of the City University of New York, New York, NY, USA. <sup>2</sup>Department of Chemical Engineering, The City College of New York, New York, NY, USA. <sup>3</sup>PhD Program in Chemistry, The Graduate Center of the City University of New York, New York, NY, USA. <sup>4</sup>PhD Program in Physics, The Graduate Center of the City University of New York, New York, NY, USA. <sup>5</sup>Department of Chemistry, Hunter College, City University of New York, New York, NY, USA. <sup>6</sup>PhD Program in Biochemistry, The Graduate Center of the City University of New York, New York, NY, USA. ✉e-mail: [rulijn@gc.cuny.edu](mailto:rulijn@gc.cuny.edu)

## Introduction

Repurposing ‘chemistry-of-life’ building blocks to produce synthetic materials combines the sustainability of biological systems with the versatility of supramolecular design. Compared with the behaviour of natural biological systems, the performance of laboratory-designed biomimetic or supramolecular structures still lags far behind in many areas. For example, most designed structures are very slow to assemble or respond to changes in conditions, whereas biological systems can do so near-instantaneously<sup>1,2</sup>. Moreover, designed systems have only very limited (and pre-programmed) response versatility<sup>3</sup>, whereas biology leverages multiple components to autonomously adapt, including to conditions that it has not encountered before<sup>4</sup>. We propose that using the entire non-covalent interaction space available to the 20 genetically encoded amino acids might provide access to biology’s vast molecular design space, even beyond the comparatively restricted space that evolution has explored<sup>5,6</sup>.

The self-assembly of peptides and proteins is intimately tied to the composition and sequence of amino acids<sup>7</sup>. Amino acids in a peptide are connected by the amide bond, which has a strong dipole moment and is conformationally rigid owing to the partial double bond character. As a result, the peptide backbone is stiff and polar, facilitating its assembly into secondary structures through highly directional hydrogen bonds and dipolar interactions. Synthetic peptides have typically been designed to fold using patterns inspired by protein secondary structures, such as the  $\alpha$ -helix and  $\beta$ -sheet, that are based on such cooperative backbone hydrogen bonding<sup>8</sup>. These designs reliably follow simple arrangements of amino acids, such as alternating polar and nonpolar amino acids, or small and large amino acids, and usually give rise to 1D or 2D structures<sup>9,10</sup>. Higher-order 3D structure is introduced by side chain interactions<sup>11</sup> and is frequently supported by computer simulations and predictions<sup>12–15</sup>.

Every amino acid has a unique side chain that can undergo a particular combination of supramolecular interactions, including hydrogen bonding,  $\pi$ -based, dipole–dipole, electrostatic and/or hydrophobic interactions (Fig. 1). These interactions are the primary determinants for variations in dynamics, packing and physical properties in self-assembled peptides. Compared with using secondary structure templates as a starting point for assembly, less research has focused on the design of peptides that use the physicochemical interactions between side chains to create supramolecular networks. The side chain interactions that occur within or between peptides are increasingly recognized as the most important contributors to disorder, and non-directional assembly is understood to be a critical feature of the adaptive chemistry of life, such as in liquid condensates<sup>16–19</sup>. Using side chain interactions in concert with backbone hydrogen-bond patterning allows the self-assembled structure to be tailored in many ways: its internal organization, its interactions with water, its hierarchical organization through interfaces and its responsiveness to the environment. Peptides designed to assemble into functional forms with both side chain and backbone interactions in mind might look quite different and potentially can be much simpler in sequence and composition, compared with those that evolution has provided. Given the systematic, code-like sequence space, machine learning/artificial intelligence (AI) approaches are particularly suited to accelerate discovery of such peptides<sup>20</sup> and have been increasingly applied to this problem<sup>21–26</sup>. We argue that, for these approaches to be truly effective and predictive, amino acids must be considered in the context of their sequences and environmental conditions.

The diversity of interaction modes is unique to peptides in comparison to other biomolecules such as DNA and RNA, carbohydrates and

lipids. Self-assembled peptides in water ultimately balance directional backbone interactions and side chain interactions of varying degrees of directionality; competing attractive and repulsive interactions; and interactions with water molecules that are part of the supramolecular network and either guide or compete with peptide–peptide interactions<sup>27,28</sup>. By contrast, the nucleic acids in DNA and RNA are mostly limited to binary encoding possibilities (AT and CG pairs) that are predetermined by their highly directional hydrogen-bonding interactions, not strongly influenced by their neighbours. This relatively simple code results in highly predictable, ordered structures<sup>27</sup> and underpins the versatile design approach of DNA nanotechnology<sup>29,30</sup>. The design principles of carbohydrates, on the other end of the spectrum, are only in a very early stage. The monosaccharide units can hydrogen bond in multiple directions<sup>31</sup>, adding a degree of complication, but cannot undergo a broad range of interaction types. The peptide assembly code is therefore fundamentally different from those of the other biomolecules in that it is context-dependent and capable of many interaction modes: even for simple peptides of just a few amino acids, the (combination of) interactions that it selects can depend on both sequence and environment<sup>32</sup>. As a consequence, this code is not only infinitely more versatile but also much more challenging to rationalize.

In this Review, we introduce the term ‘sequence context’ to describe how environmental conditions and peptide sequence collectively influence peptide self-assembly outcomes (Fig. 1a). Environmental conditions include factors such as pH, ionic strength, co-solutes, temperature and mechanical forces (such as shear and pressure). The peptide sequence refers to how variations in the amino acid order and composition affect self-assembly. We explore how focusing on combinations of side chain interactions in context of sequence can aid the design of functional complex adaptive systems. We also elaborate on the role of dynamics in peptide self-assembly by highlighting how interactions are frequently controlled or directed by water. We believe that sequence context is a valuable concept to incorporate into the design of supramolecular peptide materials, because it underscores the importance of environmental conditions and sequence in drawing out complex structures and functions.

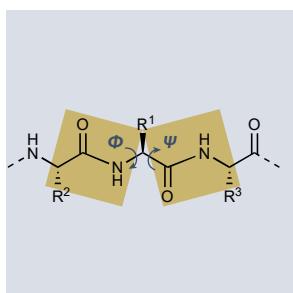
## Interaction modes, conformational flexibility and sterics of side chains

Peptides self-assemble through a plethora of cooperative interactions. The balance between intermolecular and intramolecular interaction strengths, molecular alignment and solvability of components controls the formation of supramolecular structures on a spectrum of ordered to disordered<sup>33–35</sup>.

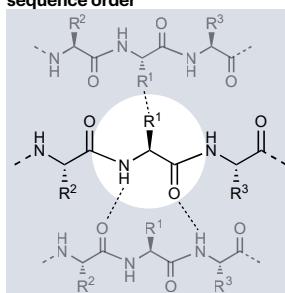
There are 20 genetically encoded L-amino acids common to all life forms, which we denote by their single letter codes. Each amino acid has a characteristic charge and polarity, depending on the chemical properties of the side chains and environmental conditions, and a characteristic conformational flexibility, defined by the Ramachandran angles  $\phi$  and  $\psi$  (ref. 36) (Fig. 1a). Quantitative parameters such as charge and hydrophobicity are traditionally used to describe the chemical character of a side chain<sup>37</sup>. We use here the experimentally determined Wimley–White octanol–water partition coefficients<sup>38</sup> and the net charge at neutral pH to colour-code the 20 amino acids based on their primary chemical nature – aromatic, aliphatic, polar, basic, acidic or ‘special’ (for those whose dominant feature is not captured by these categories) – using depth of colour to represent the degree of that characteristic (Fig. 1b). However, these categories are broad and often

## a Sequence context

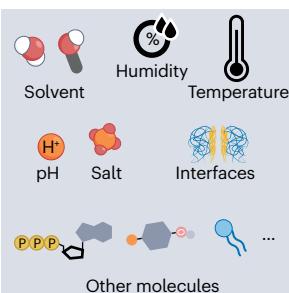
### Conformational flexibility



### Intramolecular and intermolecular sequence order



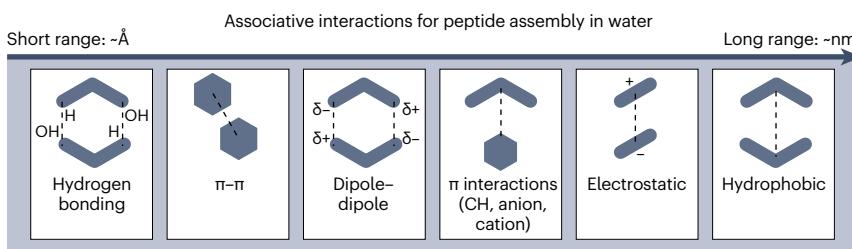
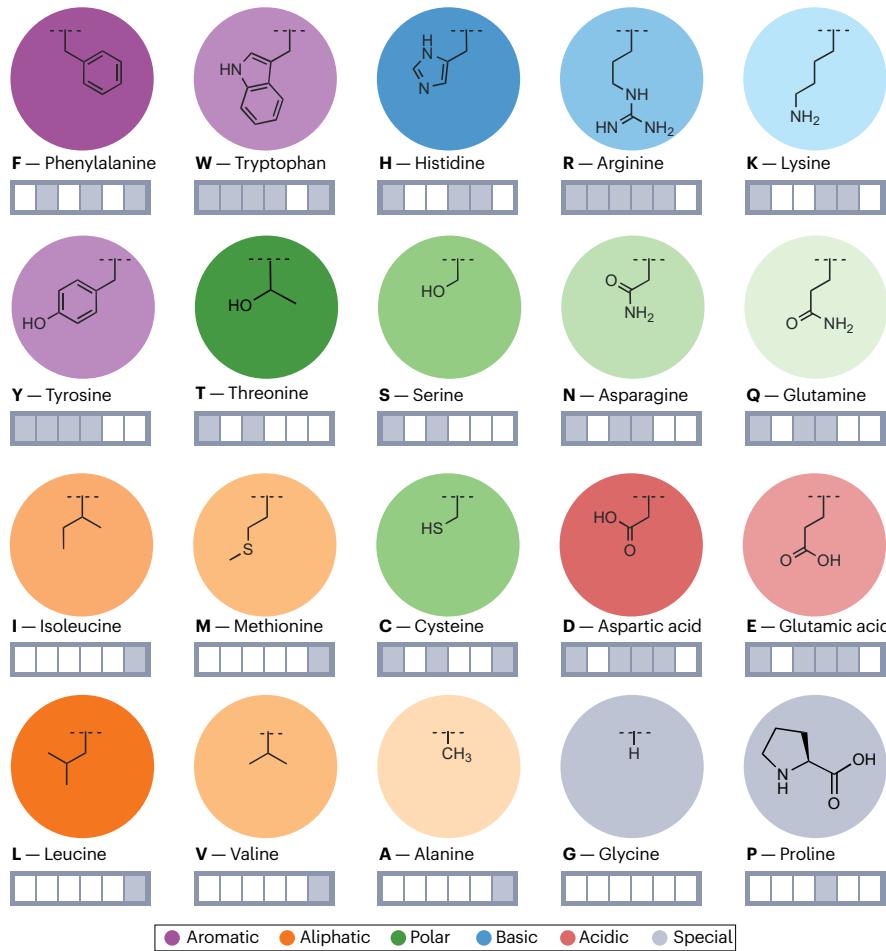
### Environmental conditions



## Fig. 1 | Biology's universal assembly code.

**a**, Supramolecular peptide interactions depend on the intramolecular and intermolecular sequence context of the amino acid side chain ( $R^{1-3}$ ). Sequence context is determined by conformational flexibility and the environmental conditions, such as solvent, pH, temperature and salt ions. **b**, The 20 gene-encoded amino acids. The colour code represents the predominant chemical nature of the side chains: polar (green), basic (blue), acidic (red), aromatic (purple) and aliphatic (orange); ‘special’ (grey) side chains are those whose dominant nature is not captured by these categories. We acknowledge that oversimplification of amino acids into categories fails to capture their unique chemistry, which is why we supplement this categorization with colour shading and a six-box code to account for chemical nature and interaction modes, respectively. The colour shading indicates that every amino acid has a unique chemical identity and a deeper or lighter colour indicates stronger or weaker predominant chemical nature, respectively. The box code indicates which interactions each amino acid side chain is capable of forming in water.

**b**



fail to capture the full character of an amino acid, especially those that fall into multiple and sometimes conflicting classes<sup>39</sup> (Box 1). Side chain electronic density is another relevant feature of amino acids that is not always represented by the aforementioned categories. Several amino acids have surprisingly similar electronic structures despite their different chemical properties and can systematically replace each other in G protein-coupled receptors; the polar amino acids Q, N, T and Y can replace the nonpolar amino acids L, I, V and F, respectively, without limiting the function of the protein<sup>40</sup>.

To capture the specific interactions an amino acid can participate in, which include hydrogen bonding;  $\pi$ – $\pi$ ; dipole–dipole; CH– $\pi$ , -cation and -anion; electrostatic; and hydrophobic interactions, we assign a six-panel legend/barcode to every amino acid to designate the interaction profile of each in water (Fig. 1b). These interaction modes can work cooperatively with each other, affecting the strength of the overall interaction. For example, aromatic side chains can undergo short-ranged directional  $\pi$ – $\pi$ , anion– $\pi$  or cation– $\pi$  interactions<sup>41</sup> that support the self-assembly into ordered structures<sup>42</sup>. Although Y and F both interact strongly via  $\pi$ – $\pi$  interactions, the Y–Y interaction is stronger compared with Y–F or F–F interactions owing to additional contributions from hydrogen bonding between hydroxyl groups in Y<sup>43</sup>.

Because the interactions between amino acids depend on sequence context, it is not straightforward to determine which properties of an amino acid matter most to its assembly behaviour, both in primary and in higher order senses. By using statistical methods from large data sets of peptides, it is possible to reach a level of generalization<sup>44</sup>, but any conclusions should be treated with caution in the absence of sequence context. For example, the hydrophobic residues W, Y, F, I, V and L are found predominantly in ordered structures, and the

polar amino acids A, R, G, Q, S, P, E and K are preferentially found in disordered structures<sup>45</sup>. However, although T, N and D are polar, they can be found in ordered structures if the sequence context permits their polar side chain groups located at the  $\beta$ -carbon to take part in the hydrogen-bonding network with the backbone, which decreases their disorder-promoting character<sup>45</sup>. Similarly, other charged side chains (R, K, E and D) can act associatively or dissociatively depending on sequence and environment<sup>43</sup>. A high conformational flexibility such as in K, E, Q and R commonly leads to higher disorder, whereas rigid side chains in F, Y, V and T promote order in the assembly<sup>46</sup>, but only if they are able to partition into a hydrophobic domain.

There are multiple factors that can influence conformational flexibility and geometrical constraints on self-assembly. In this context, G and P each have a special role in the amino acid assembly code, enhancing or restricting flexibility of the sequence, respectively<sup>47</sup>. There can be apparent contradictions in the relationship between the side chain interactions and the dynamics of assembly. For example, in a protein structure, W has the least flexible side chain as it is large and planar, lacking the degrees of conformational flexibility found in, for example, L and I<sup>48</sup>, but is frequently found in phase-separated liquid condensates and in highly dynamic disordered soluble assemblies, the so-called molecular dispersions. This unexpected role of W is related to the dipole moment in the indole ring, which allows W to interact not only via the hydrophobic effect and  $\pi$ – $\pi$  interactions similar to other aromatic side chains but also via intermolecular hydrogen bonds with water and dipole–dipole interactions<sup>49</sup>. In fact, it is the most versatile amino acid in terms of its interaction potential. Introduction of geometrical constraints can alter morphology of the assembly, as demonstrated by the uniform, spherical nanostructures formed by the

## Box 1 | Oversimplification of side chains fails to capture experimental observations

Amino acid interactions depend on their position in a sequence, and sequences with the same composition but different arrangements of amino acids can have similar polarities but different structural properties<sup>35</sup>. Reducing the chemical characteristics of amino acids solely to their polarity and charge oversimplifies the wide range of possible supramolecular interaction modes (Fig. 1b) and clearly can lead to inaccuracies when trying to model peptide self-assembly. One common simplification is to group all amino acids of similar hydrophobicity and charge into one category. For example, Y, F and W are often considered simply as aromatic amino acids owing to their hydrophobicity value and preferred interaction mode via  $\pi$ – $\pi$  stacking. However, the side chains of these amino acids are able to engage in different intermolecular interactions such as the hydrophobic effect (F, W), dipole–dipole interactions (Y, W) and hydrogen-bonding interactions (Y, W), plus the directionality of these hydrogen-bonding interactions as they may compete with backbone–backbone hydrogen bonds<sup>61,170</sup>. Aliphatic and aromatic amino acids can also exhibit different side chain interactions, even though both are hydrophobic, leading to distinct macroscopic properties and phase behaviour. The sequence GLYGGYGX yields a liquid–liquid phase separation-type assembly when X=W and a gel–solid ordered assembly when X=I or V<sup>61</sup> (Fig. 2c).

Such common simplification of side chain interactions underpins coarse-grained computational simulations. Although these types

of simulations are useful for understanding interactions to predict self-assembly tendencies while reducing computational cost<sup>219,220</sup>, they sometimes fail to capture experimental observations. This is mainly because coarse-graining the physicochemical properties of side chains is accompanied by loss of information on the molecular level, such as  $\pi$ -stacking, hydrogen-bonding and peptide–water interactions. Discrepancies between coarse-grained simulations and experiment can often be avoided by applying atomistic simulations that are able to capture the side chain chemistry and sequence context more accurately. For example, the peptides KYW, KYF, KFF and KYY, which have a similar overall hydrophobicity, have been modelled in coarse-grained simulations to show similar aggregation propensities<sup>91</sup>. However, in experiment and atomistic simulations, the aggregation propensities of peptides increased from KFF < KYF < KYY < KYW, owing mainly to hydrogen-bonding interactions of Y and W side chains that could not be represented in coarse-grained models but were accounted for in the atomistic representation<sup>163</sup>. In particular, atomistic simulations revealed that KYW<sup>163</sup> forms non-directional soluble dispersions owing to substantial contributions of backbone–side chain hydrogen-bonding interactions, in contrast to fibre-forming KYF, KFF and KYY in which backbone–backbone hydrogen bonds dominate<sup>91</sup>. Overall, W is commonly observed to introduce multidirectional interaction modes.

## Box 2 | Order, disorder and self-assembly dynamics

Peptide self-assembly is driven by thermodynamic equilibrium, with two contributions from enthalpic change ( $\Delta H$ ) and entropic change ( $\Delta S$ ), in which the most organized structures have the lowest free energy<sup>65</sup>. Increasing structural order coincides with local decrease in entropy, that is, a reduction in number of states. This process is driven by an increase in enthalpy owing to favourable intermolecular interactions and an overall increase in system entropy, for example, with the release of bound solvent molecules<sup>221</sup>.

Supramolecular dynamics (both internal and exchange) stem from the unique energy landscape of a supramolecular structure, which combines thermodynamic stability (well depth) with kinetic factors (activation energy)<sup>28</sup>. For instance, initial exchange dynamics can help a kinetically trapped assembly make small adjustments in molecular packing within a few hours, eventually leading to a more stable, ordered structure and thereby reducing overall exchange dynamics<sup>222</sup>. Supramolecular dynamics in peptides can be fine-tuned by adjusting various attractive and repulsive forces, such as hydrogen bonds between amino acids<sup>51</sup>, electrostatic interactions, steric effects and polarity<sup>28</sup>, and can be typically quantified via super-resolution microscopy<sup>223</sup>, NMR<sup>224</sup> and electron paramagnetic resonance spectroscopy<sup>51</sup>. Monomer exchange between self-assembled structures occurs when the interactions among monomers are strong enough to form stable assemblies, yet not so strong that they inhibit dynamic exchange. This balance is typically achieved when the interaction strength is about 5–10 times the thermal energy ( $k_B T$ )<sup>225</sup>.

Several factors influence the assembly of peptides into ordered or disordered structures (Fig. 2). The interactions between amino acids depend on the environmental conditions, neighbouring intramolecular amino acids within the same sequence and intermolecular amino acids from proximal peptides — all of which strongly affect the degree to which certain interactions are favoured. Moreover, water organization in and around a peptide sequence has a crucial role. Dynamic, disordered peptide assemblies are highly exposed to water<sup>226</sup> (Fig. 2ai–iii) and impose fewer restrictions on the mobility of interfacial water compared with ordered, non-dynamic structures<sup>227</sup> (Fig. 2aiiv). The increased peptide dynamics observed in disordered structures are related to nonpolar residues facing the aqueous exterior, which result in less-organized hydrogen bonding with interfacial water<sup>160</sup>, and to polar residues within the structure's interior that can retain water<sup>228,229</sup> (Fig. 2aiii). By contrast, ordered, stable assemblies typically have nonpolar residues buried inside and polar residues on the exterior that are capable of forming a hydrogen-bonding network with interfacial water, which overall stabilizes the structure<sup>230</sup> (Fig. 2aiiv).

The transition towards equilibrium relies on a complex combination of dynamics across length scale and timescale. Supramolecular dynamics underpin the structural diversity of peptide self-assembly, both for ordered and for disordered structures. Dynamics arise from both intermolecular interactions and intramolecular backbone conformations<sup>231,232</sup> and can be categorized into several types based on length scale and timescale: the rotational and vibrational freedom of monomers (monomer dynamics), the rate of formation of a supramolecular structure (assembly dynamics), rate of monomers exchanging

positions within (internal dynamics) and in and out of (exchange dynamics) an assembled structure and interactions between assembled structures (interassembly dynamics)<sup>233</sup>. Each of these requires different approaches for experimental measurement and simulation<sup>234</sup>. Computational and experimental data indicate that internal dynamics occur on a microsecond scale, whereas exchange dynamics happen on a millisecond scale<sup>223,235</sup>. Assembly dynamics, which describe the rate of intermolecular interactions, can occur over a timescale ranging from seconds to days<sup>236</sup>, depending on the interaction types and thermodynamic energy landscape involved<sup>233</sup>. For example, hydrogen bonding is directional and less dynamic compared with the hydrophobic effect, which is non-directional with high assembly dynamics<sup>237</sup>. Furthermore, these assembly dynamics can be influenced by mechanical forces, such as shaking versus stirring<sup>202</sup>. Influencing assembly dynamics through changing conditions enables access to supramolecular structures that do not present the thermodynamic minimum, but kinetic ensembles<sup>231</sup>. For example, thermal annealing of the peptide FF results in the thermodynamically stable structure, hexagonal crystals, whereas rapid and uncontrolled self-assembly through dilution from a good to a bad solvent yields the kinetic product, nanotubular structures<sup>238</sup>. Free-energy contributions from enthalpy arise from intermolecular interactions and are generally studied via microscopic methods, whereas contributions from entropy, that is, the number of states, are more complex and can be determined by temperature-dependent calorimetric experiments<sup>221</sup> and theoretical models<sup>239</sup>.

Pathway complexity (and the associated assembly dynamics) is important to the structural diversity of peptide self-assembly<sup>240,241</sup>. This term describes the variety of routes a peptide monomer can take within a free-energy landscape to form a supramolecular structure<sup>240,241</sup>. Unlike simple nucleation and elongation processes, complex pathways may involve additional steps of kinetically trapped and out-of-equilibrium structures<sup>191,242</sup>, which allow access to different positions in the energy landscape that ultimately dictate the structure formed<sup>240,243–245</sup>. Different positions in the energy landscape can influence the nanostructure and potentially the function of a supramolecular peptide structure, without the need to alter the molecular sequence itself. This impact of pathway complexity on function has been demonstrated in studies of peptide fibres prepared under varying environmental conditions. By selectively switching attractive and repulsive side chain forces on or off, varied morphologies of peptide fibres (such as length and dispersity) have been achieved, and the effects of different preparation conditions — such as incubation time, temperature and ionic strength — on their toxicity have been examined<sup>28</sup>. For example, in a bioactive peptide amphiphile comprising the sequence V<sub>3</sub>A<sub>3</sub>K<sub>3</sub> and a 16-carbon alkyl chain, a correlation was found among thermodynamically stable, short fibres prepared at a low ionic strength and high cellular toxicity, in contrast to less toxic, long fibres prepared at high ionic strength<sup>28</sup>. Furthermore, temperature gradients can be used for pathway-dependent self-assembly by balancing which interaction dominates in a peptide; this strategy was shown for Fmoc-YL, which interacts mostly via π–π stacking at high temperature (>333 K) and via hydrogen bonding below the transition point, leading to differential assembly of the same peptide through competing interactions,

(continued from previous page)

even resulting in chiral inversion that could be kinetically locked through gelation<sup>243</sup>. Nature leverages a special type of assembly in which building blocks undergo exchange and assembly through pathway complexity by using transient, out-of-equilibrium assembly in active matter, such as in protein assemblies in microtubules and in synthesized peptide systems via chemically fuelled reactions driven by exchange and assembly dynamics<sup>246–249</sup>. The effective gain

in energy depends on the sequence context — that is, the specific amino acid order and environmental conditions whose contributions to supramolecular dynamics mostly remain elusive at this time<sup>233</sup>. More systematic exploration of the effects of sequence context on peptide assembly could also reveal its effects on dynamics across length scale and timescale, ultimately guiding the design of instantaneously responsive yet stable structures.

non-natural diphenylglycine, in contrast to the nanotubes formed by the natural diphenylalanine<sup>30</sup>. Diphenylglycine has molecular properties similar to diphenylalanine but is more rigid, owing to higher steric hindrance and lack of rotational freedom around its additional C–C bond<sup>30</sup>. Such geometrical constraints from methyl-group removals and insertions have been found to have a high impact at certain positions. For example, adding a methyl group to a nitrogen atom in an amide bond close to a hydrophobic core in a peptide nanotube can disrupt packing and increase internal and exchange dynamics<sup>51</sup>, and moving a terminal methyl group of the isobutyl side chain of L to yield the linear side chain regioisomer norleucine improves the packing in crystals of dipeptides with F, which, in turn, increases the tunability of crystal water channels simply through amino acid order and chirality<sup>52</sup>. Some studies demonstrate that β-carbon removal<sup>50,53,54</sup> decreases packing and assembly order in peptide crystals, indicating that this position is important for self-assembly. In addition, this finding suggests that I-to-L substitution retains hydrophobicity while enhancing dynamics; however, more studies are needed to systematically understand this relationship.

## Designing sequence context-dependent order and disorder

Because a peptide's function and structure are inherently intertwined with its context, the self-assembly state can be tuned by considering conditions and sequence holistically. Under different conditions, similar sequences can self-assemble into a multitude of structures with varying degrees of dynamicity, order and disorder<sup>55</sup> (Box 2 and Fig. 2).

Placement of aromatics in a specific sequence can be used to enhance order, or disorder, depending on sequence context. Combinations of aromatic and cationic amino acids can act as weak associative cation–π pairs and yield disordered structures<sup>33,56</sup>. Increasing the hydrophobic content generally enhances the assembly propensity, with clusters of hydrophobic groups leading to solid-type aggregates as hydrophobic side chains preferentially bury in the interior of a

structure to avoid the enthalpic penalty of unsatisfied hydrogen bonds with water<sup>57</sup>. Notably, the solvent interactions and contributions from the hydrophobic effect vary depending on block-type versus alternating arrangements of hydrophobic residues, which consequently influences the degree of order in self-assembly<sup>33</sup>. This is evident even in very short peptides. For example, the impact on (dis-)order of clustering or spacing out polar and hydrophobic residues is demonstrated with amidated D/Y/F tripeptides<sup>58</sup>. When aromatic amino acids (X = Y or F) are separated by the polar side chain D (XXD), their conformational arrangement is such that the aromatics self-stack and the D side chain is water-exposed, forming amphiphilic conformations that remain highly soluble at 20 mM (ref. 58). By contrast, XXD or DXX gives rise to high degrees of aggregation. The sequence also dictates the dihedral angle between the aromatic groups, with XXD predominantly leading to self-stacking of neighbouring side chains (*syn* conformation), similar to the well-known self-stacked FF conformation in dipeptide nanotubes<sup>59</sup>. By contrast, DXX gives rise to the *anti* conformation with both aromatic side chains on opposite sides of the peptide backbone, leading to better intermolecular stacking and higher order and in turn to well-defined nanofibres and crystals<sup>58</sup> (Fig. 2b). Remarkably, these conformations could be observed in molecular dynamics simulations and in crystal structures<sup>58</sup>, suggesting that for these tripeptide systems side-chain-induced conformers in solution dictate crystal architectures, which suggests opportunities for computational prediction. The generality of such a predictive approach requires additional studies.

Furthermore, ordered structures often contain directional hydrogen bonds that are further stabilized by π–π interactions<sup>60</sup>. These cooperative interactions may be governed by surprisingly subtle changes in the peptide sequence, as demonstrated by the self-assembly of the glycine-rich 8-mer peptide GLYGGYGX (X = G, A, S, K, E, P, F, V, W, I or L)<sup>61</sup> (Fig. 2c). Although the assembly is mainly driven by π–π stacking and hydrophobic interactions of Y, the strength of X's hydrogen-bonding interactions with water (V < W < S) regulates the macroscopic properties (solid, viscoelastic or liquid, respectively)<sup>61</sup>.

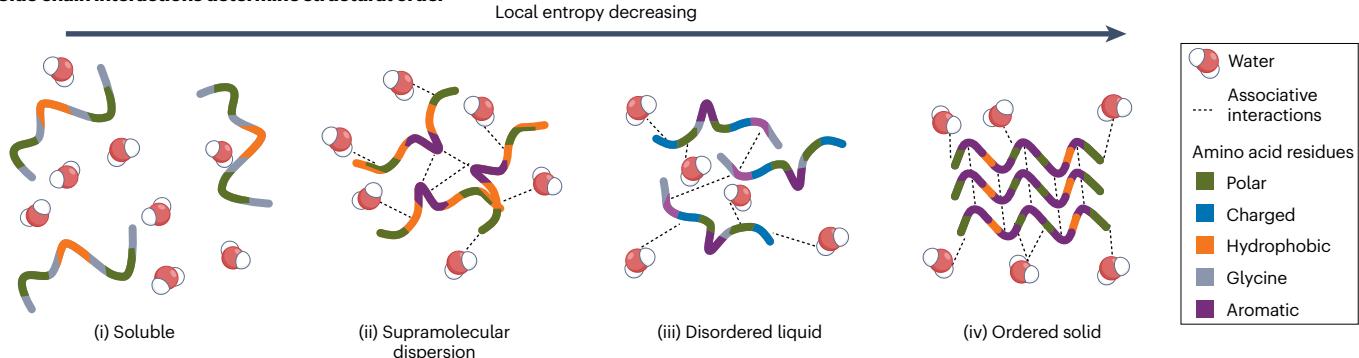
**Fig. 2 | Sequence determines the degree of order in peptide structures.**

**a**, Representation of balance between backbone and side chain molecular interactions in peptide structuring in (i) soluble, (ii) dispersed, (iii) disordered liquid and (iv) ordered solid structures. Soluble peptides are fully hydrated, whereas molecular dispersions of peptides are only partially hydrated. Peptides forming disordered structures such as in liquid–liquid phase separation show an extensive multidirectional intermolecular interaction network with water, in contrast to ordered structures that are stabilized by directional peptide–peptide interactions. **b**, Side chain flexibility can be controlled by amino acid arrangement, exemplarily shown with tripeptide isomers composed of D, F and Y. Side chain flexibility decreases when there are neighbouring aromatic amino acids (Y, F), as dihedral angles are locked into *anti* (DXX) or *syn* (XXD) conformation, which promotes efficient crystal packing. Scale bar, 100 nm.

**c**, Side chain flexibility can be controlled by amino acid mutations, exemplarily shown with the peptide Ac-GLYGGYGX-NH<sub>2</sub>. Single amino acid mutations influence structural order and physical properties from soluble peptide (X = K, A, S, G, E, P) to disordered liquid assembly (W), loose fibrillar network (F, V, I) and solid densely packed structure (L). It is important to note that these amino acids facilitate the above structural ordering through preferred intermolecular interactions in this particular peptide sequence, but that the weighting of preferred interactions can change in a different sequence. Thus, amino acids cannot be universally assigned to a certain order-promoting or disorder-promoting category but need to be considered on an individual basis. Panel **b** adapted with permission from ref. 58, AAAS. Panel **c** adapted with permission from ref. 61, American Chemical Society.

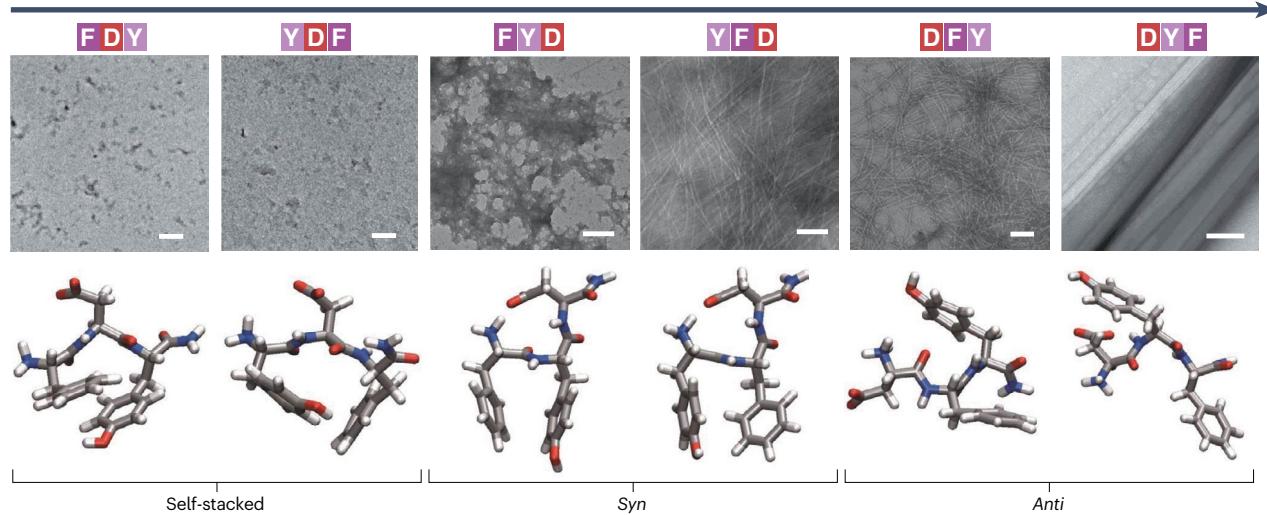
# Review article

## a Side chain interactions determine structural order

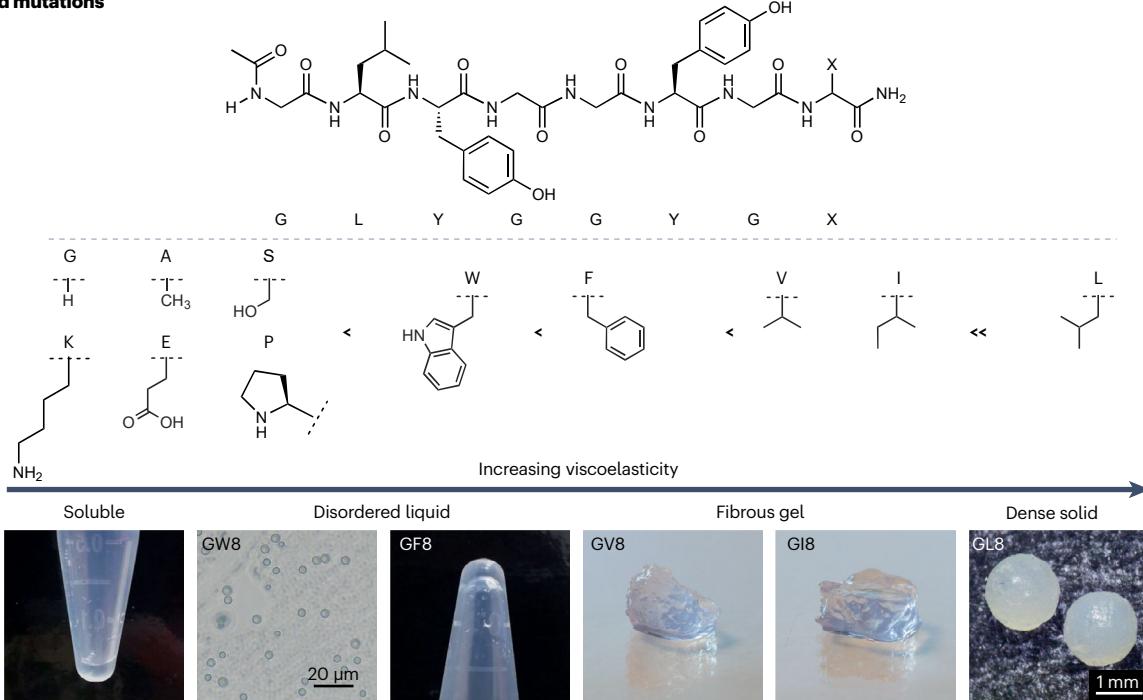


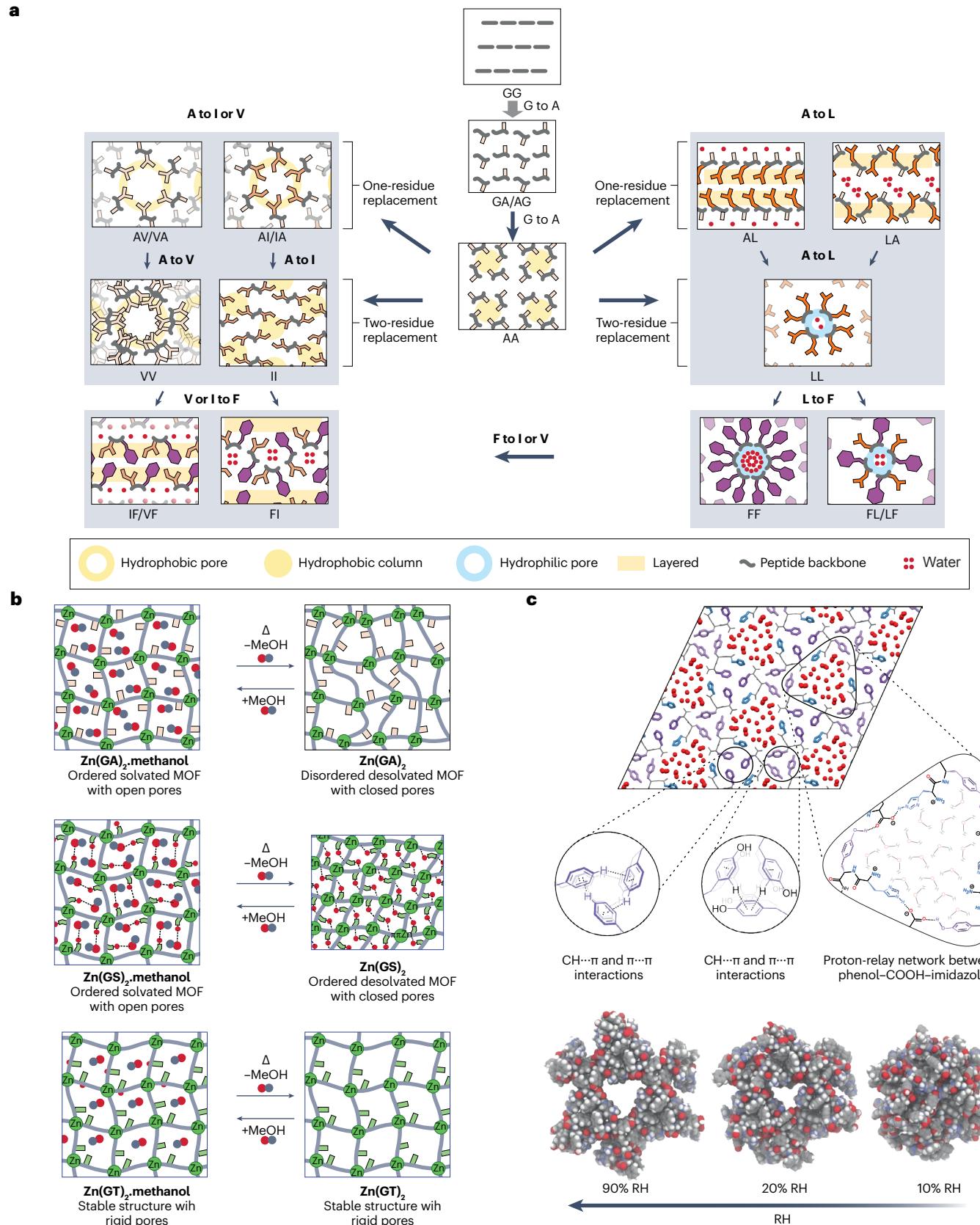
## b Amino acid arrangement

Increasing packing →



## c Amino acid mutations





**Fig. 3 | Ordered systems.** **a**, Side chain interactions of peptide molecules influencing the topology of crystal structure. Substituting one glycine of GG with A (to make GA or AG) causes the bending of the molecule owing to the steric demand of the methyl group, resulting in a zig-zag packing<sup>79</sup>. Replacing both G with A (AA) results in the formation of closed hydrophobic columns packed with methyl groups<sup>80</sup>. These columns transform into hydrophobic pores when nonpolar aliphatic amino acids I or V are introduced, as observed in AV, VA, AI and IA. Whereas VV maintains the structure of VA or AV, II retains AA structure. Conversely, the nonpolar aliphatic amino acid L disrupts the packing

drastically: LA and AL have a layered structure and LL forms hydrophilic pores. Replacing the aliphatic amino acid L with aromatic F (LF, FL and FF) does not alter packing. However, introducing branched aliphatic amino acids (I or V) disrupts packing, forming layered 2D H-bonded structures as seen in FV, VF, FI and IF<sup>54,81</sup>. **b**, Metal-organic frameworks (MOFs) composed of dipeptides GA, GS and GT as ligands demonstrate varying adaptability under reversible solvent inclusion/exclusion conditions. **c**, Reversible expansion and contraction of aqueous pore in HYF facilitated by adaptable aromatic interactions of side chains. RH, relative humidity. Panel **c** adapted from ref. 89, Springer Nature Limited.

Overall, the contribution of a specific amino acid to peptide interactions should be understood in sequence context, which has a dramatic impact on self-assembly, nanostructure formation and phase transitions<sup>62</sup>. Computational models that consider how interactions of amino acids change depending on their arrangement in a sequence and on environmental conditions could<sup>63</sup> allow universal predictions about assembly and interaction modes. Current limitations are due to the difficulty of parametrizing multiple, collective interactions that shape a complex free-energy landscape of possible self-assembled structures<sup>64,65</sup> and of accounting for changes in side chain pK<sub>a</sub> upon changes of environments, as in during self-assembly<sup>66</sup>. Using interpretable machine learning/AI to rationalize self-assembly, by combining experimental and simulated multidimensional parameter spaces, could provide insights into these interactions that surpass human information processing capabilities<sup>23,24,67,68</sup>. One way to gain insights into the multiparameter dependency is to perform systematic investigations, such as changing amino acids or conditions one at a time for self-assembly. Having discussed some of the key features that generally impact order versus disorder, we next discuss literature examples that have systematically investigated changes in peptide sequence, categorized by the general classes of ordered (fibres, crystals with specific nanoscale topology as well as dynamic components) and disordered (condensates and glasses) assemblies.

## Ordered structures with nanoscopic and dynamic domains

Ordered structures of peptides can manifest as thermodynamically stable 1D fibrous structures or 3D crystals<sup>69</sup> and, rarely, extended 2D structures<sup>70</sup>. Each of these morphologies shows hierarchically long-range ordered molecular packing. Crystals offer the advantage of unambiguous structure elucidation through single-crystal X-ray diffraction. Although sequence-specific peptide crystals have been extensively examined elsewhere<sup>71,72</sup>, we focus on how side chain interactions dictate packing and enhance dynamics within the globally ordered phase.

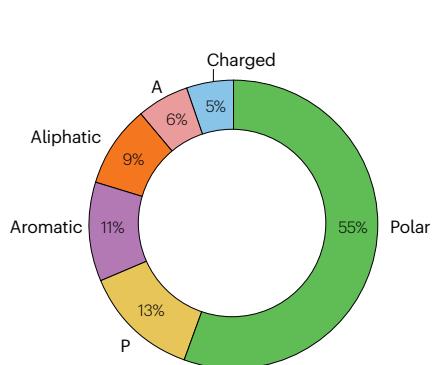
Because most studies on peptide assembly are conducted in water, and the solvent is a critical environmental condition for self-assembly<sup>73</sup>, hydrophobic peptides have been reported extensively in studies aiming to create ordered structures. Usually, these peptides are assembled upon dilution into water from a concentrated organic solvent that suppresses intermolecular interactions, as shown for the well-known amyloidogenic assembly motif FF<sup>74</sup> and AAKLVFF in water and methanol<sup>75</sup>. Besides FF, other hydrophobic dipeptides have been demonstrated to form structures with crystalline order<sup>53,76</sup>. In such minimalistic peptide systems, the influence of different sequence contexts on side chain interactions can be studied. For example, compared with the dipeptide VF, the additional side chain methyl group of IF increases the hydrophobicity of the peptide and promotes its self-assembly and gelation in water<sup>53</sup>. Chirality can also determine self-assembly, as shown for

VFF, FFV and LFF<sup>76,77</sup>. These peptides produce ordered structures in water without the need for organic solvents only when the N-terminal amino acid is switched from the D-configuration to L-configuration. This change facilitates a steric-zipper type assembly (leading to F-F interdigitated amyloid β-sheets) by extending the molecular length owing to the opposite chirality of the N-terminal amino acid<sup>76,77</sup>.

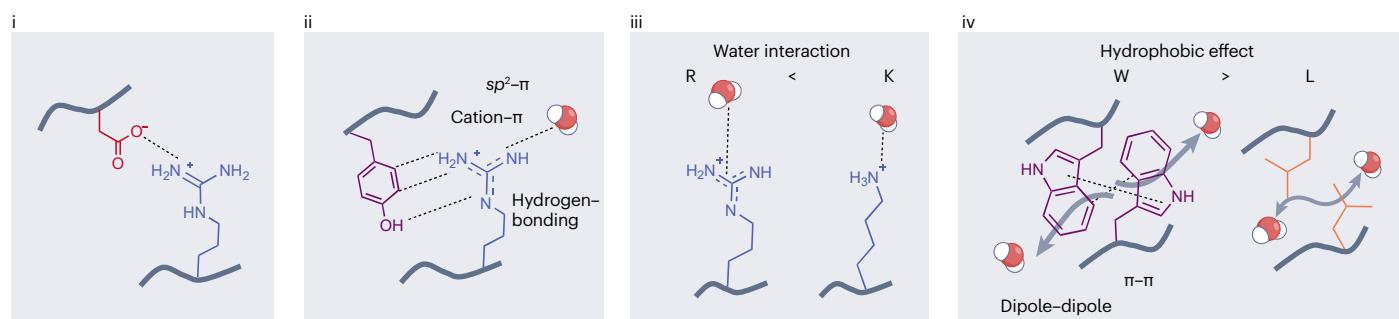
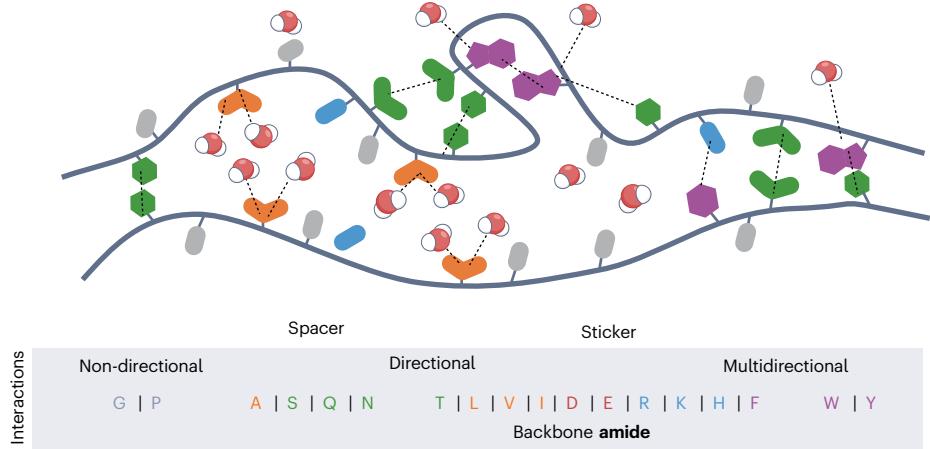
Peptide side chains and water collectively influence crystal packing. Along with water, the side chains in dipeptides can give rise to remarkably diverse architectures with distinct nanoscale domains, including those composed of hydrophobic pores, hydrophilic pores or closed columns packed with side chains (Fig. 3a). The peptide GG (as well as GGG, GGGG and GGGGG) can be viewed as a bare scaffold because its non-bulky hydrogen side chains, through efficient backbone-backbone alignment, form antiparallel β-sheet structures<sup>78</sup>. Systematic substitution of these amino acids with others of varying steric demand and hydrophobicity changes the crystal packing<sup>54,79–81</sup>. For example, substituting one G of GG with A (as in GA or AG) adds the steric demand of a methyl group, resulting in a zig-zag packing. AA forms hydrophobic columns in a tetragonal space group through electrostatic interactions between terminal ammonium and carboxylic acid. Dipeptides comprising a longer aliphatic side chain with branching close to the backbone, as in VA, AV, AI, IA and VV, form a 3D structure of hydrophobic micropores<sup>54</sup>. This microporous structure is a result of a honeycomb-shaped hydrogen-bonding network of dipeptides that arises because the steric demand of the branched side chains favours hydrogen bonding between terminal carboxylate lone pairs and the peptide backbone N–H and C–H (ref. 54). The microporous structure transitions into a layered structure with either of the following substitutions: (1) a residue with a small methyl side chain is replaced with an aromatic side chain (as in VF, FV, FI and IF); or (2) V or I, which has side chain branching close to the backbone, is replaced with L, which has a side chain branching farther from the backbone (as in LA and AL). In such structures, the hydrophobic and hydrophilic interfaces form distinct layers. The hydrophobic micropores of VA, AV, AI and IA change to a hydrophilic micropore when both of these substitutions (1 and 2) are done simultaneously – that is, A is replaced by F, and V or I is replaced by L, as in LF or FL. In such structures, the aromatic and aliphatic side chains align in a circular arrangement around a hydrophilic pore filled with water<sup>54</sup>. Strikingly, the size of these pores can be adjusted by altering the steric demand of the side chains. In extreme cases, such as in the dipeptide crystal FF, the two hydrophobic side chains with high steric demand force the dipeptide torsion angles to adapt to an uncommon high-energy conformation, with torsion angles close to 0°, that can accommodate a regular packing with a nanoporous structure<sup>54</sup>. Clearly, simple variation of aliphatic and aromatic side chains in dipeptides can give rise to nanoporous materials with variable architectures and hydrophobic and hydrophilic domains and channels.

# Review article

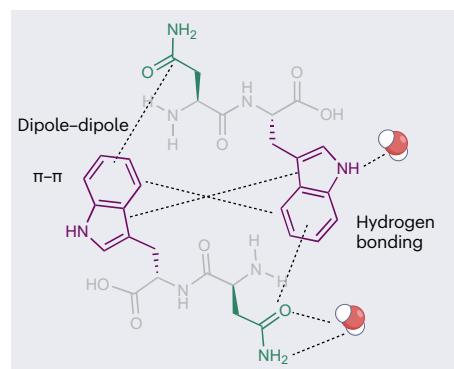
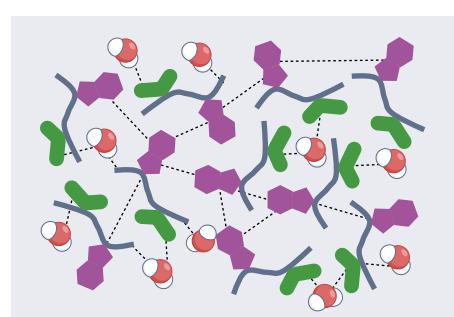
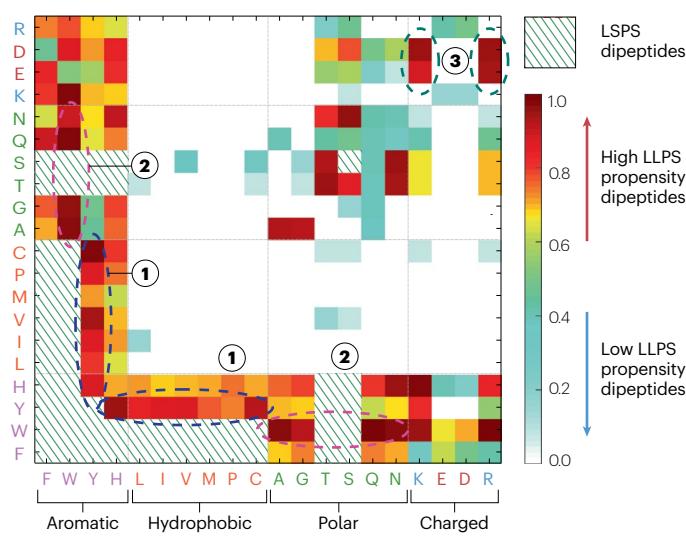
## a Statistical distribution in IDPs



## b Sticker-spacer model: peptides and proteins



## c Multivalency model: dipeptides



**Fig. 4 | Disordered systems.** **a**, Percentage of amino acid content in low-complexity domains of intrinsically disordered proteins (IDPs)<sup>105</sup>. **b**, Schematic illustration of sticker-and-spacer-type interactions with amino acid side chains categorized by overall contributions to directionality. Liquid–liquid phase separation (LLPS)-forming peptides are mostly stabilized by multidirectional and water interactions: (i) ionic interactions such as between D and R and (ii)  $sp^2$ – $\pi$ , cation– $\pi$  and hydrogen-bonding interactions such as between Y and R. (iii) Cation– $\pi$  interactions are stronger for R compared with K, which is owing to less favourable water interactions of the delocalized electrons

in the guanidinium group of R. (iv) An increase in the number of interaction types enhances the total associative engagement, as exemplified in W–W compared with L–L interactions. **c**, Heatmap (left) indicating LLPS propensity of amino acid pairs in minimalistic dipeptides. In contrast to the types of amino acids in IDPs, hydrophobic and charged amino acids are more prevalent in LLPS-prone dipeptides. Schematic illustration (right) of LLPS-forming peptides based on multivalent interactions. The selected example highlights QW, the dipeptide with the highest calculated LLPS tendency<sup>123</sup>. LSPS, liquid–solid phase separation. Panel **c** adapted with permission from ref. <sup>123</sup>, Elsevier.

Although crystals are mostly perceived as static owing to the rigid packing of molecules, externally applied changes in environmental conditions can induce dynamic properties<sup>82,83</sup>. This can be observed, for example, with an exchange in solvent as demonstrated for the metal-peptide framework Zn(GX)<sub>2</sub> with X = A or S (refs. <sup>84,85</sup>) (Fig. 3b). The metal maintains the packing by coordinating with the peptide, whereas weak van der Waals interactions of A side chains or hydrogen bonds of S side chains, combined with the flexibility of the G backbone, drive a dynamic adjustment in pore size when exposed to methanol<sup>84,85</sup>. Interestingly, when X = T, the additional hydrogen bonding introduced by the hydroxyl groups stabilizes the pore and maintains its structure even after removal of methanol<sup>86</sup>. Metallo-peptide systems are also able to exploit hydrophobic properties of amino acids to enable water responsiveness in the solid state. In metallo-peptide fibres prepared in methanol–water mixtures and dried into a solid fibre network, increasing the hydrophobicity from GH/Zn to AH/Zn decreases backbone hydrogen bonding with water, allowing the fibres to be reversibly destabilized and disassembled in the presence of water vapour<sup>87</sup>. WHHW can act as a stimuli-responsive emulsifier for water insoluble drugs: in the presence of metal ions, the amino acid residues rearrange to form an amphiphilic complex with a polar head (H–metal–H) and two hydrophobic tails (W), stabilizing oil–water emulsions<sup>88</sup>.

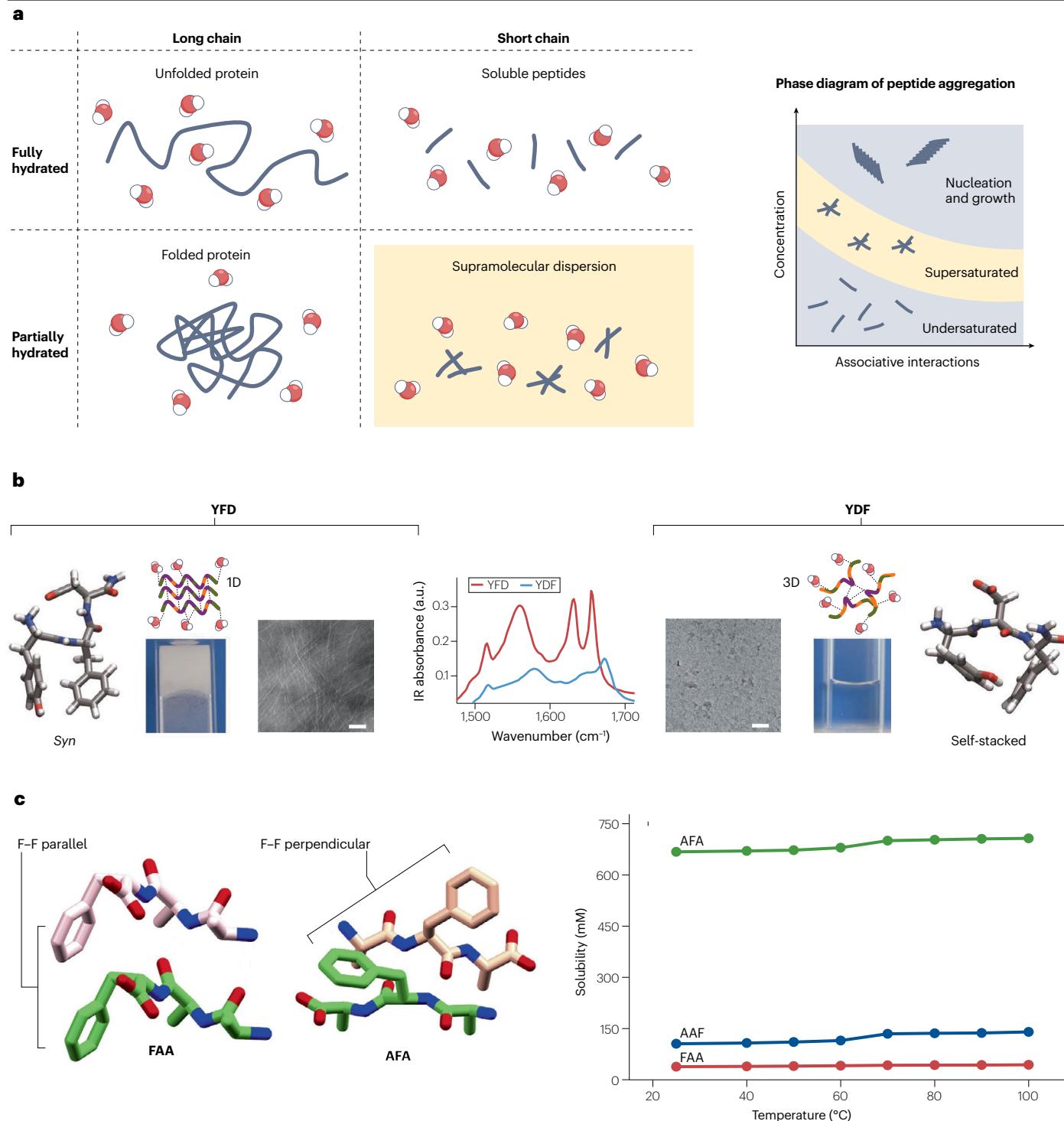
Solid-state crystals can also adapt to different humidity levels through molecular motion. For example, the tripeptide HYF forms a dual supramolecular network of aromatic regions that are interspersed with water channels (Fig. 3c). Distinct, self-segregated triangular columns, comprising the aromatic side chains from Y and F, are arranged around a central 10 Å pore filled with water molecules<sup>89</sup>. The structural integrity of the system is supported by the strengthening of the proton-relay network around the lining of the pores between proton donor–acceptor residues (phenol–COOH–imidazole) upon dehydration<sup>90</sup>. The deformability of aromatic regions in such hydrated crystals strongly dictates the water responsiveness. This is evident when comparing HYF, FF and F:Fcrytals, which feature less interactive and more deformable aromatic regions, are more water responsive than HYF and FF crystals and which have strongly interacting aromatic regions<sup>69</sup>. Amino acids of different polarities determine solvent interaction, and a combination of two paired aromatic amino acids (as a strong self-assembling motif) and a sufficiently polar residue can give rise to gelation, as observed, for example, for KFF or KYF (hydrogel) in contrast to PFF (suspension)<sup>91</sup>. Similarly, fibre-forming peptides, similar to most aromatic dipeptides such as FF, can be turned into gelators by appropriately positioning polar residues that make the fibre surface interact favourably with water. These gel fibres can be thermodynamically preferred over crystals depending on the amphiphilicity of the peptide<sup>69</sup>.

One-dimensional fibres are readily formed in various protein and peptide designs, including self-assembling amphiphilic peptides

functionalized with aliphatic tails at the N terminus. These structures offer tunable dynamics while maintaining  $\beta$ -sheet-like order<sup>51</sup>, as demonstrated for C<sub>16</sub>XXXE<sub>4</sub>GIKVAV, in which combinations of amino acids (X = G, A or V) determine the hydrophobicity and flexibility of the fibres<sup>92</sup>. Intriguingly, these tunable dynamics lead to differential neuronal cell responses cultured on these materials, with neurons performing better when exposed to dynamic fibres, indicating their potential as scaffold materials<sup>92</sup>. This effect is also relevant in other biomedical applications such as viral delivery, where amino acids A and G induce moderate hydrophobicity and flexibility, and charged amino acids such as K, R and E, placed near the hydrophobic aliphatic region, enhance dynamics and prevent large aggregation of fibres<sup>93</sup>. Increased internal dynamics and exchange dynamics, in turn, facilitate better cell interaction and degradation, making dynamic fibres advantageous for various applications<sup>92,93</sup>, including formulations for storage and release of proteins in peptide–amphiphile hydrogels<sup>94</sup>. Although it is still elusive how these small changes in the sequence affect mesoscopic fibre–fibre interactions<sup>95</sup>, atomistic molecular dynamics simulations suggest that both core and surface hydrophobicity contribute to the monomer dynamics<sup>96</sup>. Designing structures with strategically positioned hydrophobic and polar amino acids can thus be a feasible way to control dynamics and aggregation propensity. For example, the sequence XLVIY has a higher aggregation propensity when X = V compared with when X = K (ref. <sup>97</sup>), and FFYX has a higher aggregation propensity when X = G than when X = R (ref. <sup>98</sup>). In either case, K or R is preferentially solvated by water and disrupts self-assembly that is driven by the hydrophobic effect<sup>98</sup>.

## Disordered structures, glasses and condensates

The functional properties of disordered structures have been widely acknowledged in the context of liquid–liquid phase separation (LLPS), and solid disordered glass states have been demonstrated as the kinetically trapped states of supercooled peptides. These so-called peptide glasses have been found for various cyclic (such as  $\text{cyclo}^{\text{a}}\text{FP}$ )<sup>99</sup>, modified (such as acetyl-F, cyclobenzene-FFG)<sup>100</sup> and non-modified (YYY)<sup>6</sup> short peptide sequences upon thermal heating and fast (10 K min<sup>-1</sup>) cooling. A requirement for glass formation is thermal stability of the peptides, to ensure they do not decompose near their melting temperature. It has been found that the glass transition temperature inversely correlates with glass network connectivity, which decreases with side chain  $\pi$ – $\pi$  aromatic interactions (glass transition for  $\text{cyclo}^{\text{a}}\text{PY} < \text{cyclo}^{\text{a}}\text{WY} < \text{cyclo}^{\text{a}}\text{WW}$ )<sup>99</sup>. These breakthroughs in biomolecular glass preparation may have implications for biomedical implant research, such as subcutaneous controlled long-term (weeks to months) drug release<sup>100</sup>. Future research may even reveal alternative methods for achieving peptide glass transition beyond thermal control, such as using designed sequences and selecting environmental conditions such as mechanical forces or pressure.



In recent years, there has been an increased interest in LLPS formation from short peptides. These studies take their inspiration from observations that a substantial percentage of proteins are partly or predominantly disordered so that associative interactions exist, but do not dominate<sup>101</sup>. The resulting peptides or proteins are neither fully hydrated nor fully aggregated, and they are dynamic and take on regulatory roles that differ depending on their supramolecular context.

Distinct from their ordered counterparts, segments of these proteins are in fact best described as a dynamic ensemble of transiently interconvertible configurations, sampling multiple Ramachandran angles<sup>102,103</sup>. In the context of biology, these structures are known as intrinsically disordered proteins, some of which form LLPS structures<sup>104</sup>. Usually, intrinsically disordered peptide regions are rich in polar amino acid side chains such as G, S, T, N, Q and C that are arranged with

**Fig. 5 | Supramolecular dispersions.** **a**, Schematic representation of molecular dispersions from short-chained peptides. Folded proteins that have a partially hydrated sequence chain can be compared with molecular dispersions that are similarly partially hydrated but consist of an ensemble of several peptide chains. Fully hydrated soluble peptides are compared to a fully hydrated unfolded protein chain<sup>168</sup>. **b**, Comparison of fibre-forming YFD (left) and dispersion-forming YDF (right). YFD participates in directional side chain interactions owing to its locked *syn* conformation. By contrast, YDF participates in multidirectional intermolecular interactions owing to its higher side chain flexibility and self-stacked conformation, and it forms a macroscopically transparent solution. Micrographs of each tripeptide show the morphologies

resulting from the respective flexibilities of the side chains, providing order and disorder at the intermolecular level as observed via infrared (IR) spectroscopy. Scale bar, 100 nm. **c**, Comparison of solid nanosphere-forming FAA and dispersion-forming AFA. The long-range ordered morphology of FAA versus the disordered aggregates of AFA can be explained by torsion angles of F–F in molecular dynamics simulations (left). Intermolecular F side chains in FAA are aligned in parallel, whereas they are aligned perpendicularly in AFA. This subtle difference in amino acid arrangement leads to a large difference in water interaction, that is, solubility (right). a.u., arbitrary units; IR, infrared. Panel **b** adapted with permission from ref. 58, AAAS. Panel **c** adapted with permission from ref. 164, Elsevier.

associative, hydrophobic amino acids such as F, W, Y, H, L, I and V (refs. 105,106) (Box 1 and Fig. 4a). Owing to its inability to form directional hydrogen-bonding interactions and conformational restraints of its cyclic backbone, P is highly abundant in intrinsically disordered peptides as a disorder-promoting amino acid<sup>107</sup>. However, the design rules of LLPS structures are less well understood compared with structured proteins, in large part because most experimental techniques are better suited to measure ordered structures that are precisely defined. We note that this area is developing rapidly, with new papers appearing daily, and has been reviewed numerous times, some examples of which we refer the interested reader to refs. 108–110. In the current paper, we focus on selected studies that systematically investigate short peptides where sequence context is apparent.

**Side chain interactions in LLPS systems.** The sticker-and-spacer model – in which amino acids that interact attractively are termed ‘stickers’, and those that do not undergo any specific interactions are ‘spacers’ – is a useful framework for understanding and aiding the rational design of LLPS peptides<sup>111</sup>. For example, free-energy potentials can be assigned for the ‘stickiness’ of side chains<sup>112,113</sup>. Common sticker-categorized amino acids have charged (R, K, E, D) and hydrophobic (such as Y, F, W, I, L, H) side chains that interact via ionic<sup>114</sup> (Fig. 4bi), cation–π<sup>115</sup> and  $sp^2$ –π, hydrogen bonding<sup>116</sup> (Fig. 4bii,iii), dipole–dipole<sup>117</sup>, π–π and hydrophobic interactions<sup>118</sup> (Fig. 4biv). They are separated by disorder-promoting spacers such as P and G<sup>106</sup>.

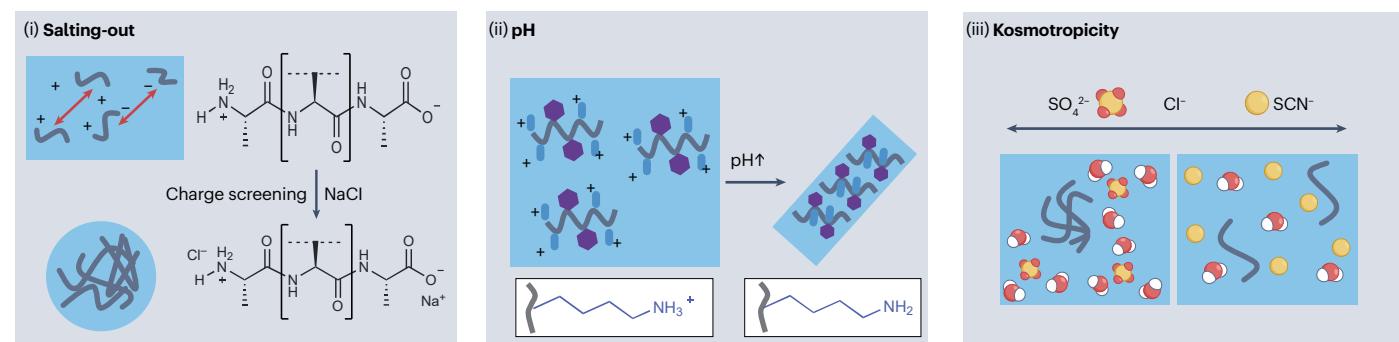
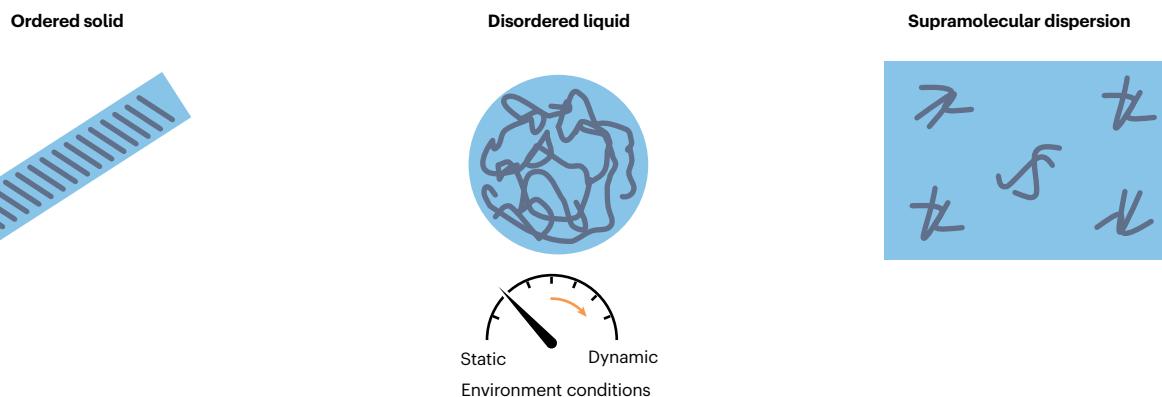
Although this sticker-and-spacer model is helpful to explain many observations in biology, it does not exclusively explain LLPS. It has been supplemented with a multivalency model that describes non-sticker-classified amino acids (S, A, N, Q; Fig. 4c) that drive phase separation by collective interactions of side chains and backbones<sup>119</sup>, such as van der Waals interactions in GGXGG peptides (X = G, Q, N, V)<sup>120</sup>, hydrogen-bonding interactions in G<sub>10</sub>, the hydrophobic effect in V<sub>10</sub> and salt bridges in the presence of high salt concentrations in R<sub>10</sub> or D<sub>10</sub> (ref. 121). The impact of multivalency is especially pronounced for very short (di and tri) peptides. The main reason for this is that minimalistic peptides inherently cannot be designed with repeating adhesive and flexible amino acids. Additionally, the high entropic costs associated with small peptide self-assembly must be compensated by more interactions per molecule. In these cases, peptides form LLPS structures via multivalent intermolecular interactions, as shown using Martini<sup>122</sup> coarse-grained molecular dynamics simulations of simple<sup>123</sup> and complex coacervates<sup>63</sup> made of dipeptides (Box 1 and Fig. 4c). Usually, this type of simulations performs well in identifying solid self-assembling structures but cannot distinguish between LLPS and soluble dipeptides<sup>124</sup>. However, by considering cluster size, number and exchange rate of molecules between two phases, several dipeptides that

display LLPS could be identified. A, G, Q and N side chains were found to undergo LLPS in combination with W, whereas the corresponding Y, F and H variants are unable to phase separate<sup>123</sup>. The dipeptide with the strongest tendency to undergo LLPS was found to be QW<sup>123</sup>. Although LLPS formation for this particular peptide was confirmed experimentally, the multivalent W–W and Q–W interactions, including hydrogen bonding, aromatic stacking, and anion–π and cation–π stacking, have to date only been shown computationally (Fig. 4cv) and remain to be experimentally verified, including beyond this single example<sup>123</sup>. We note that formation of liquid droplets has also been observed previously in short peptides, but not always labelled as LLPS<sup>125</sup>. For example, Fmoc-YQ peptides form liquid droplets, whereas Fmoc-YN, Fmoc-YT and Fmoc-YS give rise to fibre morphologies, demonstrating that the additional CH<sub>2</sub> in Q compared with N causes a steric demand that prevents molecular packing into fibres. However, although T has a greater steric demand than S, both Fmoc-YT and Fmoc-YS peptides form fibres because the hydroxyl group in T and S contributes additional hydrogen bonding that stabilizes the morphology<sup>125</sup>. This and other studies show that hydrogen-bonding interactions from backbones and side chains contribute strongly to the stabilization of ordered supramolecular peptide assemblies<sup>70,126–128</sup>.

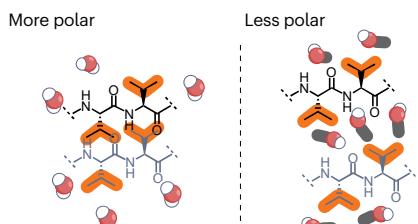
**Mapping the LLPS boundaries of peptides.** Aside from the recent surge in interest in protein-based LLPS, the influence of various conditions on LLPS has long been recognized and systematically studied for polymers<sup>129–132</sup>. Theoretical frameworks for LLPS were originally established for charged and flexible polymers, which assign interaction coefficients<sup>133</sup> and consider intermolecular and intramolecular interaction strength, the local electrostatic environment and even the chain connectivity, that is, the sequence context, of the polyelectrolytes<sup>33–35,134,135</sup>. We refer the interested reader to a review on LLPS in polymers<sup>136</sup>.

The theoretical frameworks established for LLPS in polymers can, to some extent, inform the design of other material classes, including peptides<sup>137</sup>. However, one limitation of using polymer LLPS models to guide peptide design is that these models are developed on polyelectrolytes consisting of simple charge distributions that form complex coacervates, such as polycationic polymers paired with poly-anionic polymers<sup>138</sup>. Currently, there are few known examples of simple coacervate-forming polyampholytes that mimic zwitterionic proteins and peptides. Developing such polyampholytes presents opportunities for multiple interaction modes beyond electrostatics that allow better control over structure under various environmental conditions and, in turn, could guide design of LLPS-capable peptides<sup>139</sup>. In the following, we summarize efforts to systematically investigate changes in amino acids on LLPS boundaries<sup>140</sup>. The impact of changes in environmental

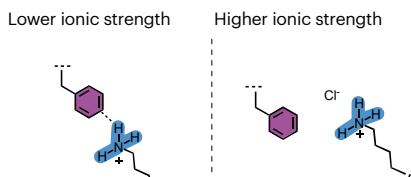
a



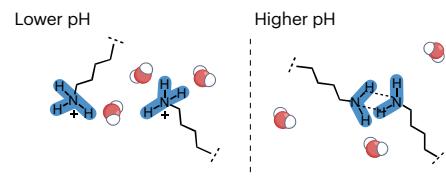
b Solvent apolarity



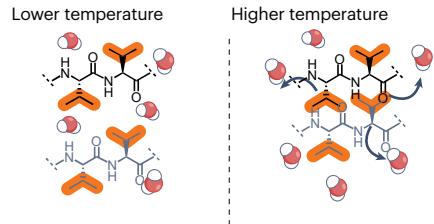
c Ionic strength



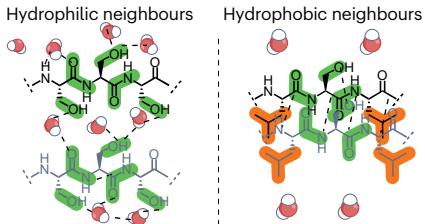
d pH



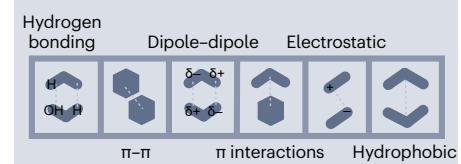
e Temperature



f Hydrophobic neighbours



Non-covalent interactions



conditions, which is also a key factor for LLPS boundaries<sup>141–149</sup>, is discussed in more detail in a later section.

As multiple interactions between amino acids collectively drive phase separation, systematic single amino acid mutations can shed

light on what interaction modes are important to navigate the phase space. Although in general aromatic residues, often Y<sup>33</sup>, and charged residues such as R drive phase separation through hydrogen-bonding,  $\text{sp}^2-\pi$  and cation- $\pi$  interactions<sup>116,150</sup> (Fig. 4bii), this tendency to

**Fig. 6 | Environment-dependent assembly.** **a**, Factors that influence peptide assemblies from ordered to disordered. (i) Charge screening from high salt-ion concentration in charged peptides can promote phase separation, as shown for N-terminal and C-terminal charged peptides<sup>149</sup>. (ii) Self-assembly from disordered into ordered structures can be favoured for K-containing and R-containing peptides upon deprotonation of the side chains at elevated pH<sup>179</sup>. (iii) Different salt-ions have different kosmotropism (water ordering), which influences the order in the peptide self-assembly<sup>149</sup>. **b–f**, Side chain interactions dependent on environmental conditions, illustrated by the most affected interaction modes and amino acids. The grey-scale six-box code indicates which side chain interactions are affected most at illustrated conditions. **b**, Switching the solvent of self-assembly changes almost all interaction modes compared

with water-based systems. The most prominent changes are in hydrophobic effect and hydrogen-bonding interactions. **c**, Electrostatic, cation–π or anion–π interactions decrease at increased ion concentrations owing to charge screening. **d**, Electrostatic interactions are highly pH-dependent. For example, positively charged amine residues carry the same charge and are repulsive. However, if the pH exceeds their  $pK_a$  value, they are deprotonated and can interact via hydrogen-bonding interactions. **e**, The hydrophobic effect dominates at elevated temperatures owing to increase of the entropic term (release of water). **f**, In hydrophilic sequence environments, hydrogen-bonding interactions with water dominate, whereas in hydrophobic environments water is excluded and intermolecular hydrogen-bonding interactions become favourable.

attribute LLPS to only selected amino acids has been challenged in the past year<sup>119</sup>. By systematically exchanging the amino acids G, R and Y in the peptide (GRGDSPYS)<sub>25</sub> with polar uncharged (S, Q, N), charged (K) and nonpolar (A, F, W) amino acids, it was revealed that besides the expected R–Y interaction, the unconventional and often overlooked pairwise contacts R–P, R–S, S–Y and G–Y between charged and uncharged, non-aromatic or aromatic, uncharged and non-aromatic amino acids, yield LLPS<sup>119</sup>.

In both the sticker-and-spacer and multivalency models, two interaction modes are frequently identified as drivers for LLPS assembly: hydrophobicity and cation–π interactions. The impact of the hydrophobic effect on assembly is strongly determined by sequence context. For example, consecutive arrangements of polar amino acids Q and N exhibit a surprisingly low solubility in water, rendering them actively functional as non-directional stickers through the hydrophobic effect<sup>151</sup>. Similarly, in elastin-like peptides with the characteristic repeat unit XPGVG<sup>152</sup>, replacing X with a hydrophobic amino acid and increasing its number of interaction modes (from V to F to W) requires fewer repeat motifs for droplet formation<sup>153–156</sup>.

In the presence of aromatic side chains (Y, F, W), charged side chains (R, K, H, D, E) form short-ranged, directional cation–π or anion–π interactions that often have a critical role in LLPS<sup>43</sup>. The importance of these cation–π interactions was studied in a systematic library of short peptides (WGRGRGRGWPGVGY), where removing Y or W from the sequence or replacing all R with K was sufficient to disrupt LLPS<sup>157</sup>. Compared with K, R interacts more strongly with π-systems<sup>119</sup> because it is less prone to interact with water, owing to the delocalized electrons in its guanidinium group<sup>158</sup> (Fig. 4biii). These delocalized electrons allow R to associate with aromatic amino acids through both cation–π and π–π interactions<sup>150,159</sup> (Fig. 4bii), which might explain its abundance in proteins and peptides that form LLPS. However, K's electrostatic interactions are more susceptible than R's to changes in environmental conditions. In a high salt concentration, K's amine side chains can be readily screened by ions, which, in turn, reduces K's electrostatic repulsion and promotes cation–π bonds with aromatic residues<sup>144</sup>. Nature takes full advantage of these condition-dependent associative interactions, as observed for the decapeptide motif (AKPSYPPTYK)<sub>12</sub> in the mussel foot protein 1, which is rendered adhesive only at seawater conditions<sup>144</sup>. Overall, R-aromatic cation–π interactions imply a hierarchy in the ability of aromatic groups (W > Y > F > H) to regulate the upper critical solution temperature<sup>156</sup>. As all these aromatic groups are equally prevalent in intrinsically disordered peptides, it can be speculated that this diversity is important to allow tunable disorder depending on the sequence and solute environment. For example, W is a potent sticker

in the presence of R, whereas H's associative impact changes at physiologically relevant pH between 6.0 and 7.4 (ref. 156). Besides cation–π interactions between R and W, the enhancement in W's attractive interactions is also traced back to an increase in hydrophobicity as studied in peptides with systematic amino acid mutations<sup>57</sup>. Notably, an increase in hydrophobicity (from L to F to W) increases contributions of associative interactions and eventually leads to solid-type aggregation<sup>57</sup> (Fig. 4biv and Box 1). In a broader context, hydrophobic residues in LLPS systems may be forced to the water-facing interface, which, in turn, weakens the hydrogen bonding and decreases the organization of interfacial water (Fig. 2aii). This was shown in modular assemblies of alternating R/Y adhesive amino acids, separated by tripeptide spacers with varying tendencies of backbone hydration and structuring (SGS, GSG, GLG). When S was exchanged with L in GXG repeat motifs in the general sequence [R(GXG)Y(GXG)], R(GXG)YG, the change of the intramolecular hydrogen-bonding content between side chains and backbones and hydrophobicity dictated the phase between solution (GSG), liquid condensate (GLG) and solid aggregate (SGS)<sup>160</sup>. This example also highlights that side chain can influence backbone orientation; in this case, GLG enhances the  $n \rightarrow p^*$  transition in the backbone, thus reducing its overall flexibility to favour LLPS.

## Soluble supramolecular dispersions

A number of short peptides have been shown to form assemblies that have unexpectedly high solubility without the formation of specific structures<sup>161–164</sup>. This behaviour is strongly sequence-dependent. We propose that peptides with a certain degree of backbone disorder and adhesive side chains combined with polar residues can form these dynamic and fully soluble assemblies. They are currently not separately classified in the literature, but we propose that these are molecular dispersions of nanometre-sized supramolecules, analogous to folded proteins, that are not fully hydrated<sup>161</sup> (Fig. 5a). Supramolecular dispersions are soluble assemblies that are highly dynamic and lack sufficient long-range order that could lead to phase separation. Whereas some studies in the field of protein misfolding consider them to be possible intermediate states to solid, crystalline assemblies<sup>165</sup>, other studies on supersaturated glycine nanoaggregates report them to persist for an indefinite time owing to the equilibrium between molecular and assembled states and conclude that the small size is not sufficient as a nucleation site for crystal growth<sup>162</sup>. Even though it remains elusive, the formation of supramolecular dispersions from peptides could be associated with a non-equilibrium state of metastable intermolecular interactions that prevents peptide aggregation, similar to a supersaturation observed in protein folding<sup>166–169</sup> (Fig. 5a). We therefore believe that understanding the driving force for supramolecular

## Box 3 | The role of water in peptide assembly

Because peptide self-assembly typically takes place in water, peptide–water interactions and the hydrophobic effect are important to consider as they can both enhance and reduce peptide–peptide interactions<sup>87,181</sup> (Fig. 6). The role of water can be seen when comparing the self-assembly of peptides in different (co-)solvents. For example, in methanol, electrostatic, hydrogen-bonding and hydrophobic interactions between peptides or between peptides and solvents are weakened compared with those in water (Fig. 6b), thus changing the aggregated morphology, as demonstrated for peptides acetyl-KIIIIK, acetyl-RIIIR and acetyl-HIIIIH<sup>250</sup>. As methanol cannot (de-)protonate K, R and H, it decreases the impact of electrostatic and polar interactions on assembly and promotes packing of I side chains. This promotes fibre formation but decreases lateral fibre interactions because of missing hydrogen-bonding and electrostatic interactions, resulting in layered systems<sup>250</sup>. To control self-assembly, peptides can be first dissolved in a polar aprotic solvent such as dimethyl sulfoxide or

hexafluoroisopropanol, which suppresses peptide–peptide and water interactions. Subsequent addition of the cosolvent water restores favourable intermolecular hydrogen-bonding and hydrophobic interactions<sup>251</sup>. Because interaction with water is so important in self-assembly processes, salt ions with different kosmotropencies (water ordering) can shift the phase boundaries<sup>149</sup>. For example, sulfate ions compete with peptides to interact with water molecules and thereby promote peptide–peptide interactions, which lowers the critical liquid–liquid phase separation (LLPS) concentration. Conversely, chaotropic bromide ions interact preferentially with peptides, shifting coacervation to higher critical concentrations<sup>149</sup> (Fig. 6aiii). Similarly, ATP, a multivalent metabolite that commonly participates in LLPS of peptides, can act as a kosmotrope by decreasing peptide hydrogen-bonding interactions, which overall increases peptide solubility<sup>252,253</sup>.

dispersion formation in peptides and proteins could establish a unifying framework for the role of amino acid interactions in various aggregation states.

Short peptides that form molecular dispersions are usually rich in hydrophobic multidirectional side chain interactions, such as Y and W, which can accommodate several interaction types including hydrogen-bonding,  $\pi$  and dipole–dipole interactions (Fig. 1b). The multidirectional interactions suppress long-range ordered crystalline assembly, and weak water interactions suppress dynamic exchange of peptides with the solvent. Tetrapeptides (WWWW)<sup>170</sup> and tripeptides (KYW)<sup>163</sup> have been shown to form molecular dispersions, and crystal structures comparing a fibre-forming (YFD) peptide with a dispersion-forming (YDF) peptide revealed that the latter lacks the directional backbone hydrogen-bonding interactions of the former and instead involves competition between side chain and backbone interactions<sup>58</sup> (Fig. 5b).

In fact, screening the entire space of tripeptides for aggregation propensities found that peptides with aromatic side chains are the most aggregation-prone<sup>91</sup>. Studies have demonstrated that the solubility of aromatic tripeptides strongly depends on the positioning of amino acids within the sequence<sup>164</sup>. For example, the tripeptide AFA forms a visibly clear solution up to a concentration of 650 mM, more than fivefold higher compared with its sequence isomers AAF and FAA<sup>164</sup>. The molecular packing of side chains in the aggregated state, which is determined by the sequence-dependent torsion angles of the side chains, correlates with the solubility of the peptide. Specifically, no intermolecular F–F aromatic interactions were observed for AFA because the F side chains point away from each other (30°–90°), leading to relatively higher hydrogen-bonding interactions with water. By contrast, tripeptides AAF and FAA formed solid nanospheres mainly via intermolecular hydrogen bonds and  $\pi$ – $\pi$  interactions from a parallel F side chain arrangement<sup>164</sup> (Fig. 5c). Similar sequence-dependent solubilities with respect to changes in torsion angles have been observed for the tripeptide series composed of F, D and Y<sup>58</sup> (Fig. 2b).

### Environment-dependent assembly

As peptide self-assembly is dependent on environmental conditions, peptides are ideal materials for stimuli-responsive applications in

which external factors such as solvent, pH (charge screening), metal-ion exchange and ionic strength (salting-in/salting-out) can be tuned<sup>149,171,172</sup> (Fig. 6a). Existing stimuli-responsive biomolecular materials are usually slow to respond, especially if their assembly is governed by secondary structure-mimicking arrays of hydrogen bonding, with hours or days of incubation time typical to ensure complete transitions<sup>173,174</sup>. A focus on side chain interactions that can reorganize on shorter timescales in response to changes in their environment provides access to versatile context-responsive assembly. Water interactions can have a key role in stimuli-responsiveness of a peptide (Box 3), as these interactions can hydrate the peptide structure, induce flexibility in the intermolecular arrangement and contribute to disordered structures. For example, interaction strengths of charged and polar amino acids are weakened by dipole–charge interactions and hydrogen bonding of side chains with water, whereas apolar solvents promote intermolecular interactions in polar amino acids<sup>175</sup> (Fig. 6b).

In general, charge-screening effects on Coulombic and polar interactions can either hinder or facilitate the self-assembly process, depending on the charge and local environment of the peptide<sup>176</sup>. For example, a high salt-ion concentration results in charge screening (Fig. 6c) and can either promote self-assembly through non-electrostatic associative interactions<sup>144</sup> or reduce interactions of oppositely charged polyelectrolytes<sup>145</sup>. At high salt concentrations, cation– $\pi$  interactions increase and can compensate for reduced electrostatic interactions<sup>147</sup>. Charge screening from a high salt-ion concentration in the GHGLYGAGFAGHGLHGFGAGHGLY peptide shifts the critical phase separation concentration to lower values owing to salting-out effects<sup>148,149</sup> (Fig. 6ai). Using different ion types can further markedly change the assembly order, as shown, for example, with kosmotropics that support hydrophobic interactions and facilitate the assembly of ordered structures<sup>149,177</sup> (Fig. 6aiii and Box 3). Overall, liquid-to-solid transitions can be readily achieved by changing the ionic environment.

Furthermore, changes in pH can easily protonate and deprotonate the side chains of basic and acidic amino acids, which, in turn, can alter hydrogen-bonding, electrostatic and hydrophobic interactions<sup>146</sup> (Fig. 6d). Thus, pH control is a widely used method to tune the self-assembly properties of peptides<sup>178</sup>. For example, the peptide

KRRFFRRK remains monomeric at pH 7.4, in which K and R are protonated, but quickly self-assembles once the pH is increased above 9, in which K and R are partially deprotonated<sup>179</sup> (Fig. 6aii). The self-assembly is caused by an increase in hydrophobic interaction contributions from F as electrostatic repulsion between K and R is reduced<sup>179</sup>. Other notable examples are the histidine-rich beak proteins that transition between viscosity states at physiologically relevant pH ranges, owing to the p*K<sub>a</sub>* (~6.0) of the imidazole group in the H side chain<sup>180</sup>. In a squid beak protein-inspired peptide designed with modular GHGXY tandem repeats (in which X is P, V or L), histidine acts as a molecular switch: in acidic pH, it hydrogen-bonds mainly with water and does not form structures, whereas upon deprotonation of its side chain at higher pH, hydrophobic interactions, hydrogen-bonding and π–π stacking dominate and favour the formation of aggregates<sup>146</sup>. When ionizable groups become embedded into environments with reduced dielectric constants during self-assembly, the observed, or apparent, p*K<sub>a</sub>* values are known to shift by multiple pH units upon assembly, as shown for Fmoc-FF and other hydrophobic peptides<sup>66</sup>.

Increasing the temperature can either increase or decrease peptide solubility, depending on which intermolecular interactions of the sequence prevail. Peptides that interact via the hydrophobic effect decrease in solubility and favour aggregation owing to the release of intermolecular trapped water, causing a gain in the overall entropic term (Fig. 6e); polar peptides, instead, become more soluble at higher temperature. Peptides that combine polar and apolar domains may therefore be expected to undergo unusual asymmetric reconfigurations in response to temperature change.

Hydrogen-bonding interactions in water are highly susceptible to sequence context because peptide–water hydrogen-bonds compete with intermolecular peptide–peptide hydrogen-bonds<sup>181</sup>. Consequently, hydrogen-bonding-capable or charged side chains that are buried in a hydrophobic environment are less accessible for water and

can interact more strongly with other charged side chains (Fig. 6f). This is because hydrophobic regions have a reduced dielectric constant, which strengthens salt bridges and hydrogen-bonding. For example, arranging K and F in an alternating manner (FKFKFKF) dehydrates intermolecular contact sites and enables intermolecular cation–π interactions that are 3.5 and 30 times stronger than those in sequences composed of K and F blocks (FFFFKKKK and FFKKKFFK, respectively)<sup>182</sup>. The hydrophobic context provided by the sandwiching F residues prevents hydration of the amine side chains of K, rendering them highly available for cation–π interactions to the extent that the peptide can be applied as an underwater glue<sup>182</sup>.

Overall, the balance of interactions in a peptide sequence can be full of opposing tendencies, and it is consequently difficult, even for simple peptidic systems, to predetermine whether certain arrangements will dissolve, liquid demix or aggregate into a gel or solid state. At the same time, the susceptibility of this intricate balance to molecular or environmental changes opens opportunities for stimuli-responsive transitions between ordered and disordered states<sup>141–143</sup>. The importance of environmental conditions on the self-assembly of peptides has been recognized in protein polymorphism<sup>183,184</sup>, such as misfolding, in which a protein assembles into a dysfunctional form<sup>185</sup> rather than its functional native form that has been selected by evolution<sup>186</sup>. The physical basis for the polymorphism is hypothesized to be conformational variance in dihedral angles and competing intermolecular interactions that cause multiple different local minima in an energy landscape<sup>187</sup> – all of which are inherently linked to the existence of diverse non-covalent interaction types. Even though collective contributions to supramolecular interactions deterministically change their strength and range of interactions, the local environment can vary considerably from the bulk conditions (pH, salt and dielectric moment), making the self-assembly of peptides a chaotic process<sup>188</sup>. Therefore, other than statistical analysis of samples, so far it is not possible to predict interactions of

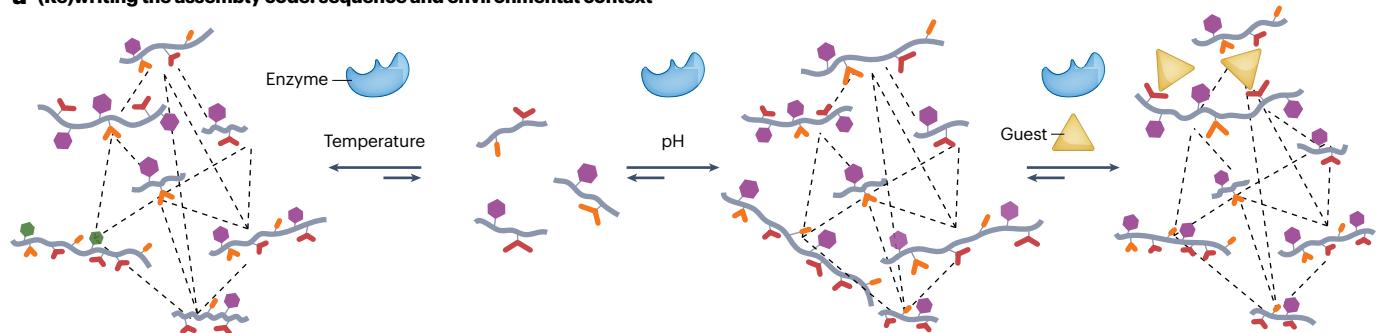
## Box 4 | Systems and emergence

Collective interactions between molecules underpin life's characteristics with living cells as an example par excellence. Systems chemistry attempts to understand and create these emergent properties<sup>254,255</sup>. Dynamic combinatorial peptide libraries are composed of networks of interconverting molecules<sup>256</sup>, and the network information is encoded in the sequence or in the structural arrangement<sup>257</sup>. Thereby, dynamic peptide libraries can demonstrate molecular adaptation<sup>209,210</sup>, mirroring nature's complex adaptive systems<sup>258</sup>. Notably, in such dynamic libraries, the formation and the self-organization of newly formed sequences are interwoven and cannot be separated from each other. In contrast to traditional chemical approaches, the system itself gives rise to properties that emerge in the ensemble. Thus, it is futile to try to capture the properties of a system by describing its components only. We believe that building complex systems from peptides requires a focus on side chain interactions, as the diversity found in intermolecular interactions is the basis for networks with complex and emergent functions<sup>213</sup>.

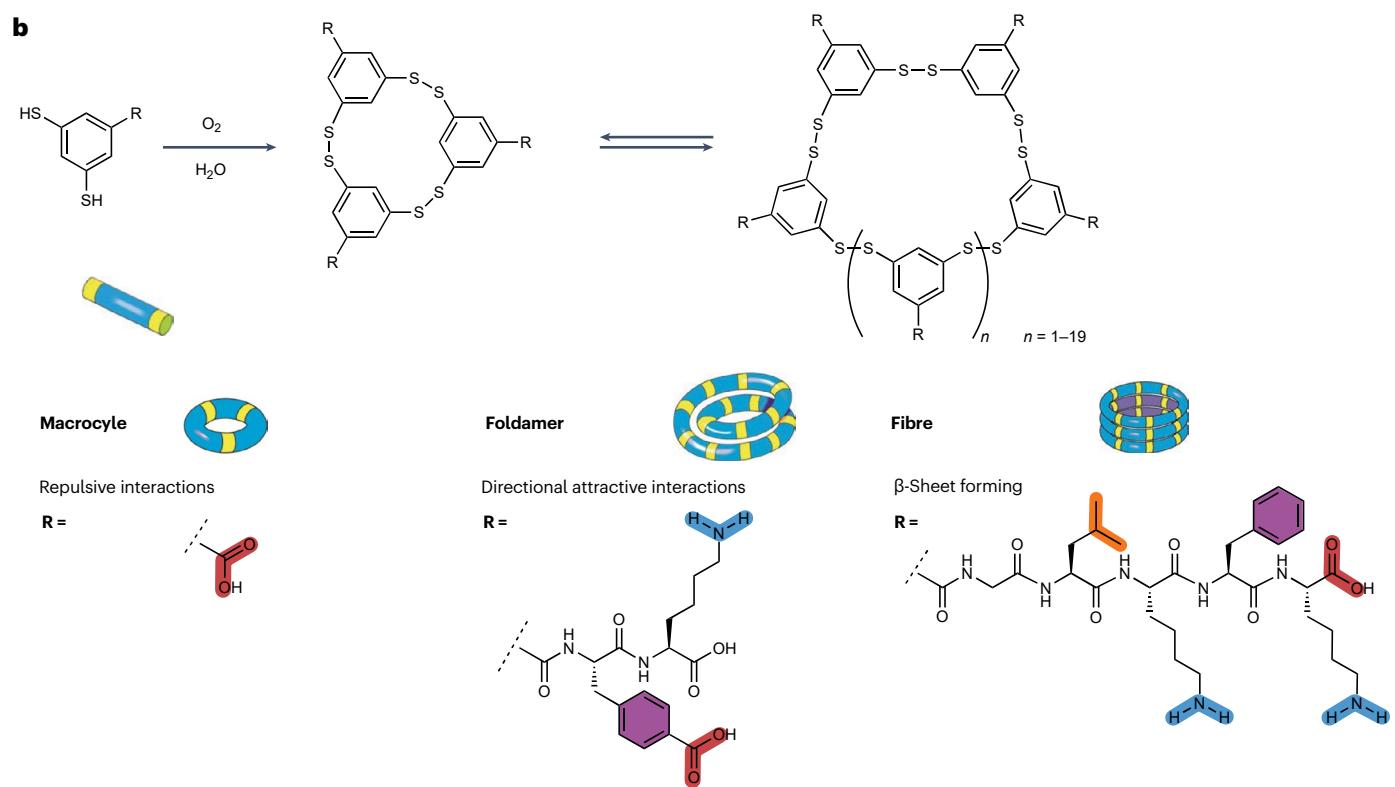
System complexity is a key concept in understanding the emergence of order and disorder in supramolecular assemblies, where intriguing properties arise from components even when their intermolecular interactions follow simple rules<sup>254</sup>. This requires

a holistic approach that embraces the concept of emergence<sup>259</sup>. A complex system is characterized by its sensitivity to initial environmental conditions, the large number of interacting components and multiple pathways that are possible for the system to evolve<sup>254</sup>. The pathway complexity can cause a system to exist in a state between high order (low internal entropy) and high disorder (high internal entropy)<sup>260</sup>. For example, ordered systems, such as crystal structures, are highly deterministic and well defined, where each atom's position can be described with just a few parameters. By contrast, entirely disordered systems are characterized by randomness and are typically described using statistical methods and random networks<sup>260</sup>. Recent research shows that chemical design of peptide materials enables the coexistence of disordered regions within overall ordered structures that give rise to unique functions such as evaporation-induced actuation of nanoporous tripeptides<sup>89</sup>, as well as the promotion of cellular growth<sup>92</sup> and changing phenotype of cells<sup>225</sup>, when used as a biomatrix scaffold. Importantly, context-dependent molecular interactions as part of system complexity cannot be simply reduced to the interactions between individual components<sup>254</sup>, making analysis and prediction inherently challenging<sup>261</sup>.

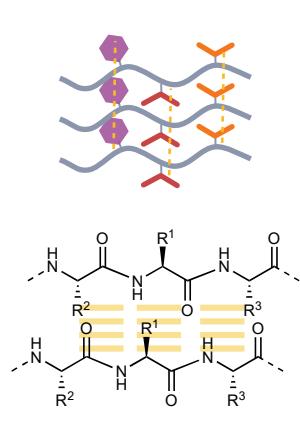
**a (Re)writing the assembly code: sequence and environmental context**



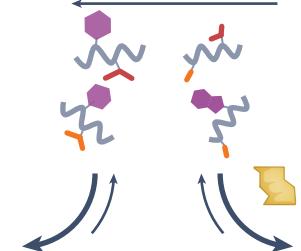
**b**



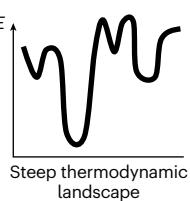
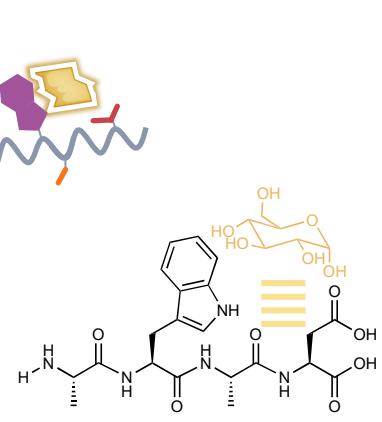
**c** Sequence formation driven by backbone-backbone interaction



Aggregation propensity



Sequence formation driven by cooperative side chain interactions



**Fig. 7 | Sequence evolution via dynamic combinatorial libraries.** **a**, The reaction equilibrium, and thus the production of peptide sequences, is governed by intermolecular interactions that change with environmental conditions, such as the temperature, pH or presence of a guest molecule that binds only to certain peptide sequences of the dynamically exchanging network<sup>195,200</sup>. **b**, Dynamic combinatorial libraries based on aromatic dithiols. Peptide side chains dictate the formation of isolated macrocycles<sup>204</sup>, foldamers or rigid β-sheet-assembled macrocycles<sup>170</sup>. **c**, Sequence evolution of peptides is enabled

by dynamic and reversible peptide bond shuffling, whose equilibrium is shifted if the newly formed soluble peptide is removed from the reaction owing to structure formation. Structure formation can be driven by cooperative side chain interactions that bind to molecules such as glucose, shifting the equilibrium to the kinetic product<sup>209</sup>, whereas backbone interactions favour formation of peptides that assemble into thermodynamically favoured states<sup>170</sup>. Panel **b** adapted from ref. 170, Springer Nature Limited.

single monomers in mixtures of polymorphic peptide assemblies or theoretically calculate the intricacies of free-energy landscapes<sup>187,189</sup>. In rationally designed peptides with less complex sequences, it is to some degree possible to study the impact of select conditions on side chain interactions and self-assembled morphologies<sup>190</sup>. However, the current lack of knowledge and databases for peptide self-assembly does not allow a predictive approach to polymorphism<sup>188</sup>. Future peptide designs can leverage the full potential of sequence context for assembly by considering cooperative effects of multiple types of non-covalent interactions and ensuring low levels of overall backbone order to give rise to instantaneously responsive materials. Assemblies that utilize the entire supramolecular interaction space will likely be challenging to characterize using spectroscopy, and atomistic simulations are expected to keep having a key role in design<sup>20</sup>.

## In situ editing of peptide sequence

Collectively, peptides can navigate intricate energy surfaces through subtle side chain interactions, providing insights for the design of sophisticated supramolecular structures with customizable properties<sup>191</sup> (Box 4). The sometimes-subtle energy differences between peptide sequences discussed in previous sections can be exploited in combinatorial screenings to identify supramolecular interactions<sup>192–194</sup>. A variation of these are dynamic combinatorial libraries in which bonds between peptide backbones or side chains are reversibly formed, enabling selection of the most stable structures from various options depending on the sequence conditions and environment conditions<sup>195–198</sup> (Fig. 7a).

Because side chain interactions influence dynamic sequence evolution, networks of interacting and exchanging peptides can access new and often unexpected distributions in peptide-based systems<sup>199,200</sup>. Short peptides functionalized at the N terminus with aromatic dithiols have been found to reversibly exchange with each other through disulfide exchange, forming cyclic oligomers of variable dimensions<sup>170,201</sup> (Fig. 7b). The distribution of sizes of cyclic oligomers is controlled by intermolecular side chain interactions that can yield β-sheet-stabilized stacked macrocycles<sup>202</sup>. For example, in a mixture of GLKFK-functionalized and GKLKL-functionalized dithiol building blocks, the steric zipper-like interdigititation of the F side chain of the former block between the two L residues of the latter results in the preferential formation of a hexamer sequence comprising alternating, equal amounts of both peptides<sup>203</sup>. The 3D assembly of macrocycles, meanwhile, is controlled by the types of side chain interactions the peptide macrocycles can undergo. Whereas electrostatically repulsive interactions favour the formation of isolated macrocycles<sup>204</sup>, dipeptides with directional attractive sequences produce foldamers and alternating amphiphilic pentapeptides prone to β-sheet interactions form rigid stacks of macrocycles<sup>170</sup>.

Some combinatorial libraries have been created in which the sequence is edited in situ in response to changing conditions<sup>205</sup>. In one

such dynamic peptide library, amide bonds in the peptide backbone are reversibly formed and broken in situ in the presence of an endopeptidase catalyst<sup>206</sup>. The enzyme favours hydrolysis over peptide bond formation, except when it leads to sequences that assemble into nanostructures or are otherwise stabilized, for example, through complexation with a polymer template<sup>205</sup> or long range supramolecular interactions driven by energy<sup>207</sup> or charge transfer<sup>208</sup>. Therefore, the energy landscape can be navigated to discover peptides that form thermodynamically preferred nanostructures through autonomous selection<sup>170</sup> (Fig. 7c). Starting from binary dipeptide (FD and FS) mixtures, longer sequences emerge that eventually favour the dominance of a specific sequence (FDFSFDFS) driven by the self-assembly of directional backbone interactions in π-sheet-like motifs<sup>170</sup>. The self-assembly is governed by the free energy of the resulting sequence, which is mainly driven by contributions from hydrogen-bonding and hydrophobic interactions. Because these interactions depend on the sequence context, they can be tuned by selecting the environmental conditions, such as introducing co-solvents and ions, that alter the distribution of the resulting sequences<sup>170</sup>. Whereas the backbone interactions are important for the formation of the nanostructure, including the shape of the binding site, the side chain interactions determine the site's molecular recognition and binding to a target molecule<sup>209,210</sup>. Weak side chain interactions adapt to target molecules especially if the side chains have a low aggregation propensity that would instead drive peptide self-assembly (Fig. 7c). For example, by deliberately selecting dipeptide systems that are rich in a combination of low-aggregation-prone amino acids such as A, W, D<sup>209</sup> and V, H, K<sup>210</sup>, a shallow energy landscape can be navigated in which side chain hydrogen-bonding and CH–π interactions to glucose and charge–charge interactions to ATP dominate, respectively. The competition of such hydrophilic molecules for water interactions makes them not only challenging ligands but also interesting candidates to discover selective peptide binding sites<sup>211</sup>.

Amino acids grafted to synthetic supramolecular structures can act as handles for enzymatic reactions that add or remove other amino acids, enabling active editing of the supramolecular structure and function<sup>212,213</sup>. Conceptually, such platforms can be used to screen for receptor-like side chain interactions, because transiently coupled amino acids become kinetically trapped upon binding to target molecules and thus can be isolated (Fig. 7c). The amplification of specific peptide sequences resulting from binding to particular substrate molecules might represent a simplistic version of induced fit interactions in receptors and proteins<sup>214</sup>, providing a mechanistic clue to the evolution of strong ligand–receptor binding affinities in nature. Balancing both the backbone and side chain interactions is key to forming adaptive peptide structures, and we think that the system's history of cooperative interactions – that is, how the system achieved isolatable structures – might mirror how structural memory emerges in nature<sup>215</sup>.

## Glossary

### Assembly dynamics

Rate of formation of a supramolecular structure, governed by interaction strength and range of involved intermolecular interaction types.

### Environmental conditions

Factors such as pH, ionic strength, co-solutes, temperature and mechanical forces.

### Exchange dynamics

Rate at which monomers in an assembled structure exchange with their environment, such as a solution.

### Interassembly dynamics

Rate of interactions between assembled structures.

### Internal dynamics

Reversible interactions between peptides that allow for rearrangement within an assembled structure.

### Monomer dynamics

Flexibility of the peptide's backbone to populate different dihedral angles.

### Order and disorder

The regularity and irregularity of monomer arrangement in a self-assembled structure.

### Out-of-equilibrium

A state in which a system is not in a thermodynamic equilibrium.

### Pathway complexity

The variety of kinetic and out-of-equilibrium routes a monomer can take within a free-energy landscape to access different supramolecular morphologies.

### Peptide sequence

The specific order and composition of amino acids in a peptide.

### Polymorphism

The observation of multiple morphologies for the same type of molecule.

### Sequence context

The combined influence of environmental conditions and sequence on the peptide self-assembly process and outcomes.

### Supramolecular dynamics

The reversible changes in the structure of the assembly through movements and exchanges of monomers over time.

for peptide self-assembly. As highlighted in this Review, subtle changes in preparation methods and environmental conditions can substantially impact self-assembly pathways in supramolecular systems<sup>202</sup>. This sensitivity to variability has led to different experimental results even when the reported materials and methods are similar, contributing to the apparent reproducibility crisis in experimental sciences<sup>218</sup>. Automation in sample preparation might be able to control pathway complexity and help establish protocols that ensure reproducible properties in self-assembled peptide materials. To achieve this, future experimental studies should make protocols accessible not only through traditional materials and methods sections but also by incorporating metadata from automated experiments.

In conclusion, reductionist categorization schemes for peptides cannot capture their complex behaviour. The true potential of nature's assembly code requires holistic and system-level consideration of the underlying context-dependent biophysical properties of amino acids to describe sequence–structure–function relationships. Right now, most designed biomolecular materials are still fairly limited in their performance, and they do not yet capture the full spectrum of possibilities found in the living world. We are now, however, at a point in time where the problem-solving powers of AI and the automation of robotics can accelerate our understanding of the cooperative interactions available in this design space. Eventually, biology's assembly code might become a universal molecular assembly code, able to address materials design challenges beyond those that evolution required.

Published online: 13 March 2025

## References

1. Sies, H., Mailloux, R. J. & Jakob, U. Fundamentals of redox regulation in biology. *Nat. Rev. Mol. Cell Biol.* **25**, 701–719 (2024).
2. Kojima, K. & Sudo, Y. Convergent evolution of animal and microbial rhodopsins. *RSC Adv.* **13**, 5367–5381 (2023).
3. Ibusuki, R. et al. Programmable molecular transport achieved by engineering protein motors to move on DNA nanotubes. *Science* **375**, 1159–1164 (2022).
4. Scheuerl, T. et al. Bacterial adaptation is constrained in complex communities. *Nat. Commun.* **11**, 754 (2020).
5. Yang, Y. et al. Peptide programming of supramolecular vinylidene fluoride ferroelectric phases. *Nature* **634**, 833–841 (2024).
6. Finkelstein-Zuta, G. et al. A self-healing multispectral transparent adhesive peptide glass. *Nature* **630**, 368–374 (2024).
7. Guzzo, A. V. The influence of amino acid sequence on protein structure. *Biophys. J.* **5**, 809–822 (1965).
8. Levin, A. et al. Biomimetic peptide self-assembly for functional materials. *Nat. Rev. Chem.* **4**, 615–634 (2020).
9. Korendovych, I. V. & DeGrado, W. F. De novo protein design, a retrospective. *Q. Rev. Biophys.* **53**, e3 (2020).
10. Wilson, C. J. et al. Biomolecular assemblies: moving from observation to predictive design. *Chem. Rev.* **118**, 11519–11574 (2018).
11. Hoffnagle, A. M. & Tezcan, F. A. Atomically accurate design of metalloproteins with predefined coordination geometries. *J. Am. Chem. Soc.* **145**, 14208–14214 (2023).
12. Sinha, N. J., Langenstein, M. G., Pochan, D. J., Kloxin, C. J. & Saven, J. G. Peptide design and self-assembly into targeted nanostructure and functional materials. *Chem. Rev.* **121**, 13915–13935 (2021).
13. Baker, D. & Salis, A. Protein structure prediction and structural genomics. *Science* **294**, 93–96 (2001).
14. Woolfson, D. N. A brief history of de novo protein design: minimal, rational, and computational. *J. Mol. Biol.* **433**, 167160 (2021).
15. Meisenhelter, J. E. et al. Impact of peptide length and solution conditions on tetrameric coiled coil formation. *Biomacromolecules* **25**, 3775–3783 (2024).
16. Kassem, S. & Ulijn, R. V. Designed complex peptide-based adaptive systems: a bottom-up approach. *ChemSystemsChem* **5**, e202200040 (2023).
17. Kouwer, P. H. J. et al. Responsive biomimetic networks from polyisocyanopeptide Hydrogels. *Nature* **493**, 651–655 (2013).
18. de Greef, T. F. A. & Meijer, E. W. Supramolecular polymers. *Nature* **453**, 171–173 (2008).
19. Yanagisawa, Y., Nan, Y., Okuro, K. & Aida, T. Mechanically robust, readily repairable polymers via tailored noncovalent cross-linking. *Science* **359**, 72–76 (2018).
20. Ramakrishnan, M., van Teijlingen, A., Tuttle, T. & Ulijn, R. V. Integrating computation, experiment, and machine learning in the design of peptide-based supramolecular materials and systems. *Angew. Chem.* **135**, e202218067 (2023).

21. Grisoni, F. et al. Designing anticancer peptides by constructive machine learning. *ChemMedChem* **13**, 1300–1302 (2018).
22. Plisson, F., Ramírez-Sánchez, O. & Martínez-Hernández, C. Machine learning-guided discovery and design of non-hemolytic peptides. *Sci. Rep.* **10**, 16581 (2020).
23. Kaygisiz, K. et al. Inverse design of viral infectivity-enhancing peptide fibrils from continuous protein-vector embeddings. *Biomater. Sci.* **11**, 5251–5261 (2023).
24. Batra, R. et al. Machine learning overcomes human bias in the discovery of self-assembling peptides. *Nat. Chem.* **14**, 1427–1435 (2022).
25. Wang, J. et al. Deep learning empowers the discovery of self-assembling peptides with over 10 trillion sequences. *Adv. Sci.* **10**, e2301544 (2023).
26. Xu, T. et al. Accelerating the prediction and discovery of peptide hydrogels with human-in-the-loop. *Nat. Commun.* **14**, 3880 (2023).
27. Eigen, M. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–523 (1971).
28. Tantakitti, F. et al. Energy landscapes and functions of supramolecular systems. *Nat. Mater.* **15**, 469–476 (2016).
29. Jones, M. R., Seeman, N. C. & Mirkin, C. A. Programmable materials and the nature of the DNA bond. *Science* **347**, 1260901 (2015).
30. Rothmund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
31. Djalali, S., Yadav, N. & Delbianco, M. Towards glycan foldamers and programmable assemblies. *Nat. Rev. Mater.* **9**, 190–201 (2024).
32. Wang, G. Z., Chen, L. L. & Zhang, H. Y. Neighboring-site effects of amino acid mutation. *Biochem. Biophys. Res. Commun.* **16**, 531–534 (2007).
33. Martin, E. W. et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).
34. Tejedor, A. R. et al. Protein structural transitions critically transform the network connectivity and viscoelasticity of RNA-binding protein condensates but RNA can prevent it. *Nat. Commun.* **13**, 5717 (2022).
35. Alshareedah, I., Moosa, M. M., Pham, M., Potoyan, D. A. & Banerjee, P. R. Programmable viscoelasticity in protein–RNA condensates with disordered sticker-spacer polypeptides. *Nat. Commun.* **12**, 6620 (2021).
36. White, S. H. & Wimley, W. C. Hydrophobic interactions of peptides with membrane interfaces. *Biochim. Biophys. Acta Rev. Biomembr.* **1376**, 339–352 (1998).
37. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99 (1963).
38. Wimley, W. C. & White, S. H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* **3**, 842–848 (1996).
39. Bohórquez, H. J., Suárez, C. F. & Patarroyo, M. E. Mass & secondary structure propensity of amino acids explain their mutability and evolutionary replacements. *Sci. Rep.* **7**, 7717 (2017).
40. Zhang, S. et al. QTY code enables design of detergent-free chemokine receptors that retain ligand-binding activities. *Proc. Natl. Acad. Sci. USA* **115**, E8652–E8659 (2018).
41. Mahadevi, A. S. & Sastry, G. N. Cation–π interaction: its role and relevance in chemistry, biology, and material science. *Chem. Rev.* **113**, 2100–2138 (2013).
42. Martinez, C. R. & Iverson, B. L. Rethinking the term ‘pi-stacking’. *Chem. Sci.* **3**, 2191 (2012).
43. Bremer, A. et al. Deciphering how naturally occurring sequence features impact the phase behaviours of disordered prion-like domains. *Nat. Chem.* **14**, 196–207 (2022).
44. Iorger, T. R. The context-dependence of amino acid properties. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **5**, 157–166 (1997).
45. Dunker, A. K. et al. Intrinsically disordered protein. *J. Mol. Graph. Model.* **19**, 26–59 (2001).
46. Fujiwara, K., Toda, H. & Ikeguchi, M. Dependence of alpha-helical and beta-sheet amino acid propensities on the overall protein fold type. *BMC Struct. Biol.* **12**, 18 (2012).
47. Williamson, M. P. The structure and function of proline-rich regions in proteins. *Biochem. J.* **297**, 249–260 (1994).
48. Viihni, M., Torkkila, E. & Riikinen, P. Accuracy of protein flexibility predictions. *Proteins Struct. Funct. Bioinforma.* **19**, 141–149 (1994).
49. Andersen, O. S., Greathouse, D. V., Providence, L. L., Becker, M. D. & Koeppe, R. E. Importance of tryptophan dipoles for protein function: 5-fluorination of tryptophans in gramicidin a channels. *J. Am. Chem. Soc.* **120**, 5142–5146 (1998).
50. Reches, M. & Gazit, E. Formation of closed-cage nanostructures by self-assembly of aromatic dipeptides. *Nano Lett.* **4**, 581–585 (2004).
51. Ortony, J. H. et al. Internal dynamics of a supramolecular nanofibre. *Nat. Mater.* **13**, 812–816 (2014).
52. Bellotto, O. et al. Supramolecular hydrogels and water channels of differing diameters from dipeptide isomers. *Biomacromolecules* **25**, 2476–2485 (2024).
53. 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).
54. Görbitz, C. H. Microporous organic materials from hydrophobic dipeptides. *Chem. Eur. J.* **13**, 1022–1031 (2007).
55. Ulijn, R. V. & Lampel, A. Order/disorder in protein and peptide-based biomaterials. *Isr. J. Chem.* <https://doi.org/10.1002/ijch.201900051>.
56. Holehouse, A. S., Ginell, G. M., Griffith, D. & Böke, E. Clustering of aromatic residues in prion-like domains tunes the formation, state, and organization of biomolecular condensates. *Biochemistry* **60**, 3566–3581 (2021).
57. Abbas, M., Lipiński, W. P., Nakashima, K. K., Huck, W. T. S. & Spruijt, E. A short peptide synthon for liquid–liquid phase separation. *Nat. Chem.* **13**, 1046–1054 (2021).
58. Lampel, A. et al. Polymeric peptide pigments with sequence-encoded properties. *Science* **356**, 1064–1068 (2017).
59. Görbitz, C. H. Nanotube formation by hydrophobic dipeptides. *Chem. Eur. J.* **7**, 5153–5159 (2001).
60. Lipiński, W. P. et al. Fibrils emerging from droplets: molecular guiding principles behind phase transitions of a short peptide-based condensate studied by solid-state NMR. *Chemistry* **29**, e202301159 (2023).
61. Hiew, S. H. et al. Modulation of mechanical properties of short bioinspired peptide materials by single amino-acid mutations. *J. Am. Chem. Soc.* **145**, 3382–3393 (2023).
62. Chan, K. H., Xue, B., Robinson, R. C. & Hauser, C. A. E. Systematic moiety variations of ultrashort peptides produce profound effects on self-assembly, nanostructure formation, hydrogelation, and phase transition. *Sci. Rep.* **7**, 12897 (2017).
63. Joseph, J. A. et al. Physics-driven coarse-grained model for biomolecular phase separation with near-quantitative accuracy. *Nat. Comput. Sci.* **1**, 732–743 (2021).
64. Yu, T. & Schatz, G. C. Free energy profile and mechanism of self-assembly of peptide amphiphiles based on a collective assembly coordinate. *J. Phys. Chem. B* **117**, 9004–9013 (2013).
65. Zhang, X. et al. The entropy-controlled strategy in self-assembling systems. *Chem. Soc. Rev.* **52**, 6806–6837 (2023).
66. Tang, C., Smith, A. M., Collins, R. F., Ulijn, R. V. & Saiani, A. Fmoc-diphenylalanine self-assembly mechanism induces apparent pK<sub>a</sub> shifts. *Langmuir* **25**, 9447–9453 (2009).
67. Vargo, E. et al. Using machine learning to predict and understand complex self-assembly behaviors of a multicomponent nanocomposite. *Adv. Mater.* **34**, e202203168 (2022).
68. Abdin, O. & Kim, P. M. Direct conformational sampling from peptide energy landscapes through hypernetwork-conditioned diffusion. *Nat. Mach. Intell.* **6**, 775–786 (2024).
69. Ramos Sasselli, I., Halling, P. J., Ulijn, R. V. & Tuttle, T. Supramolecular fibers in gels can be at thermodynamic equilibrium: a simple packing model reveals preferential fibril formation versus crystallization. *ACS Nano* **10**, 2661–2668 (2016).
70. Hughes, M. et al. Biocatalytic self-assembly of 2D peptide-based nanostructures. *Soft Matter* **7**, 10032 (2011).
71. Yuan, C. et al. Hierarchically oriented organization in supramolecular peptide crystals. *Nat. Rev. Chem.* **3**, 567–588 (2019).
72. Vijayakanth, T. et al. Peptide hydrogen-bonded organic frameworks. *Chem. Soc. Rev.* **53**, 3640–3655 (2024).
73. Jonkheijm, P., van der Schoot, P., Schenning, A. P. H. J. & Meijer, E. W. Probing the solvent-assisted nucleation pathway in chemical self-assembly. *Science* **313**, 80–83 (2006).
74. Reches, M. & Gazit, E. Casting metal nanowires within discrete self-assembled peptide nanotubes. *Science* **300**, 625–627 (2003).
75. Castelletto, V., Hamley, I. W., Harris, P. J. F., Olsson, U. & Spencer, N. Influence of the solvent on the self-assembly of a modified amyloid beta peptide fragment. i. Morphological investigation. *J. Phys. Chem. B* **113**, 9978–9987 (2009).
76. Marchesan, S., Easton, C. D., Kushkaki, F., Waddington, L. & Hartley, P. G. Tripeptide self-assembled hydrogels: unexpected twists of chirality. *Chem. Commun.* **48**, 2195–2197 (2012).
77. Marchesan, S. et al. Unzipping the role of chirality in nanoscale self-assembly of tripeptide hydrogels. *Nanoscale* **4**, 6752 (2012).
78. Smith, A. J., Ali, F. I. & Soldatov, D. V. Glycine homopeptides: the effect of the chain length on the crystal structure and solid state reactivity. *CrystEngComm* **16**, 7196–7208 (2014).
79. Capelli, S. C., Bürgi, H.-B., Dittrich, B., Grabowsky, S. & Jayatilaka, D. Hirshfeld atom refinement. *IUCrJ* **1**, 361–379 (2014).
80. Fletterick, R. J., Tsai, C., Hughes, R. E. & Tsai, C.-C. The crystal and molecular structure of L-alanyl-L-alanine. *J. Phys. Chem.* **75**, 918–992 (1971).
81. Görbitz, C. H. Hydrophobic dipeptides: the final piece in the puzzle. *Acta Crystallogr. Sect. B Struct. Cryst. Eng. Mater.* **74**, 311–318 (2018).
82. Catalano, L. & Naumov, P. Exploiting rotational motion in molecular crystals. *CrystEngComm* **20**, 5872–5883 (2018).
83. Naumov, P., Chizhik, S., Panda, M. K., Nath, N. K. & Boldyreva, E. Mechanically responsive molecular crystals. *Chem. Rev.* **115**, 12440–12490 (2015).
84. Rabone, J. et al. An adaptable peptide-based porous material. *Science* **329**, 1053–1057 (2010).
85. Martí-Gastaldo, C. et al. Side-chain control of porosity closure in single- and multiple-peptide-based porous materials by cooperative folding. *Nat. Chem.* **6**, 343–351 (2014).
86. Martí-Gastaldo, C., Warren, J. E., Stylianou, K. C., Flack, N. L. O. & Rosseinsky, M. J. Enhanced stability in rigid peptide-based porous materials. *Angew. Chem. Int. Ed.* **51**, 11044–11048 (2012).
87. Ulijn, R. V. et al. Water-vapor responsive metallo-peptide nanofibers. *Angew. Chem. Int. Ed.* **63**, e202409391 (2024).
88. Boas, D. et al. A multifunctional drug delivery system based on switchable peptide-stabilized emulsions. *Chem* **10**, 1821–1838 (2024).
89. Piotrowska, R. et al. Mechanistic insights of evaporation-induced actuation in supramolecular crystals. *Nat. Mater.* **20**, 403–409 (2021).
90. Sheehan, F. K. et al. Aromatic zipper topology dictates water-responsive actuation in phenylalanine-based crystals. *Small* **19**, 2207773 (2023).
91. Frederix, P. W. J. M. et al. Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels. *Nat. Chem.* **7**, 30–37 (2015).
92. Álvarez, Z. et al. Bioactive scaffolds with enhanced supramolecular motion promote recovery from spinal cord injury. *Science* **374**, 848–856 (2021).

93. Kaygisiz, K. et al. Peptide amphiphiles as biodegradable adjuvants for efficient retroviral gene delivery. *Adv. Healthc. Mater.* **13**, 2301364 (2024).
94. Bianco, S. et al. Mechanical release of homogenous proteins from supramolecular gels. *Nature* **631**, 544–548 (2024).
95. Kaygisiz, K. et al. Data-mining unveils structure–property–activity correlation of viral infectivity enhancing self-assembling peptides. *Nat. Commun.* **14**, 5121 (2023).
96. Iscen, A., Kaygisiz, K., Synatschke, C. V., Weil, T. & Kremer, K. Multiscale simulations of self-assembling peptides: surface and core hydrophobicity determine fibril stability and amyloid aggregation. *Biomacromolecules* **25**, 3063–3075 (2024).
97. Prabakaran, R. et al. Effect of charged mutation on aggregation of a pentapeptide: insights from molecular dynamics simulations. *Proteins Struct. Funct. Bioinforma.* **90**, 405–417 (2022).
98. MacPherson, D. S. et al. Tuning supramolecular chirality in iodinated amphiphilic peptides through tripeptide linker editing. *Biomacromolecules* **25**, 2277–2285 (2023).
99. Yuan, C. et al. High-entropy non-covalent cyclic peptide glass. *Nat. Nanotechnol.* **19**, 1840–1848 (2024).
100. Xing, R., Yuan, C., Fan, W., Ren, X. & Yan, X. Biomolecular glass with amino acid and peptide nanoarchitectonics. *Sci. Adv.* **9**, eadd8105 (2023).
101. Forman-Kay, J. D. & Mittag, T. From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure* **21**, 1492–1499 (2013).
102. Van Der Lee, R. et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **9**, 6589–6631 (2014).
103. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **1**, 285–298 (2017).
104. Brangwynne, C. P. et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).
105. Farag, M., Borchers, W. M., Bremer, A., Mittag, T. & Pappu, R. V. Phase separation of protein mixtures is driven by the interplay of homotypic and heterotypic interactions. *Nat. Commun.* **14**, 5527 (2023).
106. Romero, P. et al. Sequence complexity of disordered protein. *Proteins Struct. Funct. Genet.* **42**, 38–48 (2001).
107. Theillet, F.-X. et al. The alphabet of intrinsic disorder. *Intrinsically Disord. Proteins* **1**, e24360 (2013).
108. Holehouse, A. S. & Kragelund, B. B. The molecular basis for cellular function of intrinsically disordered protein regions. *Nat. Rev. Mol. Cell Biol.* **25**, 187–211 (2024).
109. Abbas, M., Lipiński, W. P., Wang, J. & Spruijt, E. Peptide-based coacervates as biomimetic protocells. *Chem. Soc. Rev.* **50**, 3690–3705 (2021).
110. Yuan, C., Li, Q., Xing, R., Li, J. & Yan, X. Peptide self-assembly through liquid–liquid phase separation. *Chem* **9**, 2425–2445 (2023).
111. Akahoshi, Y. et al. Phase-separation propensity of non-ionic amino acids in peptide-based complex coacervation systems. *Biomacromolecules* **24**, 704–713 (2023).
112. Rubinstein, M. & Dobrynin, A. Solutions of associative polymers. *Trends Polym. Sci.* **5**, 181–186 (1997).
113. Villegas, J. A. & Levy, E. D. A unified statistical potential reveals that amino acid stickiness governs nonspecific recruitment of client proteins into condensates. *Protein Sci.* **31**, e4361 (2022).
114. Pak, C. W. et al. Sequence determinants of intracellular phase separation by complex coacervation of a disordered protein. *Mol. Cell* **63**, 72–85 (2016).
115. Qamar, S. et al. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation–π interactions. *Cell* **173**, 720–734.e15 (2018).
116. Vernon, R. M. et al. Pi–Pi contacts are an overlooked protein feature relevant to phase separation. *eLife* **7**, e31486 (2018).
117. Chandler, D., Weeks, J. D. & Andersen, H. C. van der Waals picture of liquids, solids, and phase transformations. *Science* **220**, 787–794 (1983).
118. Hong, Y. et al. Hydrophobicity of arginine leads to reentrant liquid–liquid phase separation behaviors of arginine-rich proteins. *Nat. Commun.* **13**, 7326 (2022).
119. Rekhi, S. et al. Expanding the molecular language of protein liquid–liquid phase separation. *Nat. Chem.* **16**, 1113–1124 (2024).
120. Workman, R. J. & Pettitt, B. M. Thermodynamic compensation in peptides following liquid–liquid phase separation. *J. Phys. Chem. B* **125**, 6431–6439 (2021).
121. Poudyal, M. et al. Intermolecular interactions underlie protein/peptide phase separation irrespective of sequence and structure at crowded milieu. *Nat. Commun.* **14**, 6199 (2023).
122. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
123. Tang, Y. et al. Prediction and characterization of liquid–liquid phase separation of minimalistic peptides. *Cell Rep. Phys. Sci.* **2**, 100579 (2021).
124. Frederix, P. W. J. M., Uljin, R. V., Hunt, N. T. & Tuttle, T. Virtual screening for dipeptide aggregation: toward predictive tools for peptide self-assembly. *J. Phys. Chem. Lett.* **2**, 2380–2384 (2011).
125. Hughes, M. et al. Differential supramolecular organisation of fmoc-dipeptides with hydrophilic terminal amino acid residues by biocatalytic self-assembly. *Soft Matter* **8**, 11565 (2012).
126. Kim, J. et al. Role of water in directing diphenylalanine assembly into nanotubes and nanowires. *Adv. Mater.* **22**, 583–587 (2010).
127. Knowles, T. P. et al. Role of intermolecular forces in defining material properties of protein nanofibrils. *Science* **318**, 1900–1903 (2007).
128. Sassielli, I. R. et al. Using experimental and computational energy equilibration to understand hierarchical self-assembly of fmoc-dipeptide amphiphiles. *Soft Matter* **12**, 8307–8315 (2016).
129. Spruijt, E., Westphal, A. H., Borst, J. W., Cohen Stuart, M. A. & van der Gucht, J. Binodal compositions of polyelectrolyte complexes. *Macromolecules* **43**, 6476–6484 (2010).
130. Fuoss, R. M. & Sadek, H. Mutual interaction of polyelectrolytes. *Science* **110**, 552–554 (1949).
131. Azzari, P. & Mezzenga, R. LLPS vs. LLCPS: analogies and differences. *Soft Matter* <https://doi.org/10.1039/D2SM01455F> (2023).
132. Overbeek, J. T. G. & Voorn, M. J. Phase separation in polyelectrolyte solutions. Theory of complex coacervation. *J. Cell. Comp. Physiol.* **49**, 7–26 (1957).
133. Bray, A. J. Theory of phase-ordering kinetics. *Adv. Phys.* **51**, 481–587 (2002).
134. Lytle, T. K., Chang, L. W., Markiewicz, N., Perry, S. L. & Sing, C. E. Designing electrostatic interactions via polyelectrolyte monomer sequence. *ACS Cent. Sci.* **5**, 709–718 (2019).
135. Danielsen, S. P. O., McCarty, J., Shea, J.-E., Delaney, K. T. & Fredrickson, G. H. Molecular design of self-coacervation phenomena in block polyampholytes. *Proc. Natl. Acad. Sci. USA* **116**, 8224–8232 (2019).
136. Naz, M. et al. Self-assembly of stabilized droplets from liquid–liquid phase separation for higher-order structures and functions. *Commun. Chem.* **7**, 79 (2024).
137. Brangwynne, C. P., Tompa, P. & Pappu, R. V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **11**, 899–904 (2015).
138. Dompé, M. et al. Thermoresponsive complex coacervate-based underwater adhesive. *Adv. Mater.* **31**, 1808179 (2019).
139. Capasso Palmiero, U. et al. Programmable zwitterionic droplets as biomolecular sorters and model of membraneless organelles. *Adv. Mater.* **34**, 2104837 (2022).
140. Sathyavageeswaran, A., Bonesso Sabadini, J. & Perry, S. L. Self-assembling polypeptides in complex coacervation. *Acc. Chem. Res.* **57**, 386–398 (2024).
141. Riback, J. A. et al. Stress-triggered phase separation is an adaptive, evolutionarily tuned response. *Cell* **168**, 1028–1040.e19 (2017).
142. Franzmann, T. M. et al. Phase separation of a yeast prion protein promotes cellular fitness. *Science* **359**, eaao5654 (2018).
143. Krainer, G. et al. Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nat. Commun.* **12**, 1085 (2021).
144. Kim, S. et al. Salt triggers the simple coacervation of an underwater adhesive when cations meet aromatic π electrons in seawater. *ACS Nano* **11**, 6764–6772 (2017).
145. Wang, Q. & Schlenoff, J. B. The polyelectrolyte complex/coacervate continuum. *Macromolecules* **47**, 3108–3116 (2014).
146. Gabryelczyk, B. et al. Hydrogen bond guidance and aromatic stacking drive liquid–liquid phase separation of intrinsically disordered histidine-rich peptides. *Nat. Commun.* **10**, 5465 (2019).
147. Hazra, M. K. & Levy, Y. Cross-talk of cation–π interactions with electrostatic and aromatic interactions: a salt-dependent trade-off in biomolecular condensates. *J. Phys. Chem. Lett.* **14**, 8460–8469 (2023).
148. Li, L. et al. Phase behavior and salt partitioning in polyelectrolyte complex coacervates. *Macromolecules* **51**, 2988–2995 (2018).
149. Wu, X., Sun, Y., Yu, J. & Miserez, A. Tuning the viscoelastic properties of peptide coacervates by single amino acid mutations and salt kosmotropicity. *Commun. Chem.* **7**, 5 (2024).
150. Fisher, R. S. & Elbaum-Garfinkle, S. Tunable multiphase dynamics of arginine and lysine liquid condensates. *Nat. Commun.* **11**, 4628 (2020).
151. Das, R. K., Ruff, K. M. & Pappu, R. V. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **32**, 102–112 (2015).
152. Quiroz, F. G. & Chilkoti, A. Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. *Nat. Mater.* **14**, 1164–1171 (2015).
153. Maeda, I. et al. Comparison between coacervation property and secondary structure of synthetic peptides, Ile-containing elastin-derived pentapeptide repeats. *Protein Pept. Lett.* **20**, 905–910 (2013).
154. Maeda, I. et al. Design of phenylalanine-containing elastin-derived peptides exhibiting highly potent self-assembling capability. *Protein Pept. Lett.* **22**, 934–939 (2015).
155. Taniguchi, S., Watanabe, N., Nose, T. & Maeda, I. Development of short and highly potent self-assembling elastin-derived pentapeptide repeats containing aromatic amino acid residues. *J. Pept. Sci.* **22**, 36–42 (2016).
156. Dzuricky, M., Rogers, B. A., Shahid, A., Cremer, P. S. & Chilkoti, A. De novo engineering of intracellular condensates using artificial disordered proteins. *Nat. Chem.* **12**, 814–825 (2020).
157. Baruch Leshem, A. et al. Biomolecular condensates formed by designer minimalistic peptides. *Nat. Commun.* **14**, 421 (2023).
158. Fossat, M. J., Zeng, X. & Pappu, R. V. Uncovering differences in hydration free energies and structures for model compound mimics of charged side chains of amino acids. *J. Phys. Chem. B* **125**, 4148–4161 (2021).
159. Mason, P. E. et al. The structure of aqueous guanidinium chloride solutions. *J. Am. Chem. Soc.* **126**, 11462–11470 (2004).
160. Sementa, D. et al. Sequence-tunable phase behavior and intrinsic fluorescence in dynamically interacting peptides. *Angew. Chem. Int. Ed.* **62**, e202311479 (2023).
161. Hirsch, A. K. H., Buhler, E. & Lehn, J. M. Biodynamers: self-organization-driven formation of doubly dynamic proteins. *J. Am. Chem. Soc.* **134**, 4177–4183 (2012).
162. Jawor-Baczynska, A., Sefcik, J. & Moore, B. D. 250 nm glycine-rich nanodroplets are formed on dissolution of glycine crystals but are too small to provide productive nucleation sites. *Cryst. Growth Des.* **13**, 470–478 (2013).

163. Dave, D. R. et al. Adaptive and space-filling peptide self-assembly upon drying. Preprint at <https://doi.org/10.26434/chemrxiv-2024-cmwjx> (2024).
164. Bera, S. et al. Solid-state packing dictates the unexpected solubility of aromatic peptides. *Cell Rep. Phys. Sci.* **2**, 100391 (2021).
165. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
166. Baldwin, A. J. et al. Metastability of native proteins and the phenomenon of amyloid formation. *J. Am. Chem. Soc.* **133**, 14160–14163 (2011).
167. Noji, M. et al. Breakdown of supersaturation barrier links protein folding to amyloid formation. *Commun. Biol.* **4**, 120 (2021).
168. Portugal Barron, D. & Guo, Z. The supersaturation perspective on the amyloid hypothesis. *Chem. Sci.* **15**, 46–54 (2024).
169. Louros, N., Schymkowitz, J. & Rousseau, F. Mechanisms and pathology of protein misfolding and aggregation. *Nat. Rev. Mol. Cell Biol.* **24**, 912–933 (2023).
170. Pappas, C. G. et al. Emergence of low-symmetry foldamers from single monomers. *Nat. Chem.* **12**, 1180–1186 (2020).
171. Duan, C. & Wang, R. A unified description of salt effects on the liquid–liquid phase separation of proteins. *ACS Cent. Sci.* **10**, 460–468 (2024).
172. Ji, W. et al. Metal-ion modulated structural transformation of amyloid-like dipeptide supramolecular self-assembly. *ACS Nano* **13**, 7300–7309 (2019).
173. Abul-Haija, Y. M., Scott, G. G., Sahoo, J. K., Tuttle, T. & Ulijn, R. V. Cooperative, ion-sensitive co-assembly of tripeptide hydrogels. *Chem. Commun.* **53**, 9562–9565 (2017).
174. Dawson, W. M. et al. Structural resolution of switchable states of a de novo peptide assembly. *Nat. Commun.* **12**, 1530 (2021).
175. Rehm, T. H. & Schmuck, C. Ion-pair induced self-assembly in aqueous solvents. *Chem. Soc. Rev.* **39**, 3597 (2010).
176. von Gröning, M., de Feijter, I., Stuart, M. C. A., Voets, I. K. & Besenius, P. Tuning the aqueous self-assembly of multistimuli-responsive polyanionic peptide nanorods. *J. Mater. Chem. B* **1**, 2008 (2013).
177. Roy, S. et al. Dramatic specific-ion effect in supramolecular hydrogels. *Chem. Eur. J.* **18**, 11723–11731 (2012).
178. Li, Z., Zhu, Y. & Matson, J. B. pH-responsive self-assembling peptide-based biomaterials: designs and applications. *ACS Appl. Bio Mater.* **5**, 4635–4651 (2022).
179. Shen, Z. et al. Biomembrane induced in situ self-assembly of peptide with enhanced antimicrobial activity. *Biomater Sci.* **8**, 2031–2039 (2020).
180. Tan, Y. et al. Infiltration of chitin by protein coacervates defines the squid beak mechanical gradient. *Nat. Chem. Biol.* **11**, 488–495 (2015).
181. Krieg, E., Bastings, M. M. C., Besenius, P. & Rybtchinski, B. Supramolecular polymers in aqueous media. *Chem. Rev.* **116**, 2414–2477 (2016).
182. Chang, H. et al. Short-sequence superadhesive peptides with topologically enhanced cation–π interactions. *Chem. Mater.* **33**, 5168–5176 (2021).
183. Paravastu, A. K., Leapman, R. D., Yau, W.-M. & Tycko, R. Molecular structural basis for polymorphism in Alzheimer’s β-amyloid fibrils. *Proc. Natl Acad. Sci. USA* **105**, 18349–18354 (2008).
184. Ostermeier, L., de Oliveira, G. A. P., Dzwolak, W., Silva, J. L. & Winter, R. Exploring the polymorphism, conformational dynamics and function of amyloidogenic peptides and proteins by temperature and pressure modulation. *Biophys. Chem.* **268**, 106506 (2021).
185. Fändrich, M., Fletcher, M. A. & Dobson, C. M. Amyloid fibrils from muscle myoglobin. *Nature* **410**, 165–166 (2001).
186. Antfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
187. Close, W. et al. Physical basis of amyloid fibril polymorphism. *Nat. Commun.* **9**, 699 (2018).
188. Wang, F. et al. Deterministic chaos in the self-assembly of β sheet nanotubes from an amphipathic oligopeptide. *Matter* **4**, 3217–3231 (2021).
189. Rüter, A. et al. Tube to ribbon transition in a self-assembling model peptide system. *Phys. Chem. Chem. Phys.* **22**, 18320–18327 (2020).
190. Merg, A. D. et al. Shape-shifting peptide nanomaterials: surface asymmetry enables pH-dependent formation and interconversion of collagen tubes and sheets. *J. Am. Chem. Soc.* **142**, 19956–19968 (2020).
191. Mattia, E. & Otto, S. Supramolecular systems chemistry. *Nat. Nanotechnol.* **1**, 111–119 (2015).
192. Swann, P. A. et al. Nonspecific protease-catalyzed hydrolysis/synthesis of a mixture of peptides: product diversity and ligand amplification by a molecular trap. *Biopolymers* **40**, 617–625 (1996).
193. Schmuck, C., Heil, M., Scheiber, J. & Baumann, K. Charge interactions do the job: a combined statistical and combinatorial approach to finding artificial receptors for binding tetrapeptides in water. *Angew. Chem. Int. Ed.* **44**, 7208–7212 (2005).
194. Schmuck, C. & Geiger, L. Efficient complexation of N-acetyl amino acid carboxylates in water by an artificial receptor: unexpected cooperativity in the binding of glutamate but not aspartate. *J. Am. Chem. Soc.* **127**, 10486–10487 (2005).
195. Lam, R. T. S. et al. Amplification of acetylcholine-binding catenanes from dynamic combinatorial libraries. *Science* **308**, 667–669 (2005).
196. Chen, C. et al. Design of multi-phase dynamic chemical networks. *Nat. Chem.* **9**, 799–804 (2017).
197. Rodrigues, A., Rocard, L. & Mourné, R. Peptide and peptidomimetic assemblies in dynamic combinatorial chemistry, *ChemSystemsChem* <https://doi.org/10.1002/syst.202300011> (2023).
198. Liu, K. et al. Light-driven eco-evolutionary dynamics in a synthetic replicator system. *Nat. Chem.* **16**, 79–88 (2024).
199. Li, J., Nowak, P. & Otto, S. Dynamic combinatorial libraries: from exploring molecular recognition to systems chemistry. *J. Am. Chem. Soc.* **135**, 9222–9239 (2013).
200. Martin, C. et al. Water-based dynamic depsipeptide chemistry: building block recycling and oligomer distribution control using hydration–dehydration cycles. *JACS Au* **2**, 1395–1404 (2022).
201. Liu, B. et al. Complex molecules that fold like proteins can emerge spontaneously. *J. Am. Chem. Soc.* **141**, 1685–1689 (2019).
202. Carnall, J. M. A. et al. Mechanosensitive self-replication driven by self-organization. *Science* **327**, 1502–1506 (2010).
203. Eleveld, M. J. et al. Departure from randomness: evolution of self-replicators that can self-sort through steric zipper formation. *Chem* <https://doi.org/10.1016/j.chempr.2024.11.012> (2024).
204. Sadownik, J. W., Mattia, E., Nowak, P. & Otto, S. Diversification of self-replicating molecules. *Nat. Chem.* **8**, 264–269 (2016).
205. Abul-Haija, Y. M. & Ulijn, R. V. Sequence adaptive peptide-polysaccharide nanostructures by biocatalytic self-assembly. *Biomacromolecules* **16**, 3473–3479 (2015).
206. Williams, R. J. et al. Enzyme-assisted self-assembly under thermodynamic control. *Nat. Nanotechnol.* **4**, 19–24 (2009).
207. Nalluri, S. K. M., Berdugo, C., Javid, N., Frederix, P. W. J. M. & Ulijn, R. V. Biocatalytic self-assembly of supramolecular charge-transfer nanostructures based on N-type semiconductor-appended peptides. *Angew. Chem. Int. Ed.* **53**, 5882–5887 (2014).
208. Berdugo, C. et al. Dynamic peptide library for the discovery of charge transfer hydrogels. *ACS Appl. Mater. Interfaces* **7**, 25946–25954 (2015).
209. Kassem, S., McPhee, S. A., Berisha, N. & Ulijn, R. V. Emergence of cooperative glucose-binding networks in adaptive peptide systems. *J. Am. Chem. Soc.* **145**, 9800–9807 (2023).
210. Jain, A. et al. Tractable molecular adaptation patterns in a designed complex peptide system. *Chem* **8**, 1894–1905 (2022).
211. Davis, A. M. & Teague, S. J. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem. Int. Ed.* **38**, 736–749 (1999).
212. Zagiel, B. et al. Dynamis amino acid side-chains grafting on folded peptide backbone. *Chem. Eur. J.* **28**, e202200454 (2022).
213. Kumar, M. et al. Amino-acid-encoded biocatalytic self-assembly enables the formation of transient conducting nanostructures. *Nat. Chem.* **10**, 696–703 (2018).
214. Lehn, J. M. Constitutional dynamic chemistry: bridge from supramolecular chemistry to adaptive chemistry. *Top. Curr. Chem.* **322**, 1–32 (2012).
215. Kaygisiz, K. & Ulijn, R. V. Can molecular systems learn? *ChemSystemsChem* <https://doi.org/10.1002/syst.202400075> (2025).
216. Rha, A. K. et al. Electrostatic complementarity drives amyloid/nucleic acid co-assembly. *Angew. Chem. Int. Ed.* **59**, 358–363 (2020).
217. Nijrijk, M. et al. Reshaping the discovery of self-assembling peptides with generative AI guided by hybrid deep learning. *Nat. Mach. Intell.* **6**, 1487–1500 (2024).
218. Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452–454 (2016).
219. Lampel, A., Ulijn, R. V. & Tuttle, T. Guiding principles for peptide nanotechnology through directed discovery. *Chem. Soc. Rev.* **47**, 3737–3758 (2018).
220. van Teijlingen, A., Smith, M. C. & Tuttle, T. Short peptide self-assembly in the martini coarse-grain force field family. *Acc. Chem. Res.* **56**, 644–654 (2023).
221. Syamala, P. P. N., Soberats, B., Görl, D., Gekle, S. & Würthner, F. Thermodynamic insights into the entropically driven self-assembly of amphiphilic dyes in water. *Chem. Sci.* **10**, 9358–9366 (2019).
222. Schoenmakers, S. M. C. et al. Structure and dynamics of supramolecular polymers: wait and see. *ACS Macro Lett.* **11**, 711–715 (2022).
223. Albertazzi, L. et al. Probing exchange pathways in one-dimensional aggregates with super-resolution microscopy. *Science* **344**, 491–495 (2014).
224. Ni, Q. Z. et al. Peptide and protein dynamics and low-temperature/DNP magic angle spinning NMR. *J. Phys. Chem. B* **121**, 4997–5006 (2017).
225. Freeman, R. et al. Reversible self-assembly of superstructured networks. *Science* **362**, 808–813 (2018).
226. Zaslavsky, B. Y. & Uversky, V. N. In aqua veritas: the indispensable yet mostly ignored role of water in phase separation and membrane-less organelles. *Biochemistry* **57**, 2437–2451 (2018).
227. Rani, P. & Biswas, P. Diffusion of hydration water around intrinsically disordered proteins. *J. Phys. Chem. B* **119**, 13262–13270 (2015).
228. Mukherjee, S. & Schäfer, L. V. Thermodynamic forces from protein and water govern condensate formation of an intrinsically disordered protein domain. *Nat. Commun.* **14**, 5892 (2023).
229. Arya, S. & Mukhopadhyay, S. Ordered water within the collapsed globules of an amyloidogenic intrinsically disordered protein. *J. Phys. Chem. B* **118**, 9191–9198 (2014).
230. Ribeiro, S. S., Samanta, N., Ebbinghaus, S. & Marcos, J. C. The synergic effect of water and biomolecules in intracellular phase separation. *Nat. Rev. Chem.* <https://doi.org/10.1038/s41570-019-0120-4> (2019).
231. Chang, R., Yuan, C., Zhou, P., Xing, R. & Yan, X. Peptide self-assembly: from ordered to disordered. *Acc. Chem. Res.* **57**, 289–301 (2024).
232. da Silva, R. M. P. et al. Super-resolution microscopy reveals structural diversity in molecular exchange among peptide amphiphile nanofibres. *Nat. Commun.* **7**, 11561 (2016).
233. Pavlović, R. Z., Egner, S. A., Palmer, L. C. & Stupp, S. I. Supramolecular polymers: dynamic assemblies of ‘dancing’ monomers. *J. Polym. Sci.* **61**, 870–880 (2023).

234. Piskorz, T. K., Perez-Chirinos, L., Qiao, B. & Sasselli, I. R. Tips and tricks in the modeling of supramolecular peptide assemblies. *ACS Omega* **9**, 31254–31273 (2024).
235. Bochicchio, D., Salvaglio, M. & Pavan, G. M. Into the dynamics of a supramolecular polymer at submolecular resolution. *Nat. Commun.* **8**, 147 (2017).
236. Rodon-Fores, J. et al. A chemically fueled supramolecular glue for self-healing gels. *Chem. Sci.* **13**, 11411–11421 (2022).
237. Mendes, A. C., Baran, E. T., Reis, R. L. & Azevedo, H. S. Self-assembly in nature: using the principles of nature to create complex nanobiomaterials. *WIREs Nanomed. Nanobiotechnol.* **5**, 582–612 (2013).
238. Yan, X., Li, J. & Möhwald, H. Self-assembly of hexagonal peptide microtubes and their optical waveguiding. *Adv. Mater.* **23**, 2796–2801 (2011).
239. Packwood, D. M., Han, P. & Hitosugi, T. Chemical and entropic control on the molecular self-assembly process. *Nat. Commun.* **8**, 14463 (2017).
240. Korevaar, P. A. et al. Pathway complexity in supramolecular polymerization. *Nature* **481**, 492–496 (2012).
241. Roth, P. et al. Supramolecular assembly guided by photolytic redox cycling. *Nat. Synth.* <https://doi.org/10.1038/s44160-023-00343-1> (2023).
242. Matern, J., Dorca, Y., Sánchez, L. & Fernández, G. Revising complex supramolecular polymerization under kinetic and thermodynamic control. *Angew. Chem. Int. Ed.* **58**, 16730–16740 (2019).
243. Debnath, S. et al. Tunable supramolecular gel properties by varying thermal history. *Chem. Eur. J.* **25**, 7881–7887 (2019).
244. Knowles, T. P. J. et al. An analytical solution to the kinetics of breakable filament assembly. *Science* **326**, 1533–1537 (2009).
245. Hirst, A. R. et al. Biocatalytic induction of supramolecular order. *Nat. Chem.* **2**, 1089–1094 (2010).
246. Boekhoven, J., Hendriksen, W. E., Koper, G. J. M., Eelkema, R. & Van Esch, J. H. Transient assembly of active materials fueled by a chemical reaction. *Science* **349**, 1075–1079 (2015).
247. Boekhoven, J. et al. Dissipative self-assembly of a molecular gelator by using a chemical fuel. *Angew. Chem. Int. Ed.* **49**, 4825–4828 (2010).
248. Sorrenti, A., Leira-Iglesias, J., Sato, A. & Hermans, T. M. Non-equilibrium steady states in supramolecular polymerization. *Nat. Commun.* **8**, 15899 (2017).
249. Debnath, S., Roy, S. & Ulijn, R. V. Peptide nanofibers with dynamic instability through nonequilibrium biocatalytic assembly. *J. Am. Chem. Soc.* **135**, 16789–16792 (2013).
250. Li, X. et al. Solvent modulated structural transition of self-assemblies formed by bola-form hexapeptide amphiphiles. *J. Mol. Liq.* **355**, 118940 (2022).
251. Levine, M. S. et al. Formation of peptide-based oligomers in dimethylsulfoxide: identifying the precursor of fibril formation. *Soft Matter* **16**, 7860–7868 (2020).
252. Patel, A. et al. ATP as a biological hydrotrope. *Science* **356**, 753–756 (2017).
253. Mehringer, J. et al. Hofmeister versus Neuberg: is ATP really a biological hydrotrope? *Cell Rep. Phys. Sci.* **2**, 100343 (2021).
254. Whitesides, G. M. & Ismagilov, R. F. Complexity in chemistry. *Science* **284**, 89–92 (1999).
255. Ashkenasy, G., Hermans, T. M., Otto, S. & Taylor, A. F. Systems chemistry. *Chem. Soc. Rev.* <https://doi.org/10.1039/c7cs00117g> (2017).
256. Lehn, J.-M. Dynamic combinatorial chemistry and virtual combinatorial libraries. *Chem. Eur. J.* **5**, 2455–2463 (1999).
257. Herrmann, A. Dynamic combinatorial/covalent chemistry: a tool to read, generate and modulate the bioactivity of compounds and compound mixtures. *Chem. Soc. Rev.* <https://doi.org/10.1039/c3cs60336a> (2014).
258. Otto, S. An approach to the *de novo* synthesis of life. *Acc. Chem. Res.* **55**, 145–155 (2022).
259. Lehn, J.-M. Towards complex matter: supramolecular chemistry and self-organization. *Eur. Rev.* **17**, 263–280 (2009).
260. Mao, X. & Kotov, N. Complexity, disorder, and functionality of nanoscale materials. *MRS Bull.* **49**, 352–364 (2024).
261. Davis, A. V., Yeh, R. M. & Raymond, K. N. Supramolecular assembly dynamics. *Proc. Natl Acad. Sci. USA* **99**, 4793–4796 (2002).

## Author contributions

K.K., D.S., V.A. and R.V.U. did the literature research and wrote the original manuscript. K.K. and R.V.U. conceptualized the manuscript. X.C. and R.V.U. revised, supervised and provided the funding resources. The final manuscript was approved by all authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Peer review information** *Nature Reviews Materials* thanks Tao Jiang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2025