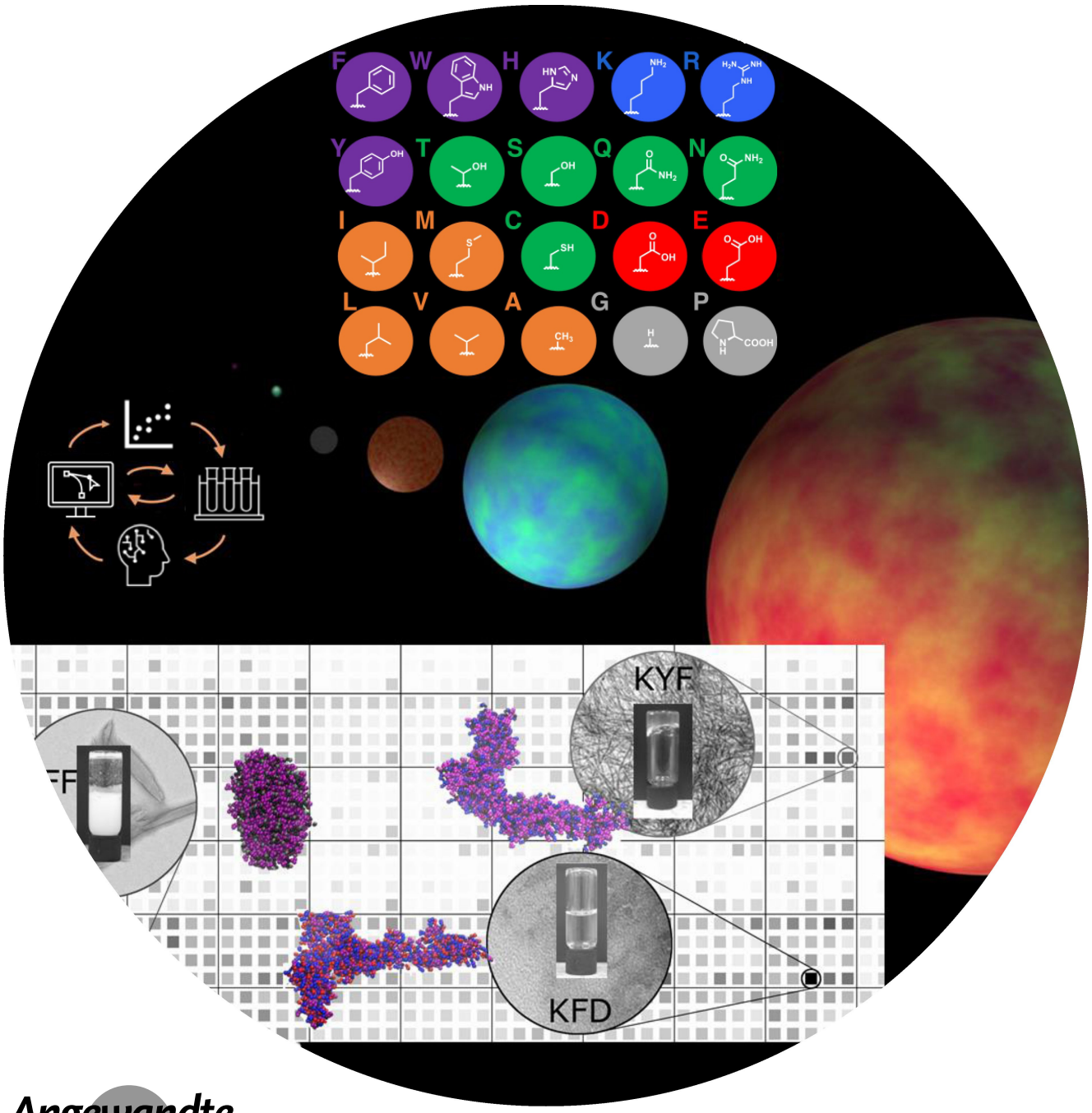
Zitierweise: *Angew. Chem. Int. Ed.* **2023,** *62,* e202218067
Internationale Ausgabe: doi.org/10.1002/anie.202218067
Deutsche Ausgabe: doi.org/10.1002/ange.202218067

Peptide-Based Materials

Integrating Computation, Experiment, and Machine Learning in the Design of Peptide-Based Supramolecular Materials and Systems

Maithreyi Ramakrishnan[+], Alexander van Teijlingen[+], Tell Tuttle, and R. V. Ulijn**



Abstract: Interest in peptide-based supramolecular materials has grown extensively since the 1980s and the application of computational methods has paralleled this. These methods contribute to the understanding of experimental observations based on interactions and inform the design of new supramolecular systems. They are also used to virtually screen and navigate these very large design spaces. Increasingly, the use of artificial intelligence is employed to screen far more candidates than traditional methods. Based on a brief history of computational and experimentally integrated investigations of peptide structures, we explore recent impactful examples of computationally driven investigation into peptide self-assembly, focusing on recent advances in methodology development. It is clear that the integration between experiment and computation to understand and design new systems is becoming near seamless in this growing field.

1. Introduction

There is a significant and growing interest in repurposing the building blocks of life to create supramolecular materials. Peptides are well-suited for this purpose, as they are composed of 20 natural amino acid building blocks (Figure 1) that cover a wide range of non-covalent interactions and chemical functionalities.[1–11] The limitless possibilities of designing sophisticated functions based on amino acid building blocks are evident from the protein structures found in the biological world. It is, however, still challenging to design peptide materials and architectures from first principles due to the inherent complexity, even when considering short peptide sequences.

There are three general approaches to peptide-based materials design. These are primarily bio-inspired or biomimetic approaches, where known biological structures or sequences are copied, modified, optimised or simplified.[12,13] Alternatively, combinatorial screening approaches are used to identify (ensembles of) peptides with desired properties through side-by-side comparison and optimisation. In these systems, there is no inherent bias from known biological designs.[14–18] Finally, based on a fundamental understanding of the combinatorial interactions of amino acid side chains, rational sequence design based purely on physical chemistry

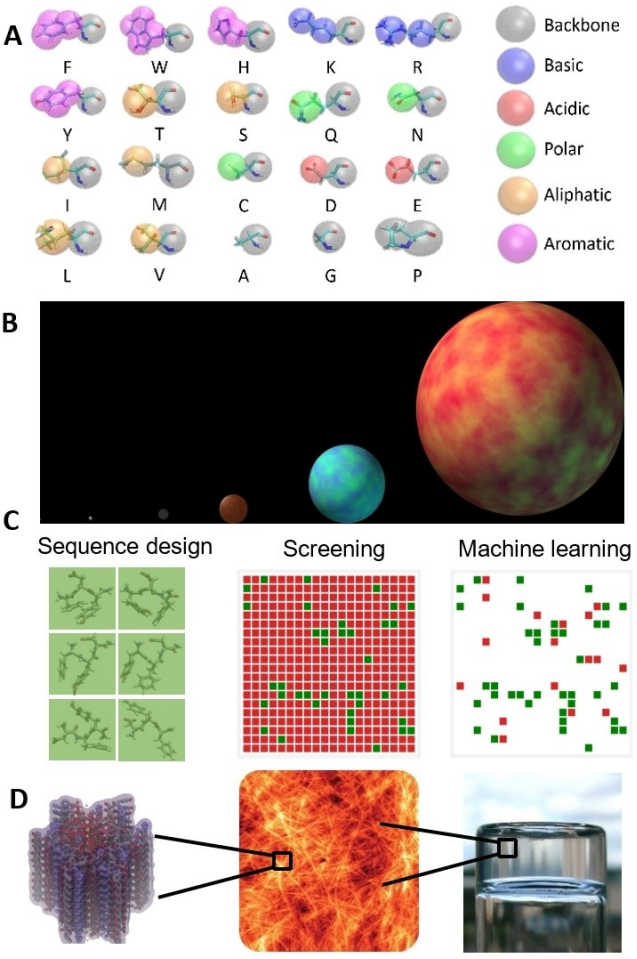


Figure 1. Integrating computation and experiment in design of peptide-based supramolecular materials and systems. (A) Atomistic gene-encoded amino acid alphabet with MARTINI coarse-grained bead overlaid on the atoms they represent.[21] (B) Representation of peptide sequence complexity with planetary volume representing sequence space spaces ranging from 400 for dipeptides to 1 280 000 000 for heptapeptides, varying as 20^n with n being the peptide chain length. (C) Computational methods such as molecular dynamics, virtual screening and machine learning can expand the accessible sequence space. (D) Schematic showing how computational design and experimental observations can be combined to design self-assembling systems, including supramolecular hydrogels.

[*] M. Ramakrishnan,[+] R. V. Ulijn
Advanced Science Research Center (ASRC) at the Graduate Center, City University of New York (CUNY)
New York, NY 10031 (USA)
and
Department of Chemistry, Hunter College, The City University of New York
New York, NY 10065 (USA)
E-mail: rulijn@gc.cuny.edu

M. Ramakrishnan[+]
Ph.D. Program in Chemistry, The Graduate Center of the City University of New York
New York, NY 10016 (USA)

A. van Teijlingen,[+] T. Tuttle
Pure and Applied Chemistry, University of Strathclyde
295 Cathedral Street, Glasgow, G1 1XL (UK)
E-mail: tell.tuttle@strath.ac.uk

R. V. Ulijn
Ph.D. Program in Chemistry and PhD program in Biochemistry, The Graduate Center of the City University of New York
New York, NY 10016 (USA)

[+] These authors contributed equally to this work.

insights into interactions becomes possible. Integrating computational and experimental approaches offers significant complementary advantages in these areas.

The current feature article focuses on recent examples of computationally-driven and integrated computational/experimental approaches to probe peptide design rules and their resulting supramolecular structures. Applications of computational chemistry methods are reviewed, including emerging methods such as machine learning. We also highlight examples that show integrated computational methods with experimental approaches which have complementary advantages in understanding bio-inspired systems, combinatorial screening, and rational design. Throughout the review, we consider a spectrum of complexity, from the self-assembly of single peptides to complex systems involving multiple interacting peptides.[1] In such a fast-moving field, it is only possible to do justice to some of the excellent research papers published in recent years, and our aim is not to provide a comprehensive overview. Instead, we focus on a smaller number of key examples that are well-suited to illustrate complementary benefits from using both experiment and computation in supramolecular peptide materials and systems design.

2. Brief historical overview: integration of experiment and computation in peptide assembly

Regarded as a pioneering example in the reductionist design of supramolecular peptide assemblies, in 1988, DeGrado et al. designed, informed by the structure and sequence of alamethicin and melittin, bilayer membrane pore-forming peptides, a 21-mer peptide composed of only L and S residues ($(LSSLLSL)_3$). This design was based on chemical intuition and understanding of protein structures and rationalised as having a water-facing (S) hydrophilic and bilayer-facing (L) hydrophobic surface held in positions by an amphipathic α-helical secondary structure. This intuition was confirmed by circular dichroism (CD) and minimisation of the helices in the AMBER forcefield, which at the time required using a Cray 1-A supercomputer (Figure 2a). It was found via the Muller-Montal planar bilayer method that this designed peptide sequence did form ionic conduction membrane pores and, when inserted into a membrane, had a conductivity similar to that of an acetyl-choline receptor via planar bilayer ion conductance experiments.[19] Further investigating α-helical design rules, Woolfson et al. inserted proline residues into known α-helical sequences,[20] and they found via molecular modelling in an implicit solvent that in every instance, the proline residue produced a helical kink in the α-helical structure, at a significant hydrophobic penalty due to the proline sidechain becoming orientated towards the solvent. This demonstrated the sensitivity of peptide secondary structure on individual residues.[20] These early studies linked specific amino acids to sequence structure, which provided a basis for designing supramolecular properties from amino acid sequences.

In 1993 Zhang et al. designed the hexadecapeptide EAK16 ($(AEAEAKAK)_2$) based on the design rule that alternating hydrophobic-hydrophilic residue patterns tend to form β-sheet structures,[22] which was confirmed by CD. They reasoned that the alternating sequence and secondary structure amplified the charge complementarity effect by directing positive and negative residues towards each other (Figure 2b). These attractive and regular interactions pro-

Alexander van Teijlingen received his M.Sc. in functional nanomaterials from the University of Bristol (UK). In 2019 he joined the group of Prof. Tell Tuttle at the University of Strathclyde (UK), where he undertook doctoral studies, followed by a postdoctoral appointment to the same group in 2022. Currently, he is working on machine learning-driven discovery of bilayer pore-forming peptides and their design rules in the Tuttle research group.

Maithreyi Ramakrishnan obtained an M.Sc. from the Centre for Excellence in Basic Sciences (India). In 2019 she became a Ph.D. student in Prof. Rein V. Ulijn's research group at the Advanced Science Research Center (ASRC), City University of New York (USA). Her research interests include sequestering metabolites using peptide self-assembly.

Rein V. Ulijn is a professor of nanoscale systems chemistry at Hunter College and founding director of the Advanced Science Research Center (ASRC) at City University New York (USA). His research interests focus on biomolecular building blocks and how designed systems can combine to produce complex and adaptive functional materials.

Tell Tuttle is professor of theoretical chemistry at the University of Strathclyde (UK) and head of department of pure and applied chemistry. His research focuses on computationally reducing molecular search space to better understand existing systems, rationalise their design and predict new systems with enhanced or new properties.

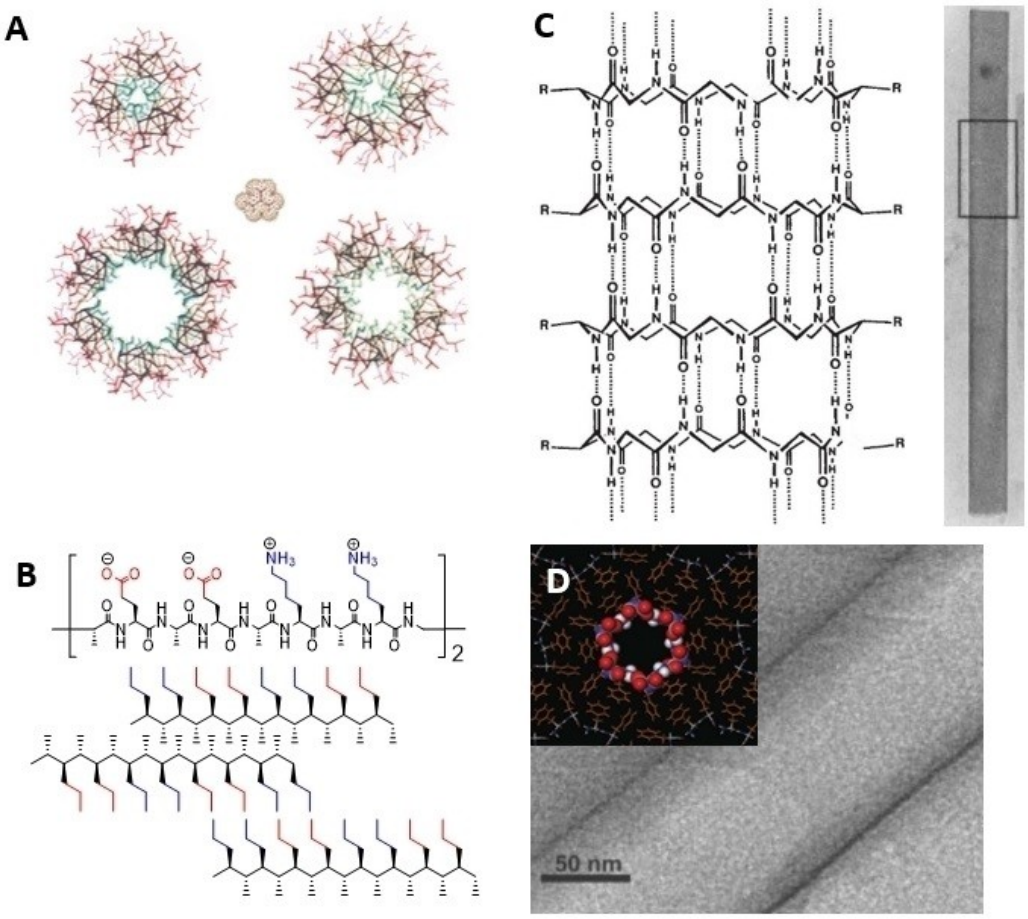


Figure 2. Early examples of reductionist design and discovery of peptide-based supramolecular materials. (A) Tritetrameric models of $(LSLLLSL)_3$ and penta- and hexameric models of $(LSSLLSL)_3$ showing increasing pore radius, a space-filling model of a guanidinium ion included for size comparison.,[19] reproduced with permission from AAAS 1988. (B) Repeating structure of $(AEAEAKAK)_2$ and the alternative electrophilic and hydrophobic stacking observed between the peptides,[22] reproduced with permission from PNAS 1993 (C) Nanotube formed of alternative D/L cyclic octapeptides with the backbone hydrogen bonds visualised, inset: hydrated single nanotube structure resolved by electron microscopy with dimensions of ≈86×1180 nm,[26] Reproduced with permission from Springer 1993. (D) Transmission electron microscopy resolved FF nanotube,[27] Reproduced with permission from AAAS 2003 inset FF nanotube pore visualised with the space-filling model within the greater atomistic framework.[28] Reproduced with permission from RSC 2006.

moted the self-assembly of the peptide into membranes formed of individually woven filaments of 10–20 nm width. This peptide design rule was found to be broadly extendable with the discovery in 1995 that a similar hexadecapeptide RAD16 ($(RARADADA)_2$) also yielded a β-sheet structure. This later study showed that EAK16 and RAD16 could self-assemble into membranes or filaments depending on the solution conditions.[23] Researchers continue to build on these designs, including charge complementary patterns, bioactive groups, shorter sequences, etc.[11,24]

This experimental work was followed by computational validation by Jun et al., who compare EAK16 with EAK16-IV ($(AE)_4(AK)_4$) in terms of energetic contributions of Lennard–Jones (LJ) exclusion, intrachain electrostatic interaction and bending energy of their coarse-grained model. Using Metropolis Monte Carlo (MC) simulations, the interaction energies were calculated for many possible conformations. The results showed an energetic penalty of +1.6 kBT for forming a β-hairpin for EAK16, while EAK16-IV had a favourable β-hairpin energetic change of −2.4 kBT. This finding was in line with their Fourier transformed infra-red spectroscopy (FTIR) spectra of EAK16 and EAK16-IV, which showed a peak for β-sheet structures in both but an additional peak from a β-turn structure in EAK16-IV only.[25]

A non-biological architecture also developed in the 1990s was the alternating D/L cyclic octapeptide cyclo-$[-(^dAE^dAQ)_2]$ designed by Ghadiri et al.[26] This cyclopeptide was rationalised by the idea that an alternating cyclic D/L sequence of an even number of residues could form a low-energy ring where the side chains lay perpendicular to the ring and that the backbones would hydrogen-bond in an alternating antiparallel pattern (Figure 2c). Under basic conditions, the monomers remain soluble and experience electrostatic repulsion, preventing the stacking of the backbones; however, under acidic conditions where the E residue

sidechains become protonated, the monomers become less soluble and no longer experience the electrostatic repulsion. This was confirmed by electron microscopy, which confirmed that these monomers spontaneously self-assemble into nanotubes in solution (Figure 2c, inset). Subsequently, by using all-atom molecular dynamics, Ghadiri et al.[29] showed that a stack of 10 of these cyclic octapeptides could hold up to 37 water molecules in a pattern of alternating density and remain stable with only a slight tendency of the end-of-tube peptides to dissociate.

Reches and Gazit pioneered a reductionist approach to identify self-assembling sequences, first demonstrated by systematic reduction of the β-amyloid peptide to its shortest aggregating sequence (FF).[27,30] This dipeptide rapidly formed nanotubes in water upon dissolution from an organic solvent (Figure 2d). These nanotubes could be loaded with silver ions and used as a sacrificial, protease-degradable template to reduce silver ions to form silver nanowires. Görbitz showed that these solvated nanotubes were of the same morphology (via X-ray diffraction (XRD)) as single crystal FF and composed of water channels surrounded by peptide backbones and separated by aromatic zippers (Figure 2d inset).[28] These FF systems were simulated in 2009 by Marrink et al. in the development of the MARTINI forcefield;[21] they found that while they were able to reproduce the degrees of freedom of all-atom forcefields and closely resemble the radial distribution functions, they did not form the expected nanotube morphology. It was subsequently shown that the number of dipeptides simulated (96) would not be sufficient to form a nanotube in this forcefield. Building on this pioneering work, computational studies were developed to explore the sequence space further. Screening of the entire di-and tripeptide sequence space has been explored by Frederix et al., who found many self-assembling peptide systems that have been further investigated, including the reproduction of the water channel arrays surrounded by stabilising F zippers in FF, confirming earlier XRD studies of Görbitz.[28,31–33]

These early examples illustrate that computation and experimental approaches have played an essential role in designing and understanding self-assembling peptides. Studies were initially rarely integrated and usually driven by experimentation, with the computational aspect providing a theoretical basis for earlier observations. Over the years, the focus gradually became more tightly integrated, using computation to understand these systems and to inform experiments confirming computational predictions when designing new systems.

3. Computationally Driven Materials Discovery

Since Gazit's observation of the possible role of π-stacking as the key driver of the formation of certain amyloid fibrils[30] and their subsequent discovery of the central dipeptide (FF) as a powerful self-assembling unit, there has been growing interest to gain insight into the underlying peptide self-assembly rules, starting with the simplest possible systems. However, even with simple systems, analysing the vast chemical space created by the potential combinations of amino acids to determine the self-assembly rules of peptides quickly becomes intractable using modern experimental approaches. An exception are the dipeptides (400 options), where the combinatorial sequence space can feasibly be fully synthesised and characterised. This is illustrated by Görbitz's analysis of a database of 160 dipeptide crystal structures, which could then be classified into different types, depending on key features involving the polar or nonpolar nature of the side chains, and backbone orientations. This observation of sequence-dependent crystal ordering among dipeptides paved the way for further studies to elucidate this behaviour.[27]

$$\mathrm{AP} = \frac{\mathrm{SASA}_{\mathrm{initial}}}{\mathrm{SASA}_{\mathrm{final}}} \tag{1}$$

Frederix et al. computationally screened the dipeptide sequence space (400 sequences) for their aggregation ability using a coarse-grained (CG) MARTINI force field.[21] The Martini forcefield uses a 4:1 atom/CG-bead mapping to represent protein backbone and side chains, and a 2:1 or 3:1 atom/CG-bead mapping for ring systems, parameterised based on the reproduction of partitioning free energies between polar and apolar phases of a large number of chemical compounds. The simulations of all 400 dipeptides were then compared by their ability to aggregate through the definition of the AP (aggregation propensity) score [Eq. (1)]. The results from these simulations could reproduce known self-assembling dipeptide sequences such as FF, FW, and IF while also distinguishing against those known not to self-assemble, such as FE, FK and VF. Remarkably, despite the coarse-grained nature of the FF simulation, more extended timescale simulations demonstrated that the CG forcefield could also provide information about the assembly process (e.g., snapshots for FF assembly are displayed in Figure 3b). In addition to revealing the molecular level detail of the formation of FF nanotubes, the CG forcefield was also able to reproduce experimentally observed properties of the final nanostructure. The diameter of the FF nanotube was in close agreement with previously reported XPRD values and the shift of dihedral angles between aromatic residues to a constrained, self-stacked conformer (Figure 3b,c).[28,33,34] This rotation was observed in the case of FF where the sidechain reorganisation during self-assembly gives rise to the hydrophobic zipper motif . In this architecture, an extended configuration enables the hydrophobic residues of one molecule to be exposed, creating a space for the hydrophobic residues of another molecule to "zip" into the 0° dihedral angle configuration of another peptide molecule (Figure 3a,c,g).[35,36]

Building on this work,[31] as well as several studies on liquid-liquid phase separation (LLPS) of proteins and peptides,[37,38] Tang et al.[39] explored the 400 dipeptide sequence space to identify minimalist peptides that can undergo LLPS. In order to do so, they quantified the microscopic fluidity known to be important in liquid condensates and used these insights to establish design rules for dipeptide condensates. Simulations were performed

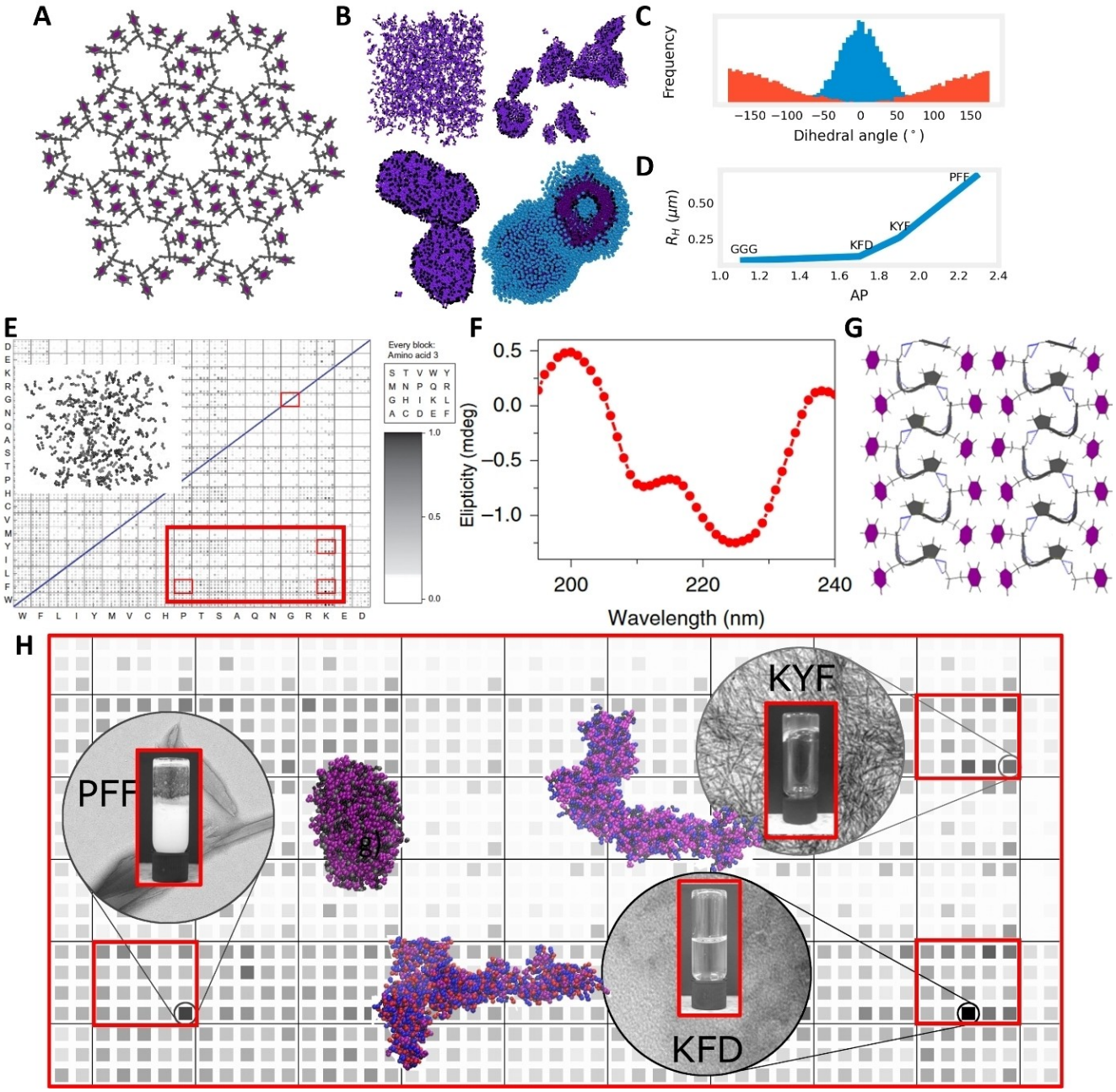


Figure 3. Computationally driven-discovery of supramolecular peptide materials (A) All-atom FF crystal structure showing a water channel of ≈10 Å. (B) Self-assembly process of FF revealed by CGMD from dispersed in solution (top left) to single nanotube with water channel of ≈10 Å (bottom right). (C) Dihedral angle distribution for a self-assembled FF nanotube (blue) and randomly distributed FF monomers in solution (red). (D) AP score for selected tripeptides vs. their measured R_H values from DLS measurements. (E) Grid showing combinations of amino acids in screened tripeptides and their AP score (N-termini on the Y-axis and C-termini on the X-axis) with a darker shade indicating a higher degree of aggregation, inset: CGMD simulation showing the dispersal of 300 triglycine monomers. (F) CD spectra of aqueous PFF displaying the presence of helical structures. (G) PFF 2×3 representations of unit cells. Grey and purple paper chain representations of P and F side chains, respectively, showing π-stacking with ribbon representation of the tripeptide backbone and dashed lines showing hydrogen bonding. (H) Close-up of a section of the grid shown in (E) indicating the positions of three experimentally confirmed self-assembling tripeptides and the hydrogels of each. Reproduced with permission from ACS 2007 and 2021, Springer 2015 and 2019.

using the MARTINI force field. They could identify LLPS systems by combining AP and clustering degree to quantify mobility and exchange of dipeptides within aggregates.[39] They identified a number of potential LLPS dipeptides and experimentally demonstrated the LLPS of QW, which was corroborated by differential interference contrast (DIC) microscopy and turbidimetry propensity. They also predicted 346 other dipeptides capable of LLPS. The design rules that could be derived from these predictions indicated three types of sequences with high LLPS probabilities—

1) an aromatic residue of high polarity (e.g., Y) with a hydrophobic residue (e.g., L); 2) an aromatic residue (such as F, and a polar residue (e.g., Q); 3) positively charged N-terminal residue and negatively charged C-terminal residue (e.g., K, D respectively). Interaction analysis shows that the LLPS of QW is a result of the interplay among QW-QW aromatic stacking, hydrogen-bonding, anion- and cation-π interactions, and the QW-water interactions. The work showed that, perhaps surprisingly, there is no lower limit to the size of peptides that can phase separate and provided a basis for understanding the key interactions that lead to it.

For longer peptides, the sequence space is too large to fully, or representatively, assess experimentally. Therefore, computational and, more recently, machine learning algorithms have been used to increase the accessible sequence space by several orders of magnitude.[31,32,40] After the aforementioned analysis of the dipeptide self-assembly space, Frederix et al. screened the entire tripeptide sequence space using coarse-grained molecular dynamics (CGMD) (Figure 3e,h).[32] In this case, the aim was to identify tripeptides that could form hydrogels, which require high self-assembly propensity combined with favourable solvent interactions.[41] This study, therefore, included a logP-weighted measure of self-assembly [AP_H, Eq. (2)], which could discriminate between the top aggregators (PFF, WFL, etc.) and the top hydrophilicity-weighted aggregators (KFD, KWD, etc.). Thirteen target peptides were synthesised, and their assembly behaviour was experimentally assessed to verify these predictions, with four examples studied in more detail. Dynamic light scattering (DLS) was used to measure the hydrodynamic radii (R_H) of the aggregates of GGG, KFD, KYF and PFF in water which showed that the AP was a good indicator of aggregation (Figure 3Figure 3d). FTIR absorption spectra showed that in the case of KYF and PFF, the red-shifted amide modes indicate highly ordered structures where present.[32] By analysing the entire data set, design rules for higher aggregation could be formulated, namely that aromatic amino acids are most favourable in positions 2 and 3 in a tripeptide, cationic and hydrogen-bonding residues favour position 1 (N-terminus) while anionic residues favour position 3.

$$AP_H = \left(AP'\right)^{\alpha} \cdot \log P \tag{2}$$

Where AP′ represents the AP normalised to between 0 and 1 and α the exponent that can be used to bias for peptides containing polar residues for favourable solvent interactions favoured for hydrogelation.

A number of these tripeptides were subsequently explored for oil-water emulsification propensities. Several peptides that were found to self-assemble in water to form hydrogels could also stabilise water-in-oil emulsions, either through the formation of monolayers (DFF, FFD) or interfacial fibre (KYF, KWY) networks, which could be computationally correlated with self-assembly propensity.[42]

Inspired by this tripeptide self-assembly mapping, Bera et al. investigated the tripeptide that was predicted to have the highest AP score, PFF, in more detail. Unlike most previously described FF-containing structures that are typically β-sheet-like, the PFF peptides were shown to form helical assemblies. Single crystal X-ray analysis of the self-assembled PFF crystal showed the torsion angles of the PFF residue to be in the right-handed α-helices region of the Ramachandran plot and the hydrophobic regions where oriented on the exterior, thus available for interactions with neighbouring peptide side chains through zipper motifs. FTIR and CD spectra (Figure 3f) showed a characteristic helical peak confirming that structured fibres can emerge from packed arrays of helices. These short intermolecular helices stacked head-to-tail and side-to-side (the direction of growth of the fibre), with the sides forming a highly hydrophobic phenylalanine zipper region (Figure 3g). To understand whether this is a unique feature of PFF or a more generic self-assembly mechanism, the proline residue was replaced by hydroxyproline (O), a non-gene-encoded amino acid found in collagen. This tripeptide was also found to self-assemble into a superhelical structure. The reported stiffness of PFF was more than double that of FF and more than five times greater than that of FF in the case of OFF due to the superhelical organisation and additional hydrogen bonds formed by the O residue.[35] The CG equilibrium simulation approach thus provides useful insight into the potential for short peptide structures to self-assemble and furthermore identify interesting candidates for detailed experimental analysis to elucidate details, such as the remarkable helical conformer observed in PFF crystals.

CG methods cannot change chemical stages in situ, such as those relevant to dynamic protonation or deprotonation, which is known to occur during self-assembly and could help to drive assembly, particularly in charged systems. It has been recognised that ionisation equilibria can be strongly impacted by proximity and local dielectric constants. Hence the observed (apparent) pK_a values in self-assembled structures with ionisable groups are frequently significantly shifted compared to those observed in dilute solutions. For example, the apparent pK_a of the terminal carboxylic acid in Fmoc-FF was measured to be shifted by as much as 6 units.[27] In order to appropriately accommodate these pK_a shifts, a variant of MD was developed that can handle this chemical change, namely constant pH [coarse-grained] molecular dynamic CpHMD.[43,44] This approach enables in situ changing of protonation states of titratable groups depending on the system pH and local environment. This dynamic method is preferential to fixing charge states based on pH as the theoretical pK_a of a residue is only relevant in dilute systems where the local environment is only water, i.e., the ability to switch charge states could facilitate self-assembly. Van Teijlingen et al.[44] used a method developed initially by Radak et al.[43] with a stochastic charge neutralisation mechanism whereby the system remains charge-neutral by switching on/off charges on Martini water beads. This method, which measures the likelihood of deprotonation at a given pH based on a Monte-Carlo criterion [*S*, Eq. (3)] and cooperativity (n), was validated by comparing an experimental titration of oleic acid with a CpHMD virtual titration of an oleic acid 30-mer.

$$S = \frac{1}{10^{n(\mathrm{p}K_a - \mathrm{pH})} + 1} \quad (3)$$

In simulating Fmoc-FF self-assembly, the method was found to provide results that were in agreement with the reported pK_a shifts by Tang et al. (Figure 4a,b).[45] The dynamic collapse of a nanotubular Fmoc-FF aggregate observed by Adams et al.[46] via pressure-response measurements was also simulated at different virtual pH values and these simulations showed good agreement with the experimental results. Namely, at pH 9, the nanotube is stable, and most carboxylate groups are deprotonated. In contrast, at pH 4, the tube collapses on itself (Figure 4d), and as it collapses, the local environment becomes more hydrophobic, causing further protonation, which propagates further collapse.[44] Applying CpHMD to CG simulations opens the door for more accurate modelling of the self-assembly of short peptide systems. Moreover, by incorporating the ability of the building blocks to adapt (change their chemical functionality) between the assembled and non-assembled states, the ability to design adaptive, functional materials through computational modelling becomes more feasible.

Moving beyond the di- and tripeptide sequence space requires dealing with much larger data sets. For these situations, machine learning has been used increasingly to accelerate materials discovery, with a particular focus on discovering non-intuitive peptide sequences. For example, this can include peptide sequences that self-assemble despite being soluble and those that do not rely on known self-assembling sub-sequences.[47] Van Teijlingen et al. demonstrated how an iterative machine learning algorithm (active learning) could search the sequence space of longer peptides and select for properties such as aggregation and solubility.[48] Validating this method based on the previously discussed virtual screening of tripeptides by Frederix et al.[32] the active learning algorithm was able to find all of the top 20 AP scoring tripeptides with an order of magnitude less CGMD simulations than was possible with a traditional full screening (Figure 4e). Using this method and restricting the machine learning model only towards soluble peptides, they were able to find those aggregating and soluble peptides beyond often-used intuitive residues (F/K/D/V/I/L), including peptides with residues such as Y/G/S/C/P/M such as YCGY, DKYW, SYGYF, PPPYV and WGGGGGC.[40] Independent from this work, Batra et al. applied an alternative machine learning approach using a tree search algorithm to identify soluble self-assembling pentapeptides.[49] This algorithm was also validated by identifying the tripeptides with the highest self-assembly propensities and was shown to outperform virtual screening. In comparing the machine learning approach to intuition (by asking human experts to design several self-assembling peptide sequences), this machine learning process also identified soluble aggregating peptides such as SYCGY, PPPHY, RWLDY and WKPYY that deviate from the established intuitive norms that suggest pentapeptides such as VVVVV, VKVFF and KFAFD. In conjunction with the work of Van Teijlingen et al., this research demonstrates a broader chemical space for short peptide materials development than is usually exploited.

Another example of active learning-directed peptide sequence space exploration was introduced by Shmilovich et al.,[50] who used an active learning approach to explore peptides functionalised with organic semiconductors to identify optoelectronic peptide nanostructures. They studied a tripeptide-based bola-amphiphile design, the symmetrical D-tripeptide-oligophenylenevinylene-tripeptide-D(DXXX-OPV3-XXXD) and explored the tripeptide sequence space (Figure 4f). DXXX-OPV3-XXXD-type monomers form 1D nanowires due to the linear stacking of the π-core and show promising properties for nanoelectronics within a biocompatible context.[50] The target of the active learning model was to find the monomers that yielded the most significant degree of aggregation driven by π-core stacking. The machine learning method (Figure 4g) also included an active learning component in this work. Active learning iterations performed an additional four simulations and added their results to the model, which was then used to make additional predictions. This process was repeated over 25 iterations, at which point further improvement in the predictions was obtained. This outcome suggests that the evaluation of the entire 8000 tripeptide sequence space was effectively traversed within 100 MD simulations (i.e., 80× faster than using a full screening method). The method revealed that tripeptides without bulky aromatic residues were favoured as their character impeded the formation of linear non-covalent linkages between the OPV3 moieties. Instead, smaller hydrophobic residues, particularly M in the 3rd X position from the OPV3 moiety, were promising candidates for experimental validation.[50]

Machine learning algorithms have also been trained on experimental data, Li et al. synthesized a library of 2304 peptide-like molecules[51] based on 12 Fmoc-N-Protected amino acids, of which 81 were found to be hydrogel-forming via rheology. For each of the monomers, 3109 PaDEL[52] descriptors were calculated and used to rationalize the molecular properties that promote hydrogel formation. Three machine learning binary classifiers (random forest, gradient boosting trees and logistic regression) were trained to predict from a data-resampled set (to mitigate against the large imbalance between positive and negative results) whether a monomer would form a hydrogel or not. Further, the models give an index of relative feature importance in how inputs map to classifications made, for instance, derivatives of the first ionization potentials, the Fmoc-amino acid the peptide-like structure was derived from, and other electro-topographical descriptions of hydrogen atoms proved to be important in forming hydrogels. This example reveals how machine learning techniques can be usefully implemented to explain results and observations with vast input data.

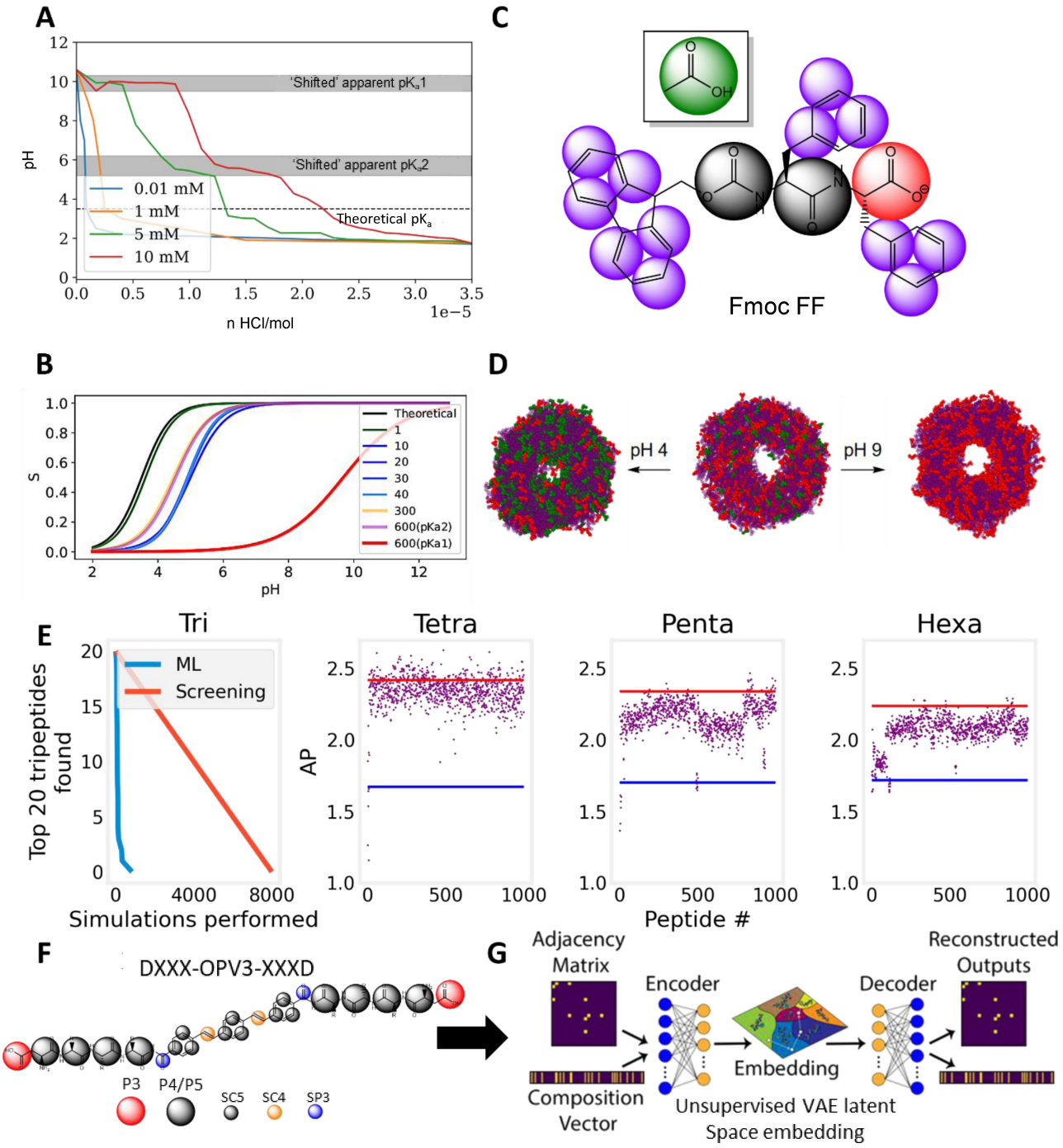


Figure 4. Development of new methods for computational analysis of supramolecular peptide assembly, (A) Experimental titration of Fmoc-FF at different concentrations showing the theoretical (dilute) pK_a and two shifted apparent pK_a's due to local environment effects that arise from self-assembly.[45] (B) CpHMD simulations of 600 Fmoc-FF molecules, which produce an increased apparent pK_a from the theoretical value, which is in line with the measurements of oleic acid pK_a at concentrations of 0.01–10 mM. (C) Fmoc-FF model with the titratable group (carboxylate bead) shown in red with the switchable carboxylic acid bead shown inset in green. (D) Hollow cylinders of Fmoc-FF at pH 4, 7 and 9, CpHMD equilibrates to a mixture of protonated/deprotonated carboxylate groups and a stable structure, pH 9 is almost entirely deprotonated and a stable hollow tube while at pH 4 carboxylate groups have become mostly protonated showing structure collapse. (E) Number of CGMD simulations performed to find the top 20 previously reported tripeptide aggregators for the active learning algorithm vs screening (random probabilistic) and active learning results for aggregating tetra-, penta- and hexapeptides, where the red and blue lines represent maximum and mean APs respectively. (F) CG model of the generalised sequence space searched by Shmilovish et al.[50] (G) Representation of the machine learning method used to conduct the search. The autoencoder attempts to reproduce an input molecule through a condensed informational embedding. This condensed latent space can then be used by a machine learning algorithm as a representation of the molecule and space to be searched for the optimal sequence. Reproduced with permission from ACS 2009 and 2020.

4. Integrated Approaches to Supramolecular Peptide Materials Design

The preceding section described the ability to screen large datasets and expand the search space to discover new peptide-based materials through computational methods. Traditionally, the design of new peptide-based materials is frequently inspired by motifs observed in biology, such as the minimalist amyloid FF motif, intrinsically disordered protein (IDP) inspired coacervation models, and from here involves an exploration driven by chemical intuition or testing of small sets of sequence variants. Thus, data-driven and intuitive approaches can often go hand-in-hand, enabling both predictive learning-informed experimental studies of new materials and understanding specific molecular interactions that lead to macroscopic properties in materials. To illustrate this process, we discuss examples of distinct supramolecular morphology: one-dimensional self-assembly of supramolecular fibres, followed by 2D assembly and finally, the formation of coacervates and transiently disordered peptide structures.

An important consideration when integrating computation and experiment is that MD methods typically approach thermodynamic equilibrium. At the same time, experimentally obtained structures may be kinetic aggregates and require thermal annealing to ensure they represent equilibrium structures. Notably, most reported self-assembling peptide nanostructures have very high aspect ratios (fibres and tubes). The high aspect ratio can result from growth kinetics along different axes in the supramolecular architecture. Networks of high aspect ratio fibres can form gels when fibres can interact to form a self-supporting network where the solvent is "trapped" within the fibre network. The strength of the intermolecular interactions between the fibres in the network can cause the structure to become kinetically trapped, in which case a gel-to-crystal transition may occur over time. In cases where the fibres present solvophilic interfaces, these 1D structures can also represent a thermodynamic minimum, as demonstrated through a coarse-grained approximation of a self-assembling process using 3D shapes with different solvophobic/-philic surfaces.[41]

A widely studied class of self-assembling peptides that form fibres and gels representing kinetic or thermodynamic structures are the aromatic peptide amphiphiles, short peptides functionalised with flat aromatic moieties to favour self-assembly.[53] As an early example, Zhang et al. studied 6 peptides, ^{D}A^{D}A, AA, GG, G^{D}A, GS, and GT, by applying a fluorenylmethyloxycarbonyl (Fmoc) group to the N-terminus. The formation of nanofibrous hydrogels was observed in 5 out of 6 dipeptides, with CD spectra indicating the formation of superhelical arrangements by Fmoc and alanine residues.[54] Our group and Gazit simultaneously and independently discovered that Fmoc-FF forms nanostructured hydrogels at physiological pH,[55,56] and a supramolecular architecture composed of π-interlocked beta sheets was proposed informed by spectroscopy and scattering data and optimised using computational modelling and energy minimisation.[54] Xu et al.[58] provided an enzyme-catalysed route to facilitate the formation of nanotubular hydrogels (and xerogels) of Fmoc-LLL. The morphology and supramolecular architecture of the nanotubes were studied using spectroscopy and scattering approaches, as well as molecular dynamics simulations. This xerogel was shown to be composed of nanotubes characterised by antiparallel β-sheets by the FTIR bands around 1636 and 1690 cm^{-1}. All-atom molecular dynamics found the β-sheets and π-stacking formed around 4.7 Å and 3.6 Å respectively, both values being within 0.2 Å of the equivalent wide-angle x-ray scattering (WAXS) measurements.

At this time, Fmoc is probably the most popular flat aromatic N-modification, however others, including naphthalene (Nap) and Pyrene (Py), have been studied in the context of FF self-assembly. For instance, Ghosh et al. found via UV/Vis spectroscopy that the relatively stronger Py-Py π-π interactions induced parallel stacking while the weaker Nap-Nap π-π stacking yielded parallel stacking.[59]

The integration of computation and experiment proved useful in enhancing the understanding of peptide-based materials. Fleming et al., investigated the use of FTIR in the structural assignment of supramolecular systems.[57] Fleming et al., specifically studied the interpretation of the presence of the IR peak at 1680–1695 cm^{-1} in Fmoc-AA by analysing the vibrational modes using DFT. In proteins and longer peptides, this vibrational peak is typically associated with carbonyls in an antiparallel β-sheet conformation, and it was previously assigned as such, e.g. for Fmoc-FF.[57] Evidence of decreasing intensities of this peak with increasing peptide bulkiness was previously observed.[60] The fact that this substitution left the 1685 cm^{-1} peak, which was absent in the Fmc spectrum unshifted, proved its attribution is correctly assigned to the carbamate group rather than the formation of a particular secondary structure motif upon self-assembly. However, DFT calculations of the monomers and parallel and antiparallel tetramers were able to indicate that the antiparallel orientation of the carbamates was closest to the experimental spectra.[61]

Reversible, enzymatic self-assembly of peptide nanostructures was explored as a means of annealing by Hughes et al., who reported the formation of rare 2D nanosheets composed of Fmoc-SF-OMe through a thermodynamically controlled process involving in situ condensation of Fmoc-S and F-OMe.[62] The advantage of in situ enzymatic condensation of Fmoc-S/F-OMe to form the Fmoc-peptide in situ is that it operates under thermodynamic control, thus avoiding kinetic aggregates and favouring the thermodynamically most stable supramolecular organisation.[59] Fmoc-SF-OMe, upon formation, gave rise to an extended 2D supramolecular structure. FTIR peaks at 1640 and 1680 cm^{-1} also indicated the growth of an ordered structure, while WAXS diffraction identified a repeating unit of 16 Å along the peptide backbone. Characteristic π-stacking and β-sheet spacing peaks demonstrated the ability to grow via antiparallel π-stacking in one dimension and β-sheet alignment in another.[62] All-atom MD was used to gain insight into the hydrophobic and hydrogen-bonding molecular association of monomers. This confirmed the ability of the phenyl side chains to form

extended π-β stacked bilayers, which, taken into consideration with the fluorenyl π-stacking, explains the 2D growth of the supramolecular structure.

Building on this work, to further understand hierarchical organisation in Fmoc dipeptide nanostructures, Sasselli et al., explore a similar approach combining reversible enzymatic self-assembly and molecular dynamics to supplement spectral information. They studied Fmoc-TF-NH_2, Fmoc-TF-OMe, Fmoc-TL-NH_2 and Fmoc-TL-OMe, all formed through reversible enzymatic condensation reactions. In competition during enzymatic assembly, with the precursors of all four peptides mixed, Fmoc-TF-NH_2 showed to outcompete the others, which indicated that its assembly was the most stable. This observation was likely due to the higher hydrophobicity and aromaticity of F, in conjunction with the presence of the amide as opposed to methyl ester, which results in more H-bonds. Following insights from Fmoc-FF assembly as discussed earlier,[57] two antiparallel models of Fmoc-TL-NH_2 were constructed, each having a different H-bond arrangement. It could be confirmed that one of these had a more favourable configuration based on how one of the models began to resemble the other following 150 ns MD simulations of each of them.[64] As observed in TEM analysis, the hierarchical twisting of fibres was favoured in further simulations comparing the two fibre configuration versus bilayers. Thus, computation supplements the understanding of these structures from the atomistic to the microscopic level.

In an effort to explore tripeptide sequence-dependent assembly atomistically, following tripeptide design insights.,[32] Lampel et al. studied a subset of identified tripeptides to investigate sequence isomers on solution conformations and, consequently, self-assembly. The self-assembly propensity and conformation selection of six C-terminal amidated tripeptide sequence isomers made of F, D and Y were investigated.[65] Conformational analysis of single molecule atomistic MD simulations showed that the dihedral angles between Y and F residues were strongly dependent on the sequence (Figure 5e,f). These aromatic moieties were seen to be predominantly self-stacked in both FYD and YFD (*syn* with respect to the peptide backbone), similar to the sidechain self-stacking observed in FF. By contrast, in DYF and DFY, the aromatic moieties were arranged on opposite sides of the peptide backbone in an *anti*-conformation. These conformations were experimentally observed in crystal structures obtained for DYF and YFD (Figure 5g and 5h). This observation suggests that these tripeptides pre-organise in solution and retain these conformations in supramolecular crystals. Thus, this work shows the peptide sequence dependence of the aggregation of these peptides and how they can be better understood by the variation in intramolecular interactions. This sequence dependence has been further used for the controlled formation of polymeric peptide pigments by enzymatic oxidation with varying efficiencies based on their preorganisation.

We hypothesise that supramolecular peptide crystals could serve as functional materials with water-responsive behaviour,[67] requiring aqueous pores, stiff and deformable domains, and anisotropic hierarchical architectures. Piotrowska et al. selected four peptides composed of a YF dyad coupled with a D/H amino acid at the C- and N-termini to make HYF, YFH, DYF and YFD.[66] Except for YFH, the tripeptides self-assembled to form needle-like crystals in solution (Figure 5j). X-ray powder diffraction measurements (PXRD) were performed in a humidity-controlled chamber, which showed crystal lattice changes, in the case of HYF, during water absorption/desorption cycles (Figure 5f). HYF exhibited aqueous pores that contract upon dehydration (Figure 5n) when the pore size reduces and expands on rehydration. The most significant change in pore size was registered between 10–30 % relative hydration (R. Hyd). To study the mechanism of water responsiveness at a molecular level, atomistic MD simulations were performed, starting with a single pore from the crystal structure and simulating with water amounts corresponding to varying humidities. Upon analysis of the hydrogen bonding occurrence and energy associated with the stability of these H-bonds between the internal peptide pore wall and the water molecules, it could be demonstrated that the H-bonding network between water and the pore surface was a critical feature to the water-responsive behaviour (Figure 5k). The 20–30 % transition, which was reproduced in MD simulations in close agreement with the PXRD results (Figure 5l), is highlighted by the comparison of pore width changes in the MD to the crystal lattice parameter changes at different R. Hyd (Figure 5o). This is also consistent with the amide region shifts observed in FTIR (Figure 5m), indicating a strengthening of H-bonds upon dehydration.

The Marchesan group studies the effects of introducing hetero-chirality in tri-peptides to influence self-assembly behaviour.[68–72] They deconvoluted the design rules for specific aromatic motifs that are present in heterochiral XFF and FXF systems (where X is either D or L enantiomer), showing different macroscopic and morphological behaviours. They first explored this concept by considering the effect of chirality in the formation of the FF zipper motif in the tripeptide LFF and its epimer ^{D}LFF.[68] Both epimers were observed to give rise to nanostructures with highly organised characteristics, as seen from microscopy, CD and Thioflavin T fluorescence. However, only the D-stereoisomer was observed to lead to the formation of nanofibers and gelation following the lowering of its pH to 7.4. In contrast, the L-peptide was seen to form short, thin filaments that were not sufficiently stable to form the network required to support a stable hydrogel. Energy minimisation simulations were carried out to understand how changing a single amino acid chirality could affect the self-assembly ability. These simulations showed that the D-epimer prefers a more extended molecular conformation that exposes the F rings, creating a space for the ring of another molecule to "zip" into the configuration. By contrast, the L-epimer was more disordered, with less of a conformational preference and thus more likely prone to a transient nature (Figure 6e,f).[68]

Building on this work, Garcia et al., studied the effects of amino acid chirality on the self-assembly of a set of tripeptides composed of phenylalanine dyads separated by a

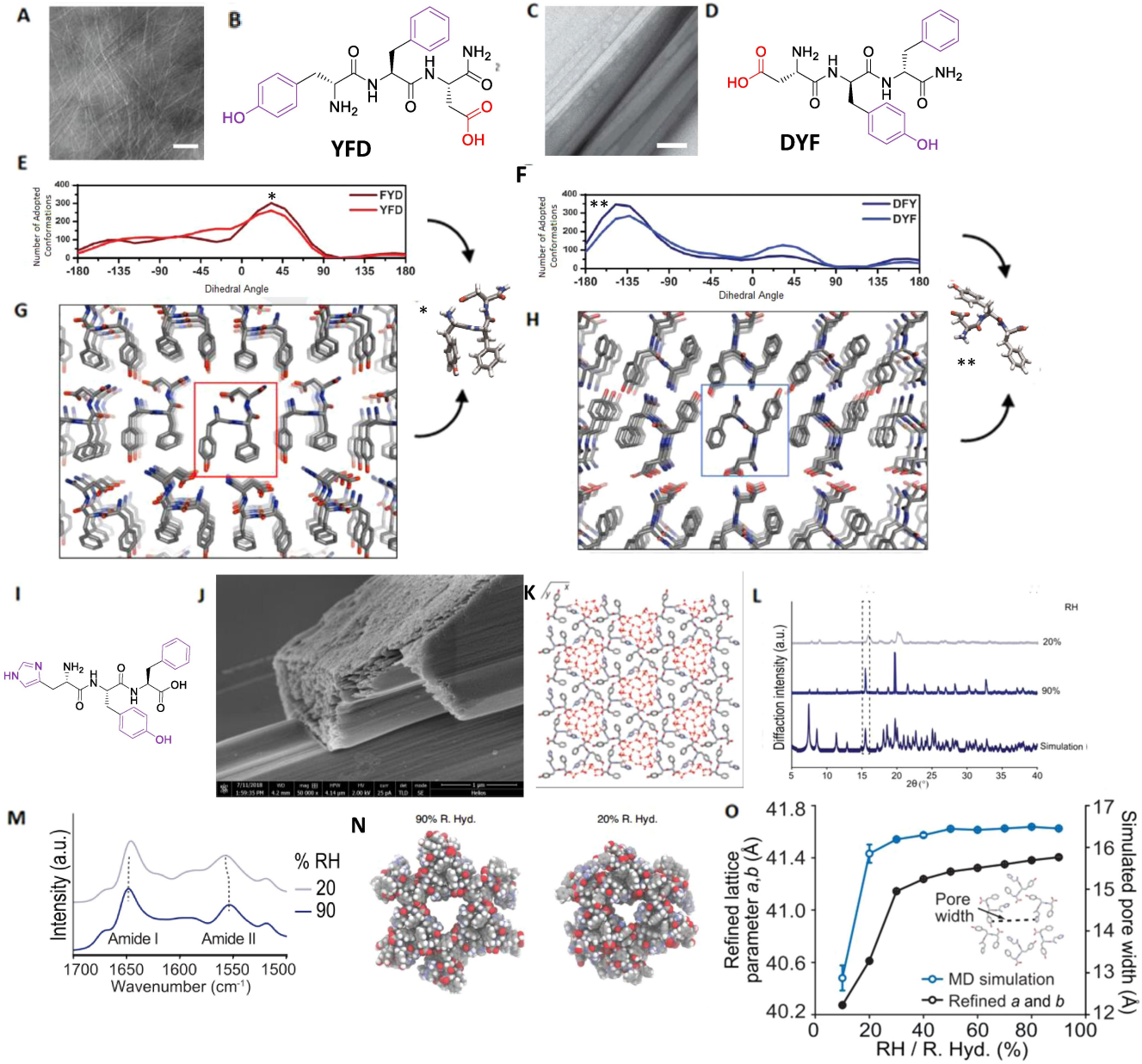


Figure 5. Integrating computation and experiment in tripeptide assemblies. Transmission electron micrographs of (A) YFD and (C) DYF showing nanofibers and crystals. Chemical structures of C terminal amidated (B) YFD (D) DYF. Preferred conformation of the single peptide simulations (E) YFD (F) DYF. Crystal structures, (G) YFD and (H) DYF showing different interfaces forming the lattice.[65] (I) Chemical structure of HYF. (J) Scanning electron microscope image of HYF crystal. (K) Interactions between the peptide and water (L) PXRD measurements of lattice parameters indicate water pore expansion upon hydration (M) FTIR spectra of HYF showing the amide bands shifting during dehydration pores in the HYF crystal. Red lines indicate water-water and water-tripeptide H-bonds. Blue lines indicate tripeptide-tripeptide H-bonds. (N) Change in the pore volume in the crystal at high (90%) and low (20%) humidities (O) lattice parameter changes derived from the Pawley refinement with MD simulated pore width in response to RH (and R. Hyd.) changes[66] Reproduced with permission from AAAS 2017 and springer 2021.

variable D or L amino acid (Figure 6a). Specifically, they investigated the differences between the assemblies of FAF, F^{D}LF, and F^{D}IF to find out why the heterochiral peptides assemble and form gels, whereas the homochiral ones do not. By using theoretical and experimental CD (Figure 6d), it could be demonstrated that the D-peptides showed an aromatic zipper contribution due to the steric hindrance introduced by the hetero-chirality. In particular, Ramachandran plots of all the peptides indicated type II β turns (Figure 6e). However, only the heterochiral peptides had the three hydrophobic side chains on the same side of the peptide backbone, where their mutual interactions favoured a turn overall by excluding water molecules (Figure 6f). The evidence of the F zippers can potentially be applied to other peptides with similar motifs, thus expanding design concepts in peptide design.[70]

The ability of heterochiral peptide-based materials to self-assemble and remain (as opposed to homochiral, all L

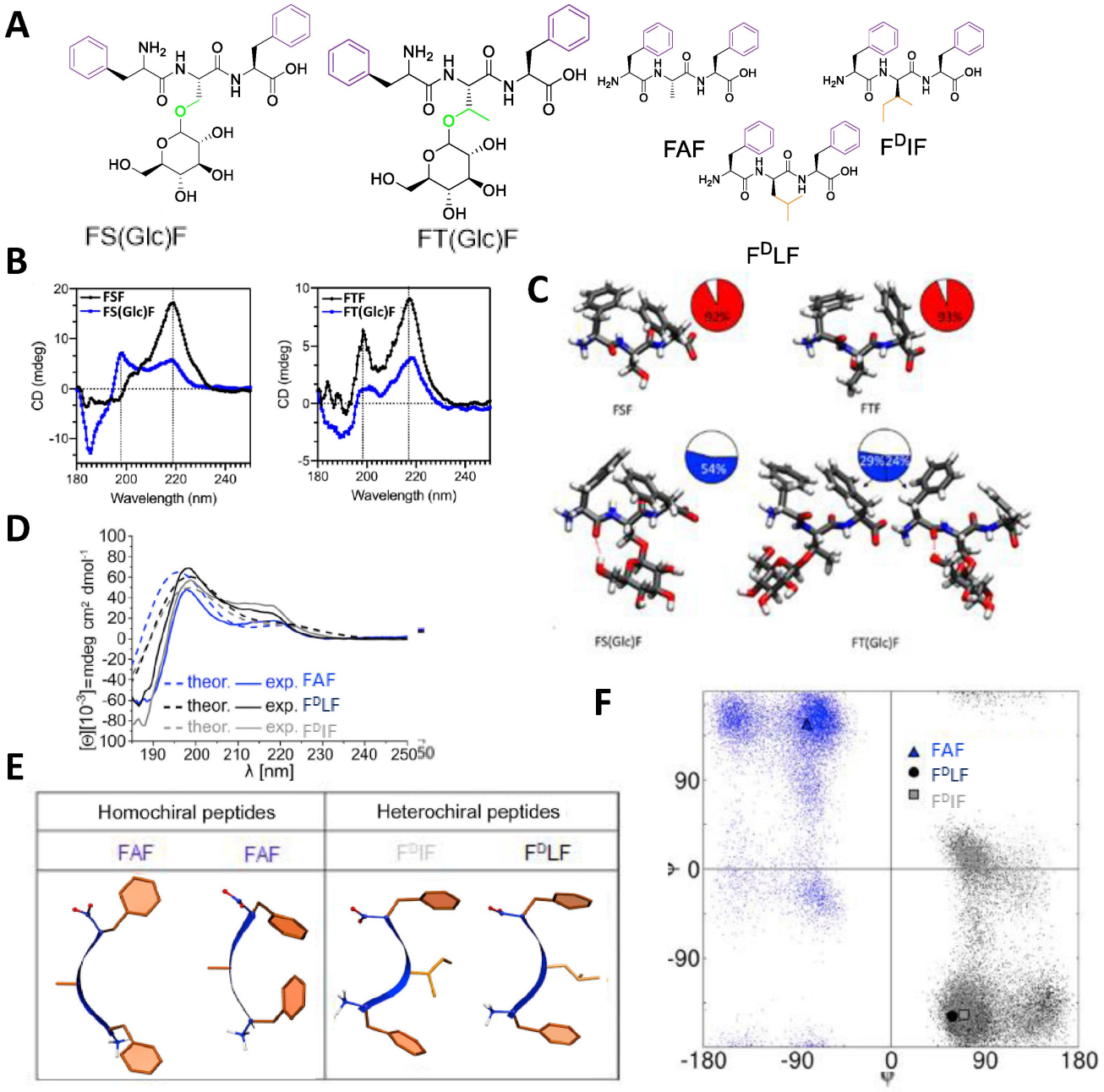


Figure 6. Conformer analysis of aromatic tripeptides and consequent self-assembly behaviour. a) Chemical structures of the o-glycosylated, homo- and heterochiral tripeptides b) CD spectra of peptides (black) and o-glycosylated peptides (blue). c) Preferred conformations observed in the MD simulations, with the pie charts showing the fraction of 5000 structures adopting these conformations using a root-mean-square deviation (RMSD) linkage cut-off of 0.1 nm to define the clusters of conformations. d) Calculated and experimental CD spectra for FAF, F^{D}IF and F^{D}LF tripeptides. e) Favoured amphiphilic conformation (hydrophobic side chain in orange and hydrophilic backbone in blue) for self-assembling tripeptides and the two equally most stable conformations for FAF. f) Ramachandran plot highlighting the most frequent conformations for the three tripeptides studied. Reproduced with permission from RSC 2012, Cell press 2018 and ACS 2021.

peptides) unmetabolised in the body led to the investigation of these materials as potential vehicles for drug release. Parisi et al., study the release of an anticancer drug, 5-fluorouracil (5-FU), from hydrogels of ^{D}LFF.[73] The peptide was found to self-assemble and give rise to hydrogel formation. A lower storage modulus was observed when the drug was incorporated into the gel structures, suggesting that the drug plays a role in cross-linking the fibres. This was corroborated by TEM images which showed that fibres with 5-FU were more heterogeneous in their median diameter. Although the drug inclusion did not significantly impact the fibre morphology, further investigation via FTIR revealed that the amide I region did not shift substantially upon the inclusion of 5-FU, both measurements showing only a weak binding to the drug molecule. ThT fluorescence and CD studies revealed that highly organised structures underpin

the fibre morphology, with minimal change in the self-assembly observed in the presence of the drug. There was no clear correlation between the rates of release among different drugs and their hydrophilicities, which led to the hypothesis that the ability of drugs to engage in specific non-covalent interactions with the peptide caused faster drug release. This hypothesis was verified by MD simulations, where parallel and antiparallel stacks of ^{D}LFF were formed in the absence and presence of the drug. These simulations revealed π-π stacking between the F and 5-FU rings as well as dynamic hydrogen bonding between the carboxylic acid group and N1-H group of the 5-FU. The weakness of the π-π stacking and the inconsistent nature of the interim hydrogen bonds explain the fleeting interactions between the drug and the hydrogel and its relatively fast release. This work paves the way for future computationally informed studies of peptide/drug combinations with specific release profiles.

Building on the F-zipper motif and adding additional functionality, Brito et al., studied the effects of o-glycosylation on self-assembling peptides, FSF and FTF, as minimalistic models of glycoproteins (Figure 6a), To understand the conformational differences upon introduction of glycans, and to serve as building blocks for future materials that incorporate both sugars and peptides. In addition to the effect of the central amino acid on self-assembly, this study also examined the role that the conformation of the glycol had on the supramolecular organisation and consequent material's properties of multi-molecule ensembles. Once the variations in the assembly behaviour of the different systems were established, MD simulations were used to help rationalise the different assemblies and revealed that O-glycosylation reduces the aggregation propensity due to the introduction of CH-π interactions, which compete with π-π interactions. Due to this difference in assembly, it was found that the glycosylated peptides formed aggregates with reduced stiffness as measured by atomic force microscopy (AFM) and higher thermal stabilities and corroborated by the calculated lower aggregation propensity. On measuring the CH-π, as well as the aromatic interactions in silico, and comparing it to the CD spectra, an aromatic zipper model similar to that suggested by Marchesan[68] was proposed for the aggregation of these tripeptides (Figure 6b). In addition to the reduced macroscopic stability of the assemblies and the lower aggregation propensity, details on the molecular level were also revealed. The XRD showed the disappearance of the β-sheet peak in the glycosylated FSF and that of the aromatic interactions in FTF. This was in line with the reduction of H-bonds in the backbones of the glycosylated peptides observed in the MD studies (Figure 6c). This study of the specific role of glycosylation in aggregation, thus, enables the understanding of the molecular origin of glycoprotein interactions as the basis for functional glycopeptide materials.[74]

Using the knowledge of aromatic dyads but looking to modify the peptides with non-peptidic functional biomolecules, Baek et al. studied 16 nucleopeptides, where the four nucleobases were attached to the N-termini of FFF, AFF, GFF and KFF.[75] In this work, TEM analysis demonstrated that most nucleopeptides could form nanofibres, and CD revealed helically dominant structures. The MD simulation of these systems showed that both peptide secondary structure (through backbone H-bonding) and DNA-like nucleobase stacking contributed to the stable, helically dominant fibres that formed hydrogels. This elucidation of interactions could feed into the further design of hybrid peptide/DNA materials that combine the advantages of both types of biomolecular assemblies.

Peptide amphiphiles (PAs),[76] short peptides functionalised with alkyl tails, show predictable and robust self-assembly behaviours and are therefore well-suited to further expand the functionality of these material classes, informed by integrated computation and experiment. Freeman et al. studied supramolecular fibres of alkylated peptides with the alkyl chain forming the core of the fibre and the non-apeptide ($V_3A_3K_3$) presented at the fibre surface (Figure 7a,b).[77] These PAs were further modified with complimentary units of DNA base pairs or charged amino acids which cross-link fibres via non-covalent linkages (Figure 7c).[34] The inclusion of the DNA base pairs causes the fibres to form twisted bundles with a residue-dependent pitch of twist. Scanning electron microscopy (SEM) shows fibres, and TEM having higher resolution shows them to be formed of multiple fibres twisted into bundles. A CGMD model where each DNA base is represented by a single bead was used to further investigate this phenomenon. Using CGMD, the pitch was measured as a function of the length of oligonucleotides and as a function of the fraction of monomers with attached oligonucleotides. This showed that increasing the length of the DNA base pairs increased the pitch, while increasing the percentage of monomers with oligonucleotides decreased the pitch (Figure 7d). These results agreed with experimental TEM observations and opened the door for further modification and tuning of the fibres′ growth rate. The MARTINI forcefield was used to calculate the intramolecular energy for the formation of the fibres (E_{intra}, based on the calculated energy for the β-sheet formation and hydrophobic collapse). In contrast, the intermolecular energy associated with the bundling of the fibres (E_{inter}) was calculated as the DNA hybridisation energy (estimated from an online calculator, OligoCalc).[78] Mapping these values to fibres of various compositions and their growth rates, a 3D map showed a very narrow range of $5\,k_BT < E_{inter} < 10\,k_BT$ is required for twisted bundle growth (Figure 7e,f). Alternatively, a charge complimentary tripeptide (E_3) was used in place of the DNA and combinations of complementary peptides nucleic acids (PNAs), thus showing the findings translate to other biomolecular conjugates. This decrease in electrostatic repulsion leads to more twisting resulting in a rod-like structure, compared to the more twisted-ribbon-like structure of the DNA-DNA systems.

Computational approaches have informed the design of materials inspired by nature's protein materials. Wu et al., investigated a method to produce biocompatible, spatiotemporally controlled microstructures composed of elastin-like polypeptides (ELK).[79,80] To build additional functionality into these systems, they investigated the introduction of graphene oxide (GO, combined with (MESLLP-(VPGIG VPGIG VPGKG VPGIG VPGIG)$_{24}$, ELK1)(Figure 8a,b).

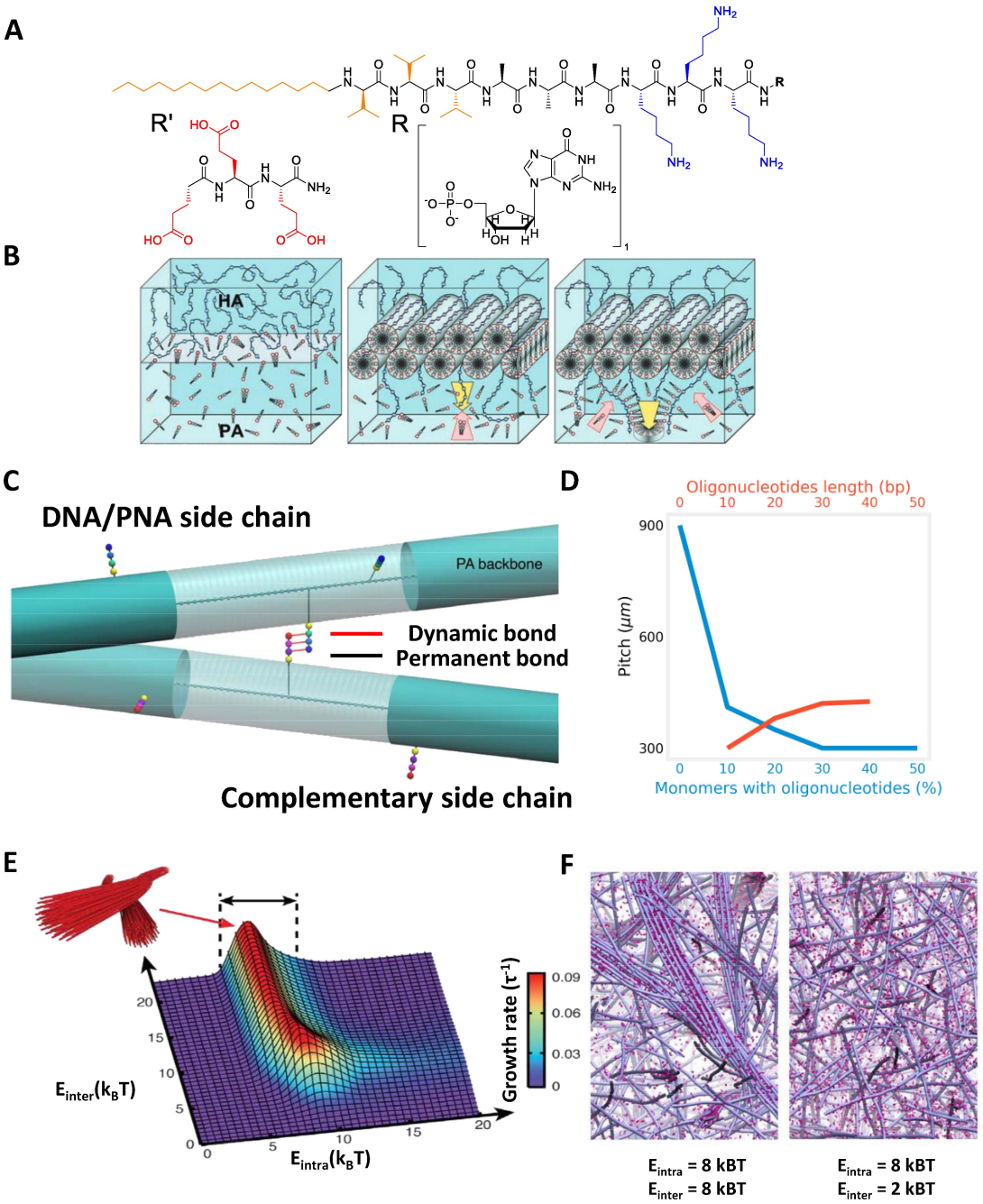


Figure 7. Reversible cross-linking and fibre twisting of modified PAs, (A) C_{16}-$V_3A_3K_3$-R, where R may be a DNA or peptide sequence (or a combination of the two). (B) membrane formation between high molecular weight polysaccharide hyaluronic acid (HA) and PA showing interface formation, followed by parallel nanofibre formation, which in turn leads to perpendicular nanofibre formation due to HA crossing the membrane via reptation[77] (C) The twisted bundle macrostructure is resolved by electron microscopy however using CGMD a much higher "resolution" is achievable revealing individual DNA base pairs or complimentary amino acid side chains interacting non-covalently and promoting fibre twisting. (D) The pitch of the twisted bundles can be controlled by the length of the linker group or a number of monomers containing a linker group (oligonucleotides in this example). (E) Twist bundle group rate as a function of E_{inter} and E_{intra} showing formation within a narrow window, this part of the Figure is taken directly from the publication. (F) Thick bundles form in the red area of (E) where the growth rate is high and E_{intra} and $E_{inter}=8$ (right). When the growth rate is low and only weak inter-fibre ($E_{inter}=2$) exists, larger bundles are not sustained. Reproduced with permission from AAAS 2008 and 2018.

The resulting non-covalent bioconjugate composites have been found to reduce the cytotoxicity of GO critical for the potential uses of GO in medical treatments where it can be used as (conducting) biological scaffolding.[81] A MARTINI-compatible coarse-grained GO model was developed to gain insights into the mode of co-assembly. This model revealed that the aggregation between ELK1 and GO was largely electrostatically driven by the interaction between negatively

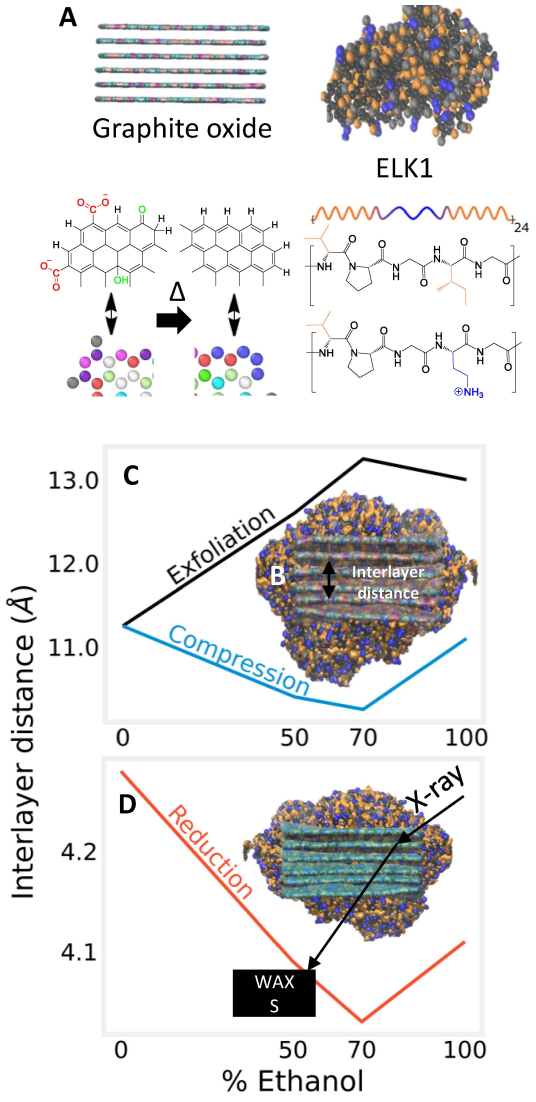


Figure 8. Co-assembly of graphene oxide and elastin-like peptides. (A) GO and ELK1 models with the atomic structure of representative examples and repeating units of ELK1 and GO, respectively. (B) WAXS was used to determine interlayer spacing on the rGO sheets (blue), which matched the pattern of results found from (C) CGMD simulations of GO (black line) and GO-protein conjugate (blue line) in increasing concentrations of ethanol. GO simulations are in line with known phenomena of graphene oxide exfoliation whereby intercalation of drives exfoliation of the sheets.[91] Whereas GO sheets wrapped in ELK1 protein are compressed with increasing ethanol percentage. Compression/exfoliation is measured via Euclidean distances between the mean Z-positions of the sheets. (D) Interlayer distance of the rGO as measured by WAXS, the results match the trend of the pre-reduction conditions elucidated by CGMD in (C). Reproduced with permission from RSC 2012.

charged GO residues and positively charged lysine side-chains. Once the GO bioconjugate was stabilised, the effect of introducing ethanol to the system at varying concentrations was investigated in order to reduce the GO (rGO) in order to increase conductance.[81] Wu et al. observed that the highest degree of reduction was found at 70 % ethanol, and by simulating the effect of the addition of an increasing percentage of ethanol in the solution, CGMD revealed that this concentration caused the largest degree of reorganisation of the protein, which resulted in pushing the GO sheets together to increase their reactivity. This effect is counter-intuitive as adding ethanol typically exfoliates GO; indeed, this effect was observed in the protein-free simulation. Following this structural revelation from the computational analysis of the system, this observation was confirmed by WAXS (Figure 8c), demonstrating how an iterative integrated collaborative effort between experimental and computational approaches can lead not only to new materials but a better understanding of existing materials.[81]

Another mode of self-assembly involving IDPs is coacervation.[82,83] Coacervates are generally understood to appear due to LLPS, resulting in an aqueous phase loaded with liquid macrostructures typically that may have high- and low-density regions. LLPS is a rapidly growing field where the rules for designing efficient systems are only now emerging. Critical distinctions between the self-assembly of ordered structures and LLPS peptides is that the formation of coacervates is strongly driven by the sidechain of the amino acid with little reliance on backbone H-bonding interactions, resulting in a globally disordered structure. One model that has been proposed and widely adopted is the *sticker-and-spacer* model, where stickers refer to segments of the peptide (typically hydrophobic and/or aromatic residues) that drive self-association and spacers connect the stickers (typically soluble polar residues) together.[1,84,85] The use of computational methods to help understand the sequence and patterning rules of LLPS (including the mapping mentioned above of the dipeptide sequence space)[39] is on the increase.

Experimental work on coacervation has shown the correlation effect of polypeptide length, charge strength, and charge pattern on the degree of ionic coacervation.[86–89] In order to systematically study pattern design, Chang et al., studied the coacervation behaviour of 50-mers of E (poly-Glu) with 50-mers composed of G and K with the same net charge but different sticker-spacer patterns, e.g., repeating units of (KG), (KKGG) etc. They found that the entropic contribution of the coacervation matched qualitatively with the isothermal titration calorimetry (ITC) data, proving that the larger the "blocks" of charge in the copolymer sequence were, the more confined counterions were to these blocks. Their release resulted in a more significant entropy change. This sequence effect was due to differences in entropic confinement of condensed counterions along the polymer. Using a restricted primitive coarse-grained representation, Chang et al. measured the effective energy and entropy for a given salt concentration via Monte Carlo simulations. These results were found to be in line with experimental trends. The coacervation was determined by assessing whether the calculation of excess free energy due to the mixing of the peptides with the salt volume was less than the free energy of a non-phase-separated system. Thus, this work establishes design considerations that link sequence to the physical properties of complex electrostatic coacervates. This effect can be used to understand the effect of sticker-spacer lengths and propose more complex assemblies.[90]

Another example of dynamic structures that are dominated by sidechain interactions is the concept of induced fit, where the presence of a ligand leads to structural reorganisation. Induced fit is typically associated with local reorganisation in proteins and has not received much attention in short peptides. Using phage display,[92,93] Kroiss et al. identified the unstructured peptide ADARYKS to exhibit a sequence-specific binding affinity for ATP using phage display techniques. This was the first known example of an ATP complexing short peptide identified through de novo selection and, remarkably, did not resemble any known nucleotide binder. It was found using MD H-bond analysis and NMR studies that the peptide interacts with ATP via the charged phosphate moiety and the aromatic ring of adenine. All the residues except the 3rd A in the peptide interact with ATP. It was also seen that the interactions between peptide and ATP were very dynamic, a finding that was further explored in a subsequent study.[94] In 2020, Kroiss et al. used a similar approach to develop computational predictions of heptapeptide binder to uridine triphosphate (UTP) from a large number of sequences. In this case, an improved selection method was used, using streptavidin-coated paramagnetic beads and UTP-biotin ligands together with a magnetic rack to facilitate selection compared to the previous centrifugation approach. The peptide KAIHPMR-NH_2 was demonstrated to form a stable, induced loop upon recognition of UTP. Binding was found to decrease the 1 and 7 residue distance of the investigated peptide KAIHPMR over time by a higher amount compared to other weaker binders. This illustrated the "induced fit" mode of binding that is often found in ligand binders in biology. In both studies, the binding constant values saw a strong correlation between experimental and computational values.[95] Subsequently, phage-display screening on self-assembled ligands was used to identify oligo-peptides that selectively bind supramolecular targets over their unassembled counterparts by using 9-Fluorenylmethoxy-phenylalanine-tyrosine-phosphate (Fmoc-FYp) micelles as targets. The lead peptide, KVYFSIPWRPM-NH_2, was found to bind to the Fmoc-FYp ligand exclusively in its self-assembled state with $K_D = 74 \pm 3\ \mu M$. Circular dichroism, NMR, and molecular dynamics simulations revealed that the peptide interacts through the KVYF terminus and this binding event disrupts the assembled structure.[96]

In order to study the ability to produce disordered coacervates on demand, as well as the ability to include folded domains into overall disordered structures, Scott et al. studied a coacervation system of poly(S-alkyl-L-homocysteine) functionalised with different amino acids and poly(L-methionine)$_{60}$ that each contained α-helical structures, and flexible side chain orientations, that could reversibly form coacervates in response to the oxidising or reducing environments. These template systems were specifically chosen for having a high level of tunability as they can be reversibly converted to a disordered conformation upon mild oxidation to the sulfoxide derivative. Moreover, the side chains can undergo protonation and deprotonation within a physiological pH range, which is advantageous for the biologically relevant adjustment of physical properties. These polymers were observed via CD spectra to form α-helical structures based on their CD spectra. Through MD simulations of a 15-mer of polyleucine, the variability in the observed RMSD values indicated that the side chains provide the required disorder in conformation despite the backbones showing helical structures. Further, experimentally, several compounds were found to phase separate in a pH and temperature-dependent manner with cloud point temperatures that decreased with increasing hydrophobicity of sidechain amino acids in the order $V < L < F$. These systems were also found to reversibly oxidise through cystine-cysteine interconversion, enabling the switchable coacervation. The authors demonstrated a series of tuneable and addressable coacervates where MD studies could be used to deconvolute the entropy of disorder that led to coacervation.[84]

Schnitzer et al., explore design of peptide inspired metal–organic frameworks (MOFs) using oligoprolines complexing with zinc ions. Different lengths of oligoprolines were seen to form different combinations of cis and trans arrangements, leading to differences in the resultant nanosheets. This differs from conventional MOFs since the oligoproline ligands interact amongst themselves, as opposed to solely serving as spacers for the ion. Through energy decomposition and density functional theory analysis, they found the specific non-covalent interactions responsible for these differences, paving the way for the design of further peptide-based MOF materials.[97]

Overall, the integration of computation and experiment can be informative at many levels in ordered and disordered systems. Examples discussed that would be impossible to study with the same level of confidence by relying exclusively on computation or experiment includes (i) prediction of assembly propensity, (ii) rational selection of sequence isomers and sequence edits, (iii) enhancing understanding of spectroscopy, (iv) connecting spectroscopy to microscopy across length scales, (v) understanding of design rules based on dynamic amino acid side chain interactions, as relevant to self-assembly, coacervation and induced fit.

5. Where do we go from here

Supramolecular chemists have recognised for decades that reversible bonds can replace covalent bonds, yielding dynamic functional materials with applications in materials science and biomedicine.[98–100] Peptides are excellent molecules for this purpose. However, identifying the sequences most suited to a particular problem is challenging because of the vastness of the biomolecular sequence space. Nature samples only a small subset by using natural selection to find and optimise the functional proteins in organisms. It is becoming clear that useful self-assembly modalities and materials with advanced non-biological functionalities are also present in this same overall sequence space. Integration of computation and experiment provides ways to find and characterise them.

While early work frequently focused on stable folds based on protein's secondary structures, there is increased

interest in materials that occupy shallow energy landscapes so that they can more readily adapt to new conditions. In these cases, there is more emphasis on side chain interactions compared to stable H-bonded backbone folds. Indeed, among the materials studied, the entire spectrum from supramolecular crystals and amyloids to structures that combine ordered and disordered domains to coacervates that have only transient structures have a high dependency on side chain-specific interactions. Each of these systems either benefits or have the potential to benefit from the emerging computational methods which can simulate systems from a few to millions of atoms.

Moreover, the focus is increasingly on multicomponent systems with increased functionality and adaptiveness, overcoming potential barriers of homotypic assemblies and reducing synthesis costs. For example, the search space of tripeptide co-assembly contains approximately the same combinatorial space as that of self-assembled hexapeptides. The possibility of including stimuli-responsive and induced fit moieties to a self-assembling peptide without significantly altering the supra structure simplifies the design of new stimuli-responsive materials and medical treatments. Another area of interest is the development of self-assembling systems under non-aqueous conditions or heterogeneous media, including solid-liquid and liquid-gas interfaces, elevated temperatures, non-aqueous solvents and dry conditions. Computational models are currently primarily based on either vacuum or aqueous conditions; however, they can be parameterised to simulate any environmental conditions leading to broader applicability. These models are being employed more and more to drive systems investigation as their cost is typically far below that of the equivalent experiment.

Materials and systems based on peptides have the potential to be produced using biotechnology approaches, ultimately offering the potential to avoid the use of petrochemical sources.[101,102] Thus, with the aid of computational approaches, machine learning, and increased experimentally derived insights, these materials are increasingly rationally accessible, and they will likely play a role in future green materials and nanotechnologies based on circular economy concepts.[103]

Acknowledgements

R.V.U. acknowledges funding from the Office of Naval Research for the Vannevar Bush Faculty Fellowship (grant N00014-21-1-2967). We thank the Air Force Office of Scientific Research (AFoSR) for funding of R.V.U., and M.R. (grant FA9550-21-1-0091. R.V.U. and M.R also acknowledge Dr. Deborah Sementa for the design of the amino acid chart in the frontpiece graphic.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: Computation · Molecular Dynamics · Peptides · Self-Assembly · Supramolecular Chemistry

[1] F. Sheehan, D. Sementa, A. Jain, M. Kumar, M. Tayarani-Najjaran, D. Kroiss, R. V. Ulijn, *Chem. Rev.* **2021**, *121*, 13869–13914.
[2] D. W. P. M. Löwik, E. H. P. Leunissen, M. van den Heuvel, M. B. Hansen, J. C. M. van Hest, *Chem. Soc. Rev.* **2010**, *39*, 3394.
[3] J. B. Matson, R. H. Zha, S. I. Stupp, *Curr. Opin. Solid State Mater. Sci.* **2011**, *15*, 225–235.
[4] X. Zhao, F. Pan, H. Xu, M. Yaseen, H. Shan, C. A. E. Hauser, S. Zhang, J. R. Lu, *Chem. Soc. Rev.* **2010**, *39*, 3480.
[5] H. Wang, Z. Feng, B. Xu, *Angew. Chem. Int. Ed.* **2019**, *58*, 10423–10432; *Angew. Chem.* **2019**, *131*, 10532–10541.
[6] A. Levin, T. A. Hakala, L. Schnaider, G. J. L. Bernardes, E. Gazit, T. P. J. Knowles, *Nat. Chem. Rev.* **2020**, *4*, 615–634.
[7] X. Hu, M. Liao, H. Gong, L. Zhang, H. Cox, T. A. Waigh, J. R. Lu, *Curr. Opin. Colloid Interface Sci.* **2020**, *45*, 1–13.
[8] J. L. Beesley, D. N. Woolfson, *Curr. Opin. Biotechnol.* **2019**, *58*, 175–182.
[9] T. Tuttle, *Isr. J. Chem.* **2015**, *55*, 724–734.
[10] A. Lampel, R. V. Ulijn, T. Tuttle, *Chem. Soc. Rev.* **2018**, *47*, 3737–3758.
[11] R. Qing, S. Hao, E. Smorodina, D. Jin, A. Zalevsky, S. Zhang, *Chem. Rev.* **2022**, *122*, 14085–14179.
[12] E. Dujardin, S. Mann, *Adv. Eng. Mater.* **2002**, *4*, 461–474.
[13] S. Ghosh, M. Reches, E. Gazit, S. Verma, *Angew. Chem. Int. Ed.* **2007**, *46*, 2002–2004; *Angew. Chem.* **2007**, *119*, 2048–2050.
[14] C. W. Gruber, M. Muttenthaler, M. Freissmuth, *Curr. Pharm. Design* **2010**, *16*, 3071–3088.
[15] L. J. Martin, B. R. Sculimbrene, M. Nitz, B. Imperiali, *QSAR Comb. Sci.* **2005**, *24*, 1149–1157.
[16] M. A. Gallop, R. W. Barrett, W. J. Dower, S. P. A. Fodor, E. M. Gordon, *J. Med. Chem.* **1994**, *37*, 1233–1251.
[17] R. Rathinakumar, W. C. Wimley, *J. Am. Chem. Soc.* **2008**, *130*, 9849–9858.
[18] D. M. Raymond, B. L. Nilsson, *Chem. Soc. Rev.* **2018**, *47*, 3659–3720.
[19] J. D. Lear, Z. R. Wasserman, W. F. DeGrado, *Science* **1988**, *240*, 1177–1181.
[20] D. N. Woolfson, D. H. Williams, *FEBS Lett.* **1990**, *277*, 185–188.
[21] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. De Vries, *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
[22] S. Zhang, T. Holmes, C. Lockshin, A. Rich, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3334–3338.
[23] S. Zhang, T. C. Holmes, C. M. DiPersio, R. O. Hynes, X. Su, A. Rich, *Biomaterials* **1995**, *16*, 1385–1393.
[24] M. A. Elsawy, J. K. Wychowaniec, L. A. Castillo Díaz, A. M. Smith, A. F. Miller, A. Saiani, *Biomacromolecules* **2022**, *23*, 2624–2634.
[25] S. Jun, Y. Hong, H. Imamura, B.-Y. Ha, J. Bechhoefer, P. Chen, *Biophys. J.* **2004**, *87*, 1249–1259.
[26] M. R. Ghadiri, J. R. Granja, R. A. Milligan, D. E. McRee, N. Khazanovich, *Nature* **1993**, *366*, 324–327.
[27] M. Reches, E. Gazit, *Science* **2003**, *300*, 625–627.
[28] C. H. Görbitz, *Chem. Commun.* **2006**, 2332–2334.
[29] M. Engels, D. Bashford, M. R. Ghadiri, *J. Am. Chem. Soc.* **1995**, *117*, 9151–9158.
[30] E. Gazit, *The FASEB Journal* **2002**, *16*, 77–83.
[31] P. W. J. M. Frederix, R. V. Ulijn, N. T. Hunt, T. Tuttle, *J. Phys. Chem. Lett.* **2011**, *2*, 2380–2384.
[32] P. W. J. M. Frederix, G. G. Scott, Y. M. Abul-Haija, D. Kalafatovic, C. G. Pappas, N. Javid, N. T. Hunt, R. V. Ulijn, T. Tuttle, *Nat. Chem.* **2015**, *7*, 30–37.

[33] C. H. Görbitz, *Chem. Eur. J.* **2007**, *13*, 1022–1031.
[34] R. Freeman, M. Han, Z. Álvarez, J. A. Lewis, J. R. Wester, N. Stephanopoulos, M. T. McClendon, C. Lynsky, J. M. Godbe, H. Sangji, E. Luijten, S. I. Stupp, *Science* **2018**, *362*, 808–813.
[35] S. Bera, S. Mondal, B. Xue, L. J. W. Shimon, Y. Cao, E. Gazit, *Nat. Mater.* **2019**, *18*, 503–509.
[36] A. M. Garcia, M. Melchionna, O. Bellotto, S. Kralj, S. Semeraro, E. Parisi, D. Iglesias, P. D'Andrea, R. De Zorzi, A. V. Vargiu, S. Marchesan, *ACS Nano* **2021**, *15*, 3015–3025.
[37] S. Boeynaems, E. Bogaert, D. Kovacs, A. Konijnenberg, E. Timmerman, A. Volkov, M. Guharoy, M. De Decker, T. Jaspers, V. H. Ryan, A. M. Janke, P. Baatsen, T. Vercruysse, R.-M. Kolaitis, D. Daelemans, J. P. Taylor, N. Kedersha, P. Anderson, F. Impens, F. Sobott, J. Schymkowitz, F. Rousseau, N. L. Fawzi, W. Robberecht, P. Van Damme, P. Tompa, L. Van Den Bosch, *Mol. Cell* **2017**, *65*, 1044–1055.
[38] F. G. Quiroz, A. Chilkoti, *Nat. Mater.* **2015**, *14*, 1164–1171.
[39] Y. Tang, S. Bera, Y. Yao, J. Zeng, Z. Lao, X. Dong, E. Gazit, G. Wei, *Cell Rep. Phys. Sci.* **2021**, *2*, 100579.
[40] A. Van Teijlingen, T. Tuttle, *J. Chem. Theory Comput.* **2021**, *17*, 3221–3232.
[41] I. Ramos Sasselli, P. J. Halling, R. V. Ulijn, T. Tuttle, *ACS Nano* **2016**, *10*, 2661–2668.
[42] G. G. Scott, P. J. McKnight, T. Tuttle, R. V. Ulijn, *Adv. Mater.* **2016**, *28*, 1381–1386.
[43] B. K. Radak, C. Chipot, D. Suh, S. Jo, W. Jiang, J. C. Phillips, K. Schulten, B. Roux, *J. Chem. Theory Comput.* **2017**, *13*, 5933–5944.
[44] A. van Teijlingen, H. W. A. Swanson, K. H. A. Lau, T. Tuttle, *J. Phys. Chem. Lett.* **2022**, *13*, 4046–4051.
[45] C. Tang, A. M. Smith, R. F. Collins, R. V. Ulijn, A. Saiani, *Langmuir* **2009**, *25*, 9447–9453.
[46] D. J. Adams, L. M. Mullen, M. Berta, L. Chen, W. J. Frith, *Soft Matter* **2010**, *6*, 1971–1980.
[47] S. K. Kauwe, J. Graser, R. Murdock, T. D. Sparks, *Comput. Mater. Sci.* **2020**, *174*, 109498.
[48] H. Moriwaki, Y.-S. Tian, N. Kawashita, T. Takagi, *J. Cheminf.* **2018**, *10*, 4.
[49] R. Batra, T. D. Loeffler, H. Chan, S. Srinivasan, H. Cui, I. V. Korendovych, V. Nanda, L. C. Palmer, L. A. Solomon, H. C. Fry, S. K. R. S. Sankaranarayanan, *Nat. Chem.* **2022**, *14*, 1427–1435.
[50] K. Shmilovich, R. A. Mansbach, H. Sidky, O. E. Dunne, S. S. Panda, J. D. Tovar, A. L. Ferguson, *J. Phys. Chem. B* **2020**, *124*, 3873–3891.
[51] F. Li, J. Han, T. Cao, W. Lam, B. Fan, W. Tang, S. Chen, K. L. Fok, L. Li, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 11259–11264.
[52] C. W. Yap, *J. Comput. Chem.* **2011**, *32*, 1466–1474.
[53] S. Fleming, R. V. Ulijn, *Chem. Soc. Rev.* **2014**, *43*, 8150–8177.
[54] Y. Zhang, H. Gu, Z. Yang, B. Xu, *J. Am. Chem. Soc.* **2003**, *125*, 13680–13681.
[55] V. Jayawarna, M. Ali, T. A. Jowitt, A. F. Miller, A. Saiani, J. E. Gough, R. V. Ulijn, *Adv. Mater.* **2006**, *18*, 611–614.
[56] A. Mahler, M. Reches, M. Rechter, S. Cohen, E. Gazit, *Adv. Mater.* **2006**, *18*, 1365–1370.
[57] A. M. Smith, R. J. Williams, C. Tang, P. Coppo, R. F. Collins, M. L. Turner, A. Saiani, R. V. Ulijn, *Adv. Mater.* **2008**, *20*, 37–41.
[58] H. Xu, A. K. Das, M. Horie, M. S. Shaik, A. M. Smith, Y. Luo, X. Lu, R. Collins, S. Y. Liem, A. Song, P. L. A. Popelier, M. L. Turner, P. Xiao, I. A. Kinloch, R. V. Ulijn, *Nanoscale* **2010**, *2*, 960.
[59] G. Ghosh, K. K. Kartha, G. Fernández, *Chem. Commun.* **2021**, *57*, 1603–1606.
[60] V. Castelletto, G. Cheng, B. W. Greenland, I. W. Hamley, P. J. F. Harris, *Langmuir* **2011**, *27*, 2980–2988.
[61] S. Fleming, P. W. J. M. Frederix, I. Ramos Sasselli, N. T. Hunt, R. V. Ulijn, T. Tuttle, *Langmuir* **2013**, *29*, 9510–9515.
[62] M. Hughes, H. Xu, P. W. J. M. Frederix, A. M. Smith, N. T. Hunt, T. Tuttle, I. A. Kinloch, R. V. Ulijn, *Soft Matter* **2011**, *7*, 10032–10038.
[63] R. J. Williams, A. M. Smith, R. Collins, N. Hodson, A. K. Das, R. V. Ulijn, *Nat. Nanotechnol.* **2009**, *4*, 19–24.
[64] I. R. Sasselli, C. G. Pappas, E. Matthews, T. Wang, N. T. Hunt, R. V. Ulijn, T. Tuttle, *Soft Matter* **2016**, *12*, 8307–8315.
[65] A. Lampel, S. A. McPhee, H. A. Park, G. G. Scott, S. Humagain, D. R. Hekstra, B. Yoo, P. W. J. M. Frederix, T. De Li, R. R. Abzalimov, S. G. Greenbaum, T. Tuttle, C. Hu, C. J. Bettinger, R. V. Ulijn, *Science* **2017**, *356*, 1064–1068.
[66] R. Piotrowska, T. Hesketh, H. Wang, A. R. G. Martin, D. Bowering, C. Zhang, C. T. Hu, S. A. McPhee, T. Wang, Y. Park, P. Singla, T. McGlone, A. Florence, T. Tuttle, R. V. Ulijn, X. Chen, *Nat. Mater.* **2021**, *20*, 403–409.
[67] Y. Park, X. Chen, *J. Mater. Chem. A* **2020**, *8*, 15227–15244.
[68] S. Marchesan, L. Waddington, C. D. Easton, D. A. Winkler, L. Goodall, J. Forsythe, P. G. Hartley, *Nanoscale* **2012**, *4*, 6752–6760.
[69] M. Melchionna, K. E. Styan, S. Marchesan, *Curr. Top. Med. Chem.* **2016**, *16*, 2009–2018.
[70] A. M. Garcia, D. Iglesias, E. Parisi, K. E. Styan, L. J. Waddington, C. Deganutti, R. De Zorzi, M. Grassi, M. Melchionna, A. V. Vargiu, S. Marchesan, *Chem* **2018**, *4*, 1862–1876.
[71] O. Bellotto, S. Kralj, R. D. Zorzi, S. Geremia, S. Marchesan, *Soft Matter* **2020**, *16*, 10151–10157.
[72] O. Bellotto, S. Kralj, M. Melchionna, P. Pengo, M. Kisovec, M. Podobnik, R. De Zorzi, S. Marchesan, *ChemBioChem* **2022**, *23*, e202100518.
[73] E. Parisi, A. M. Garcia, D. Marson, P. Posocco, S. Marchesan, *Gels* **2019**, *5*, 5.
[74] A. Brito, D. Dave, A. Lampel, V. I. B. Castro, D. Kroiss, R. L. Reis, T. Tuttle, R. V. Ulijn, R. A. Pires, I. Pashkuleva, *J. Am. Chem. Soc.* **2021**, *143*, 19703–19710.
[75] K. Baek, A. D. Noblett, P. Ren, L. J. Suggs, *ACS Appl. Bio Mater.* **2019**, *2*, 2812–2821.
[76] J. D. Hartgerink, E. Beniash, S. I. Stupp, *Science* **2001**, *294*, 1684–1688.
[77] R. M. Capito, H. S. Azevedo, Y. S. Velichko, A. Mata, S. I. Stupp, *Science* **2008**, *319*, 1812–1816.
[78] W. A. Kibbe, *Nucleic Acids Res.* **2007**, *35*, W43–W46.
[79] Y. Wu, G. M. Fortunato, B. O. Okesola, F. L. P. D. Brocchetti, R. Suntornnond, J. Connelly, C. D. Maria, J. C. Rodriguez-Cabello, G. Vozzi, W. Wang, A. Mata, *Biofabrication* **2021**, *13*, 035027.
[80] Y. Wu, B. O. Okesola, J. Xu, I. Korotkin, A. Berardo, I. Corridori, F. L. P. di Brocchetti, J. Kanczler, J. Feng, W. Li, Y. Shi, V. Farafonov, Y. Wang, R. F. Thompson, M.-M. Titirici, D. Nerukh, S. Karabasov, R. O. C. Oreffo, J. Carlos Rodriguez-Cabello, G. Vozzi, H. S. Azevedo, N. M. Pugno, W. Wang, A. Mata, *Nat. Commun.* **2020**, *11*, 1182.
[81] Y. Wu, J. Yang, A. van Teijlingen, A. Berardo, I. Corridori, J. Feng, J. Xu, M. Titirici, J. C. Rodriguez-Cabello, N. M. Pugno, J. Sun, W. Wang, T. Tuttle, A. Mata, *Adv. Funct. Mater.* **2022**, *32*, 2205802.
[82] S. Koga, D. S. Williams, A. W. Perriman, S. Mann, *Nat. Chem.* **2011**, *3*, 720–724.
[83] M. Abbas, W. P. Lipiński, J. Wang, E. Spruijt, *Chem. Soc. Rev.* **2021**, *50*, 3690–3705.
[84] W. A. Scott, E. G. Gharakhanian, A. G. Bell, D. Evans, E. Barun, K. N. Houk, T. J. Deming, *J. Am. Chem. Soc.* **2021**, *143*, 18196–18203.

[85] A. Jain, S. Kassem, R. S. Fisher, B. Wang, T.-D. Li, T. Wang, Y. He, S. Elbaum-Garfinkle, R. V. Ulijn, *J. Am. Chem. Soc.* **2022**, *144*, 15002–15007.
[86] D. Priftis, N. Laugel, M. Tirrell, *Langmuir* **2012**, *28*, 15947–15957.
[87] K. Abe, M. Koide, E. Tsuchida, *Polym. J.* **1977**, *9*, 73–78.
[88] S. L. Perry, C. E. Sing, *Macromolecules* **2015**, *48*, 5040–5053.
[89] D. Priftis, M. Tirrell, *Soft Matter* **2012**, *8*, 9396–9405.
[90] L. W. Chang, T. K. Lytle, M. Radhakrishna, J. J. Madinya, J. Vélez, C. E. Sing, S. L. Perry, *Nat. Commun.* **2017**, *8*, 1273.
[91] M. Cai, D. Thorpe, D. H. Adamson, H. C. Schniepp, **2012**, 11.
[92] J. McCafferty, A. D. Griffiths, G. Winter, D. J. Chiswell, *Nature* **1990**, *348*, 552–554.
[93] G. P. Smith, *Science* **1985**, *228*, 1315–1317.
[94] D. Kroiss, J. M. Aramini, S. A. McPhee, T. Tuttle, R. V. Ulijn, *ChemSystemsChem* **2019**, *1*, 7–11.
[95] M. P. Hendricks, K. Sato, L. C. Palmer, S. I. Stupp, *Acc. Chem. Res.* **2017**, *50*, 2440–2448.
[96] A. S. Pina, L. Morgado, K. L. Duncan, S. Carvalho, H. F. Carvalho, A. J. M. Barbosa, B. de P Mariz, I. P. Moreira, D. Kalafatovic, B. M. M. Faustino, V. Narang, T. Wang, C. G. Pappas, I. Ferreira, A. C. A. Roque, R. V. Ulijn, *Chem. Sci.* **2022**, *13*, 210–217.
[97] T. Schnitzer, E. Paenurk, N. Trapp, R. Gershoni-Poranne, H. Wennemers, *J. Am. Chem. Soc.* **2021**, *143*, 644–648.
[98] T. D. Clemons, S. I. Stupp, *Prog. Polym. Sci.* **2020**, *111*, 101310.
[99] T. Aida, E. W. Meijer, *Isr. J. Chem.* **2020**, *60*, 33–47.
[100] T. Aida, E. W. Meijer, S. I. Stupp, *Science* **2012**, *335*, 813–817.
[101] T. Bruckdorfer, O. Marder, F. Albericio, **n.d.**, 15.
[102] S. B. Lawrenson, R. Arav, M. North, *Green Chem.* **2017**, *19*, 1685–1691.
[103] A. J. M. Van Wijk, I. Van Wijk, *3D Printing with Biomaterials: Towards a Sustainable and Circular Economy*, IOS Press, Burke, **2015**.

Manuscript received: December 7, 2022
Accepted manuscript online: February 1, 2023
Version of record online: February 14, 2023